Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Il Futuro della Terra, Cambiamenti Climatici e Sfide Sociali

Ciclo XXXV

Settore Concorsuale: 04/A4 - GEOFISICA

Settore Scientifico Disciplinare: GEO/12 – OCEANOGRAFIA E FISICA DELL'ATMOSFERA

DATA ASSIMILATION FOR ADVANCED CROSS-SCALE UNSTRUCTURED-GRID OCEAN MODELLING

Presentata da

Marco Stefanelli

<table>
<tr><td>**Coordinatore Dottorato**</td><td>**Supervisore**</td></tr>
<tr><td>Silvana Di Sabatino</td><td>Ivan Federico</td></tr>
<tr><td></td><td>**Co-Supervisori**</td></tr>
<tr><td></td><td>Eric Jansen
Ali Aydogdu
Nadia Pinardi</td></tr>
</table>

Esame finale anno 2023

Za Anju...

Academic activities carried out during the 3 years of PhD

I year

The period from November 2019 to April 2020 was dedicated to the PhD lectures and examinations. Then, I started the literature review on (i) unstructured-grid modelling (e.g SHYFEM model (Umgiesser et al., (2004); Federico et al., (2017)), used in the modelling framework of the thesis); (ii) Data Assimilation methodologies (Ash et al. (2016), Carassi et al. (2017), Dobricic & Pinardi (2008)), and (iii) attempts of data assimilation in SHYFEM (Bajo (2020), Ferrarin et al. (2020)).

An investigation on type of Data Assimilation to be developed in the thesis has been performed taking into account the possibility to use variational methods (e.g. 3Dvar, 4Dvar) or statistical (e.g. Kalman Filter, Extended Kalman Filter, Optimal Interpolation). Considering statistical methods, most operational systems in oceanography currently use optimal interpolation schemes (e.g. Dobricic et al., (2005)). However, the use of an optimal interpolation scheme requires a relatively complex implementation of the linearized observational operator, because the Kalman gain contains this operator several times, multiplying the background error covariance matrix. It may therefore become difficult to assimilate observations with complicated observational operators or make the existing observational operators more complex. Another disadvantage of optimal interpolation is that the solution is always local and discontinuities might appear in the analysis due to the presence of nearby observations (e.g. Gauthier et al., (1999)) or discontinuous regional EOFs (Bellucci et al., (2008)). The negative impact of the local operator on the analysis field may be attenuated by the usage of overlapping observations and EOFs. However, in this way the problem is

alleviated by arbitrary assumptions involving an additional computational cost. Furthermore, the computational cost of an optimal interpolation scheme is approximately proportional to the number of observations that are assimilated (Dobricic & Pinardi 2008)). Based on this analysis, we have decided to move in the direction to use 3D-Var Data Assimilation because the computational cost of that scheme mainly depends on the size of the model state vector and much less on the number observations. Therefore, the application of the 3D-Var may facilitate the use of a relatively large number of observations. Furthermore, 3D-Var represents the assimilation scheme currently applied in the parent model of SHYFEM (MFS system in CMEMS based on NEMO model which use OceanVar (Dobricic & Pinardi (2008)). (3D-Var scheme developed at CMCC; it will be used for the thesis work)

In order to decide the geographical domain of interest, we have studied and mapped the in-situ CMEMS products (ARGO, Glider, XBT, Tide Gauges, etc.) in the Italian Mediterranean Sea in the period 2017-2019. This task is made building a Python tool which allows to search measurements in different worldwide areas and periods.

Looking at the spatial distribution of the observations we have selected the South Adriatic Northern Ionian Sea (SANI) basin where we are already implemented a modelling system based on unstructured grid approach [Federico et al., 2017]. To validate the simulation performed with SHYFEM model I have adopted a Python tool capable to select the ARGO profile and compare with model results. This tool has been extended to read unstructured grid outputs.

OcenaVar currently is implemented and used with regular mesh oceanographic model (NEMO).

II year

During the II year the main activities were the adaptation of the OceanVar in order to make it able to work with SHYFEM oceanographic model which is implemented using unstructured mesh.

The 3DVar scheme implemented in OceanVar compute the minimum of a cost function built starting from model and observations information. The cost function mainly depends on vertical and horizontal covariance matrix. The horizontal covariance matrix is implemented using Recursive Filter, instead vertical covariance is implemented using Empirical Orthogonal Function (EOF).

In order to compute EOF I performed a 4 years (2017-2020) simulation using SHYFEM and I adapted a Fortran tool writing different subroutines to read model output and grid information. Furthermore, I built a tool using bash script and CDO libraries to perform SST validation. For the results visualization I wrote Python scripts.

After the task about EOF, in order to adapt OceanVar to read model output and grid information. I wrote different Fortran subroutines and merged them in OceanVar software.

Finally, the first experiment using SHYFEM and OceanVar on the SANI (South Adriatic Northen Ionian) domain has been implemented without Recursive Filter assimilating Argo observations.

III year

Following the work of the previous year, during the III year the main activities were about the adaptation of OceanVar in the recursive filter (RF) routine and experiment performing.

RF is formulated and used on regular mesh to model the horizontal component of the background error covariance matrix. Considering that the data assimilation application in this project use unstructured mesh, a novel first order RF algorithm has been developed. The main problems to be solved in this work were about the geometrical structure of the unstructured mesh. Indeed, the advantages of applying RF on regular grids is in the intrinsic ordering of the nodes and a strong symmetry feature inherited by the RF formulation on an infinite line. Both the problems have been solved ordering the triangles edges respect longitude and latitude and developing an algorithm that emulates forward and backward pass of RF on regular grid. Test on regular grid shows that the novel algorithm is equivalent to the regular one. Testing the new algorithm on unstructured mesh shows that it produces a gaussian shape spreading in areas with irregular grid tessellation, like in coastal regions, and it can be used to model the horizontal component of B matrix on unstructured mesh.

Once the assimilation system OceanVar-SHYFEM has been fully adapted for unstructured mesh 2 different experiments have been performed:

1. Assimilating T and S from Argo floats using 25 bivariate EOF.
2. Assimilating T, S from Argo floats and SLA from satellite observations using 25 trivariate EOF.

Despite the lack of observations on SANI domain, the experiments show that we are able to improve the model state using the assimilation scheme implemented. Activities related to the abroad period have been performed during this year.

In the period September – November 2022 I visited the Imperial College in London in the Data Learning group of prof. Rossella Arcucci. Here the main goal was to apply Machine Learning techniques in order to train a LSTM (Long-Short-Term-Memory) model that is able to forecast the analysis increments. Here the idea is to correct model forecast using predicted increments, since the observations are not available in the future and data assimilation can't be applied.

Evaluating the LSTM model skills is clear that such a model can generate a good prediction of analysis increments.

Finally, during this last year I presented my work to: EGU general assembly, Interreg Italy-Croatia AdriaClim workshop, University of Reading and Imperial College London. And participate to ECMWF-ESA workshop on Machine learning for Earth Observation and Prediction.

List of talks

1. EGU general assembly
2. Interreg Italy-Croatia AdriaClim workshop
3. University of Reading
4. Imperial College London

List of supplementary training activities

1. NASA's Applied Remote Sensing Training Program: Remote Sensing of Coastal Ecosystems.
2. Participation at ECMWF-ESA Workshop on Machine Learning for Earth System Observation and Prediction (I Edition).
3. NASA's Applied Remote Sensing Training Program: MODIS to VIIRS Transition for Air Quality Applications
4. NASA's Applied Remote Sensing Training Program: Introduction to NASA's "Black Marble" Night Lights Data
5. ECMWF virtual training course: Data Assimilation
6. ECMWF/Ocean Predict workshop: Advances in Ocean Data Assimilation.
7. Participation at ECMWF-ESA workshop on Machine learning for Earth Observation and Prediction (III Edition)

Abstract

The coastal ocean is a complex environment with extremely dynamic processes that require a high-resolution and cross-scale modeling approach in which all hydrodynamic fields and scales are considered integral parts of the overall system. In the last decade, unstructured-grid models have been used to advance in seamless modeling between scales. On the other hand, the data assimilation methodologies to improve the unstructured-grid models in the coastal seas have been developed only recently and need significant advancements. Here, we link the unstructured-grid ocean modeling to the variational data assimilation methods. In particular, we show results from the modeling system SANIFS based on SHYFEM fully-baroclinic unstructured-grid model interfaced with OceanVar, a state-of-art variational data assimilation scheme adopted for several systems based on a structured grid. OceanVar implements a 3DVar DA scheme. The combination of three linear operators models the background error covariance matrix. The vertical part is represented using multivariate EOFs for temperature, salinity, and sea level anomaly. The horizontal part is assumed to be Gaussian isotropic and is modeled using a first-order recursive filter algorithm designed for structured and regular grids. Here we introduced a novel recursive filter algorithm for unstructured grids. A local hydrostatic adjustment scheme models the rapidly evolving part of the background error covariance. We designed two data assimilation experiments using SANIFS implementation interfaced with OceanVar over the period 2017-2018, one with only temperature and salinity assimilation by Argo profiles and the second also including sea level anomaly. The results showed a successful implementation of the approach and the added value of the assimilation for the active tracer fields. While looking at the broad basin, no significant improvements are highlighted for the sea level, requiring future investigations. Furthermore, a Machine Learning methodology based on an LSTM network has been used to predict the model SST increments.

Contents

1. Introduction

Eight of the world's top ten most populated cities are located by the coast. Around 10% of the world's population live in coastal areas less than 10 m above sea level. Since coastal areas are closely connected to human activities, improving the coastal ocean representation is crucial to understanding the present and near-future ocean state and predicting its evolution under climate change conditions. In the climate change framework, catastrophic events are increasing in frequency and intensity causing enormous damage along the coastal regions. So, developing a good oceanographic forecasting model integrated with a DA system and coupled with a Numerical Weather Prediction (NWP) model, which can predict efficiently and carefully small-scale events like storm surges and medicane, can give us the possibility to face them and reduce the damage.

Since the presence of complex coastlines, interaction with inland waters, rapid changes in topography, and high space-time variability of the phenomena involved, modeling the coastal ocean is a complex task. During recent decades, finite element models along with unstructured mesh (FESOM (Danilov, Kivman, and Schröter (2004) and Wang, Danilov, and Schröter (2008)), FLUIDITY (Ford et al. (2004) and Piggott et al. (2008)), SCHISM (Zhang and Baptista (2008)), FVCOM (Chen, Liu, and Beardsley (2003)) and SHYFEM (Umgiesser et al. (2004)) to name a few) has been used to represent the complex features of the coastlines (Chiggiato et al. (2012), Zhu et al. (2017), Federico et al. (2017), Aydoğdu et al. (2018), Bajo (2020), Shi, Tang, and Myers (2020), Gunduz, Özsoy, and Hordoir (2020), and Ferrarin, Bajo, and Umgiesser (2021)). In the recent years, models that use an unstructured-grid, like FESOM2 (Danilov et al. (2017)) and MPAS (Petersen et al. (2019)), have been employed to simulate the global ocean.

The coastal model generally is a cross-scale representation of the interested area of study describing regional and local dynamics. The model is nested in a large-scale parent model to have information from large-scale dynamics. So, the coastal model connects large-scale and coastal dynamics through a dynamic downscale approach.

Using finite element models allows us to describe small ocean scale; however, it does not resolve the problem of the deterministic chaos introduced by Lorenz (1963). Indeed, it is well known that geophysical systems like the ocean and atmosphere are highly sensitive to the initial conditions and express chaotic behavior. That feature makes the system prediction using numerical models challenging. In fact, supposing that a numerical model

is initialized by an observational dataset, reanalysis data, or the output of another numerical model, the predicted integrated trajectory gradually will move far from the "true state" of the system. This is a consequence that the numerical model equations are solved in a discrete space, and any dataset used to initialize a numerical model is imperfect and has errors. A good product that can be used to avoid the drifting of the model from the true state are observations. Using them to correct the model integration and produce the **best estimation** of the system's state would be useful.

In the last three decades, much effort has been made to build a mathematical theory, named **data assimilation (DA)**, that acts as a bridge between **numerical model** and **observations**. DA combines a model state with observations to estimate a state and its uncertainty better. This process and the resulting state is called **analysis**. The analysis state can be used to initialize a forecast or produce re-analyses over long periods. Alternatively, DA can be used to estimate the model parameters or infer the best characterization of the model forcing or controls.

Using the words in Fletcher (2017), we can say that the main goal of a DA method is to:

optimize the strengths of the models and observations while simultaneously minimizing their weaknesses

Many different DA methods and algorithms enable the bridging between model and observations, which can be grouped into two main branches: **statistical** and **variational** DA methods. These two approaches can also be combined, and the resulting methodology is named **hybrid** DA (Fig. 1.1). In both statistical and variational methods, we seek an optimal solution. Statistically, we will seek a solution with minimum variance, whereas variationally, we will seek a solution that minimizes a suitable cost (or error) function. The two approaches are identical and provide the same solution in the exceptional cases of assumption of Gaussian probability distribution. However, the statistical approach, though often more complex and time-consuming, can provide a richer information structure: an average solution and some characteristics of its variability (probability distribution) (Asch, Bocquet, and Nodet, 2016). Given the nature of the modeling and observation infrastructure in geosciences, DA is conveniently formalized as a discrete-model/discrete-observation estimation problem (Carrassi et al. (2018)). Considering the problems' dimensionality, DA can be identified as a big data problem. Indeed, if n is the dimensionality of the physical problem, the matrices involved have dimensionality of $\mathcal{O}^{n^2}$. At Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC), which is also the institution where this Ph.D. work has been done, the Mediterranean Forecasting System (MedFS) is used. MedFS is a numerical ocean prediction system that produces analyses, reanalyses, and short-term forecasts for the Mediterranean Sea and its adjacent Atlantic ocean areas. Since 2015, MedFS is part of the Copernicus Marine Service. It has been developed and operationally maintained by CMCC since 2018 providing reg-
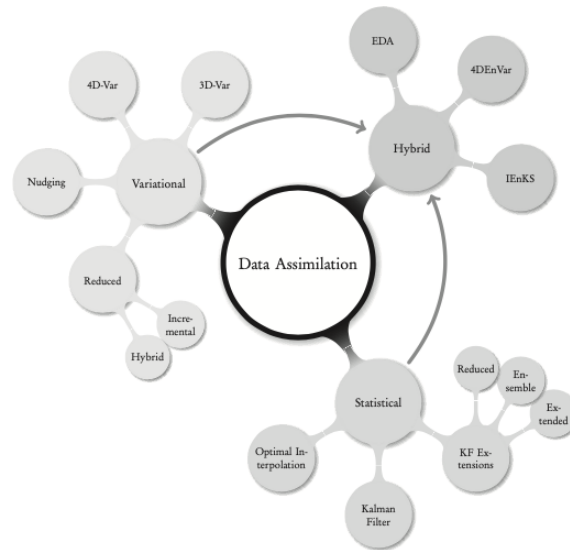
Figure 1.1: DA methods and algorithms. From: Asch, Bocquet, and Nodet, 2016.

ular and systematic information about the physical state of the Mediterranean Sea (for more information, see www.medfs.cmcc.it). The system uses a 3DVar assimilation scheme named **OceanVar** (Dobricic and Pinardi (2008) and Storto, Masina, and Navarra (2016)). The assimilated data includes: along-track satellite Sea Level Anomaly (SLA), in situ temperature profiles by VOS XBTs (Voluntary Observing Ship-eXpandable Bathyther-mograph), in situ temperature and salinity profiles by ARGO floats and from different survey CTD profiles. Satellite objectively analyzed Sea Surface Temperature (SST) is used to correct surface heat fluxes. All the assimilated observations are provided by the Copernicus Marine Service (*marine.copernicus.org*). Also, to model cross-scale ocean, the unstructured-grid finite-element three-dimensional hydrodynamic model **SHYFEM** (Umgiesser et al. (2004) and Micaletto et al. (2021)) is used at CMCC. An application of the SHYFEM model is **SANIFS** (Southern Adriatic Northern Ionian coastal Forecasting System, Federico et al. (2017)). SANIFS is a coastal-ocean operational system providing short-term forecasts. The operational chain is based on a downscaling approach starting from the large-scale system, MedFS, which provides the open-sea fields. Thanks to the high and proper horizontal resolution, ranging from 3-4 km in the open sea to 50-500 m in coastal areas (Fig. 1.2, www.sanifs.cmcc.it), the model configuration has been outlined to provide reliable hydrodynamics and active tracer forecasts in mesoscale-shelf-coastal waters of Southern Eastern Italy (Apulia, Basilicata, and Calabria regions). Since MedFS includes an assimilation cycle, SANIFS inherits this correction at the open boundaries. During the last few years, machine learning and neural network approaches have been applied in combination with DA methods to allow non-linearity in DA problems. For example, in Amendola et al. (2020) and Peyron et al. (2021) is presented a formulation of a new methodology that combines machine learning and DA called **Latent Space Assimilation**. It consists in performing the Optimal Interpolated Kalman Filter in the

latent space obtained by a Convolutional AutoEncoder with non-linear encoder functions and non-linear decoder functions. In the latent space, the dynamic system is represented by a surrogate model built by a Long-Short-Term-Memory (LSTM) network to train a function that emulates the dynamic system in the latent space.

This thesis aims to interface SHYFEM (using SANIFS configuration) with OceanVar to introduce DA on unstructured grid nested models. The approach will be cross-scale. Temperature and salinity from Argo profiling floats and sea level anomaly from satellite altimetry observations will be assimilated in the open sea. The model project the analysis in space reaching coastal regions and forward in time to the next assimilation step. No coastal observations (e.g., tide gauge) will be assimilated.

Applying DA techniques in models with an unstructured-grid is a new branch in oceanography. Mainly, in ocean modeling, statistical methods are used (e.g., nudging, Kaman filter or Ensemble Kalman filter Zhu et al. (2017), Aydoğdu et al. (2018), Bajo (2020), Shi, Tang, and Myers (2020), and Ferrarin, Bajo, and Umgiesser (2021)) and to our knowledge, variational DA has never been applied in combination with an unstructured-grid finite element ocean model.

This thesis is organized as follows. Chapters 2 and 3 will introduce the 3DVar mathematical formulation and a novel first-order recursive filter algorithm developed to interface the OceanVar DA scheme with the ocean model. Next, in Chapter 4, we will describe the modeling setting, the observation datasets used in the assimilation process, and the experiment design. Chapter 5 will discuss experiment results and the added value of introducing a 3DVar scheme in coastal modeling. Chapter 6 presents an application of the LSTM network for sea surface temperature increment prediction. Finally, in Chapter 7, we will see the conclusions and future perspectives.
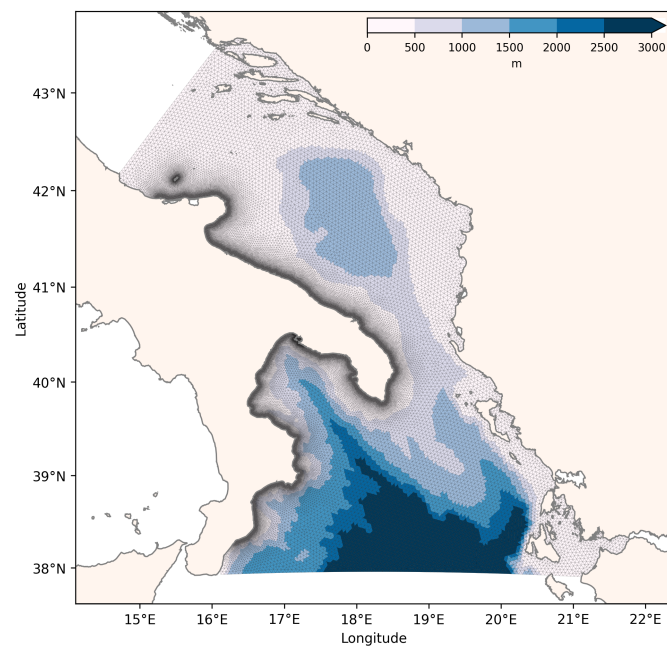
Figure 1.2: SANIFS numerical domain, grid, and bathymetry.
Coastal resolution: from 50m to 500m.
Open sea resolution: from 3Km to 4Km.
Number of vertical levels: 92

2. A 3DVar scheme for ocean data assimilation

Variational DA methods have been used in operational weather prediction for quite some time. National Center for Environmental Prediction (NCEP) was the first to go operational with a 3DVar system (Derber (1992)), followed not too far behind by the European Center for Medium-range Weather Forecasting (ECMWF) using the formulation described in Heckley (1992) and the United Kingdom's Meteorological Office (Lorenc et al. (2000)), along with most of the operational centers for global numerical weather prediction. The reason for introducing 3DVar was that DA schemes could fully model three-dimensional covariances for the errors of the background field globally but also use nonlinear observation operators, especially those associated with satellites. This was a problem for the optimum interpolation-based methods, as they rely on the linear observation operator.

The introduction of the time component in the analysis scheme enables the assimilation of observations in a window, which is the length of time between two consecutive analysis times. Schemes that do not include the time dimension are referred to as **filters**, while those that do contain the time components are **smoothers**. Filter-based methods usually only use observations available at the analysis time. As such, the data assimilation scheme filters the observations and the dynamical scales to find the best estimate of the geophysical state at that time that fits the observations. The observations here are all those in a specific assimilation window. The assimilation window can be (in most NWP centers) 6h,12h, or 24h. Once all the observations are collected, they are all referred to at the same time: the analysis time. However, in circumstances where the system has a fast dynamic (e.g., tide assimilation) or the observations has a wide time distribution, considering all the observation valid at analysis time can lead to an incorrect analysis. To overcome this problem, the first guess at the appropriate time (FGAT), a method for obtaining the instantaneous background field at the correct time, can be combined with 3DVar.

This chapter will first introduce the mathematical formulation of the 3DVar scheme used in this work. Next, we will describe the formulation of the three linear operators used to model the different components of the background error covariance matrix:

1. empirical orthogonal functions (EOFs) using singular value decomposition (SVD) is used to model the vertical component;

2. a first-order recursive filter is used to model the horizontal component;

3. a formulation of the dynamic height operator, which calculates the sea surface height errors covariance from temperature and salinity fields

2.1 Mathematical Formulation

We discussed that one of the aims of DA is to introduce the observations in a forecasting system to reduce the initialization errors and eventually improve the forecast. So, considering the forecast as our prediction of the system's state that presents errors and is a chaotic deterministic system, the observations cover the role of priori information to constrain the problem. In this framework, the strategy to tackle the problem is to use the **Bayesan** probability approach. The following general formulation can be applied to any geophysical or not system.

- Let the **state vector** of the system be $\mathbf{x}$ where $\mathbf{x} = (x_1, x_2, \ldots, x_N)^{\mathrm{T}}$ and N is total number of state variable. $\mathbf{x}$ defines the physical/model space;

- Let the **observational vector** $\mathbf{y}$ where $\mathbf{y} = (y_1, y_2, \ldots, y_{N_0})^{\mathrm{T}}$ and N_0 is total number of observations. $\mathbf{y}$ defines the observational space;

- $N_0 < N$.

Now, considering that observations are an evaluation of the model state, a relationship between $\mathbf{x}$ and $\mathbf{y}$ should exist. This relationship is given by

$$\mathbf{y} = \mathcal{H}(\mathbf{x}) \tag{2.1}$$

where $\mathcal{H}(\mathbf{x})$ is a non linear operator, named **observational operator**, from the physical space to the observational space given by

$$\mathcal{H}(\mathbf{x}) = \begin{pmatrix} h_1(x_1, x_2, \ldots, x_N) \\ h_2(x_1, x_2, \ldots, x_N) \\ \vdots \\ h_{N_0}(x_1, x_2, \ldots, x_N) \end{pmatrix} \tag{2.2}$$

The state estimation problem is a discrete inverse problem (Menke, 2018), which aim is to find the best $\mathbf{x}$ that inverts Eq. (2.1) for a given $\mathbf{y}_\mathrm{o}$, where $\mathbf{y}_\mathrm{o}$ is a physical observation which contains errors and the best value of $\mathbf{x}$ is the **analysis**, $\mathbf{x}_\mathrm{a}$. The method to set up this problem is considering a Bayesian probability approach and using Bayes' theorem.

BAYES' THEOREM

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)} \qquad P(B) \neq 0 \tag{2.3}$$

where A and B are events.

- $P(A \mid B)$ is a conditional probability: the probability of event A occurring given that B is true. It is also called the posterior probability of A given B.

- $P(B \mid A)$ is also a conditional probability: the probability of event B occurring given that A is true. It can also be interpreted as the likelihood of A given a fixed B.

- $P(A)$ and $P(B)$ are the probabilities of observing A and B without any given conditions, known as the marginal or prior probability.

In our case, we have that:

- A is the event that $\mathbf{x} = \mathbf{x}_{\mathrm{t}}$, where the subscript stands for "true".

- B is the event that $\mathbf{y} = \mathbf{y}_{\mathrm{o}}$, where the subscript stands for observed value.

Bayes' theorem is:

$$P(\mathbf{x} = \mathbf{x}_{\mathrm{t}} \mid \mathbf{y} = \mathbf{y}_{\mathrm{o}}) \propto P(\mathbf{y} = \mathbf{y}_{\mathrm{o}} \mid \mathbf{x} = \mathbf{x}_{\mathrm{t}})P(\mathbf{x} = \mathbf{x}_{\mathrm{t}}) \tag{2.4}$$

Thus Eq. (2.4) defines an N-dimensional probability density function (PDF). Let's now evaluate the various terms of Eq. (2.4) starting from the last term on the right-hand side.

- $P(\mathbf{x} = \mathbf{x}_{\mathrm{t}})$

As we know, the true state, $\mathbf{x}_{\mathrm{t}}$, is not known. However, we have an estimation of it through the background state, $\mathbf{x}_{\mathrm{b}}$ which deviates from $\mathbf{x}_{\mathrm{t}}$ by a quantity ε_{b}

$$\varepsilon_{\mathrm{b}} \equiv \mathbf{x}_{\mathrm{t}} - \mathbf{x}_{\mathrm{b}} = \mathbf{x} - \mathbf{x}_{\mathrm{b}} \tag{2.5}$$

So, the prior probability $P(\mathbf{x} = \mathbf{x}_{\mathrm{t}})$ can be expressed in terms of the error of the background state with respect to the true state.

$$P(\mathbf{x} = \mathbf{x}_{\mathrm{t}}) = P_{\mathrm{b}}(\mathbf{x} - \mathbf{x}_{\mathrm{b}}) = P_{\mathrm{b}}(\varepsilon_{\mathrm{b}}) \tag{2.6}$$

$\mathbf{x}_b$ is propagated from time-step T-1 to T through the model operator, $\mathcal{M}_{\mathrm{T-1,T}}$

$$\mathbf{x}_{\mathrm{b_T}} = \mathcal{M}_{\mathrm{T-1,T}}(\mathbf{x}_{\mathrm{b_{T-1}}}) \tag{2.7}$$

and $\mathcal{M}_{\mathrm{T-1,T}}$ is a non-linear operator.

The first term on the right-hand side of Eq. (2.4) is

- $P(\mathbf{y} = \mathbf{y}_{\mathrm{o}} \mid \mathbf{x} = \mathbf{x}_{\mathrm{t}})$

Here we can apply the same reasoning used before and Eq. (2.1) on the conditional part. So, if we know that the true values of $\mathbf{y}$ is given by $\mathbf{y}_{\mathrm{t}}$, Eq. (2.1) becomes:

$$\mathbf{y}_{\mathrm{t}} = \mathcal{H}(\mathbf{x}_{\mathrm{t}}) \tag{2.8}$$

If ε_{o} is the deviation of $\mathbf{y}_{\mathrm{o}}$ from $\mathbf{y}_{\mathrm{t}}$ than

$$
\begin{aligned}
P(\mathbf{y} = \mathbf{y}_o \mid \mathbf{x} = \mathbf{x}_{\mathrm{t}}) &\equiv P(\mathbf{y} = \mathbf{y}_{\mathrm{o}} \mid \mathbf{y} = \mathbf{y}_{\mathrm{t}}) \\
&= P_{\mathrm{o}}(\mathbf{y}_{\mathrm{o}} - \mathcal{H}(\mathbf{x})) \\
&= P_{\mathrm{o}}(\varepsilon_{\mathrm{o}})
\end{aligned}
\tag{2.9}
$$

Eq. (2.9) can be read as: the probability to obtain $\mathbf{y} = \mathbf{y}_{\mathrm{o}}$ knowing that $\mathbf{y} = \mathbf{y}_{\mathrm{t}}$ is equivalent to the probability of the deviation of $\mathbf{y}_{\mathrm{o}}$ from $\mathbf{y}_{\mathrm{t}}$.
ε_{o} is the observational error. It takes into account two sources of errors: instrumental or measurement error, ε_{I}, and the representativeness error, ε_{r}

$$\varepsilon_{\mathrm{o}} = \varepsilon_{\mathrm{I}} + \varepsilon_{\mathrm{r}} \tag{2.10}$$

ε_{r} represents the error introduced by the model in representing phenomena at a scale smaller than the grid resolution. So the representativeness error arises due to a mismatch between the scales represented in the observations and the model fields. For instance, an observation may represent the value of a geophysical variable at a single point in space and time. In contrast, the model will represent a spatial and temporal average, depending on the model's discretization. The observation and the prior will then differ, depending on the true geophysical variability at scales different from those represented by the model (Janjić et al., 2018).

The term on the left-hand side of Eq. (2.4) is

- $P(\mathbf{x} = \mathbf{x}_{\mathrm{t}} \mid \mathbf{y} = \mathbf{y}_{\mathrm{o}})$

This is the probability to obtain $\mathbf{x}$ knowing the observed values $\mathbf{y}$, which is the analysis probability that we denote with $P_{\mathrm{a}}(\mathbf{x})$. Lorenc (1986) asserts that the best estimate of $\mathbf{x}$, is either the mean or the mode of $P_{\mathrm{a}}(\mathbf{x})$ and these two quantities are the minimum variance and the maximum likelihood states respectively.
So, neglecting the denominator of Bayes' theorem, which can be seen as a normalization factor, Eq. (2.4) became

$$P_{\mathrm{a}}(\mathbf{x}) = P_{\mathrm{o}}(\mathbf{y}_{\mathrm{o}} - \mathcal{H}(\mathbf{x}))P_{\mathrm{b}}(\mathbf{x} - \mathbf{x}_{\mathrm{b}}) = P_{\mathrm{o}}(\varepsilon_{\mathrm{o}})P_{\mathrm{b}}(\varepsilon_{\mathrm{b}}) \tag{2.11}$$

Now we assume that these PDFs are multivariate Gaussian, MG, such that $P_{\mathrm{o}}(\varepsilon_{\mathrm{o}}) = MG(\mathbf{0}, \mathbf{R})$ and $P_{\mathrm{b}}(\varepsilon_{\mathrm{b}}) = MG(\mathbf{0}, \mathbf{B})$, where both have $\mathbf{0}$ mean and covariance matrix $\mathbf{R}$ for observational error and $\mathbf{B}$ for the background error. A common assumption is that observational errors are uncorrelated, which implies that the observational error covariance matrix, $\mathbf{R}$, is diagonal. This assumption could not be true for satellite observations. Indeed, observations from different channels on a sensor could be correlated. Another important assumption that we did implicitly, and that simplifies the mathematical description, as well as the operational work, is that ε_{b} and ε_{o} are unbiased

$$\langle \varepsilon \rangle_{\mathrm{b}} = 0 \tag{2.12}$$

$$\langle \varepsilon \rangle_{\mathrm{o}} = 0 \tag{2.13}$$

The two multivariate Gaussian distributions are then defined as

$$P_{\mathrm{b}}(\varepsilon_{\mathrm{b}}) \propto exp\left\{ -\frac{1}{2}\varepsilon_{\mathrm{o}}^{\mathrm{T}}\mathbf{B}^{-1}\varepsilon_{\mathrm{o}} \right\} \equiv exp\left\{ -\frac{1}{2}\Big(\mathbf{x} - \mathbf{x}_{\mathrm{b}}\Big)^{\mathrm{T}}\mathbf{B}^{-1}\Big(\mathbf{x} - \mathbf{x}_{\mathrm{b}}\Big) \right\} \tag{2.14}$$

$$P_{o}(\varepsilon_{\mathrm{o}}) \propto exp\left\{ -\frac{1}{2}\varepsilon_{\mathrm{b}}^{\mathrm{T}}\mathbf{B}^{-1}\varepsilon_{\mathrm{b}} \right\} \equiv exp\left\{ -\frac{1}{2}\Big(\mathbf{y}_{\mathrm{o}} - \mathcal{H}(\mathbf{x})\Big)^{\mathrm{T}}\mathbf{R}^{-1}\Big(\mathbf{y}_{\mathrm{o}} - \mathcal{H}(\mathbf{x})\Big) \right\} \tag{2.15}$$

Therefore substituting Eq. (2.14) and Eq. (2.15) into Eq. (2.11) yelds

$$P_{a}(\mathbf{x}) \propto exp\left\{ -\frac{1}{2}\Big(\mathbf{x} - \mathbf{x}_{\mathrm{b}}\Big)^{\mathrm{T}}\mathbf{B}^{-1}\Big(\mathbf{x} - \mathbf{x}_{\mathrm{b}}\Big) - \frac{1}{2}\Big(\mathbf{y}_{\mathrm{o}} - \mathcal{H}(\mathbf{x})\Big)^{\mathrm{T}}\mathbf{R}^{-1}\Big(\mathbf{y}_{\mathrm{o}} - \mathcal{H}(\mathbf{x})\Big) \right\} \tag{2.16}$$

Applying $-ln$ to Eq. (2.16) yields the nonlinear cost function that characterize the 3DVar scheme

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}\Big(\mathbf{x} - \mathbf{x}_{\mathrm{b}}\Big)^{\mathrm{T}}\mathbf{B}^{-1}\Big(\mathbf{x} - \mathbf{x}_{\mathrm{b}}\Big) + \frac{1}{2}\Big(\mathbf{y}_{\mathrm{o}} - \mathcal{H}(\mathbf{x})\Big)^{\mathrm{T}}\mathbf{R}^{-1}\Big(\mathbf{y}_{\mathrm{o}} - \mathcal{H}(\mathbf{x})\Big) \tag{2.17}$$

The first term in cost function $\mathcal{J}(\mathbf{x})$ measures the difference between the initial model state and our a priori expectation of this state in the model space. The second term encodes the difference between the observations and the model states in the observational space.

In the time-dependent formulation of Eq. (2.17) (4DVar), $\mathcal{M}$ enter in the equation through

Eq. (2.7). However, in the 3DVar formulation, Eq. (2.17) is not time-dependent and $\mathcal{M}$ will not enter in the minimization process.

To maximize $P_a(\mathbf{x})$ is equivalent of minimizing $\mathcal{J}(\mathbf{x})$. However, the non-linearity of $\mathcal{H}$ complicates this computation. Indeed, to minimize $\mathcal{J}(\mathbf{x})$, a computer algorithm needs to compute also $\nabla \mathcal{J}(\mathbf{x})$ to know in which direction to proceed to minimize it. Computing the gradient with non-linear $\mathcal{H}(x)$ introduce in $\nabla \mathcal{J}(\mathbf{x})$ a term which is the gradient of $\mathcal{H}(x)$. So, it is not possible to write a general formulation of $\mathcal{J}(\mathbf{x})$ since $\nabla \mathcal{H}(x)$ depends on the problem considered. To tackle this problem, we can approximately linearize these operators around the background state, $\mathbf{x}_b$, by formulating the problem in terms of perturbations.

$$\delta \mathbf{x} := \mathbf{x} - \mathbf{x}_b$$
$$\mathcal{H}(\mathbf{x}) \approx \mathcal{H}(\mathbf{x}_b) + \mathbf{H}\delta \mathbf{x} \tag{2.18}$$

Substituting Eq. (2.18) into Eq. (2.17), the cost function in the **increment formulation** is

$$\begin{aligned} \mathcal{J}(\delta \mathbf{x}) &= \frac{1}{2}(\delta \mathbf{x})^{\mathrm{T}}\mathbf{B}^{-1}(\delta \mathbf{x}) + \frac{1}{2}\Big(\mathbf{y}_o - \mathcal{H}(\mathbf{x}_b) - \mathbf{H}\delta \mathbf{x}\Big)^{\mathrm{T}}\mathbf{R}^{-1}\Big(\mathbf{y}_o - \mathcal{H}(\mathbf{x}_b) - \mathbf{H}\delta \mathbf{x}\Big) \\ &= \frac{1}{2}(\delta \mathbf{x})^{\mathrm{T}}\mathbf{B}^{-1}(\delta \mathbf{x}) + \frac{1}{2}\Big(\mathbf{d} - \mathbf{H}\delta \mathbf{x}\Big)^{\mathrm{T}}\mathbf{R}^{-1}\Big(\mathbf{d} - \mathbf{H}\delta \mathbf{x}\Big) \end{aligned} \tag{2.19}$$

where $\delta \mathbf{x}$ is called **increment**, $\mathbf{H}$ is the **linearized observational operator** around the background state (mathematically speaking it is the $\mathcal{H}$'s Jacobian evaluated at $\mathbf{x} = \mathbf{x}_b$) and $\mathbf{d}$ is the **misfit or innovation**

$$\mathbf{d} = \mathbf{y}_o - \mathcal{H}(\mathbf{x}_b) \tag{2.20}$$

Now the problem becomes to find the increment, $\delta \mathbf{x}_a$, such that the cost function in Eq. (2.19) is minimized.

The minimum of the cost function $\mathcal{J}(\delta \mathbf{x})$ on the increment space may be justified by posing $\nabla \mathcal{J}(\delta \mathbf{x}) = 0$. Then we obtain the following preconditioned linear system.

$$(\mathbf{I} + \mathbf{B}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})\delta \mathbf{x} = \mathbf{B}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{d} \tag{2.21}$$

To solve the linear equation system Eq. (2.21), iterative methods able to converge toward a practical solution are needed. At this point, we have to face two problems:

1. The system Eq. (2.21) is highly ill conditioned (Haben, Lawless, and Nichols, 2011).

2. The matrix $\mathbf{B}$ has dimensionality of $\mathcal{O}^{14}$ or higher. This implies that solving Eq. (2.21) has a high computational cost.

It is convenient to write Eq. (2.19) in the **control space** to tackle these problems. The

basic assumption is that $\mathbf{B}$ is Gaussian symmetric, block diagonal, and positive definite matrix. So, there exists a Cholesky decomposition such that

$$\mathbf{B} = \mathbf{V}\mathbf{V}^{\mathrm{T}} \tag{2.22}$$

and Eq. (2.19) becames

$$\mathcal{J}(\delta\mathbf{x}) = \frac{1}{2}(\delta\mathbf{x})^{\mathrm{T}}\left(\mathbf{V}\mathbf{V}^{\mathrm{T}}\right)^{-1}(\delta\mathbf{x}) + \frac{1}{2}\left(\mathbf{d} - \mathbf{H}\delta\mathbf{x}\right)^{\mathrm{T}}\mathbf{R}^{-1}\left(\mathbf{d} - \mathbf{H}\delta\mathbf{x}\right) \tag{2.23}$$

Introducing the control vector

$$\mathbf{v} = \mathbf{V}^{-1}\delta\mathbf{x} \tag{2.24}$$

finally, we have the cost function defined on the control space

$$\mathcal{J}(\delta\mathbf{v}) = \frac{1}{2}\mathbf{v}\mathbf{v}^{\mathrm{T}} + \frac{1}{2}\left(\mathbf{d} - \mathbf{H}\mathbf{V}\mathbf{v}\right)^{\mathrm{T}}\mathbf{R}^{-1}\left(\mathbf{d} - \mathbf{H}\mathbf{V}\mathbf{v}\right) \tag{2.25}$$

Eq. (2.25) is said the **dual problem** of Eq. (2.19). Eq. (2.25) is used in the OceanVar minimization cycle. Once the convergence criteria are satisfied, the result is written in the increment space inverting Eq. (2.24)

$$\delta\mathbf{x} = \mathbf{V}\mathbf{v} \tag{2.26}$$

Due to its large size, in OceanVar, the transformation matrix $\mathbf{V}$ is modeled at each minimization iteration as a sequence of linear operators

$$\mathbf{V} = \mathbf{V}_{\eta}\mathbf{V}_{\mathrm{H}}\mathbf{V}_{\mathrm{V}} \tag{2.27}$$

In Eq. (2.27), the linear operator $\mathbf{V}_{\mathrm{V}}$ transforms coefficients which multiply vertical EOFs into vertical profiles of temperature and salinity defined at the model vertical levels, $\mathbf{V}_{\mathrm{H}}$ applies horizontal covariances on fields of temperature, salinity, and SLA and it is modeled using a first-order recursive filter algorithm, finally $\mathbf{V}_{\eta}$ is the dynamic height operator and calculates the sea surface height error covariance from three-dimensional fields of temperature and salinity.

Most of the computational time necessary to minimize the 3DVar cost function is spent in transforming increments from the control space to the physical space by the linear operator $\mathbf{V}$ and its transpose. Therefore, the computational cost of the 3DVar is approximately proportional to the size of the control vector $\mathbf{v}$ (Dobricic and Pinardi, 2008).

The following three sections will present the three components of $\mathbf{V}$. First, the vertical component will be formulated using SVD methodology. Next, the first-order recursive filter algorithm on the regular mesh will be introduced to model the horizontal part of the horizontal error covariance matrix. Finally, we will present the mathematical formulation

of the dynamic height operator.

2.2 The vertical covariance based on Empirical Orthogonal Function

A geophysical system's variations result from exceedingly complex non-linear interactions between many degrees of freedom or modes. We saw that this system is characterized by non-linearity and high dimensionality. Consequently, a challenging task is to find ways to reduce the system's dimensionality to a few modes. A further yet challenging task is to link these modes to the dynamics/physics of the system.

Given any oceanographic space-time field, EOFs analysis finds a set of orthogonal spatial patterns and associated uncorrelated time series or principal components (PCs). In this thesis, we refer to EOFs and PCs for the spatial and temporal pattern, following the literature terminology. The geometrical constraints characterizing EOFs and PCs can be beneficial in practice since the covariance matrix of any subset of retained PCs is always diagonal. However, Navarra and Simoncini (2010) shows that EOFs have a domain dependence. Indeed, since normal modes derived from dynamical/physical models are not necessarily orthogonal can be challenging to interpret the obtained patterns because physical modes are not necessarily orthogonal. This derives from the fact that physical processes are not necessarily uncorrelated (Hannachi, Jolliffe, and Stephenson (2007)).

EOF analysis aims to find a new set of variables that capture most of the observed variance from the data through linear combinations of the original variables. This fact allows us to reduce the large number of variables of the original data to a few variables without compromising much of the variability of the data (Hannachi and O'Neill (2001)).

In this work, as we mentioned in the previous section, we use EOFs to model the vertical component of the background error covariance matrix, $\mathbf{V}_V$. To compute EOFs, we use an SVD of an anomaly matrix.

Let $\mathbf{X}(t, \mathbf{s})$ the space-time field of a gridded dataset, such as temperature or salinity, at time t and spatial position $\mathbf{s}$. The value of the field at discrete time t_i and grid point $\mathbf{s}_j$, with $i = 1 \dots n$ and $j = 1 \dots p$ is denoted x_{ij}. The observed field is then represented by the data matrix:

$$\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^T = \begin{pmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1p} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & x_{n3} & \dots & x_{np} \end{pmatrix} \tag{2.28}$$

where the columns represent the time series of the field at position $\mathbf{s}_j$ and rows represent the map, or the value of the field, at time t.

The climatology of the field is defined by

$$\overline{\mathbf{X}} = (\overline{\mathbf{x}}_1, \overline{\mathbf{x}}_2, \ldots, \overline{\mathbf{x}}_p) = \frac{1}{n}\mathbf{1}_n^T\mathbf{X} \tag{2.29}$$

where $\mathbf{1}_n^T$ is a row vector of length n that contains only one and $\overline{\mathbf{x}}_j$ is the time average of the field at the spatial grid point j

$$\overline{\mathbf{x}}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij} \tag{2.30}$$

So, the anomaly field, or departure from the climatology, is defined as

$$\mathbf{A} = \mathbf{X} - \mathbf{1}_n\overline{\mathbf{x}} \tag{2.31}$$

and the covariance matrix of $\mathbf{A}$ is defined as

$$\mathbf{S} = \frac{1}{n}\mathbf{A}^T\mathbf{A} \tag{2.32}$$

EOFs/PCs analysis aims to find uncorrelated linear combinations of the different variables that explain the maximum variance, and the EOFs are obtained as the solution to the eigenvalue problem:

$$\mathbf{S}\mathbf{u} = \lambda^2\mathbf{u} \tag{2.33}$$

Computing the covariance matrix, $\mathbf{S}$, has a high computational cost considering the high dimensionality of the field. To reduce the computational cost, we use SVD decomposition of the anomaly matrix A.

Any n × p data matrix $\mathbf{A}$ can be decomposed as

$$\underset{[n\times p]}{\mathbf{A}} = \underset{[n\times r][r\times r][r\times p]}{\mathbf{U}\ \mathbf{\Sigma}\ \mathbf{V}^T} \tag{2.34}$$

where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices (i.e. $\mathbf{U}\mathbf{U}^T = \mathbf{V}\mathbf{V}^T = \mathbf{I}_r$) and
$r \leq \min(n, p)$ is the rank of $\mathbf{A}$.
The matrix $\mathbf{\Sigma}$ is diagonal, and the diagonal elements are the singular values of $\mathbf{A}$ sorted in decreased order, i.e., $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_r \geq 0$. The columns $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$ of $\mathbf{U}$ and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ of $\mathbf{V}$ are respectively the left and right singular vectors of the data matrix $\mathbf{A}$ and represent PCs and EOFs respectively. Using Eq. (2.34), the covariance matrix $\mathbf{S}$ (Eq. (2.32)) is

$$\mathbf{S} = \frac{1}{n}\mathbf{V}\mathbf{\Sigma}^2\mathbf{V}^T \tag{2.35}$$

where $\mathbf{\Sigma}^2 = \text{Diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_r^2)$. The EOFs are therefore orthogonal and the PCs uncorrelated, and this is a major characteristic of conventional EOFs. The orthogonality is useful since it provides a complete basis for the data matrix. Indeed, Eq. (2.34) yields the

decomposition

$$\mathbf{A} = \sum_{k=1}^{r} \lambda_k \mathbf{u}_k \mathbf{v}_k^T \tag{2.36}$$

and the map $\mathbf{a}_t = (\mathbf{a}_{t1}, \mathbf{a}_{t2}, \ldots, \mathbf{a}_{tp})^T$ of the anomaly field $\mathbf{A}$ at time t is

$$\mathbf{a}_t = \sum_{k=1}^{r} \lambda_k u_{tk} \mathbf{v}_k^T \tag{2.37}$$

Eq. (2.37) is useful when EOFs are used to reduce the dimensionality of the data. This can be achieved simply by truncating the above sum by keeping the first M terms where M is generally much smaller than the rank r of $\mathbf{A}$. However, there is no universal rule for truncation, and the choice of M is generally arbitrary. The first approach is to choose the first M EOFs representing a fixed variance, e.g., 90%. This quantity can be computed as the cumulative percentage contribution, μ_M, of the first M EOFs modes

$$\mu_M = \sum_{k=1}^{M} \mu_k = \sum_{k=1}^{M} 100 \frac{\lambda_k^2}{\sum_{i=1}^{r} \lambda_i^2} \% \tag{2.38}$$

In this work, the EOFs are derived from the variations of four years (2017-2020) long SHYFEM integration and are grouped by month for the temperature, salinity, and SLA (trivariate EOFs) in each node. Eq. (2.38) is not a resolutive and complete method for truncating the EOFs; the user experience and the specific problem for which the EOFs/PCs are computed should lead the final number of modes to be used (as we will see in chapter 7). Indeed, looking at the cumulative function in Fig. 2.1, we observe that, in the location marked with a red star in the Gulf of Taranto at coordinates 39.92°N 17.43°E (Fig. 2.1(a)), the first five modes represent almost 100% of the variance contribution in both February Fig. 2.1(b) and JulyFig. 2.1(c). This means that the field has predominant modes of frequency variability in that location (Navarra and Simoncini (2010)). Furthermore, the eigenvalues provide information on the distribution of power (energy) as a function of scale and the separation/degeneracy of the EOFs patterns. For example, high/low power is associated with low/high frequency variability. Hence, low frequency and large-scale patterns tend to capture most of the variance observed in the system (Hannachi, Jolliffe, and Stephenson (2007)). This suggests that performing cross-scale oceanographic experiments, as in this work, the choice of the leading number of EOFs should be tackled by looking at the percentage of variance and considering the scale dimension. Fig. 2.2 (note the different percentage scale between Fig. 2.2(a), Fig. 2.2(b) and Fig. 2.2(c), Fig. 2.2(d)) shows how the percentage of variance is distributed for February and July on SANI domain for the first Figs. 2.2(a) and 2.2(b) and sixth mode Figs. 2.2(c) and 2.2(d). Here, it is clear how the percentage of variance varies over the whole domain, and that six modes could not be enough to represent the variability in locations differ-
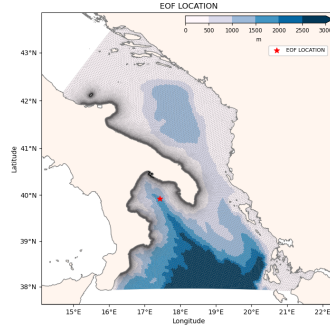
(a) EOF location
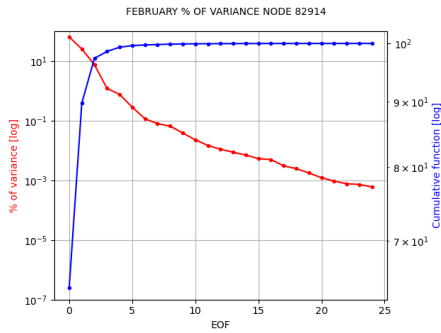
(b) February
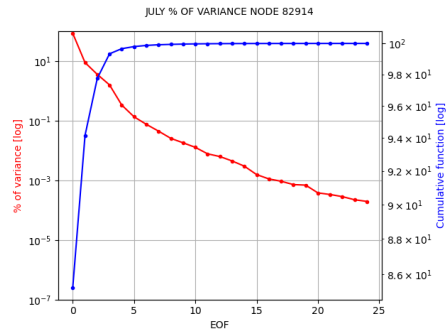
(c) July

Figure 2.1: Percentage of variance (red line) and cumulative function, μ_k (blue line) of 25 EOFs computed from 4 years long (2017-2020) SHYFEM integration in the Gulf of Taranto at coordinates 39.92°N 17.43°E (a) on February (b) and July (c).

ent than Fig. 2.1(a). For this reason, in the experiments, we will use 25 EOFs. Finally, the covariance matrix of temperature (T), salinity (S), and sea level anomaly (SLA) for February and July, in the exact location as Fig. 2.1(a), is shown in Fig. 2.3. Fig. 2.3(d) shows that the covariance between T and SLA is higher in July than in February. This is the expression of the water column stratification during the summer period. Indeed, during summer, the sea surface temperature variation has a higher impact on sea surface height (SSH) compared to the winter period when the water column is generally well mixed, and the sea surface temperature has no significant diurnal variations.

28

(a) February first mode

(b) July first mode

(c) February sixth mode
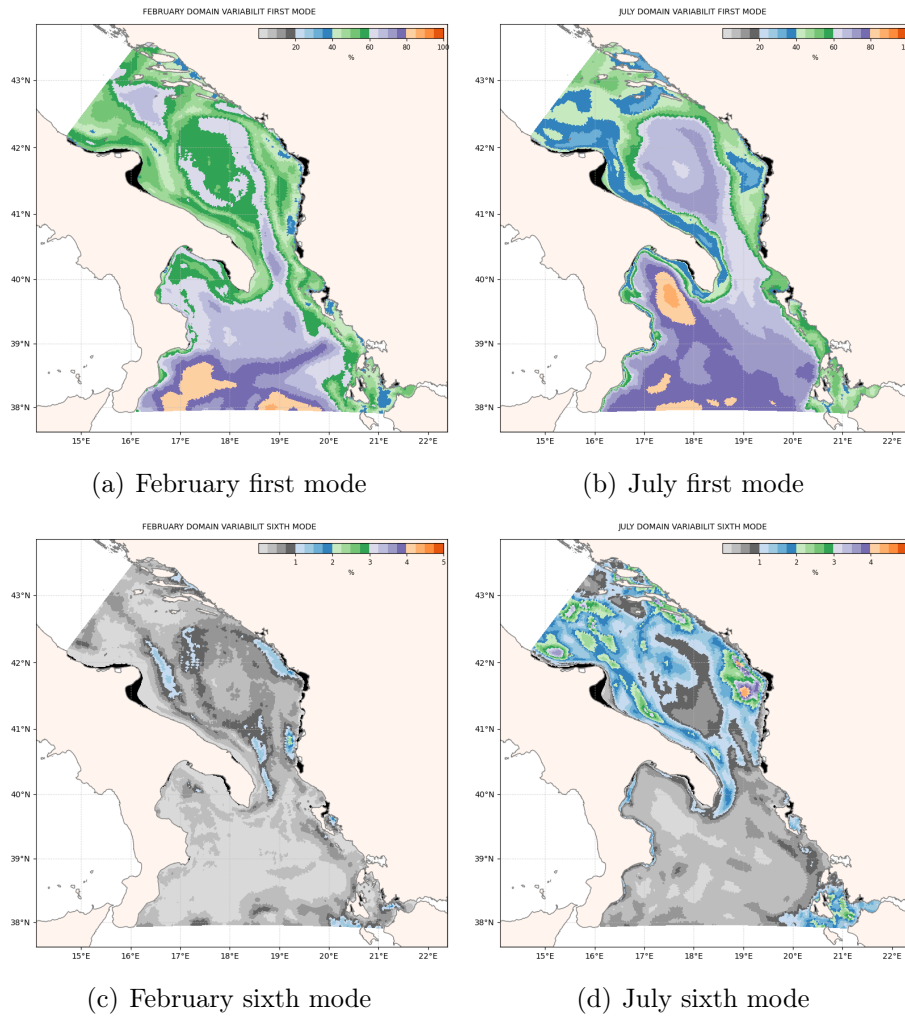
(d) July sixth mode

Figure 2.2: Percentage of variance computed from 4 years long (2017-2020) SHYFEM integration on February and July for first mode(a,b) and sixth mode (c,d) on SANI domain.
Note the different percentage scale between (a,b) and (c,d).

(a) February S, T covariance

(b) July S, T covariance

(c) February S, T, SLA covariance

(d) July S, T, SLA covariance
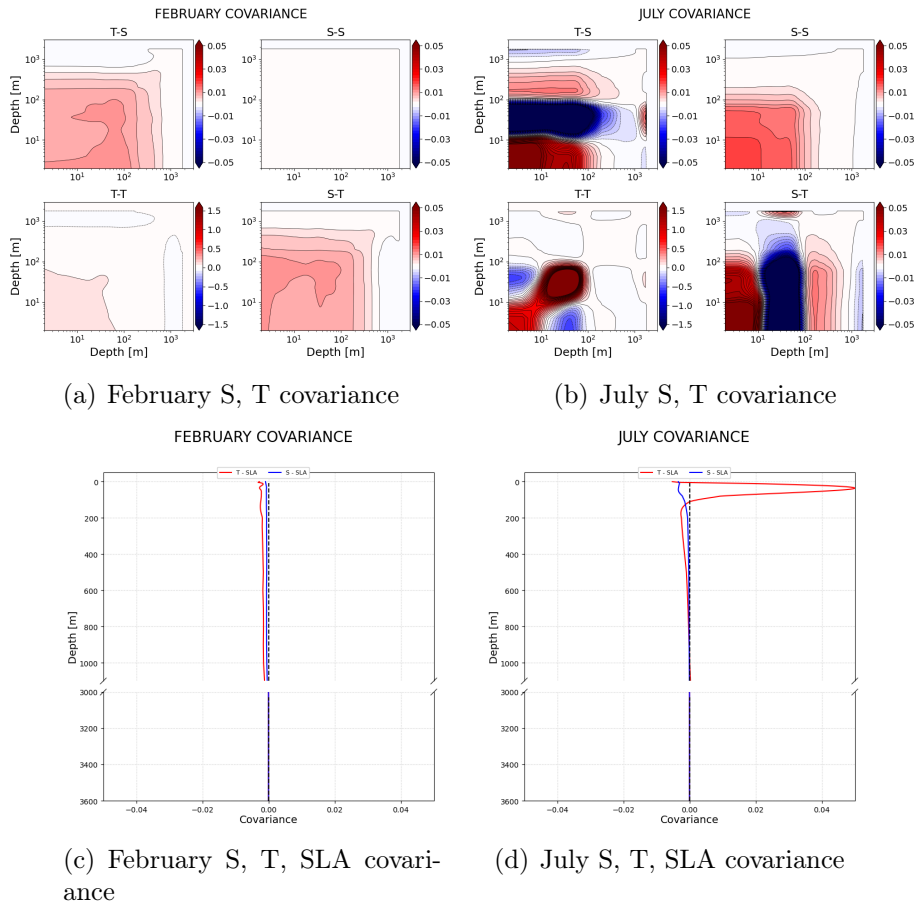
Figure 2.3: Covariance between S, T, SLA against the depth of 25 EOFs computed from 4 years long (2017-2020) SHYFEM integration on February and July in the Gulf of Taranto at coordinates 39.92N 17.43E (Fig. 2.1(a)).

2.3 The horizontal covariance based on first order recursive filter on regular mesh

It is assumed that horizontal covariances are Gaussian with a constant correlation radius. In oceanographic models, isotropic and Gaussian spatial correlations can be relatively efficiently modeled by a repeated application of the Laplacian operator, which is also the solution of the horizontal diffusion equation (Derber and Rosati (1989)). Applying the Laplacian operator reduces the computational cost by eliminating the explicit calculation of the exponential function between each pair of points and may facilitate the introduction of coastal boundaries. Weaver and Courtier (2001) investigated the application of the diffusion equation to calculate horizontal covariances in the ocean in the presence of coastal boundaries and spherical coordinates. However, although much more computationally efficient than the direct application of Gaussian correlation functions, the approximation using the explicit solution of the diffusion equation typically requires a relatively large number of iterations. An alternative to the Laplacian operator is the recursive filter (RF) (Lorenc (1992) and Hayden and Purser (1995)). It is conceptually simple, typically requires only a few iterations to approximate the Gaussian function, and its application on a horizontal grid can be split into two independent directions (Purser et al. (2003)). Furthermore, the RF has the unique feature of locally varying scaling, which gives it greater flexibility over inhomogeneous data, like unstructured mesh.

The mathematical formulation of the RF is quite simple. However, the basic assumptions about the mathematical formulation are crucial when RF is used in oceanographic applications for the presence of the coastline.

We introduce a first-order RF formulation starting from its 1D formulation, and then we generalize the concept on a 3D regular mesh.

Let's consider an **infinite line** with n grid points and an initial field A_i with $i = 1, 2, \ldots, n$. The fundamental algorithm, in the **forward** pass ("left" to "right"), is described by

$$B_i = \alpha B_{i-1} + (1-\alpha)A_i \quad i = 1, 2, \ldots, n \tag{2.39}$$

followed by the **backward** pass

$$C_i = \alpha C_{i+1} + (1-\alpha)B_i \quad i = n, n-1, \ldots, 1 \tag{2.40}$$

The application of the RF in each direction is performed to ensure zero phase change. Generally, the RF algorithm is applied recursively N times, and from $N = 2$ it starts to approximate a Gaussian function Fig. 2.3.

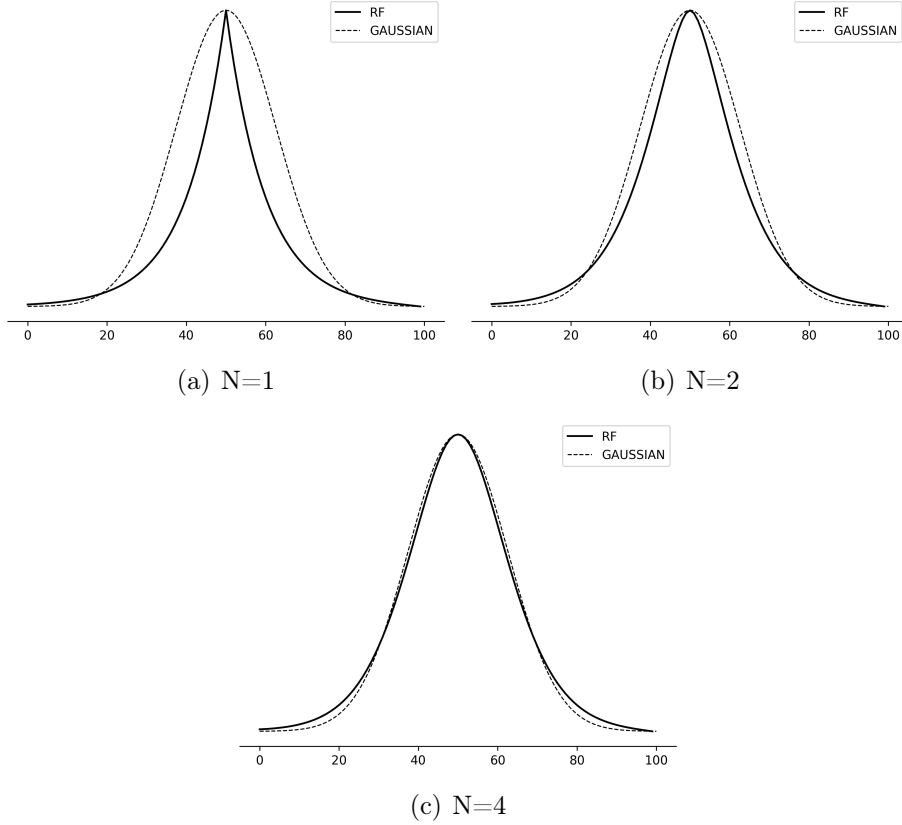In RF application, the maximum number of recursive steps is usually $N = 4$. In Eqs. (2.39)

(a) N=1

(b) N=2

(c) N=4

Figure 2.4: Comparison between RF and the Gaussian function for a different number of iterations and R=10.

and (2.40), α is the smoothing factor, and it is defined from the following equations

$$E = \frac{2N\delta^2}{4R^2}$$
$$\alpha = 1 + E - \sqrt{E(E + 2)}$$

(2.41)

Where δ is the grid resolution and R is the correlation radius set by the user, which defines the radius of application of the RF around the initial point A_i. It is possible to derive the Gaussian variance, σ, of the approximate Gaussian through the relation:

$$\sigma = \frac{R}{\delta}$$

(2.42)

Increasing R implies that the initial information is spread over more nodes, and we have a lower maximum of the RF output (Fig. 2.5). The advantage of using the RF is that we can apply the 1D algorithm separately on each dimension of a regular grid. In Fig. 2.6, we show this concept graphically on a 2D regular mesh. From Fig. 2.6, it is clear that, on a regular mesh, we profit from the RF formulation on an infinite line. Indeed, the regular meshes have the advantages of having the nodes intrinsically ordered and, introducing some boundary conditions in the presence of the coastline (see Dobricic and Pinardi (2008) for details), we can express the symmetry inherited from the formulation on an infinite
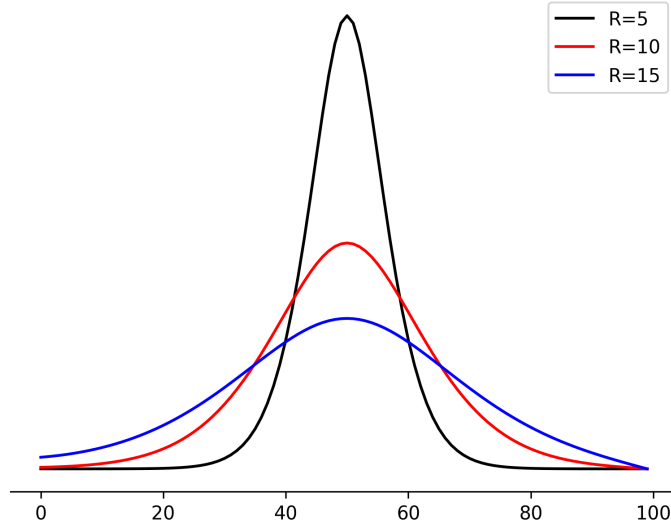
Figure 2.5: RF output with N=4 and different R.
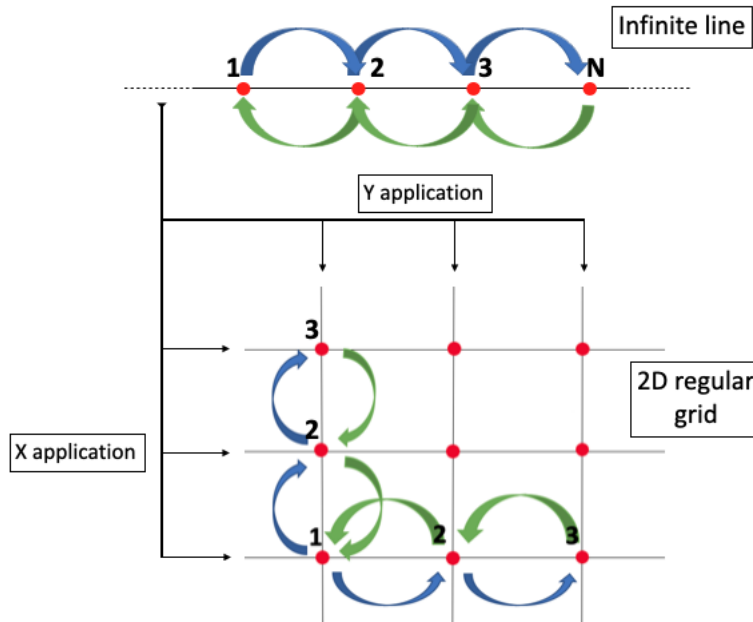
Figure 2.6: Schematic application of the first order RF on a regular mesh.
Blue arrows represent the forward pass, and the green arrows are the backward one.

line. Fig. 2.7 shows the RF application on a 2D regular mesh for different N and $R = 100$ Km. As in the 1D case, we approximate a Gaussian shape by increasing the number of iterations. From a computational point of view, a drawback of the RF algorithm is that it is inherently sequential, and with the infinite formulation, it is difficult to be parallelized. Also, using the RF algorithm, it is a complex task to describe covariance across various scales (Purser, Rancic, and Pondeca (2022)).

When we work with unstructured meshes, we lose all the advantages inherited by the infinite support. For this reason, in chapter 3, we will return to this issue and present a novel algorithm for the RF that applies to unstructured meshes. The new algorithm has the same mathematical structure given above. The key points are ordering the unstructured

(a) N=1

(b) N=2

(c) N=4
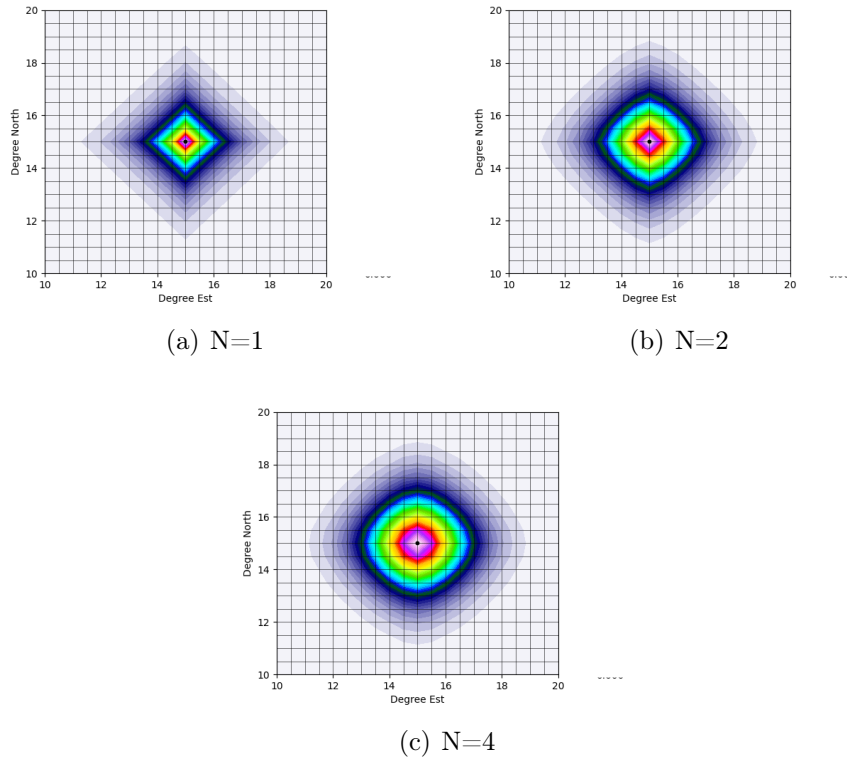
Figure 2.7: Application of the first order RF on a 2D regular grid. The black point in the center of the grid refers to the position of the initial value. R=100 km.

grid nodes using a specific algorithm and applying a scaling factor, which is necessary to use the RF on a highly varying resolution grid as SANI. Also, the new algorithm can be applied to grids with high variability in resolution (as the grid used in this work, Fig. 1.2) by formulating a local smoothing factor.

2.4 The dynamic height operator

The dynamical balance between surface elevation perturbations and the corresponding T and S perturbations may be estimated either statistically using EOFs decomposition (e.g., Dobricic et al. (2005)) as we do in this work or by applying some kind of the geostrophic constraint (e.g., Dobricic et al. (2006)). Dobricic et al. (2006) found that the vertical EOFs computed solely from the covariance between T, S, and SLA could produce sea level corrections that are not geostrophically balanced with the density corrections. If we assimilate only in situ observations of T and S, all unbalanced corrections will be rapidly removed by the fast barotropic adjustment. On the other hand, it was demonstrated in Dobricic et al. (2006) that when the SLA observations were assimilated, the enforcement of the geostrophic relationship for the sea level in the error covariance matrix had a significant positive impact on the accuracy of the analyses. However, the geostrophic relationship is usually assumed with some arbitrary hypothesis on the level

34

of no motion or the bottom pressure gradients. For example, Cooper and Haines (1996) assume that the perturbation bottom pressure gradient is zero and forms the surface elevation perturbation from the vertical integral of T and S perturbation in the whole water column. Dobricic et al. (2006) use the formula by Pinardi, Rosati, and Pacanowski (1995), which assumes that the horizontal pressure gradient at a depth of $1000m$ equals zero. We will assume the existence of a no motion level of 700 m (less than in Dobricic et al. (2006) since the region covered by SANIFS grid, Fig. 1.2) and V_η is defined through a local hydrostatic adjustment operator defined in Storto et al. (2011).

In the previous section, we derived the 3DVar cost function (Eq. (2.19)). We saw that Eq. (2.19) is composed of the background-related term (first addendum) and the observation-related term (second addendum). Satellite altimetry measures the along-track SSH (η in Eq. (2.43)) referred to the reference ellipsoid. Finally, the SLA (ξ in Eq. (2.43)) is obtained by subtracting the Mean Sea Surface (MSS, η_{MSS} in Eq. (2.43)) from SSH (see section 4.2.2 for details).

$$\xi = \eta - \eta_{\mathrm{MSS}} \tag{2.43}$$

From Eq. (2.19) and Eq. (2.20), we see that the SLA contribution to the observational term of the cost function is Eq. (2.44)

$$\mathbf{y_o} - \mathcal{H}(\mathbf{x_b}) - \mathbf{H}(\mathbf{x} - \mathbf{x_b}) = \xi_o - \xi_b - \mathbf{H}(\mathbf{x} - \mathbf{x_b}) \tag{2.44}$$

where ξ_o is the SLA observed (Eq. (2.43)) and ξ_b is the sea level anomaly predicted by the ocean model. To compute the tangent-linear increments of the ξ_b within the minimization, V_η has been formulated in terms of the Local Hydrostatic Adjustment (LHA) scheme based on the vertical integration of density increments. In other words, we can use the 1980 United Nations Educational, Scientific and Cultural Organization (UNESCO) International Equation of State (IES 80), as described in Fofonoff (1985) to link the temperature and salinity variations with the density variations and then the SSH variations. This means that we are going to compute the steric variations of the sea level. Consider a water column of unitary surface and height $z = -H(x, y, t)$. The free surface at time t is $z = \eta(x, y, t)$. Both levels are referred to a reference level $z = 0$, i.e., the geoid. The mass of the water column is:

$$m(x, y, t) = \int_{z=-H(x,y,t)}^{z=\eta(x,y,t)} \rho(x, y, z, t)\, dz \implies$$
$$\implies \frac{\partial m(x, y, t)}{\partial t} = \rho(x, y, \eta, t) \frac{\partial \eta(x, y, t)}{\partial t} + \int_{z=-H(x,y,t)}^{z=\eta(x,y,t)} \frac{\partial \rho(x, y, z, t)}{\partial t} dz \tag{2.45}$$

where $\rho(x, y, z, t)$ is the density distribution along the water column and $\rho(x, y, \eta, t)$ is the surface density that we can define as $\rho_S(x, y, t)$. So, the time variation of water column eight, $\eta(x, y, t)$, is:

$$\frac{\partial \eta(x, y, t)}{\partial t} = -\frac{1}{\rho_s(x, y, t)} \int_{z=-H(x,y,t)}^{z=\eta(x,y,t)} \frac{\partial \rho(x, y, z, t)}{\partial t} dz + \frac{1}{\rho_s(x, y, t)} \frac{\partial m(x, y, t)}{\partial t} \quad (2.46)$$

In Eq. (2.46), we have that the time variation of the sea level height depends on two components:

- **STERIC COMPONENT** $\implies -\frac{1}{\rho_s(x,y,t)} \int_{z=-H(x,y,t)}^{z=\eta(x,y,t)} \frac{\partial \rho(x,y,z,t)}{\partial t} dz$

- **MASS COMPONENT** $\implies \frac{1}{\rho_s(x,y,t)} \frac{\partial m(x,y,t)}{\partial t}$

For Stevino's law and remembering that the water column has a unitary surface, we have

$$P(x, y, z = -H, t) = P(x, y, z = \eta, t) + \rho(x, y, z, t)gH(x, y, t) \implies$$

$$(2.47)$$

$$\implies m(x, y, t) = \frac{P(x, y, z = -H, t)}{g} - \frac{P(x, y, z = \eta, t)}{g}$$

where $P(x, y, z = -H, t)$ is the pressure at the bottom of the water column, and $P(x, y, z = \eta, t)$ is the atmospheric pressure. And finally we can write Eq. (2.46) as

$$\frac{\partial \eta(x, y, t)}{\partial t} = -\frac{1}{\rho_s(x, y, t)} \int_{z=-H(x,y,t)}^{z=\eta(x,y,t)} \frac{\partial \rho(x, y, z, t)}{\partial t} dz$$

$$+ \frac{1}{\rho_s(x, y, t)g} \frac{\partial P(x, y, -H, t)}{\partial t} \quad (2.48)$$

$$- \frac{1}{\rho_s(x, y, t)g} \frac{\partial P(x, y, \eta, t)}{\partial t}$$

The mass component is divided into a term that considers processes that add or remove water mass, e.g., evaporation, precipitation, river runoff, and one that considers water mass advection due to gradients in atmospheric pressure. Neglecting the atmospheric component, which is a boundary condition in ocean models, and assuming the existence of a level of no motion, corresponding to a depth $-H^*$, where horizontal velocities are practically zero, through geostrophy, the bottom pressure variation vanishes and Eq. (2.48) in the field of increments becomes

$$\delta \eta(x, y, t) = -\frac{1}{\rho_s(x, y, t)} \int_{z=-H^*(x,y,t)}^{z=\eta(x,y,t)} \delta \rho(x, y, z, t) dz \quad (2.49)$$

Finally, our linearized observation operator is:

$$\mathbf{H}(\mathbf{x} - \mathbf{x_b}) = -\frac{1}{\rho_s(x, y, t)} \int_{z=-H^*(x,y,t)}^{z=\eta(x,y,t)} \delta\rho(\mathbf{x} - \mathbf{x_b}) dz \qquad (2.50)$$

In practice, this scheme splits the observation departure into its thermosteric and halosteric contributions over the water column using the adjoint version of the density operator linearized around the background fields of temperature and salinity. However, in areas of highly variable bottom topography (like coastal regions), the assumption of a uniform level of no motion may be wrong, and the sea level correction should be derived with more accuracy. How the sea level increment is vertically spread on the temperature and salinity increments depends upon the bivariate definition of the vertical background-error covariances, which are spatially nonuniform.

3. A novel algorithm for the first order recursive filter in unstructured grid ocean modeling

In the previous chapter, we introduced the first-order RF algorithm on a regular grid, and the entire theory was based on a strong assumption:

- The 1D RF formulation is referred to on an infinite line. This assumption leads to a strong symmetry feature of the theory. In fact, the infinite line assumption makes each point on the line a symmetry point ensuring that forward and backward steps are applied to the same geometric structure. This is suitable to have a symmetric Gaussian shape with respect to the x and y orthogonal axis.

Furthermore, a regular grid has the feature of having the nodes intrinsically ordered, and its symmetry is with respect to the x and y orthogonal axis. This means that, referring to a 2D regular grid, the "regular" RF (from here on we will denote it rRF) algorithm can be applied in its 1D formulation separately in the x and y direction, as shown in Fig. 2.6. Consequently, applying rRF on a regular n-dimensional grid is simply the application of 1D formulation on each dimension. Finally, boundary conditions are applied to avoid discontinuity due to grid open and closed boundary (Hayden and Purser (1988) and Dobricic and Pinardi (2008)).

When we move to unstructured mesh, we lose all these features. Indeed, referring to a simple triangular mesh as in Fig. 3.1, we see that there is no simple connection pattern referred to the node numbering, and we can not move through two connected nodes simply using a constant increment or decrement of the node index as in Eq. (2.39). Furthermore, the grid is no longer symmetric about the orthogonal x and y axis. For all these reasons, it is necessary to adapt the rRF algorithm to be applied to the particular case of an unstructured triangular mesh.

In the next session, the formulation of a novel RF algorithm for unstructured triangular meshes, identified as uRF in this thesis, will be shown.
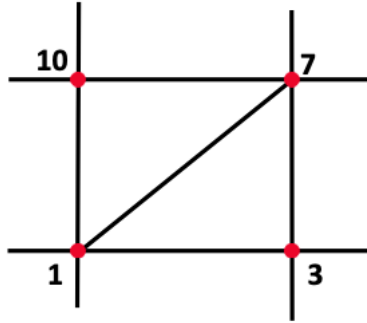
Figure 3.1: A small portion of a simple triangular grid. The connectivity node to node is not expressed by a constant increment or decrement of the node index as in Fig. 2.6. The symmetry is not with respect to the x and y orthogonal axis.

3.1   Description of the algorithm and methodology

We aim to apply the first-order RF algorithm on an unstructured triangular mesh, using the mathematical formulation shown in Eq. (2.39) and Eq. (2.40). To do that, we have to solve the node ordering problem. On a regular grid, the intrinsic ordering of the nodes leads to the intrinsic order of the connectivity node to node, i.e., node 1 is connected to node 2 by the edge both in the x and y direction. The particular square grid case is shown in Fig. 2.6. With this geometrical feature, applying Eq. (2.39) and Eq. (2.40), the information is spread from node 1 to node 2, from node 2 to node 3, and from node n-1 to node n in forward passes. The other way around for the backward passes. This reasoning suggests that it is more important to approach the problem from the connectivity node-to-node (i.e. the edges) point of view and order the edges rather than the nodes.

The idea implemented in this work is to order the edges with respect to the longitude and latitude applying different ordering criteria for the forward and backward steps both in x and y directions, as described in Algorithm 1.

Algorithm 1 produces four different matrices of dimension [k × 2], whith k the number of edges which, in combination with Eq. (2.39) and Eq. (2.40), must reproduce the Gaussian spreading defined by rRF. For each order defined in Algorithm 1, Eq. (2.39) and Eq. (2.40) are used running over the triangle's edges and taking the specific connected nodes.

In the next section, we will present results from synthetic experiments of the novel algorithm on two types of unstructured triangular mesh: Delaunay frontal equilateral triangles constant resolution unstructured mesh and SANI grid which has a variable resolution (Fig. 1.2).

Algorithm 1 Edges ordering algorithm

Consider the generic edge e_k, $k \in K \subset \mathbb{N}$. e_k is uniquely identified by a couple of connected nodes (n_{1_k}, n_{2_k}) with $n_{1_k}, n_{2_k} \in n \subset \mathbb{N}$. In a two-dimensional space, each node is uniquely identified by its coordinates, $n_{1_k} \longleftrightarrow (x_{1_k}, y_{1_k})$ and $n_{2_k} \longleftrightarrow (x_{2_k}, y_{2_k})$ with $x_{i_k}, y_{i_k} \in r \subset \mathbb{R}$. Emulating the forward and backward steps of rRF both in x and y direction, we can define four different criteria for edge ordering.

Require: Connectivity node to node $\rightarrow e_k$, $k \in K \subset \mathbb{N} \rightarrow (n_{1_k}, n_{2_k})$ with $n_{1_k}, n_{2_k} \in n \subset \mathbb{N}$
Require: Node coordinates $\rightarrow n_{1_k} \longleftrightarrow (x_{1_k}, y_{1_k})$ and $n_{2_k} \longleftrightarrow (x_{2_k}, y_{2_k})$

X-FW
1: Ordering the edges considering n_{1_k} coordinates in the following two sequential steps

1- $x_{1_1} < x_{1_2} < x_{1_3} < \ldots x_{1_k}$
2- $y_{1_1} < y_{1_2} < y_{1_3} < \ldots y_{1_k}$

This ordering results in edge steps that start in the lower-left corner and stop in the upper-right corner of the grid moving from left to right.
This ordering is used in combination with Eq. (2.39).

X-BW
2: Ordering the edges considering n_{2_k} coordinates in the following two sequential steps

1- $x_{2_1} > x_{2_2} > x_{2_3} > \ldots x_{2_k}$
2- $y_{2_1} > y_{2_2} > y_{2_3} > \ldots y_{2_k}$

This ordering results in edge steps that start in the upper-right corner and stop in the lower-left corner of the grid moving from right to left.
This ordering is used in combination with Eq. (2.40).

Y-FW
3: Ordering the edges considering n_{1_k} coordinates in the following two sequential steps

1- $y_{1_1} < y_{1_2} < y_{1_3} \ldots y_{1_k}$
2- $x_{1_1} < x_{1_2} < x_{1_3} \ldots x_{1_k}$

This ordering results in edge steps that start in the lower-left corner and stop in the upper-right corner of the grid moving from down to up.
This ordering is used in combination with Eq. (2.39).

Y-BW
4: Ordering the edges considering n_{2_k} coordinates in the following two sequential steps

1- $y_{2_1} > y_{2_2} > y_{2_3} > \ldots y_{2_k}$
2- $x_{2_1} > x_{2_2} > x_{2_3} > \ldots x_{2_k}$

This ordering results in edge steps that start in the upper-right corner and stop in the lower-left corner of the grid moving from up to down.
This ordering is used in combination with Eq. (2.40).

3.2   Validation with idealized test cases

Algorithm 1 in combination with Eq. (2.39) and Eq. (2.40) defines the novel RF algorithm for the unstructured grid, which we denote as uRF. To validate the geometrical and mathematical features of uRF we will test the uRF on a 2D Delaunay frontal equilateral triangles mesh using two different constant resolutions, 20 km, and 5 km. This type of unstructured mesh is the closer case to the SANI grid. Finally, we will test uRF on the SANI grid, which has a variable resolution from coastal to open sea regions (Fig. 1.2).

3.2.1   2D Delaunay frontal grid and local smoothing factor α

To test the algorithm on an unstructured triangular mesh, we use the grid in Fig. 3.2 that has the same structure as SANI grid, i.e., tessellation of equilateral triangles using Delaunay frontal technique.

Given a set of points $\mathbf{P}$ in the Euclidean plane, a Delaunay triangulation defines a unique triangulation of them such that no one of $\mathbf{P}$ is inside the circumcircle of any triangle. The resulting Delaunay triangulation is the dual graph of the Voronoi diagram (see Lee and Schachter (1980) and Mavriplis (1997) for details). In the advancing-front algorithm, the grid generation begins with discretizing the boundaries as a set of edges. The boundary edges form the initial front, which progressively advances into the defined domain (Mavriplis (1997)).

For our purpose, the triangular mesh is built on an idealized square domain defined on 10N-20N and 10E-20E with a resolution of 20 km. The initial field A in Eq. (2.39) is zero everywhere except in (15.1N,15.1E) where it has a value of 100. As discussed in the
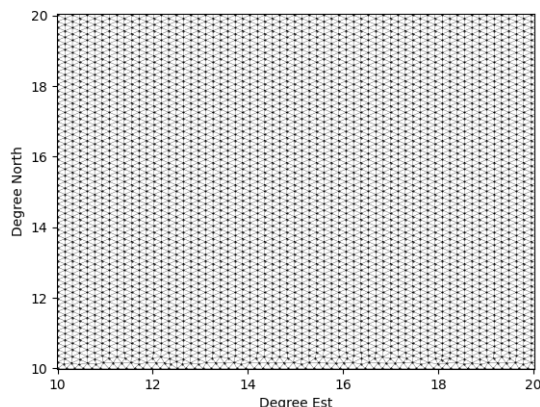
Figure 3.2: Delaunay frontal grid used to test the uRF. Resolution 20 km.

introduction of this chapter, that grid does not have the same symmetry as the regular one. Furthermore, another difference is that the symmetry of the unstructured triangular mesh is not constant on the whole domain but has a local variable symmetry. Indeed,

considering a regular grid, it inherits the local symmetry from the infinite line geometrical structure, so it has a constant local symmetry with respect to the x and y orthogonal axis. Considering that, in the case of an unstructured triangular mesh, it is necessary to introduce a local smoothing factor α to apply Eq. (2.39) and Eq. (2.40) and keep in consideration the variable local grid symmetry.

The local smoothing factor α is derived from Eq. (3.1) that links the Gaussian variance σ, the correlation radius R, and the local grid resolution δ (i.e., the length of the edges of the triangles which represent the distance between two nodes)

$$\sigma_i = \frac{R_i}{\delta_i} \quad i = 1, 2, \ldots, e \tag{3.1}$$

Where e is the number of edges. From Eq. (3.1), it is possible to design an algorithm (Algorithm 2) to compute the local correlation radius R_i and then the local α_i using Eq. (2.41).

Algorithm 2 Local smoothing factor α

R is the correlation radius, δ is the triangle's edge measure and σ is the Gaussian variance

Require: Initial correlation radius such that $R \geq MAX(\delta)$

1: Compute σ: $\sigma = \frac{R}{MAX(\delta)}$

2: Compute R_i: $R_i = \sigma \cdot \delta_i$

3: Compute α_i using the result of Step 2 along with Eq. (2.41)

From here on, all the applications of the uRF will be made using the local smoothing factor α (Algorithm 2).

The application of uRF in combination with Algorithm 2 on the grid in Fig. 3.2 produces the results shown in Fig. 3.3. In the four tests shown in Fig. 3.3, the initial correlation radius is R=60 km. Fig. 3.3(a) - Fig. 3.3(d) show the results N=4, 6, 8, 10 respectively. We observe that as N increases the Gaussian shape is better reproduced.

To approach SANI grid resolution, we can test uRF on an unstructured mesh of the same type as the one in Fig. 3.2 but with a resolution of $\delta = 5$ km closer to the maximum open sea SANI grid resolution (4 km). As in the previous case, we consider an initial field A (Eq. (2.39)) zero everywhere except in (15.0N,15.0E) where it has a value of 100.

In Fig. 3.4 we can see the results of four different tests using R=15 km and N as in the previous case. Also here, we obtained the desired Gaussian symmetry. From both these test cases we can conclude that, on a triangular grid generated using Delaunay frontal algorithm, N=6 provides a good approximation of the 2D Gaussian shape. For this reason, in the SANI grid, we are going to show just the application of N=6.
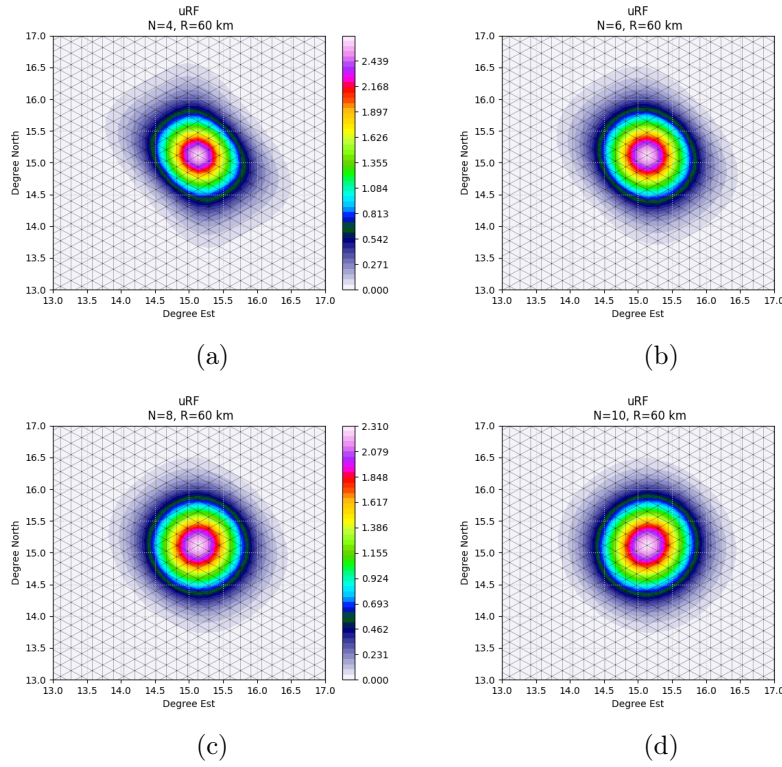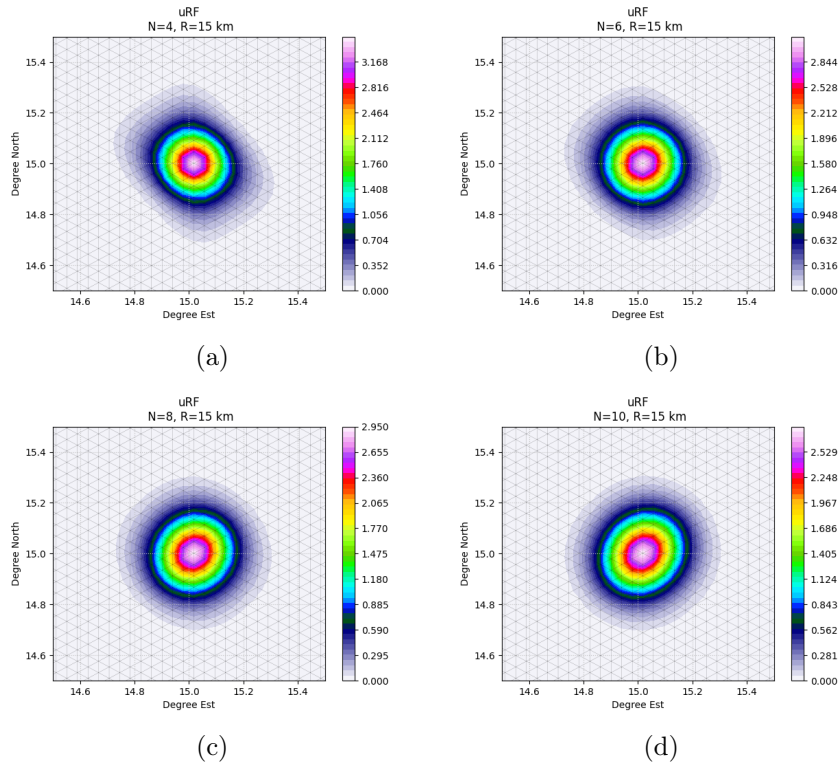
Figure 3.3: Application of uRF in combination with Algorithm 2 on the Delaunay frontal grid in Fig. 3.2 with a resolution of 20 km(note that here we show a zoom for better visualization). R=60 km. (a) N=4, (b) N=6, (c) N=8, (d) N=10. Increasing N increase also the Gaussian symmetry.

Figure 3.4: Application of uRF in combination with Algorithm 2 on the Delaunay frontal grid as in Fig. 3.2 with a resolution of 5 km (note that here we show a zoom for better visualization). R=60 km. (a) N=4, (b) N=6, (c) N=8, (d) N=10. Increasing N increase also the Gaussian symmetry.

3.3 Validation on SANIFS grid

Fig. 3.5 shows the SANIFS unstructured triangular mesh and zoom on the Gulf of Taranto coastal region and Northern Ionian open sea region. The grid has the highest resolution along Italian coasts (from 50 m to 500 m). The resolution in the open sea and along the East closed boundary is from 3 km to 4 km. Fig. 3.6 shows the results of applying
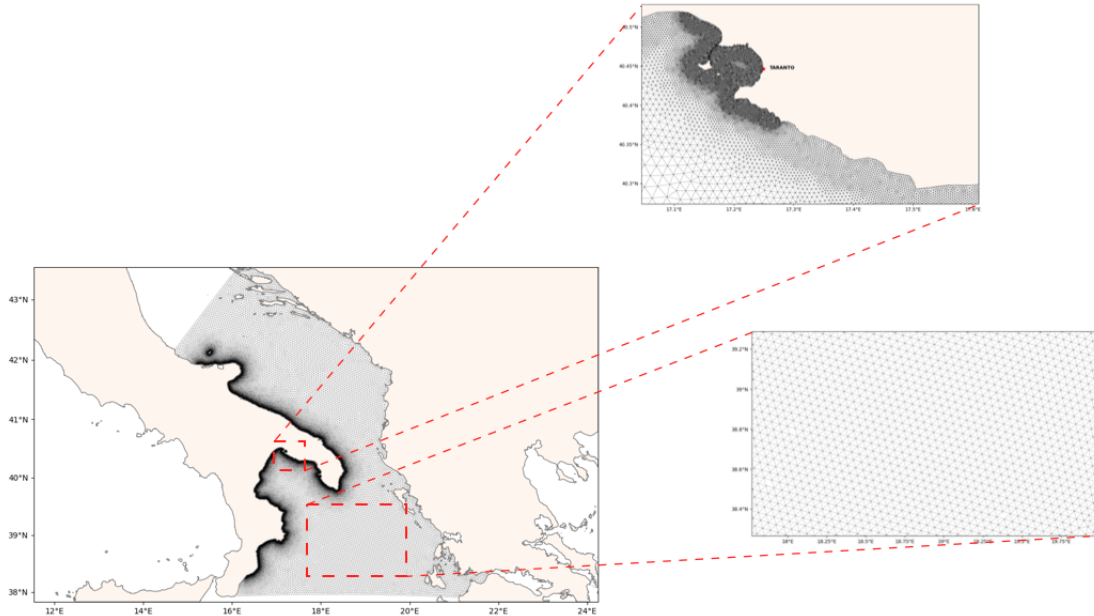
Figure 3.5: SANIFS numerical domain and grid with a zoom on Gulf of Taranto coastal region and Northern Ionian open sea region. Coastal resolution: from 50 m to 500 m. Open sea resolution: from 3 km to 4 km. The resolution is higher in Mar Grande to better model that region.

uRF on the SANIFS grid. Considering the conclusions of the previous section here we test uRF only with N=6. To approach to correlation radius used in real applications, we show results for R=10 km in Fig. 3.6(a) and Fig. 3.6(c) and R=20 km in Fig. 3.6(b) and Fig. 3.6(d). In open sea (Fig. 3.6(a) and Fig. 3.6(b)) and coastal region tests (Fig. 3.6(c) and Fig. 3.6(d)) the novel algorithm, uRF, exhibits the desired Gaussian symmetry.

In conclusion, we demonstrated that uRF in combination with a local smoothing factor algorithm (Algorithm 2) on a triangular unstructured grid generated using the Delaunay frontal algorithm with a constant resolution of 20km and 5km, and on SANIFS grid exhibits the desired Gaussian symmetry. From the tests, we also concluded that N=6 produces a suitable Gaussian shape approximation.

With the OceanVar technical components developed and adapted for unstructured mesh applications (SVD for the vertical component of the background error covariance matrix, the novel uRF for the horizontal component and the dynamical height operator for SLA covariance from T and S) SANIFS is interfaced with OceanVar to build an analysis and forecasting system.
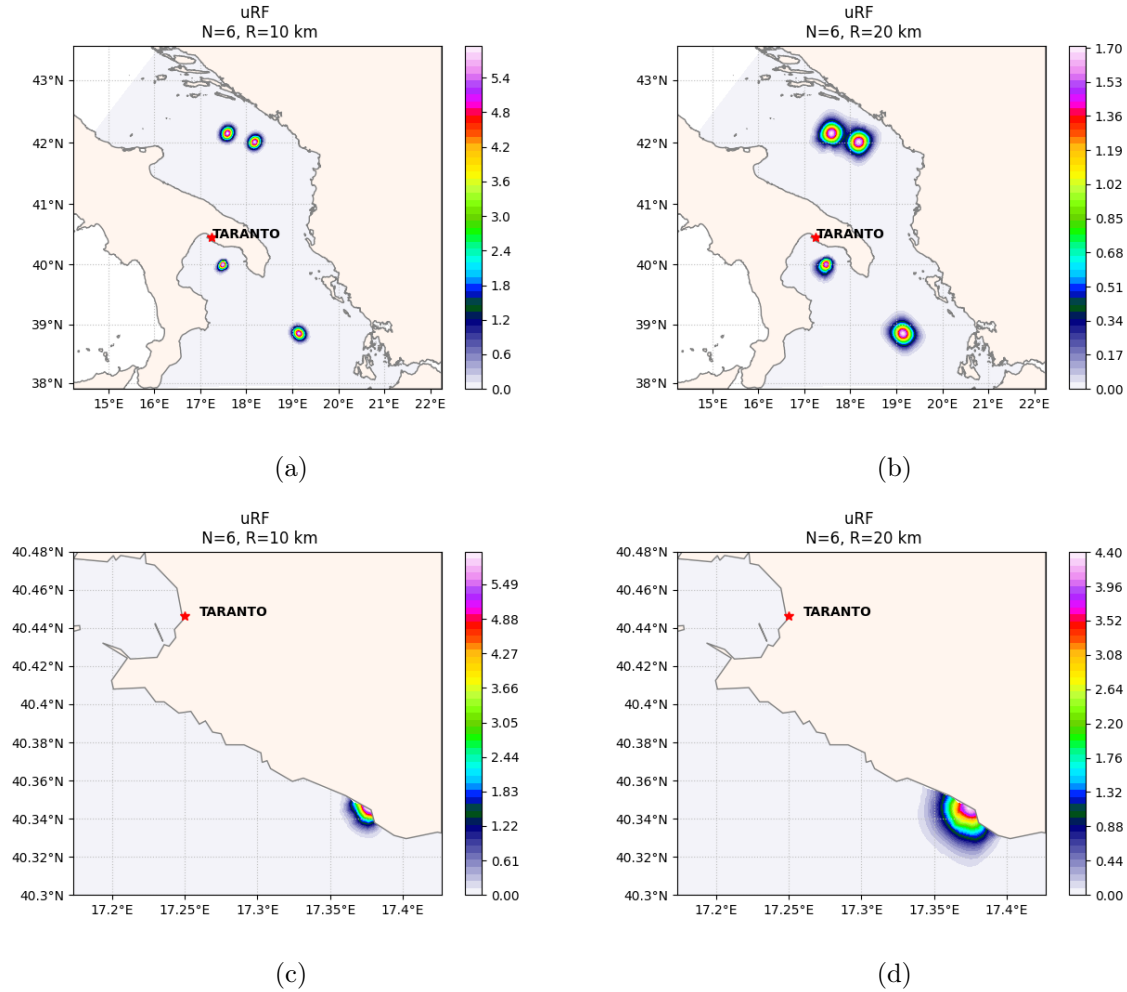
Figure 3.6: Application of uRF on SANI grid for open sea (a) and (b) and coastal regions (c) and (d). N=4. All the results exhibit the desired Gaussian symmetry.

In the next chapter, we will discuss the setting of the assimilation cycle and experiments. The first section will see the SANIFS settings, and next we will describe the observational datasets used in the assimilation process. Those components are the primary input for the 3D-Var cycle implemented using OceanVar.

4. Assimilation cycle: the case of the Southern Adriatic Northern Ionian coastal Forecasting System

With all the OceanVar technical components developed and adapted for unstructured mesh applications, SANIFS can be interfaced with OceanVar in an assimilation system. As discussed in the introduction, SANIFS is a coastal-ocean operational system based on the unstructured grid finite-element three-dimensional hydrodynamic SHYFEM model, providing short-term forecasts. Currently, the operational chain is based on a downscaling approach starting from the large-scale system for the entire Mediterranean Basin (MedFS, Mediterranean Forecasting System), which provides initial and boundary condition fields to the nested system (Federico et al. (2017)). In the introductory part of this thesis, we saw that MedFS is coupled with OceanVar and the nested model, SANIFS, inherits DA at open boundaries. SANIFS is configured to provide hydrodynamics and active tracer forecasts both in the open ocean and coastal waters using a variable horizontal resolution from the open sea (3-4 km) to coastal areas (50-500 m) (Fig. 1.2 or Fig. 3.5). OceanVar (Dobricic and Pinardi (2008) and Storto, Masina, and Navarra (2016)) is a state-of-art 3DVar assimilation scheme implemented and used with NEMO (Nucleus for European Modelling of the Ocean) regular grid ocean model at CMCC.

This chapter will first present the SHYFEM computational grid and equations. Next, we will describe the observational datasets used in the assimilation process: Argo profiling floats and satellite altimetry missions. Those components are the primary input for the 3DVar cycle that, combined with OceanVar, define the assimilation system. Finally, in the last part of this chapter, we will describe the data assimilation workflow and experiments settings.

4.1 Model description

In this study, we use the parallel implementation of SHYFEM described in Micaletto et al. (2021).

SHYFEM is a 3-D finite element unstructured mesh hydrodynamic model (Umgiesser et al. (2004) and Cucco and Umgiesser (2006)) solving the Navier-Stokes equations by applying hydrostatic and Boussinesq approximations. The unstructured grid is Arakawa-B with triangular meshes (Bellafiore and Umgiesser (2010) and Ferrarin et al. (2013)) which provides an accurate description of irregular coastal boundaries (Fig. 1.2). In Fig. 4.1(a), we see how variables are staggered in the horizontal grid. Horizontal momentum (U, V) is located in the element centers, while all the others are on the vertices (vertical velocity ω and scalars). Each vertex has a corresponding finite volume (dashed lines in Fig. 4.1(a)). The staggering of hydrodynamic variables is important to have a mass-conserving model (Jofre et al. (2014) and Fuhrer et al. (2014)). Variables are also staggered in the vertical grid, as shown in Fig. 4.1(b). The turbulent and molecular stresses and the vertical velocity are computed at the bottom interface of each layer (black dots in Fig. 4.1(b)). The free surface is at the top of the upper layer, determining the variable volume of the top cells. All the other variables are defined at the layer center (red dots in Fig. 4.1(a)). Scalar variables (red) are staggered with respect to vertical velocity (black), referenced in the middle and at layer interfaces, respectively. The sea surface elevation is a 2D field defined only in the points at the surface. The grid cells on the top layer can change their volume due to the oscillation of the free surface. The number of active cells along the vertical direction depends on the sea depth (Micaletto et al. (2021)).

(a) Horizontal grid   (b) Vertical grid
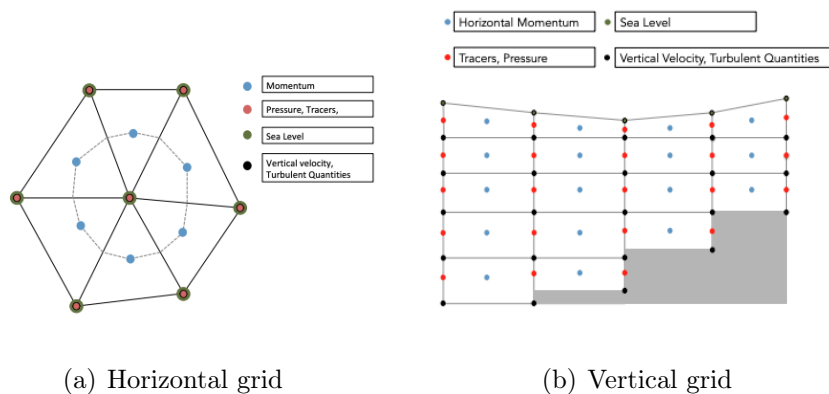
Figure 4.1: Disposition of model variables on the SANIFS grid. From Micaletto et al. (2021)

SHYFEM solves the ocean primitive equations, assuming incompressibility in the continuity equation, and advection-diffusion equation for active tracers using finite-element discretization based on triangular elements (Umgiesser et al. (2004)). A semi-implicit algorithm is used for the time integration of the free surface equation, the Coriolis term, the pressure gradient in the momentum equation, and the divergence terms in the continuity equation. Vertical eddy viscosity and vertical eddy diffusivity in the tracer equations are treated fully implicitly for stability reasons. Finally, the advective and horizontal diffusive terms in the momentum and tracer equations are treated explicitly (Micaletto et al.

(2021)). To compute the vertical viscosities is used a turbulence closure scheme adapted from the k-ϵ module GOTM (General Ocean Turbulence Model, Burchard and Petersen (1999)).

As described in Maicu et al. (2021), the momentum equations, integrated over each layer are:

$$\frac{\partial U_l}{\partial t} + u_l\frac{\partial U_l}{\partial x} + v_l\frac{\partial U_l}{\partial y} + \int_{z_l}^{z_{l-1}} w\frac{\partial u}{\partial z}dz - fV_l = -gh_l\frac{\partial \zeta}{\partial x} - \frac{gh_l}{\rho_0}\int_{z_{l_{mid}}}^{0} \frac{\partial \rho}{\partial x}dz - \frac{h_l}{\rho_0}\frac{\partial P_a}{\partial x}+$$
$$+ \nabla_h \cdot (A_H\nabla_h U_l) + \int_{z_l}^{z_{l-1}} \frac{\partial \tau_{xz}}{\partial z}dz$$
$$(4.1)$$

$$\frac{\partial V_l}{\partial t} + u_l\frac{\partial V_l}{\partial x} + v_l\frac{\partial V_l}{\partial y} + \int_{z_l}^{z_{l-1}} w\frac{\partial v}{\partial z}dz + fU_l = -gh_l\frac{\partial \zeta}{\partial y} - \frac{gh_l}{\rho_0}\int_{z_{l_{mid}}}^{0} \frac{\partial \rho}{\partial y}dz - \frac{h_l}{\rho_0}\frac{\partial P_a}{\partial y}+$$
$$+ \nabla_h \cdot (A_H\nabla_h V_l) + \int_{z_l}^{z_{l-1}} \frac{\partial \tau_{yz}}{\partial z}dz$$
$$(4.2)$$

Where $\zeta = \zeta(x,y,t)$ is the free surface, $l = 1,2\ldots N$ is the vertical layer index, $z_l = 1,2\ldots N$ are the depths of the layer interfaces at the bottom with z_0 being the free surface ζ and z_N the bottom interface of the deepest layer. z_{lmid} is the depth at the middle of layer l. u_l and v_l are the horizontal velocity components. The horizontal velocities integrated over the layer l (layer transports) are defined by $U_l = u_l h_l$ and $V_l = v_l h_l$ where h_l is the layer thickness. P_a is the atmospheric pressure at the sea surface, g is the gravitational acceleration, ρ_0 is the reference density of sea water, $\rho = \rho_0 + \rho'$ is the water density with ρ' representing the perturbation of the density from the reference value ρ_0. A_H is the horizontal eddy viscosity computed following the Smagorinsky formulation (Smagorinsky (1963) and Blumberg and Mellor (1987)). w_l is the vertical velocity for layer l defined at the bottom interface. τ_{xz} and τ_{yz} are the turbulent shear stresses defined at the bottom interface of each layer and written according to the flux-gradient theory (Maicu et al. (2021)).

Following Maicu et al. (2021), the continuity equation integrated over a vertical layer l is:

$$\frac{\partial U_l}{\partial x} + \frac{\partial V_l}{\partial y} = w_{z_l} - w_{z_{l-1}} \tag{4.3}$$

The continuity equation at $l = 1$ has an additional term representing the time variability of the top layer thickness and thus it reads as:

$$\frac{\partial h_1}{\partial t} + \frac{\partial U_1}{\partial x} + \frac{\partial V_1}{\partial y} = w_{z_1} - w_{z_0} \tag{4.4}$$

The layer-integrated salinity and temperature equations reads, respectively:

$$\frac{\partial \left(h_l S_l\right)}{\partial t} + U_l \frac{\partial S_l}{\partial x} + V_l \frac{\partial S_l}{\partial y} + \int_{z_l}^{z_{l-1}} w \frac{\partial S}{\partial z} dz = \nabla_h \cdot \left(K_H \nabla_h h_l S_l\right) +$$
$$+ \int_{z_l}^{z_{l-1}} \frac{\partial}{\partial z} \left(K_V \frac{\partial S}{\partial z}\right) dz \quad (4.5)$$

$$\frac{\partial \left(h_l \theta_l\right)}{\partial t} + U_l \frac{\partial \theta_l}{\partial x} + V_l \frac{\partial \theta_l}{\partial y} + \int_{z_l}^{z_{l-1}} w \frac{\partial \theta}{\partial z} dz = \nabla_h \cdot \left(K_H \nabla_h h_l \theta_l\right) +$$
$$+ \int_{z_l}^{z_{l-1}} \frac{\partial}{\partial z} \left(K_V \frac{\partial \theta}{\partial z}\right) dz +$$
$$+ \frac{1}{\rho_0 C_p} \left(I\left(z_{l-1}\right) - I\left(z_l\right)\right) \quad (4.6)$$

where K_H and K_V are the horizontal and vertical turbulent diffusion coefficients respectively. S_l and θ_l are salinity and temperature at layer l. The solar irradiance is expressed by the last term on the right side of Eq. (4.6). It expresses the solar irradiance at depth z, parametrized with a double exponential according to Paulson and Simpson (1977).

The hydrostatic pressure is obtained by the layer-integrated vertical momentum under the hydrostatic hypothesis:

$$p_l \left(x, y, z_{l_{\text{mid}}}, t\right) = \rho_0 g \left(\xi - z_{l_{\text{mid}}}\right) + \int_{z_{l_{\text{mid}}}}^{0} \rho' g \, dz \quad (4.7)$$

Finally, the density ρ at layer l is computed from salinity, temperature and pressure according to the UNESCO equation of state (Fofonoff (1985)).

$$p_l \left(x, y, z_{l_{\text{mid}}}, t\right) = p_l \left(S_l, \theta_l, p_l\right) \quad (4.8)$$

As we saw in (Fig. 1.2 or Fig. 3.5) in the coastal waters of the eastern Italian coastlines, SANIFS has a high spatial resolution, reaching an element size of 500 m, with higher resolution in specific areas (e.g., Mar Grande of Taranto Fig. 3.5) where the resolution reaches 50m. The horizontal resolution in the open sea areas is approximately 3-4 km (Fig. 3.5). The SANIFS bathymetry (Fig. 1.2) was derived from the US Digital Bathymetric Data Base Variable Resolutions (DBDB-1) at 1' resolution for the Mediterranean Basin and integrated with higher-resolution bathymetry for coastal areas in the Gulf of Taranto provided by the Italian Navy Hydrographic Institute. Vertical discretization has 92 levels. This is appropriate for solving the field in coastal and open-sea areas (Federico et al. (2017)). The vertical resolution is 2m from the sea surface to 40m. The resolution is then progressively (stepwise) decreasing to the bottom with a maximum layer thickness of 200 m. All the other numerical settings of the modeling systems are described in Federico et al. (2017).

SANIFS is forced using ECMWF (European Centre for Medium Weather Forecasts) dataset with 12.5 Km horizontal resolution of the following atmospheric fields:

- $2m$ air temperature;

- $2m$ dew point temperature;

- total cloud cover;

- mean sea level atmospheric pressure;

- meridional and zonal $10m$ wind components;

As in Federico et al. (2017), those fields are corrected by land-contaminated points following Kara, Wallcraft, and Hurlburt (2007) and horizontally interpolated at each ocean grid node using Cressman's interpolation technique (Cressman (1959)).

As we discussed in the introduction of this thesis, SANIFS is nested in MedFS through the two lateral open boundaries Fig. 1.2. The current MedFS implementation is based on NEMO with a horizontal resolution of 1/24° (4-5 km approximately) and 141 unevenly spaced vertical levels.

The scalar MedFS fields (non-tidal sea surface height, temperature and salinity) are imposed at the SANIFS boundary nodes. In contrast, the MedFS total velocities are specified in the barycentre of the triangular elements with two nodes attached to the boundaries. The tidal elevation derived from the OTPS (Oregon State University Tidal Prediction Software (Egbert and Erofeeva (2002))) tidal model is prescribed at each boundary node. Eight of the most significant constituents are considered: M2, S2, N2, K2, K1, O1, P1, and Q1.

4.2 The observing data

In the experiments that we will present in the next chapter, we assimilate observations of temperature (T), salinity (S), and Sea Level Anomaly (SLA). In-situ Argo profiling float observations provide T and S. SLA comes from 4 satellite missions (Saral/Altika, Cryosat-2, Jason-3, Sentinel-3A). All the products are provided by the Copernicus Marine Service (*marine.copernicus.org.*). The products are accessible through the Copernicus Marine Service web portal, which includes a structured information catalog.

4.2.1 Temperature and salinity profiles from Argo profiling floats

Argo is an international program that collects information from inside the ocean using a fleet of robotic instruments that drift with the ocean currents and move up and down

between the surface and a mid-water level (Fig. 4.2). The name Argo, from Greek mythology, was chosen to emphasize the strong complementary relationship of the global ocean float array with the space-borne Jason altimeter mission (launch Dec. 7, 2001) as well as other altimeter missions. The Argo measurements of temperature, salinity, pressure, reference velocity, and sea surface height from satellite altimetric data form a dynamic description of the upper ocean. Each instrument (float) spends almost all its life below the surface. The standard Argo float mission is a 10-day cycle, with most of the float's
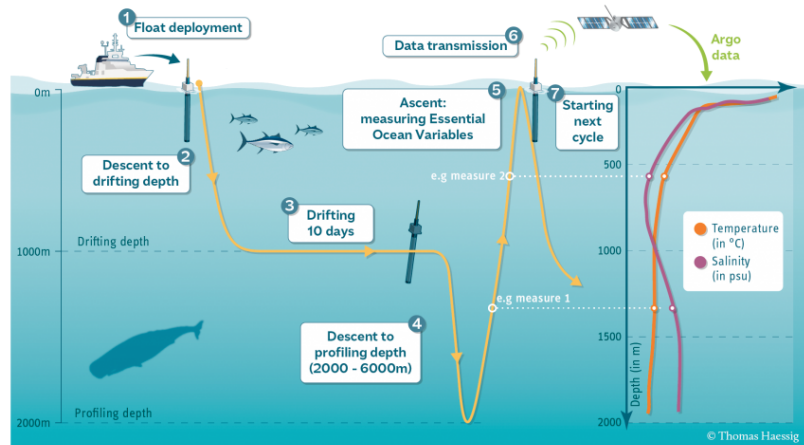
Figure 4.2: Argo measurement cycle. From *https://argo.ucsd.edu/how-do-floats-work/*

time spent in free-drifting along with deep ocean currents (Lagrangian measurement), followed by taking a series of measurements as it moves back up (profiles) to the ocean surface. Once the float is on the surface, it gets its location, often through GPS, and then communicates with a satellite to send its data and receive any new mission instructions. For most Argo fleet, the surface interval is between 15 minutes and one hour. After that, the float sinks to a drift depth of 1000 m for about 9 days. Then, the float sinks to its profile depth of 2000 m before slowly rising to the surface while measuring conductivity (salinity), temperature and pressure through CTD sensors. Depending on water depth, a standard CTD cast requires two to five hours to collect a complete data set. The lifetime of an Argo float is about 4-5 years. During this period, the instrument provides around 150 profiles.

In the next chapter, we will discuss experimental results over two years, 2017-2018. Over these years, the distribution of assimilated Argo profiling floats is shown in Fig. 4.3. Since the assimilation window used in the experiment is 24h, we show the Argo distribution on six different randomly chosen days (Fig. 4.4) and the number of profiles in a 24 h time window in the entire SANIFS domain (Fig. 4.5). From Fig. 4.5 and Table 4.1, we see that on the SANIFS domain during 2017-2018, the daily amount of Argo profiles is about 1-3 per day. The maximum number of profiles in 24h is 6. We also note that there are 122 days with no Argo observations. This is a crucial information to be considered in the DA experiments discussion in the next chapter. Finally, the total daily assimilated
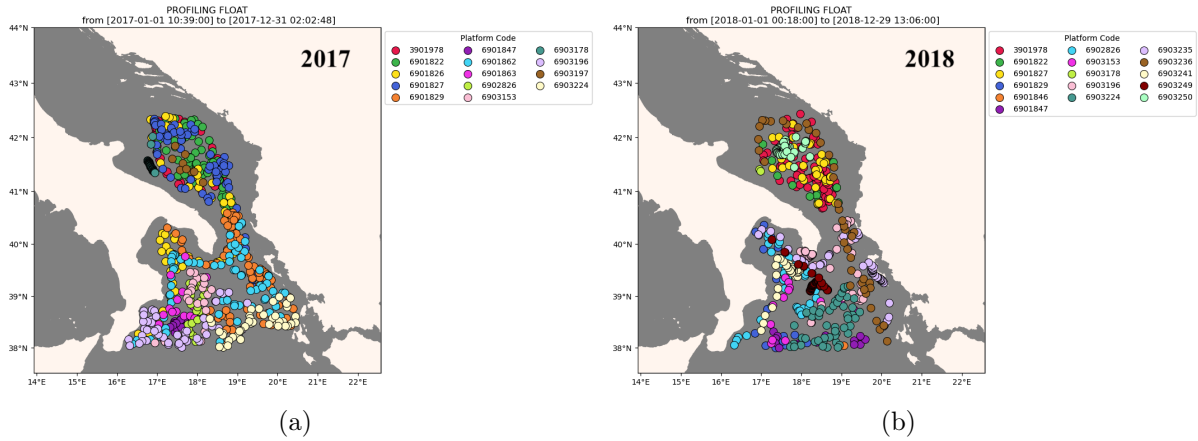
Figure 4.3: Distribution of Argo profiling floats over the years 2017 (a) and 2018 (b). Different platform code refers to different Argo float.

temperature and salinity observations are shown in Fig. 4.6. The total amount is the sum of temperature (Fig. 4.6(a)) and salinity (Fig. 4.6(b)) observations both in horizontal and vertical dimensions.
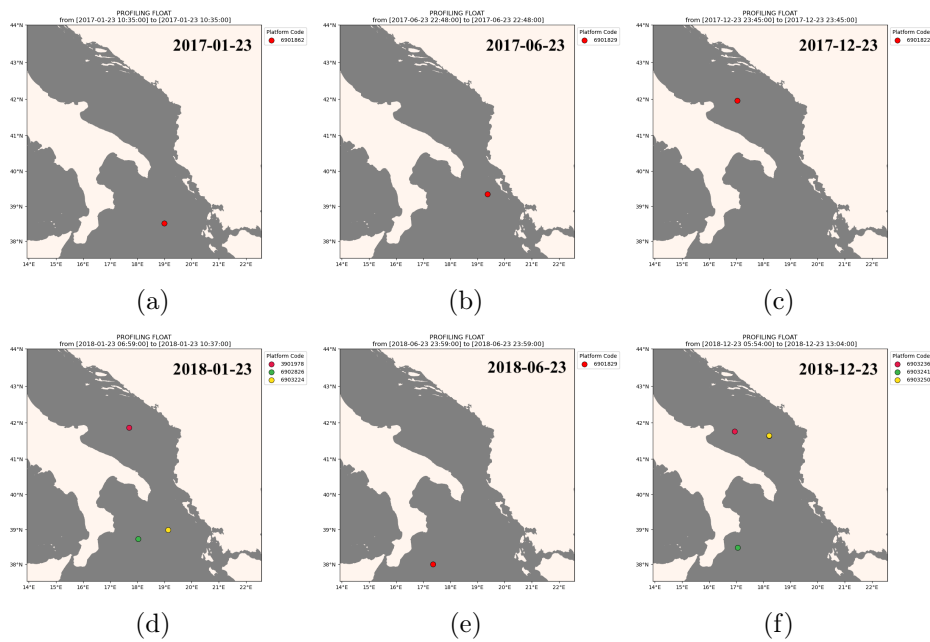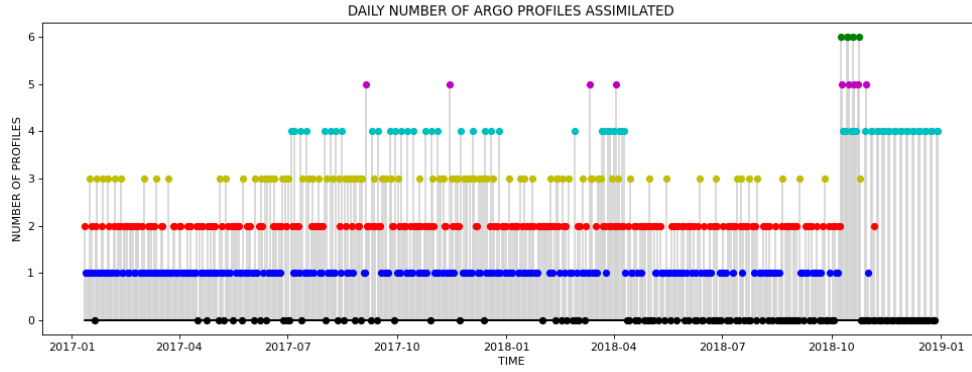

Figure 4.4: Daily distribution of Argo profiling floats over the years 2017 (a,b,c) and 2018 (d,e,f). Different platform code in (a,b,c,d,e,f) refers to different Argo float.

Figure 4.5: Daily amount of Argo profiles over the entire SANIFS domain over the period 2017-2018.

NUMBER OF PROFILES	DAYS (PERIOD 2017-2018)
0	122
1	227
2	201
3	89
4	64
5	9
6	5
TOTAL	717

Table 4.1: Number of days with 0, 1, 2, 3, 4, 5, 6 profiles over 2017-2018.
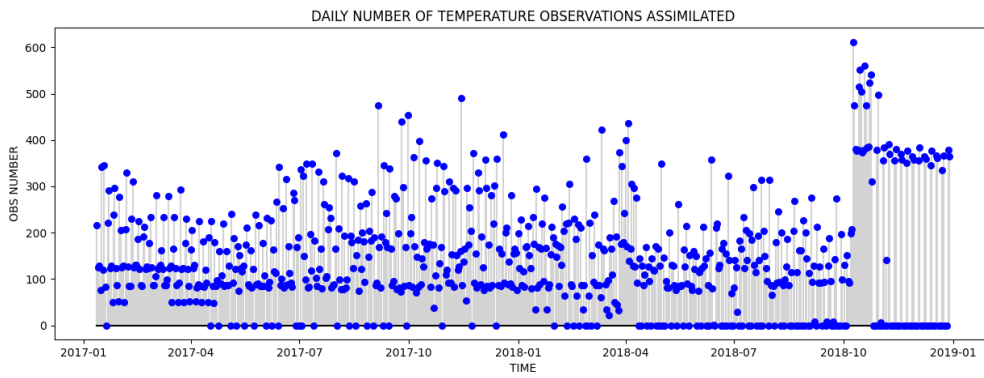
(a)
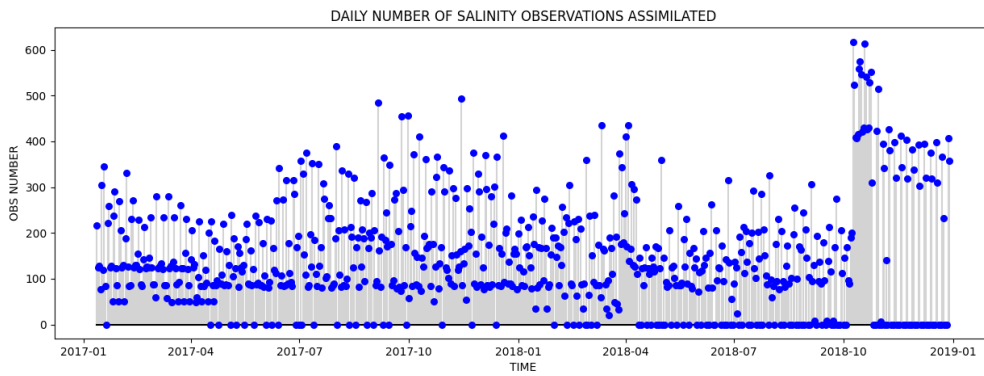
(b)

Figure 4.6: Amount of assimilated temperature and salinity observations in 24h over the entire SANIFS domain over the period 2017-2018.

4.2.2 Sea level anomaly from satellite altimetry missions

SLA is derived from satellite measurements made by a sensor called altimeter or Synthetic Aperture Radar (SAR) altimeter on board of several satellites. Radar altimeters transmit signals at a specific frequency toward the Earth and receive the return signal. This is analyzed to obtain measurements of the round-trip time (with an accuracy of approximately 0.1 ns), which multiplied by the speed of light, gives the distance R between the satellite and the surface. However, as the electromagnetic wave travels through the atmosphere, it can be decelerated by water vapor or electrons contained in the ionosphere. Measurements can be corrected for these interferences by measuring them with assistive instruments, making measurements at different frequencies, or modeling them. The satellite altitude S is another crucial parameter. It is the distance of the satellite from a reference surface (Fig. 4.7). Considering how R and S are defined, they are strictly related to the satellite's orbit. Satellite's orbit is measured with high accuracy (around 1-2 cm) with systems such as Doris and GPS (Fig. 4.7). The Sea Surface Height (SSH in
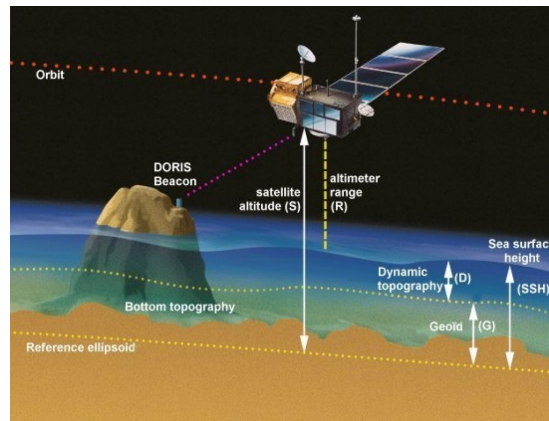
Figure 4.7: Satellite altimetry. From *www.aviso.altimetry.fr*

Fig. 4.7) is derived by subtracting R from S.

$$SSH = S - R \tag{4.9}$$

The SLA is then obtained by subtracting the Mean Sea Surface (MSS) from SSH (Eq. (6.2)). MSS is the mean SSH over a defined period. Consequently, the SLA obtained is relative to the reference period.

$$SLA = SSH - MSS \tag{4.10}$$

Related to the mission objectives, in altimetry observations different signals frequency are used:

- **K_u-band (13.6 GHz)**
  It is the band commonly used for altimetric measurements. It is the best compromise between the technology's functionality (relative to the power emitted), the available

bandwidth, sensitivity to atmospheric perturbations and perturbation due to the electrons present in the ionosphere.

- **C and S band (5.3 GHz)**
  This is more sensitive than the K_u band to disturbances in the ionosphere and less sensitive to the effects of water vapor in the atmosphere. Their main use is to allow the correction of ionospheric lag by combining measurements with K_u- band measurements.

- **K_a-band (35 GHz)**
  The signal frequency in the K_a-band allows for better observations of glaciers, rainfall amounts, coastal areas, land masses (forests, deserts, etc.) and wave heights. In this case, more bandwidth is available than for other frequencies, so higher resolution is possible, especially along coasts. However, the attenuation due to water or water vapor in the troposphere is high, which results in the impossibility of performing measurements if the precipitation is greater than 1.5 mm/h

In this study, we use SLA from 4 satellite missions: Saral/Altika, CryoSat-2, Jason-3 and Sentinel-3A.

- **SARAL/Altika** is an ISRO (Indian Space Research Organization) satellite, SARAL (Satellite with ARgos and ALtika), embarks the Altika altimeter (working in K_a- band, $35 GHz$), built by the Centre National d'Études Spatiales (CNES), as well as a DORIS instrument. Signal frequencies in the K_a-band will enable better observation of ice, rain, coastal zones, land masses (forests, etc.), and wave heights. Its central objectives are:

  1. to carry out precise, repetitive global measurements of sea surface height for:
     - developing operational oceanography;
     - improving understanding of the climate and developing forecasting capabilities;
     - operational meteorology;
     - coastal, inland waters and ice applications.
  2. to ensure, in association with Jason-2, the continuity of the ocean observation;
  3. to meet the requirements of various international ocean and climate study programs and contribute to building a global ocean observing system.

  From its launch on February 2013 to July 2016, Saral/Altika had a repetitive orbit. Due to technical issues on the reaction wheels from March 2015, CNES and ISRO have decided to pursue this mission with a new phase named "SARAL-DP, for SARAL-Drifting Orbit." From July 2016, and during this drifting phase, the

repetitive ground track is no more maintained and with the natural decay of the orbit, the ground track will drift.

- **CryoSat-2** is an altimetry satellite built by the European Space Agency dedicated to polar observation. It embarks on a three-and-a-half-year mission to determine variations in the thickness of the Earth's continental ice sheets and marine ice cover, and to test the prediction of thinning arctic ice due to global warming. Its orbit, at an inclination of about 92 degrees and an altitude of 717 Km, covers almost all polar regions. CryoSat-2 carries an altimeter/interferometer called Siral and a Doris instrument but no radiometer. Siral is a K_u-band instrument (13.575 GHz). CryoSat-2 was initially designed to study the cryosphere, but it is also an opportunity mission for oceanographic studies as it operates over the oceans for validation purposes in low-resolution mode, and in SAR mode over some specific zones. That means that the ground segments can process oceanic altimetry measurements acquired by Siral.

- **Jason-3** succeeding Topex/Poseidon, Jason-1 and Jason-2. It extends the high-precision ocean altimetry data record to support climate monitoring, operational oceanography and seasonal forecasting. Jason-3 is the result of a joint effort by CNES, NASA, EUMETSAT and NOAA, pursuing a heritage that has kept the oceans under close watch for over 20 years. The partnership is as for Jason-2, but the operational agencies (NOAA and EUMETSAT) take the lead; CNES serves as the system coordinator and all partners -including Nasa- support science team activities. From its launch to April 7th 2022, Jason-3 is on its nominal orbit (1336 Km). From April 25th, 2022 (cycle 300), Jason-3 is on the repeat interleaved orbit used by Topex from 2002 to 2005, Jason-1 from 2009 to 2012 and Jason-2 from October 2016 to July 2017. Jason-3 main mission is related to ocean topography and waves. The altimetric measurements are performed using a dual-frequency radar that measures in the K_u (13.575 GHz) and C (5.3 GHz) bands.

- **Sentinel-3A** fits into the Copernicus program, a joint project between ESA and European Union. The main objective of the Sentinel-3 mission is to measure sea surface topography, sea and land surface temperature, and ocean and land surface colour with high accuracy and reliability to support ocean forecasting systems, environmental monitoring and climate monitoring. The main frequency used for range measurements is the K_u-band (13.575 GHz), but, contrary to CryoSat, a second frequency in the C-band (5.41 GHz) is used for ionospheric correction. Its central objectives are:

  1. to carry out precise, repetitive global measurements of sea surface height for:
     − ocean, inland sea and coastal zone colour measurements;

- sea surface temperature measurements;

- sea surface topography measurements including an along-track SAR capability of CryoSat heritage for improved measurement quality in coastal zones and over sea-ice.;

2. to ensure, in association with Jason-2, the continuity of the ocean observation;

3. to meet the requirements of various international ocean and climate study programs and contribute to building a global ocean observing system.

Table 4.2 resumes the principal features of the different altimetric missions.

	SARAL/Altika	CryoSat-2	Jason-3	Sentinel-3A
Agency	ISRO/CNES	ESA	CNES/NASA/EUMETSAT/NOAA	ESA
Project	CMEMS	CMEMS	EUMETSAT Jason-3 L2P/L3 Marine Altimetry Service	EUMETSAT Sentinel-3 L2P/L3 Marine Altimetry Service
Goals	Observe the oceans	Polar Observation	Observe the oceans	Observe the oceans and deliver routine operational services to policy-makers and marine and land service users
Launch on	25/02/2013	08/04/2010	17/01/2016	16/02/2016
Altitude	800 Km	717 Km	1336 Km	814 Km
Inclination	98.5°	92.0°	66.0°	98.6°
Repetitivity	35 days	369 days with 30 day sub-cycle	10 days	27 day repeat cycle, 14+7/27 orbits per day
Lat Max	70.99N	70.99N	66.14N	70.98N
Lat Min	15.04N	27.90N	15.05N	27.49N
Lat Resolution	0.0560N	0.0560N	0.009N	0.054N
Lon Max	359.95E	354.75E	359.90E	359.98E
Lon Min	0.022E	4.54E	0.007E	0.014E
Lon Resolution	0.020E	0.008E	0.050E	0.025E
Processing Level	L3	L3	L3	L3

Table 4.2: Resume of the principal features of the 4 altimetric missions used in this study. From *www.aviso.altimetry.fr* and product NetCDF files.

The SLA variable from Copernicus Marine Service product used in this study is **sla_filtered** (SLA filtered from high-frequency signal considered as noises) from **phy_assim** products (*duacs.cls.fr*) with a reference period [1993-2012]. While the SLA (filtered or not) field is by default corrected from different geophysical effects, some of them (namely, the ocean tide, internal tide, dynamical atmospherical correction and Long Wavelength Error correction) are included in the phy_assim product to let the user compute the SLA field without these corrections (*https://duacs.cls.fr/faq/phy-and-phy_assim-l3-products/*).
Fig. 4.8 shows the daily distribution of the satellite tracks. The difference between available and assimilated observations is related to applying the no motion level at 700 m in the dynamic height operator (see chapter 3). All observations located in places with depths less than 700 m are rejected.

The daily assimilated amount of along-track SLA observations is shown in Fig. 4.9. As discussed in section 4.1, SANIFS has the tidal elevation derived from the OTPS tidal model prescribed at each boundary node. However, the SLA assimilated variable considered will not be defined as sla_filtered + ocean_tide as the model background field has. SANIFS does not yet provide an interface for FGAT (First Guess at Appropriate Time), a method for obtaining the instantaneous background field at the correct observation time.
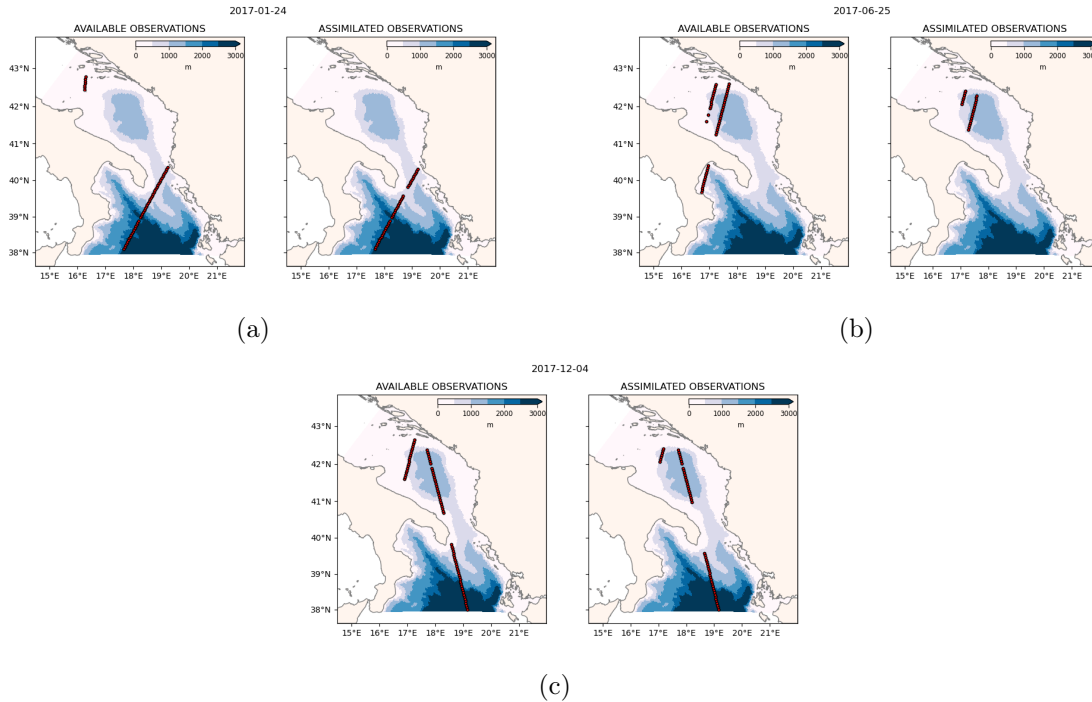
(a)   (b)

(c)

Figure 4.8: Daily distribution of SLA observation from the four altimetry missions used in this study. The figure shows the available observations and the assimilated ones. The difference is due to applying no motion level of 700 m in the dynamic height operator. All the observations located in places with depths less than 700 m are rejected.
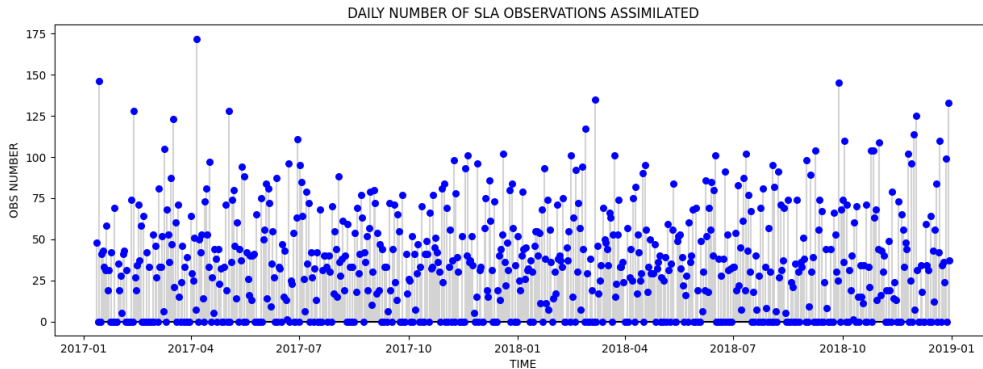

Figure 4.9: Amount of assimilated along-track SLA observations in 24h over the entire SANI domain over the period 2017-2018.

This may be a future improvement to the work presented here.

In the next section, we will discuss the data assimilation workflow and experiment setting.

4.3 The data assimilation workflow and experiments setting

Fig. 4.10 shows the assimilation workflow of coupling SANIFS and OceanVar. We saw a detailed description of all the components in Fig. 4.10 in chapter 2 and chapter 4. SANIFS provides the background state $\mathbf{x}_b$, which along with the observations $\mathbf{y}_o$, are the main input
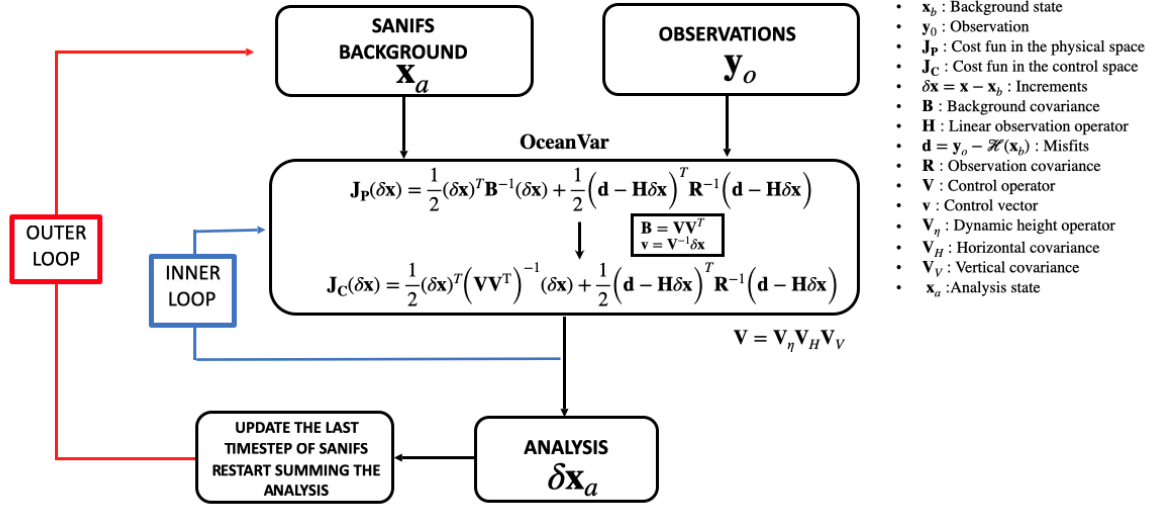
Figure 4.10: SANIFS-OceanVar workflow.

for the OceanVar algorithm. OceanVar's first step is formulating the cost function in the physical space. However, as discussed in chapter 2, it is more appropriate to minimize the cost function in the control space by applying the Cholesky transformation $\mathbf{B} = \mathbf{V}\mathbf{V}^{\mathrm{T}}$ at the background error covariance matrix. In the inner loop, the minimum of $\mathbf{J}_{\mathrm{C}}$ is computed using the L-BFGS algorithm. At this point, the minimum of the cost function in the control space is projected in the physical space and can be used to correct the model state. The solution of the minimization process is a matrix of increments referred to the variables used in the EOFs computation. Finally, the last timestep of the model background is corrected by summing the increments field to the variables to which those refer. This correction provides the analysis, which will be the new initial condition for the next SANIFS integration. This process will be repeated in the next data assimilation step forming the outer loop. In the experiments performed, the data assimilation step has a frequency of 24h, using all the observations included in the temporal range of 24h before the assimilation time. SANIFS is then used to project the analyzed fields forward to the next assimilation step.

In the next chapter, we will discuss the results of two assimilation experiments over the period 2017-2018:

1. STSLA: assimilation of salinity and temperature observations from Argo profiles and sea level anomaly from satellite altimetry missions (Saral/Altika, CryoSat-2, Jason-3 and Sentinel-3A).

2. ST: assimilation of salinity and temperature observations from Argo profiles.

The purpose of the two experiments is to evaluate quantitatively the improvement of SANIFS skills coupled with OceanVar and if the SLA observations improve the model

skills when assimilated in combination with S and T observations. In both experiments, we use 25 trivariate (S, T, SLA) EOFs derived from the variations in four years of SANIFS integration (2017 - 2020). As shown in chapter 2, the EOFs are grouped by month and computed at each node. The uRF algorithm has been implemented using a correlation radius of 8km and 4 iterations. Table 4.3 e Table 4.4 summarise the main informations related to the experiments setting.

MODEL	DOMAIN	GRID	COASTAL RESOLUTION	OPENS SEA RESOLUTION	LEVELS
SHYFEM SANIFS IMPLEMENTATION (Federico et al. (2017))	SANI (Fig. 1.2)	Unstructured Arakawa-B triangular mesh 90351 nodes	$50 - 500$ m	$3 - 4$ km	92 2 m up to 40 m stepwise up to the bottom max step 200 m

Table 4.3: Summarise of the ocean model features.

EXP NAME	PERIOD	SANIFS speed-up	ASSMILATION SCHEME	ASSIMILATION WINDOW	EOF	OBSERVATIONS USED	uRF CORR RADIUS	uRF ITERATIONS	NO MOTION LEVEL
STSLA	2017 - 2018	10 days	3DVar OceanVar (Dobricic and Pinardi (2008))	24 h before assimilation time	25 trivariate S, T, SLA	S,T,SLA	8 km	6	700 m
ST	2017 - 2018	10 days	3DVar OceanVar (Dobricic and Pinardi (2008))	24 h before assimilation time	25 trivariate S, T, SLA	S,T	8 km	6	700 m

Table 4.4: Summarise of the DA experiments.

5. Results of data assimilation experiments

In the previous chapters, we introduced and discussed the components of the assimilation system OceanVar-SHYFEM. In this chapter, we will discuss the results of the two experiments resumed in Table 4.4. Referring to the experiment STSLA, we will see how data assimilation is correcting the model temperature, salinity profiles and sea level anomaly locally in two example days, one in September 2017 and the other in December 2017. The choice of the two days is to show SANIFS skills along with data assimilation correction for temperature and salinity profiles and sea level anomaly on different seasons, which present different water column structure. Next, we will discuss the daily mean absolute error of misfits and residuals (MAE in Eq. (5.1)) for the three state variables over the entire period of experiments (2017-2018).

$$
\begin{aligned}
MAE &= \frac{1}{N} \sum_{i=1}^{N} |X_i| \\
RMSE &= \sqrt{\frac{1}{N} \sum_{i=1}^{N} X_i^2}
\end{aligned}
\tag{5.1}
$$

X_i represents the misfits or residuals and N is the number of those quantities in the defined temporal window. Also, temperature, salinity and sea level anomaly increment fields will be shown. Finally, to quantify the differences between the two experiments and the improvement introduced by the data assimilation, we will see the monthly root mean square of misfits and residuals (RMSE in Eq. (5.1)) for both the experiments at the surface layer and in the (2-60] m layer. The (2-60] m layer is so defined for the presence of the thermocline at different depths during different seasons.

5.1 Temperature, salinity and sea level anomaly misfits and residuals

To demonstrate the corrections introduced by DA, we discuss results in two days on 2017, 2017-09-05 and 2017-12-04, representing the summer and winter seasons. Fig. 5.1 and Fig. 5.2 show the Argo profiles and along-track satellite observation distribution over the two days respectively.

(a) 2017-09-05
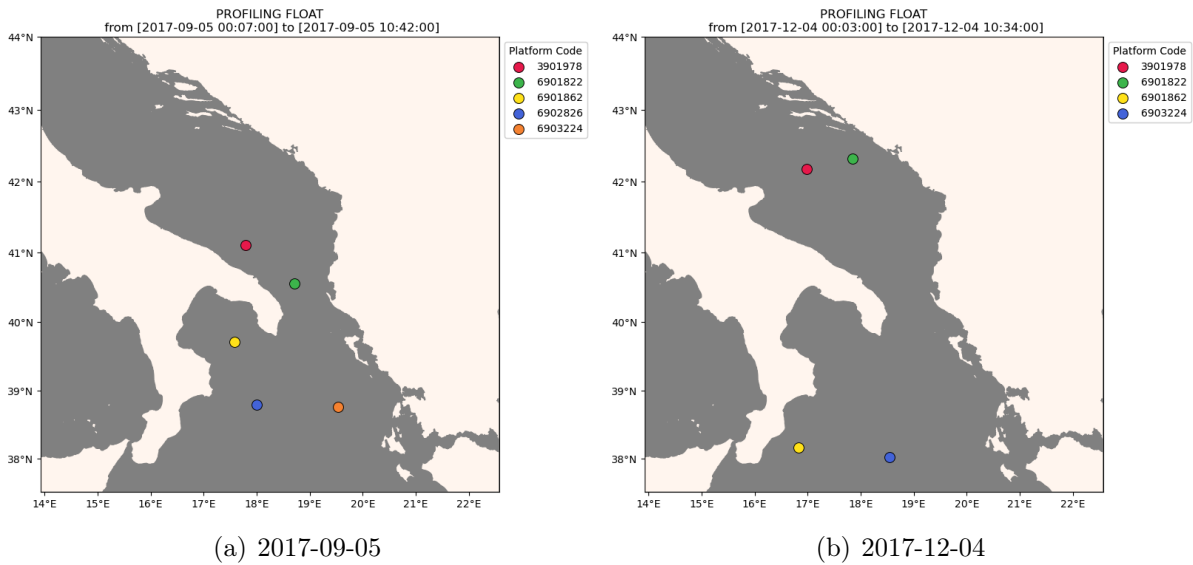
(b) 2017-12-04

Figure 5.1: Distribution of assimilated Argo profiles on 2017-09-05 (a) and 2017-12-04 (b).

(a) 2017-09-05
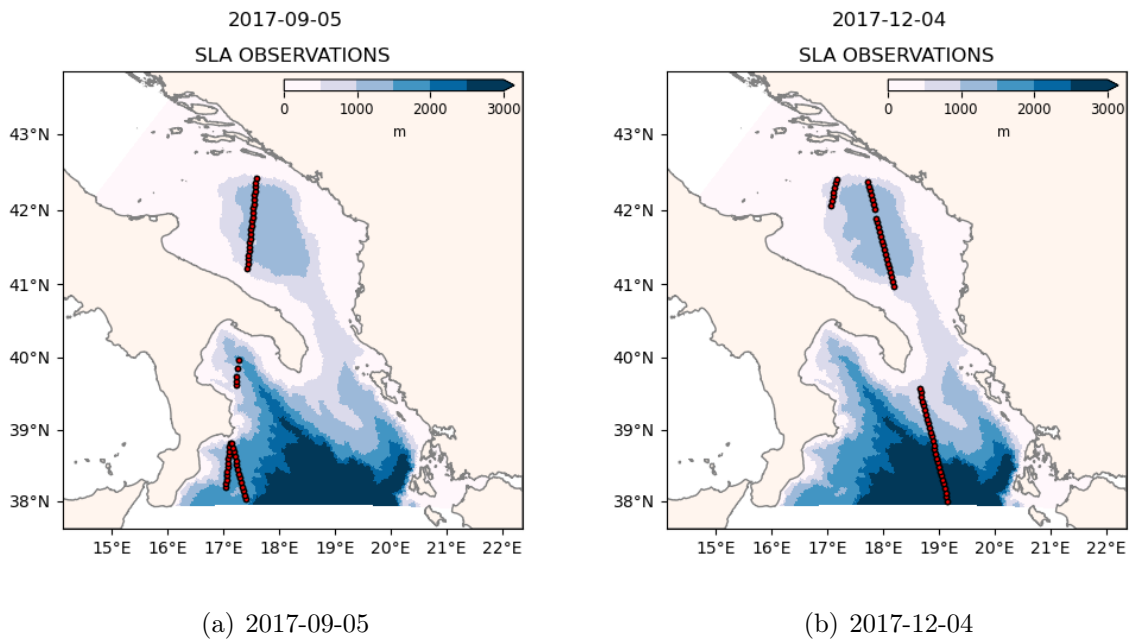
(b) 2017-12-04

Figure 5.2: Assimilated along-track satellite observation on 2017-09-05 (a) and 2017-12-04 (b).

On 2017-09-05, we have five Argo profiles (Fig. 5.1(a)). On 2017-12-04, there are four Argo profiles (Fig. 5.1(b)). On both days, the satellite track is almost crossing the entire SANIFS domain (Fig. 5.2(a) and Fig. 5.2(b)). As discussed in chapter 4, assuming a no motion level of 700 m in the dynamical height operator, only satellite along-track observations located in places with a depth greater than 700 m are assimilated.

Fig. 5.3 and Fig. 5.4 show results for temperature profiles, Fig. 5.5 and Fig. 5.6 instead show results for salinity profiles on 2017-09-05 and 2017-12-04 in the first 120 m. In the subfigures (e.g., Fig. 5.3(a)) the left side profiles refers to the value of the considered variable for model representation (blue line), observations (green line) and analysis (red line). The right side shows misfits (blue line) and residuals (red line).

All the figures show that the analysis reproduces temperature and salinity profiles in all the Argo locations better than the background in both the considered days. Indeed, looking at the profiles of misfits and residuals, we observe that over the water column, the residual profile is closer to zero with respect to the misfit profile. In particular, observing the temperature profiles, we see how the analysis corrects the sea surface temperature and the thermocline shape (e.g., Fig. 5.3 (b, d and e)). Observing misfits and residual profiles in Fig. 5.3(d), we see that around 40m the temperature correction reaches a value of about $3°C$. Referring to Fig. 5.3 (a and c) we can observe how only in some cases DA introduces a complete correction of the thermocline model representation. Indeed, the analysis thermocline is still at a shallower level than the observed one and exhibits a smoother gradient, as in the model profile.

In Fig. 5.4, we observe how the model can better reproduce the thermocline features in the representative winter day 2017-12-04. This is a consequence of the lower temperature gradient with respect to summer. During winter, the correction introduced by the analysis is almost constant along the water column depth with fewer oscillations of the residuals around zero.

Since ocean models have difficulty reproducing sea surface temperature and thermocline features, the correction introduced by DA on the temperature field is an important added value.

Observing salinity profiles Fig. 5.5 and Fig. 5.6, we see how the correction introduced is one magnitude order smaller than in the temperature. Also, in this case, DA introduces a correction in the surface layer and in the layers where there are fast variations of salinity with respect to depth. In Fig. 5.6(d), we can observe that DA introduces a small salinity correction with respect to the other salinity profiles.

To show how model and analysis reproduce temperature and salinity in deeper layers in the two days, we can observe Fig. 5.7 and Fig. 5.8. These figures are referred to Argo 6902826 for 2017-09-05 and 6901862 for 2017-12-04 in Fig. 5.1. For both the considered variables, we observe that both in summer and in winter, the model reproduces well the deep water features, and DA introduces minor corrections with respect to the upper layers,

where mixing phenomena occur.

To observe the correction introduced in sea level anomaly, we can observe Fig. 5.9. This figure shows along-track sea level anomaly misfits and residuals for the two considered days, 2017-09-05 (Fig. 5.9(a)) and 2017-12-04 (Fig. 5.9(b)). Fig. 5.9 shows that on 2017-09-05 (Fig. 5.9(a)) the misfits are higher with respect to 2017-12-04 (Fig. 5.9(b)). However, DA is partially correcting the model error. In fact, the residuals are closer to zero than the misfits in both Fig. 5.9(a) and Fig. 5.9(b) .

To have a global view of misfits and residuals of the STSLA experiment over the period 2017-2018, we plot in Fig. 5.10 the daily mean absolute of misfits and residuals for temperature and salinity at the surface (Fig. 5.10(a) and Fig. 5.10(b)) and (2-60] m (Fig. 5.10(c) and Fig. 5.10(d)) layer. Fig. 5.10(e) shows the daily mean absolute of misfits and residuals for sea level anomaly. As you can see, we have a consistent correction in temperature (by $2°C$) during the summer period on both layers (Fig. 5.10(a) and Fig. 5.10(c)). We observe that the model error decreases during winter and DA introduces a smaller correction. This is in accordance with the discussion on the profiles (Fig. 5.3 and Fig. 5.4).

Considering the salinity (Fig. 5.10(b) and Fig. 5.10(d)), we observe that there is a higher error during the last part of 2018, and there is no identified difference between the summer and winter seasons. Finally, Fig. 5.10(e) shows how DA is correcting the sea level anomaly with an average improvement of about 0.01 m.

To quantify the improvement introduced by DA and the differences between the two experiments summarized in Table 4.4, we can observe the monthly RMS of misfits over the period 2017-2018 for temperature and salinity on surface layer (Fig. 5.11(a) and Fig. 5.11(b)) and (2-60] m layer (Fig. 5.11(c) and Fig. 5.11(d)). Fig. 5.11(e) refers to sea level anomaly. The temperature on the surface and (2-60] m layers show the same improvement in both experiments with a relative gain with respect to the control run of 5% and 10% respectively. Salinity, on the considered layers, shows different improvements in the two experiments. On the surface layer, it has a relative gain of 2.4% in the STSLA experiment and 2.9% in the ST experiment. On (2-60] m layer STSLA experiment introduces an improvement of 5.3%. Instead, the ST experiment has an added value of 5.9%. Finally, the sea level anomaly in the STSLA experiment has an added value of 0.5%. Instead, the ST experiment has a negative gain of -3.2%. Table 5.1, summarizes the relative gains with respect to the control run of the two experiments for all the variables assimilated on the two considered layers.

Referring to the sea level anomaly case, we must remember that, in SANIFS, tidal boundary conditions are prescribed at boundaries. To assimilate the correct observation of sea level anomaly in the STSLA experiment, it is necessary to implement the FGAT interface in the ocean model. This allows us to consider the proper tide at the appropriate time to be added at sea level anomaly observation. Furthermore, as future improvement, also the Mean Dynamic Topography adjustment must be introduced. However, in the STSLA

experiment, the positive relative gain in the sea level anomaly RMS of misfits shows that the assimilation system is improving the model state.

To explain the small differences in temperature and salinity relative gain between the two experiments, we can observe the increment fields of the three assimilated variables in Fig. 5.12 and Fig. 5.13. We observe how, thanks to the EOFs cross-covariance matrix between temperature, salinity and sea level anomaly (introduced in chapter 2), the increments introduced in a variable also influence the others. However, the different order of magnitude is such that the sea level anomaly increment has a small impact on temperature and salinity fields. In those figures, the increment isolines of ± 0.002 are drawn. As you can observe, the temperature increments introduced by the sea level anomaly observation are relatively small compared to the magnitude of the temperature increments computed in the observation location. The same happens for salinity. However, the magnitude of salinity increments is closer to the sea level anomaly increments and in this case we have a different improvement in the two experiments. This consideration explains why the relative gain on temperature and salinity is small when sea level anomaly is also assimilated. In conclusion, the assimilation system implemented using SANIFS and OceanVar improves the model state. However, considering that SANIFS includes tides at the open boundaries, an FGAT interface is necessary to assimilate the sea level anomaly with appropriate tide observations.

EXPERIMENT	VARIABLE	LAYER	GAIN
STSLA	T	Surface	5.0%
ST	T	Surface	5.0%
STSLA	T	(2-60] m	10.0%
ST	T	(2-60] m	10.0%
STSLA	S	Surface	2.4%
ST	S	Surface	2.9%
STSLA	S	(2-60] m	5.3%
ST	S	(2-60] m	5.9%
STSLA	SLA	Surface	0.5%
ST	SLA	Surface	-3.2%

Table 5.1: Relative gain with respect to the control run in terms of RMS of misfits for the temperature and salinity on the surface and (2-60] m and sea level anomaly for the two experiments performed.
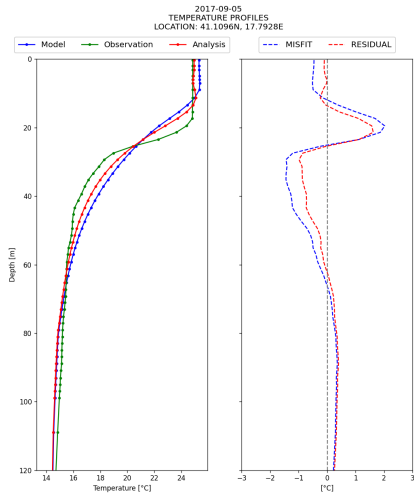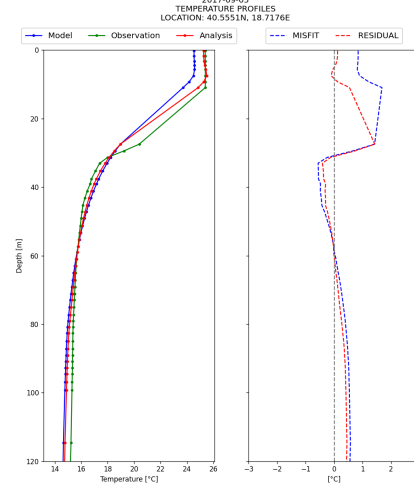
(a) Temperature
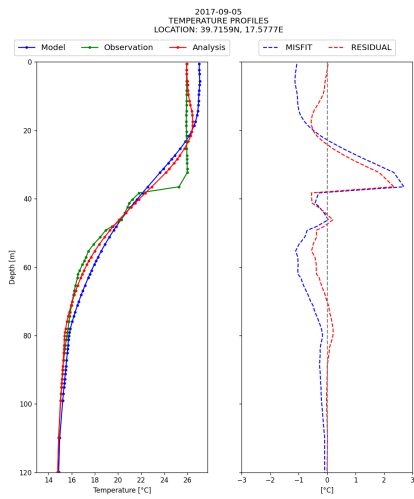
(b) Temperature
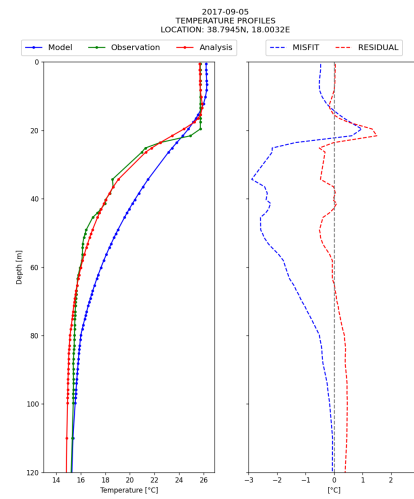
(c) Temperature
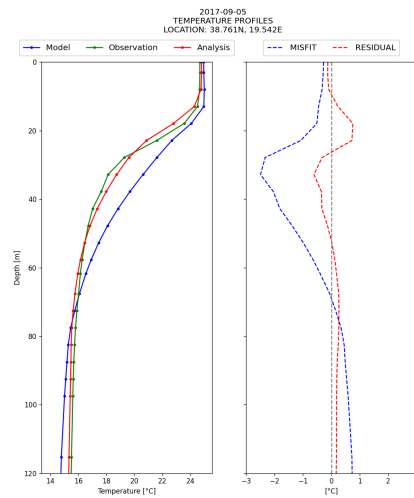
(d) Temperature

(e) Temperature

Figure 5.3: Temperature profiles on 2017-09-05.

(a) Temperature

(b) Temperature

(c) Temperature
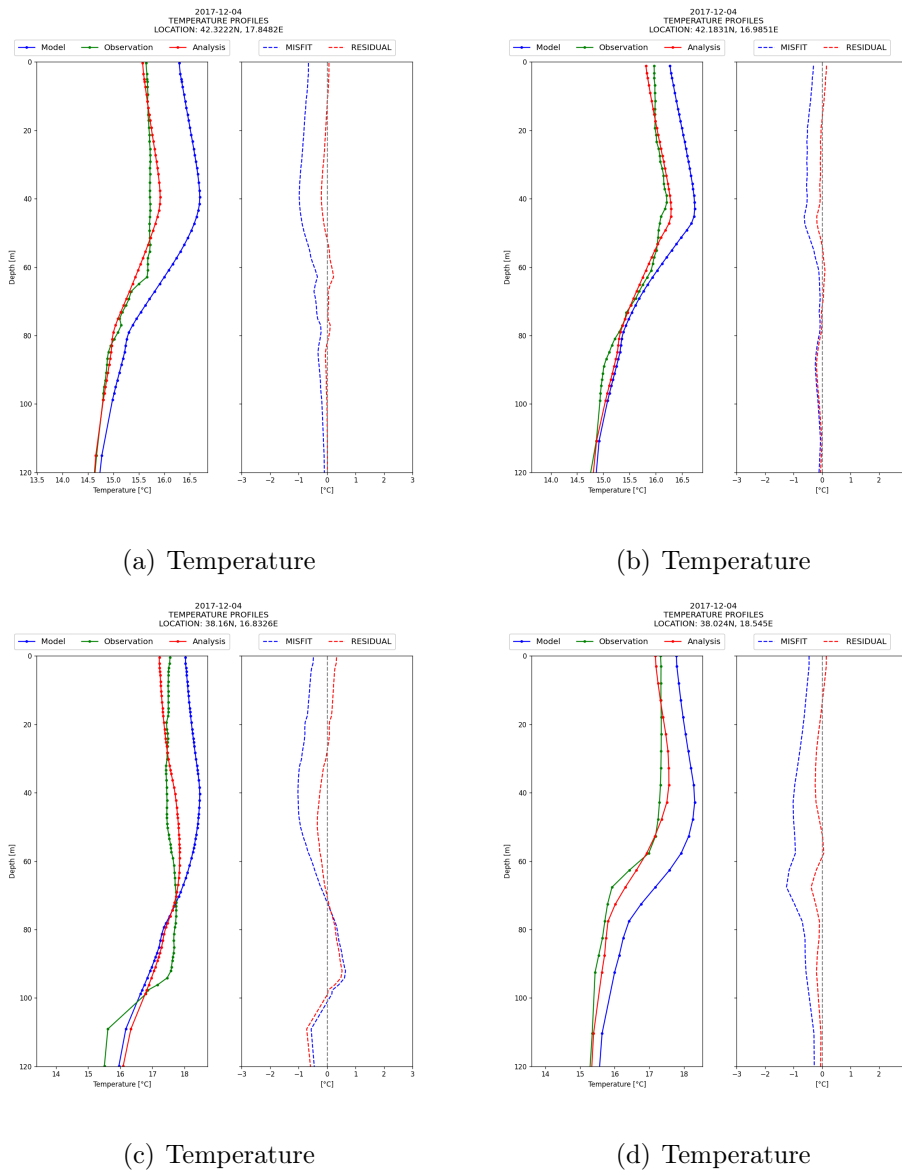
(d) Temperature

Figure 5.4: Temperature profiles on 2017-12-04.
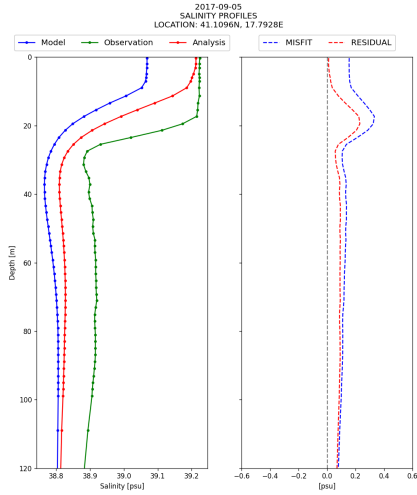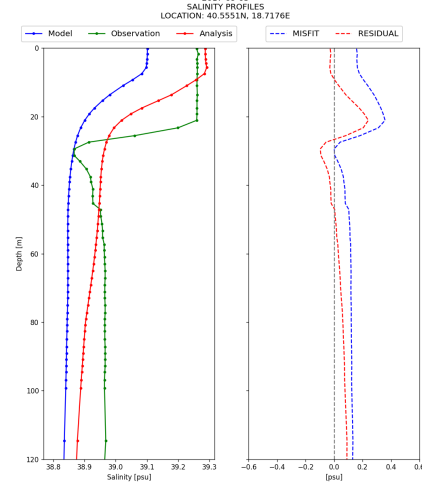
(a) Salinity
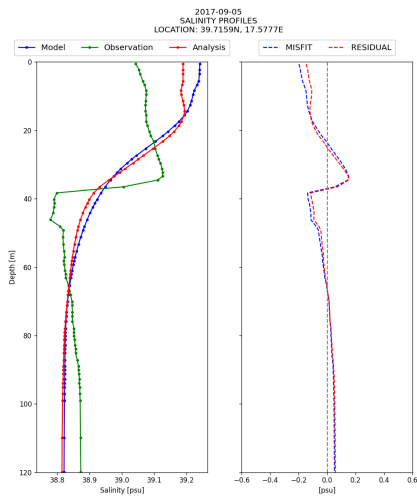
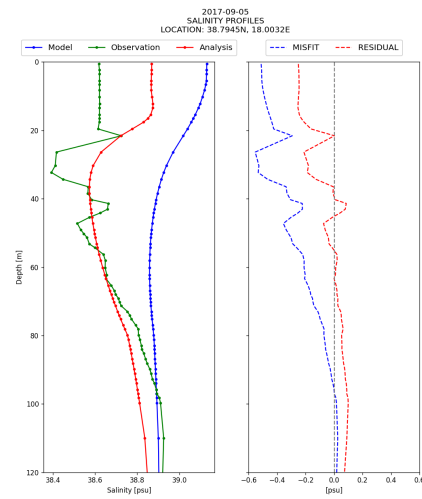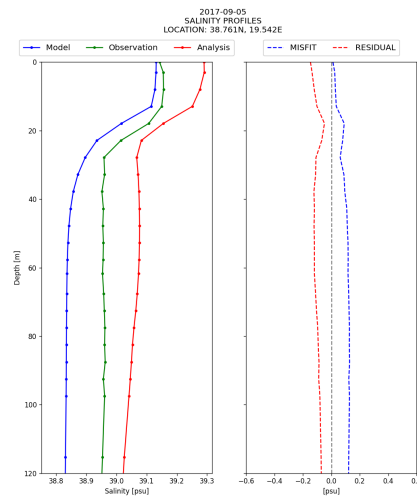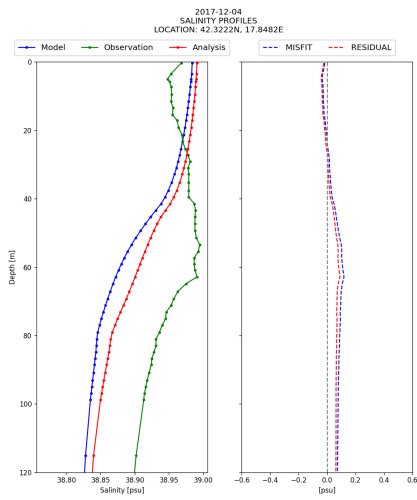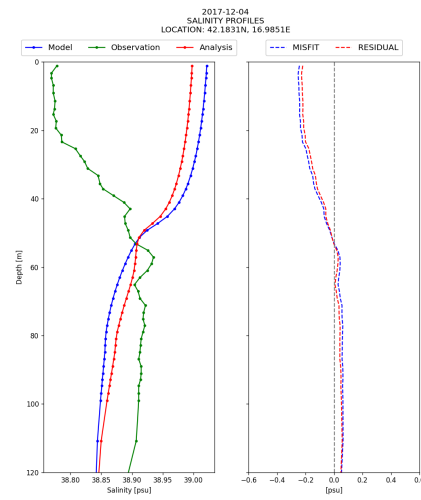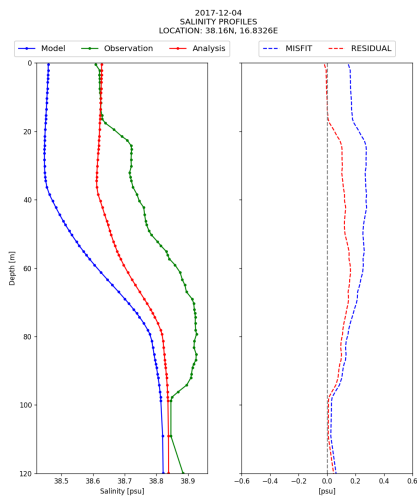(b) Salinity

(c) Salinity

(d) Salinity
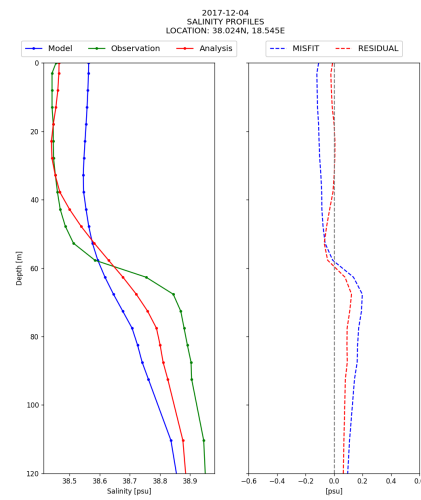

(e) Salinity

Figure 5.5: Salinity profiles on 2017-09-05.

(a) Salinity

(b) Salinity

(c) Salinity

(d) Salinity

Figure 5.6: Salinity profiles on 2017-12-04.

(a) 2017-09-05   (b) 2017-12-04
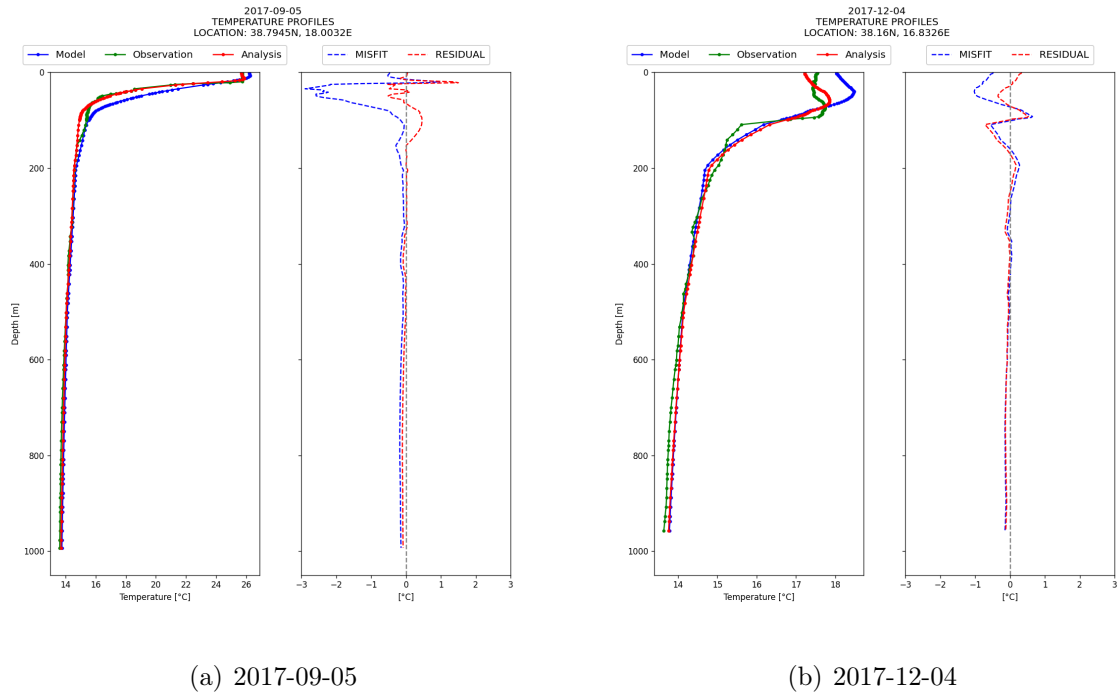
Figure 5.7: Temperature profiles on 2017-09-05 (a) and 2017-12-04 (b). The reference for platform code in Fig. 5.1 is 6902826 in (a) and 6901862 in (b). The purpose is to show deeper layers with respect to Fig. 5.3 and Fig. 5.4.

(a) 2017-09-05   (b) 2017-12-04
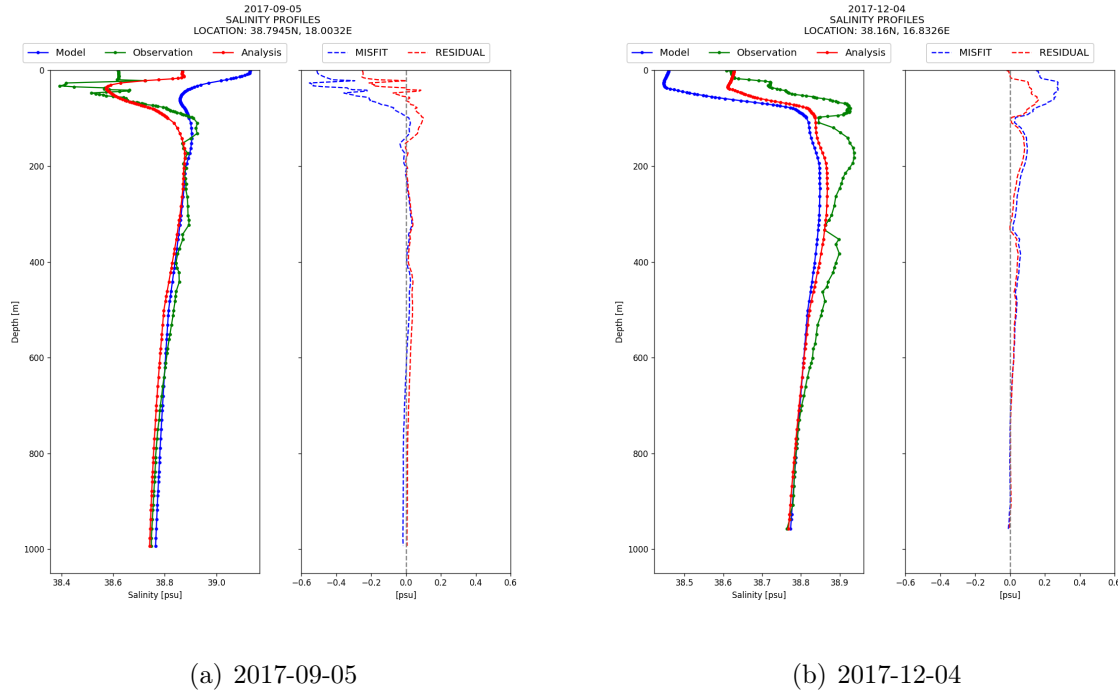
Figure 5.8: Salinity profiles on 2017-09-05 (a) and 2017-12-04 (b). The reference for platform code in Fig. 5.1 is 6902826 in (a) and 6901862 in (b). The purpose is to show deeper layers with respect to Fig. 5.6 and Fig. 5.10.

(a) 2017-09-05
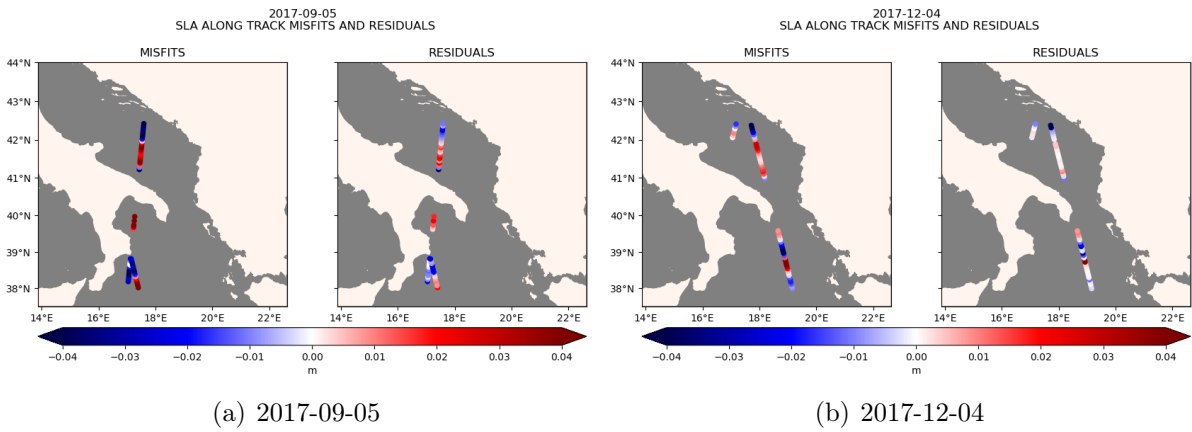
(b) 2017-12-04

Figure 5.9: Sea level anomaly along track misfits and residuals on 2017-09-05 (a) and 2017-12-04 (b).

(a) Sea surface temperature

(b) Sea surface salinity

(c) Temperature in layer (2-60] m

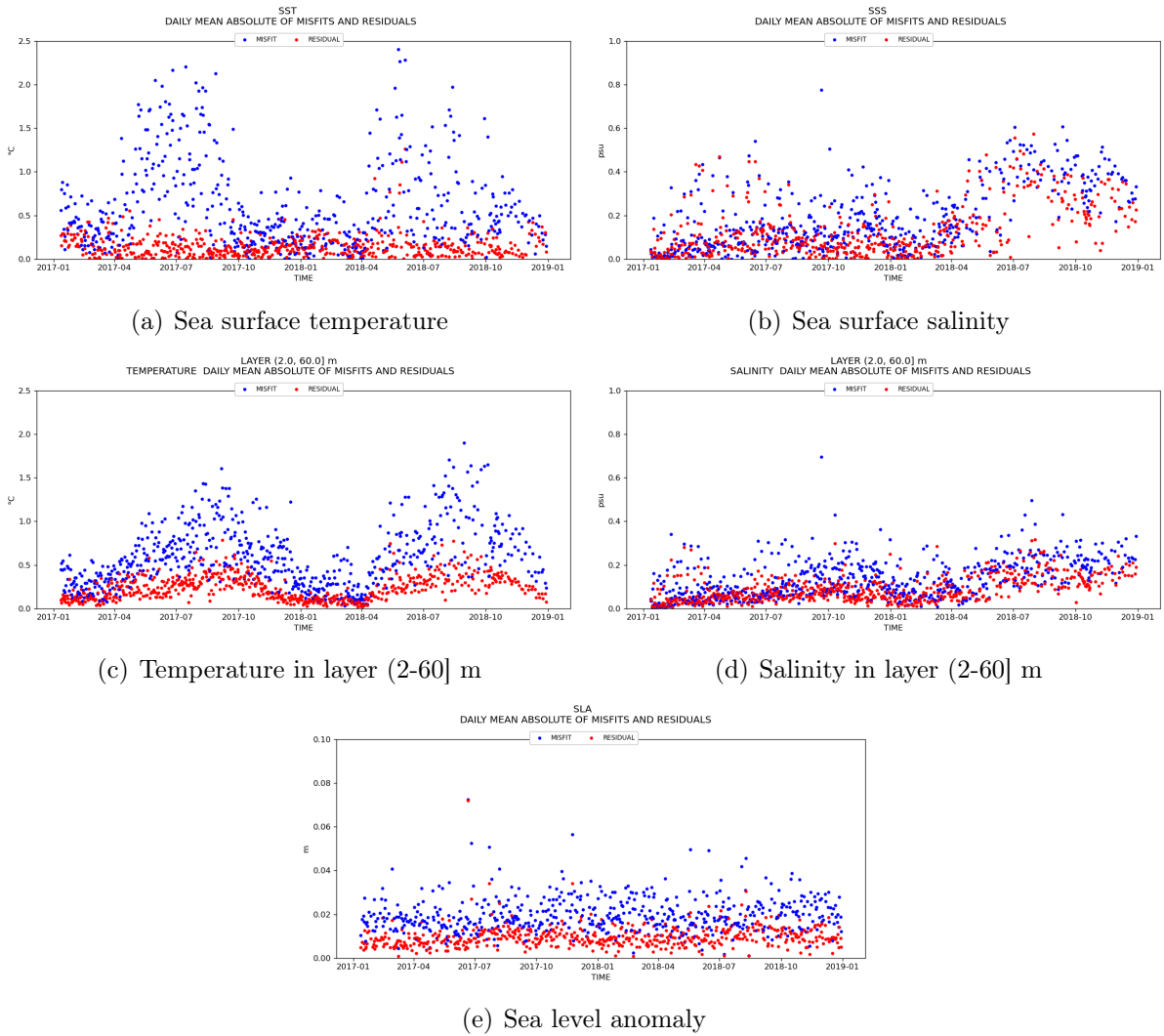(d) Salinity in layer (2-60] m

(e) Sea level anomaly

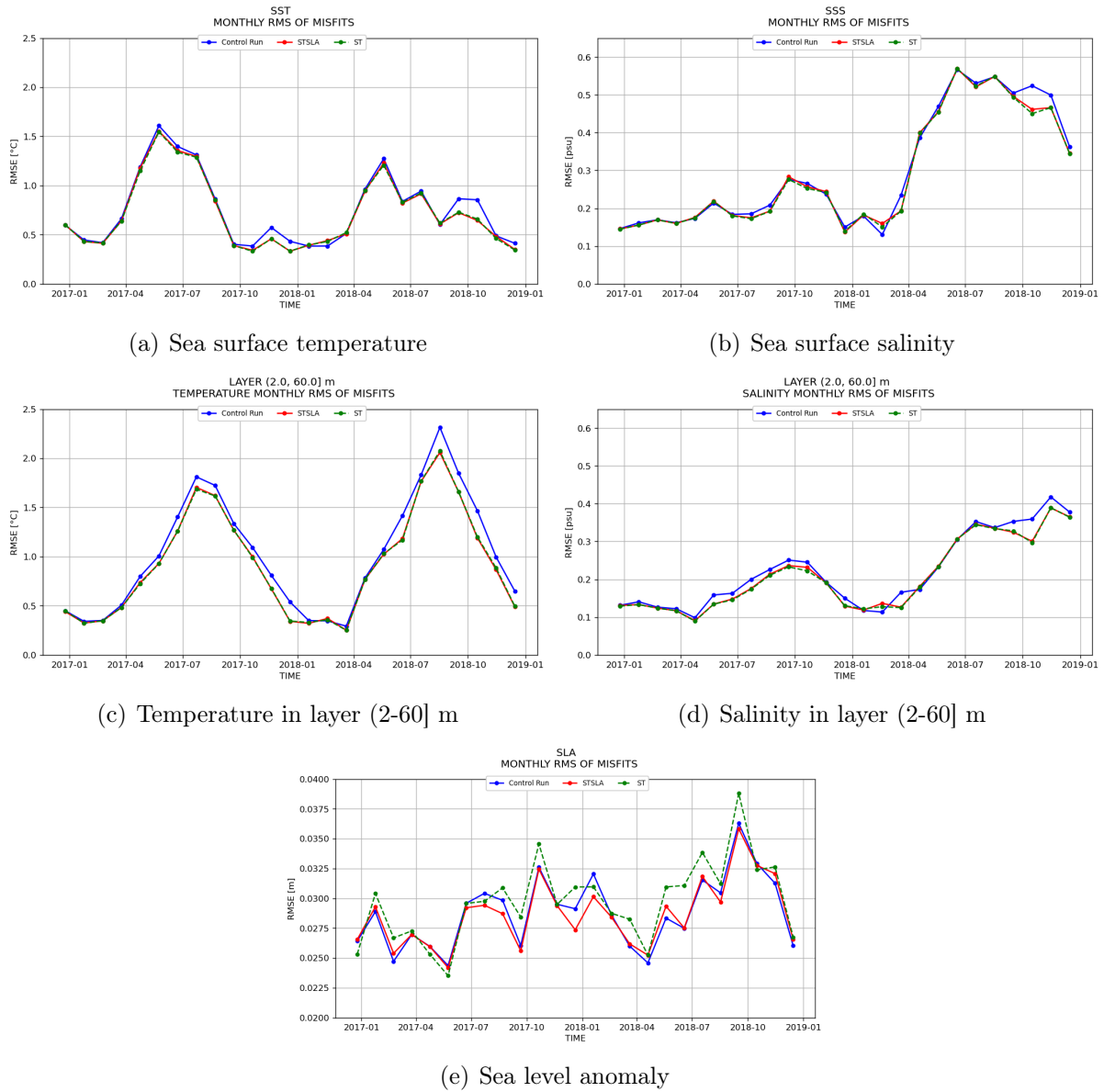Figure 5.10: Daily mean absolute of misfits and residuals over the period 2017-2018. Sea surface temperature (a), sea surface salinity (b), temperature in layer (2-60] m (c), salinity in layer (2-60] m (d), sea level anomaly (e).
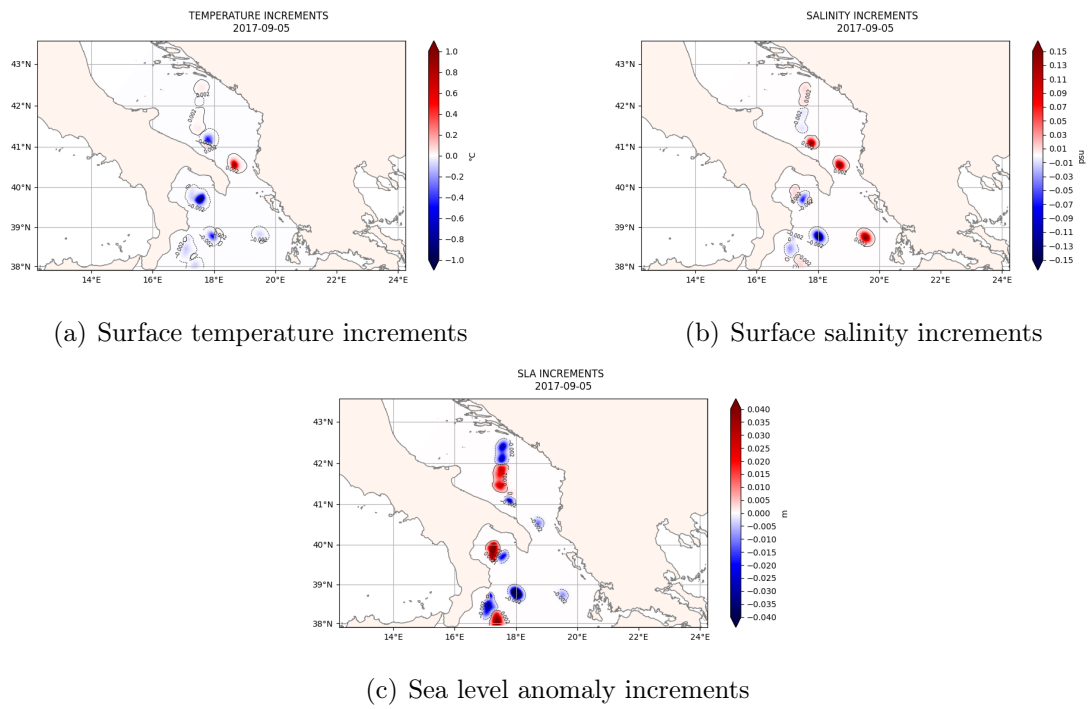
(a) Sea surface temperature

(b) Sea surface salinity

(c) Temperature in layer (2-60] m

(d) Salinity in layer (2-60] m

(e) Sea level anomaly

Figure 5.11: Monthly root mean square of misfits over 2017-2018. Sea surface temperature (a), sea surface salinity (b), temperature in layer (2-60] m (c), salinity in layer (2-60] m (d), sea level anomaly (e).

(a) Surface temperature increments

(b) Surface salinity increments

(c) Sea level anomaly increments

Figure 5.12: Increments on 2017-09-05. Surface temperature (a), surface salinity (b) and sea level anomaly along track (c).
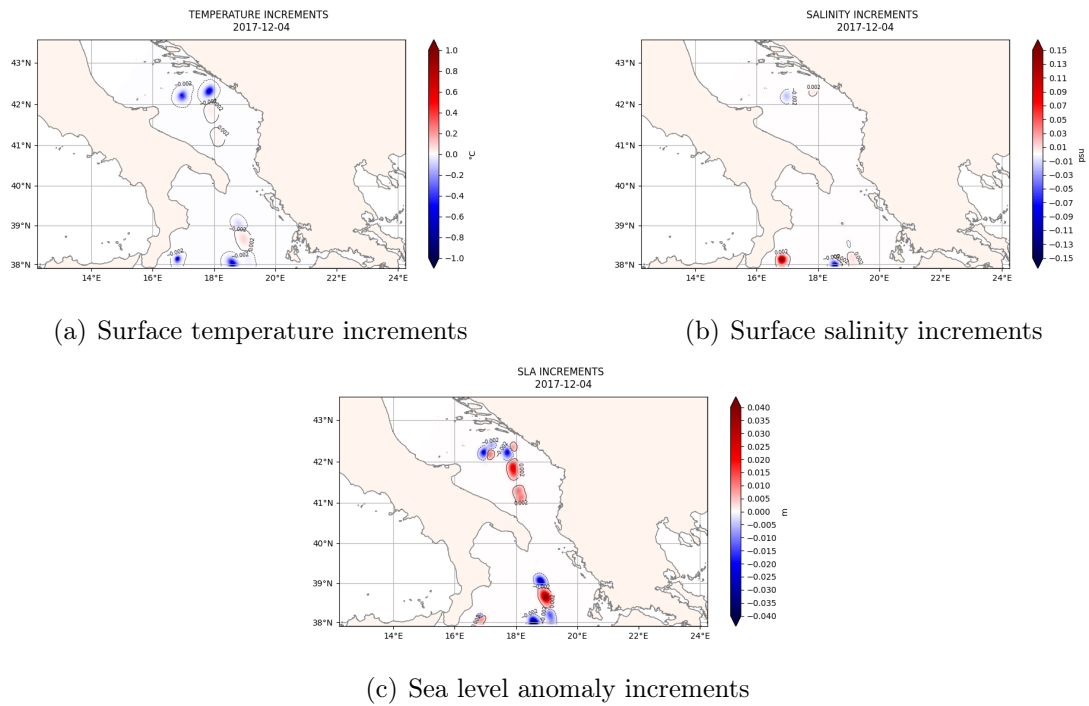

(a) Surface temperature increments

(b) Surface salinity increments

(c) Sea level anomaly increments

Figure 5.13: Increments on 2017-12-04. Surface temperature (a), surface salinity (b) and sea level anomaly along track (c).

6. A Long Short Term Memory algorithm for increment prediction

In the last decade, we have witnessed the rapid growth of Artificial Intelligence (AI) applications in different fields. When talking about AI, we are referring to the branch of computer science which aims to develop intelligent systems and machines that can carry out tasks that typically require human intelligence. A more complex family of AI is Machine Learning (ML) algorithms. ML aims to create algorithms to learn from data and make decisions based on observed patterns. ML requires human intervention when the decision is incorrect. A subset of ML is Deep Learning (DP). DP uses an artificial neural network (ANN) to reach accurate conclusions without human intervention (Fig. 6.1).
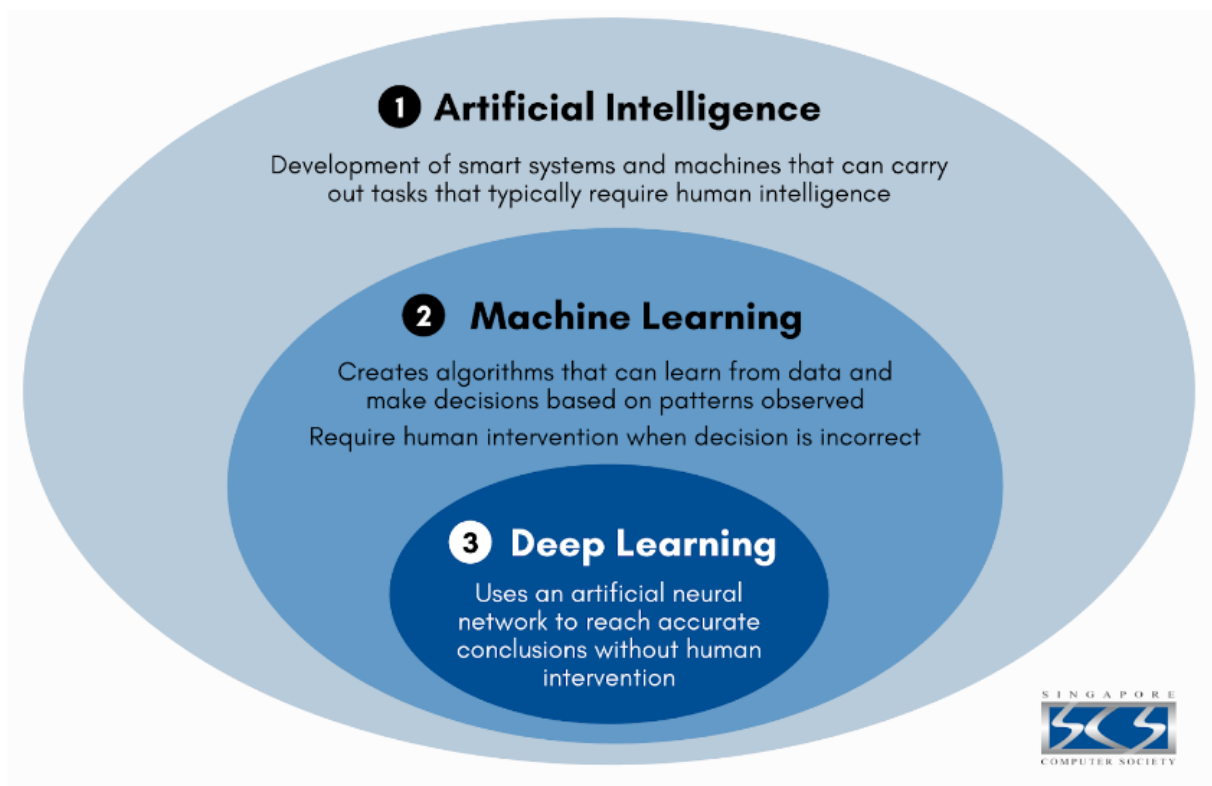

Figure 6.1: AI, ML and DL classification. From *https://www.scs.org.sg/*

In geoscience, DL applications are largely used on different systems. There are applications related to weather and climate forecasting and extreme events forecasting at a

global and local scale (e.g., Dueben and Bauer (2018), Scher and Messori (2019), Zheng et al. (2019)). In oceanography, there are ML applications for sea surface temperature prediction (e.g., Xiao et al. (2019b), Xiao et al. (2019a), Lins et al. (2013)), parameter prediction for wave models and wave height forecasting (e.g. Makarynskyy (2004), Agrawal and Deo (2002), Jain and Deo (2008), James, Zhang, and O'Donncha (2018)), eddy forecasting and classification (e.g., Bai, Wang, and Li (2019), Lguensat et al. (2018)). During the last years, since the large interest in assessing the risk of coastal areas due to the sea level rise related to climate change, the main ML applications in oceanography are focused on sea level forecast (e.g., Accarino et al. (2021), Raj et al. (2022), Wang et al. (2022)) and storm surge events forecast (e.g., Rus et al. (2022), Ayyad, Hajj, and Marsooli (2022), Zhang et al. (2022), Vincent et al. (2022)). Moreover, there is a growing interest in integrating DA with ML. The purpose of these applications is mainly to reduce the computational cost of DA problems in case of complex dynamical systems (e.g., Gao et al. (2022), Buizza et al. (2022), Lee et al. (2022)). Furthermore, Farchi et al. (2021) proposed to use machine learning to correct the model error in data assimilation and forecast applications.

All the DL applications that involve predicting time series data use the Long Short Term Memory (LSTM) technique. LSTM is a neural network methodology with a strong learning and predictive ability for time-series data.

In this thesis, we propose an LSTM Luong-attention (Luong, Pham, and Manning (2015)) model trained on available SST increments data extracted from the trivariate experiment assimilating T, S and SLA that we discussed in the previous chapter. This model can predict the SST increments. The predicted SST increments can be used to correct the model forecast, which can not benefit DA since no observations are available.

In the following sections, the LSTM Luong-attention algorithm and the results of its application on SST increments prediction will be presented.

6.1 The Long-Short Term Memory network

The Long Short-Term Memory (LSTM) model is a Recurrent Neural Networks (RNN) subtype. LSTMs have feedback connections that differentiate them from ANN, also known as feedforward neural networks. This property enables LSTMs to process entire data sequences without treating each point in the sequence independently. Rather, retaining useful information about previous data in the sequence to help process new data points. As a result, LSTMs are particularly good at processing data sequences such as text, speech and general time series.

To understand how RNN works, looking at how ANN are structured is helpful. In these, a neuron (also called node or unit) of the hidden layer is connected with the neurons from the previous layer and the neurons from the following layer (Fig. 6.2). In ANN, the

output of a neuron can only be passed forward, but never to a neuron on the same layer or even the previous layer, hence the name feedforward. This is different for RNN, also named feedback neural networks, since the output of a neuron can be used as input for a previous or current layer. For this reason, they are called recurrent. This is much closer to how our brain works than how feedforward neural networks are constructed. Indeed, while you are reading this thesis, your brain's neurons share information at the same level of learning/understanding (layer). Also, the meaning of each word is linked to the sense of the previous one. This means that the learned information is an input for the next input word, which, in turn, adds meaning to the previous one to have a comprehensive understanding. Let's n the number of neurons of the considered layer (input or hidden
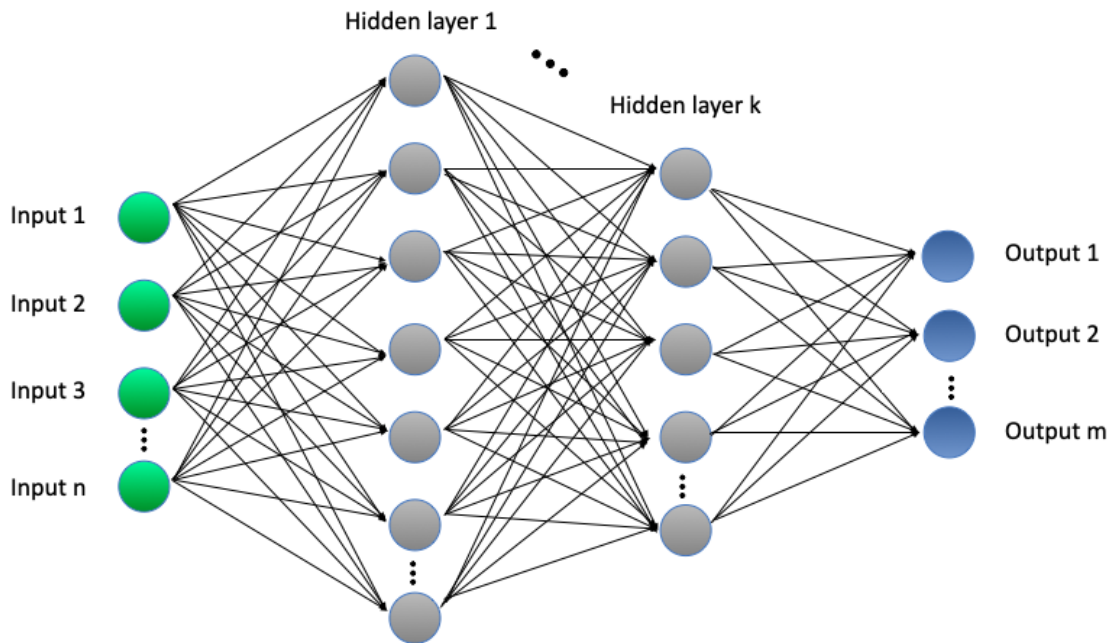

Figure 6.2: ANN or feedforward neural network.

layer) in ANN. To produce an output from a neuron, three different steps are executed at the left side of it in Fig. 6.2:

1. On each neuron the inputs, x_i, are multiplied by a weight factor, w_i; $i = 1, 2, \ldots, n$ (Eq. (6.1)). The weight factors decide how important the input is to solve the considered problem. If an input x_i is very important, the value for the factor w_i becomes larger. An unimportant input is multiplied by a value of 0.

$$x_i \rightarrow x_i \cdot w_i \quad i = 1, 2, \ldots, n \tag{6.1}$$

2. All weighted inputs of a neuron are summed and a bias is added (Eq. (6.2))

$$s_i = \sum_{i=1}^{n} (x_i \cdot w_i) + b_i \quad i = 1, 2, \ldots, n \tag{6.2}$$

3. The result is given into a so-called activation function, σ (Eq. (6.3)). The primary role of the activation function is to transform the summed weighted input from the node into an output value to be fed to the next hidden layer or as output. The purpose of the activation function is, besides introducing nonlinearity into the neural network, to bind the value of the neuron so that the neural network is not paralyzed by divergent neurons (Wang (2003)). Different activation functions can be chosen: Sigmoid, Tanh, Rectified Linear Unit (ReLU), Leaky ReLU (Fig. 6.2) to cite a few commonly used. The activation function's choice depends on the problem you are facing. For example, the sigmoid activation function is used to represent probability output since it maps values in the range between 0 and 1.

$$h_i = \sigma(s_i) = \sigma \left(\sum_{i=1}^{n} (x_i \cdot w_i) + b \right) \quad i = 1, 2, \ldots, n \tag{6.3}$$
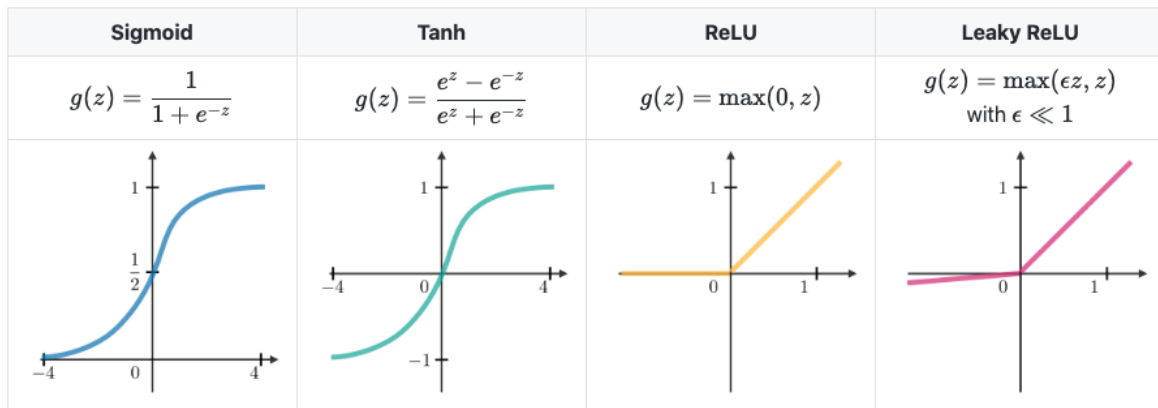
Sigmoid	Tanh	ReLU	Leaky ReLU
$g(z) = \dfrac{1}{1+e^{-z}}$	$g(z) = \dfrac{e^z - e^{-z}}{e^z + e^{-z}}$	$g(z) = \max(0, z)$	$g(z) = \max(\epsilon z, z)$ with $\epsilon \ll 1$

Figure 6.3: Examples of activation functions. From. *https://stanford.edu/*

Now, it is clear how ANN is trying to emulate our brain. Indeed, considering the example of the reader of this thesis, the process is not simply reading the text word by word and understanding the sentences, but also weighting how each word is important for understanding the text, summing all the information, and finally matching what you read in the framework of a Ph.D. thesis evaluation (the activation function purpose). However, more is needed. Reading the next word, you need to remember the information from the previous one; otherwise, you cannot understand the sentence's overall meaning. In other words, it is necessary to have a memory of the previous sequence of words to understand the sentence. The process of remembering each word of a sentence to understand the

meaning of the sequence and predict the next word is named the short-term memory feature of a neural network. Furthermore, remembering each word is not enough to predict the next one and have a complete understanding of the sequence and be able to predict the next word; it is also necessary to update the meaning of the previous word with the intention of the current one. RNN overcomes the short-term memory problem of ANN and also recursively updates the weight of each word for the prediction process.

RNN consists of multiple layers similar to ANN (Fig. 6.2). However, RNN has a recurrent unit on each neuron of the hidden layers (Fig. 6.4). A recurrent unit processes informa-
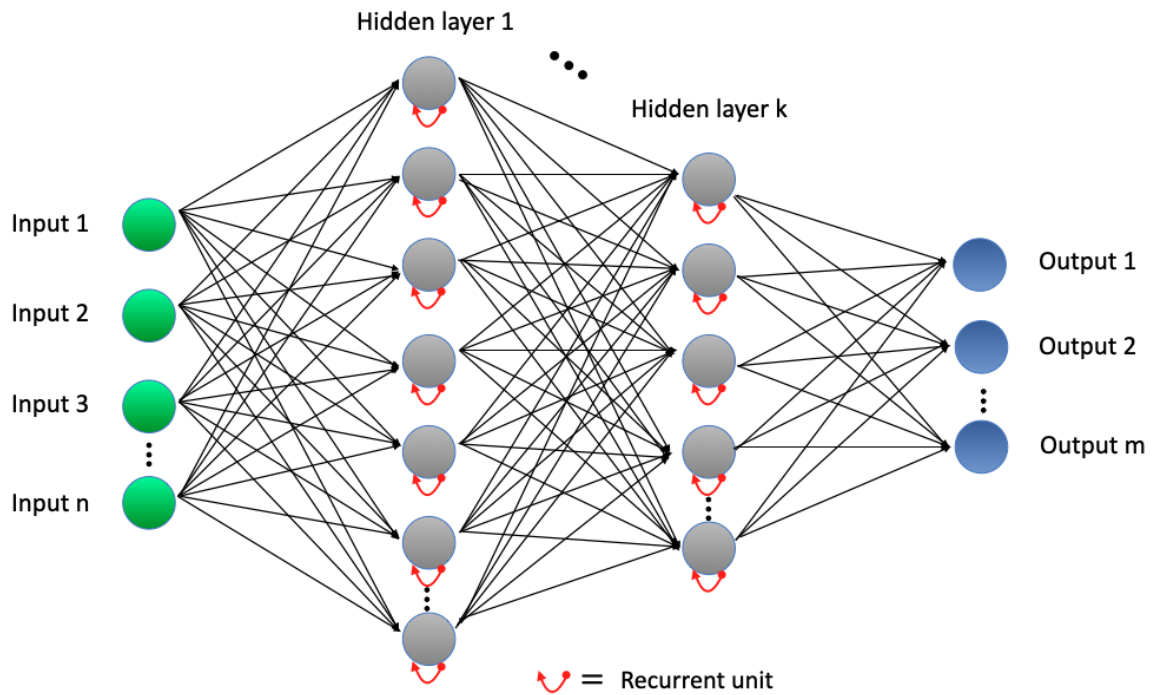
Figure 6.4: RNN or feedback neural network.

tion for a predefined number of timesteps. Let's consider the problem of predicting the following 24 timestep increment field from the previous 24 timesteps. Each input x_t is an input for the recurrent unit, producing an output that will feed in the hidden neuron h_t. h_t will be passed back, and with the following timestep input, x_{t+1} will contribute to the h_{t+1} for the same hidden neuron. Predicting n timestep from the previous n is known as many-to-many or sequence-to-sequence RNN Fig. 6.5. While ANN networks have different weights across each node, recurrent neural networks share the same weight parameter within each network layer. The weights are adjusted in the backpropagation and gradient descent processes to facilitate reinforcement learning. The algorithm of backpropagation is a backpropagation through time (BPTT) algorithm. It is used to determine the gradients. The principle of BPTT is that the model trains itself by calculating errors from its output layer to its input layer. These calculations allow us to adjust and fit the model's parameters appropriately. BPTT differs from the traditional approach in that BPTT sums errors at each time step, whereas feedforward networks do not need to sum errors
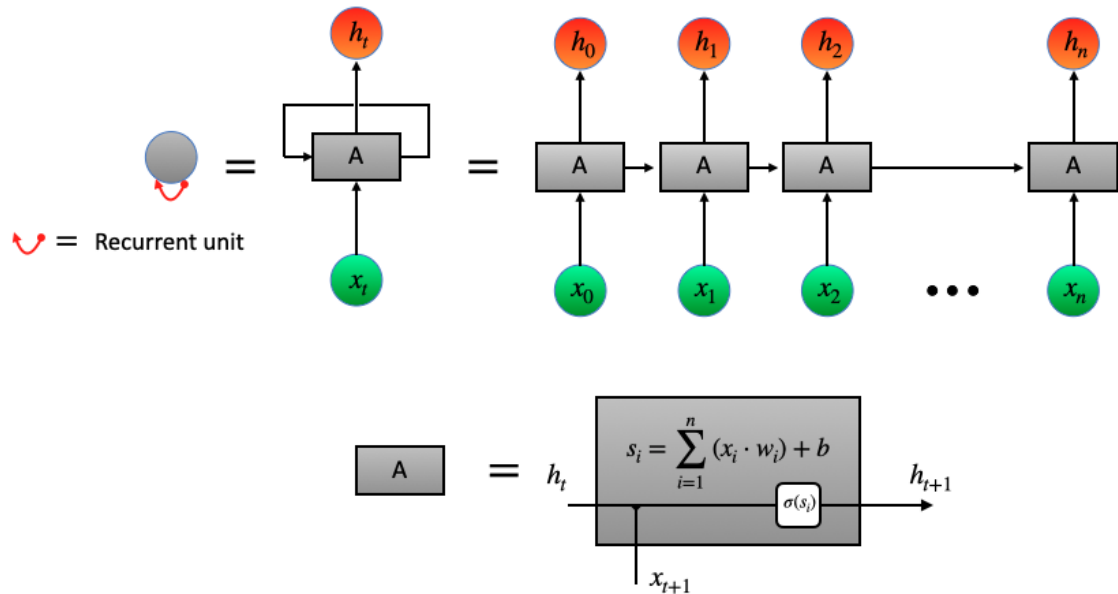
Figure 6.5: Recurrent unit structure in a many-to-many (also named sequence-to-sequence) RNN. $\sigma(s_i)$ represent the activation function. h_t is the hidden state and x_t is the input state at timestep t.

as they do not share parameters across each layer. Through this process, RNNs tend to run into two problems, known as exploding gradients and vanishing gradients (Bengio, Simard, and Frasconi (1994)). These issues are defined by the size of the gradient, which is the slope of the loss function along the error curve. When the gradient is too small, it continues to become smaller, updating the weight parameters until they become zero. When that occurs, the algorithm is no longer learning. Exploding gradients occur when the gradient is too large, creating an unstable model. In this case, the model weights will grow too large and eventually be represented as NaN. One solution to these issues is to reduce the number of hidden layers within the neural network, eliminating some of the complexity in the RNN model. However, more than a short-term memory is needed for the reader's purpose. A memory of the meaning of the previous sentences is also necessary to completely understand the text. For that reason, it is necessary to have a short-term memory combined with a long-term memory. LSTM networks have been introduced (Hochreiter and Schmidhuber (1997)) to overcome the long-term and the gradient problem of RNN.

LSTM is a particular type of RNN. LSTM has the same structure as RNN (Fig. 6.4 and Fig. 6.5); however, the recurrent unit structure is different (Fig. 6.6). The forget gate controls what information should be forgotten or retained. The input gate helps to identify essential elements that need to be added to the cell state. Note that the results of the input gate get multiplied by the cell state candidate, with only the information deemed important by the input gate being added to the cell state. To obtain the updated cell state first, the previous cell state c_t gets multiplied by the results of the forget gate. Then we add new information from [input gate × cell state candidate] to get the latest cell
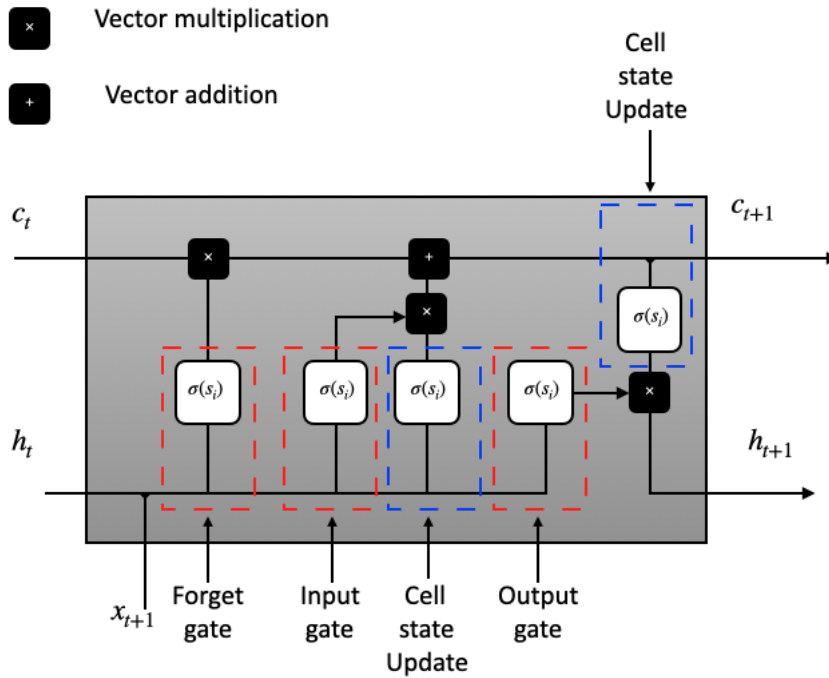
Figure 6.6: LSTM recurrent unit. $\sigma(s_i)$ represent the activation function. h_t is the hidden state, c_t is the cell state and x_t is the input state at timestep t.

state c_{t+1}. In the last part, the hidden state is updated. The newest cell state c_{t+1} is passed through the activation function and multiplied by the results of the output gate h_{t+1}. Finally, the latest cell state c_{t+1} and the hidden state h_{t+1} go back into the recurrent unit, and the process repeats at timestep $t + 2$. The loop continues until we reach the end of the last timestep of the sequence. This is a great step forward in emulating the human brain. However, for the reader's purpose, more is needed. In this thesis, some sections need more attention than others. The information in that section must be stored in memory with a higher importance value to understand the other sections and chapters. The reader purpose has the same structure as the increment SST field prediction. Indeed, looking at Fig. 6.7, we observe that the SST increments field is not homogeneous over the SANIFS domain but has some peaks corresponding with observation locations. This means that the values in these locations must be stored in memory with a higher importance value. Furthermore, the increments computed for the winter season differ from those calculated for the summer. So, the DL algorithm must remember these features for long-term memory with the appropriate attention. A simple RNN, which has only short-term memory abilities, cannot achieve this. So, to accomplish the increment prediction problem, we will use an LSTM attention-based RNN algorithm. The attention algorithm implemented in this work is the Luong attention algorithm (Luong, Pham, and Manning (2015)) used in text translation. The structure of the Luong attention is an LSTM network with the introduction of an attention layer. The goal is to derive a

SST INCREMENTS MAP OVER 6 DIFFERENT ASSIMILATION TIME ON 2017 AND 2018
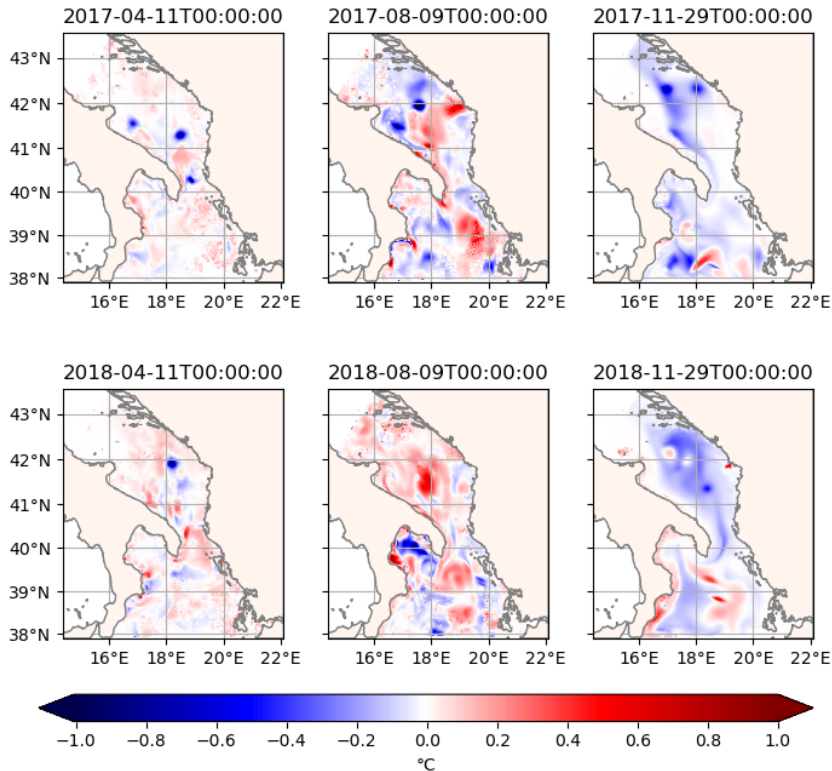

Figure 6.7: SST increments over six different assimilation times on SANIFS over 2017-2018.

cell vector c_t that captures relevant source-side information to help predict the current target feature (Luong, Pham, and Manning (2015)). The core idea is to calculate and distribute the attention weight, and the focus is placed on important content by increasing its weight. This addresses the problem of the sequence-to-sequence model losing information when the input sequence is too long. Taking a typical machine translation scenario as an example, the sequence-to-sequence model only uses the final hidden state of the input sentence to calculate the cell vector. In contrast, the attention layer calculates the individual cell vector for each word of the input sentence (Yang, Li, and Liu (2021)). In the next section, we will present the methodology used for the LSTM-Luong attention model implementation and we will discuss the results.

6.2 Methodology and results

In chapter 4 and 5, we presented the SANIFS-OceanVar assimilation system and discussed the results. We saw an experiment assimilating T, S, and SLA over 2017-2018. The ocean model produced the hourly output of the ocean state. Considering only the surface layer of the temperature model variable, we can compute the SST increments as the difference between STSLA (Table 4.4) and control run experiments. The dimensionality of the SST increment field is 17232 (hourly output from DA experiments) timesteps

times 90351 SANIFS grid nodes. From a computational point of view, this is too high dimensionality to be used to train an LSTM model. Using the methodology discussed in chapter 2, PCs and EOFs analysis is used to reduce the space dimensionality. So, we apply the LSTM model in the reduced PCs space and then back on the SST increment space by the multiplication of PCs with EOFs. Indeed, if n is the number of PCs/EOFs, T is the time dimensionality, and S is the space dimensionality of the physical space Eq. (6.4) holds.

$$\delta_{\textbf{SST}} = \underset{[T \times n]}{\textbf{PC}} \times \underset{[n \times S]}{\textbf{EOF}} \tag{6.4}$$

where $\delta_{\textbf{SST}}$ identifies the SST increments space. As discussed in chapter 2, the correct choice of n is important to represent the correct features of the considered field. Fig. 6.8 - Fig. 6.11 show the true SST increment field (first row) and the reduced space using n = 25, n=125, n=200 and n=500 PCs/EOFs (second row) respectively over three different timesteps of 2017 (a) and three different timesteps of 2018 (b). Here we observe how the chosen n affects the features representation of the six timesteps of the SST increment field shown in Fig. 6.7.   Using 25 PCs/EOFs, we observe that the increment features
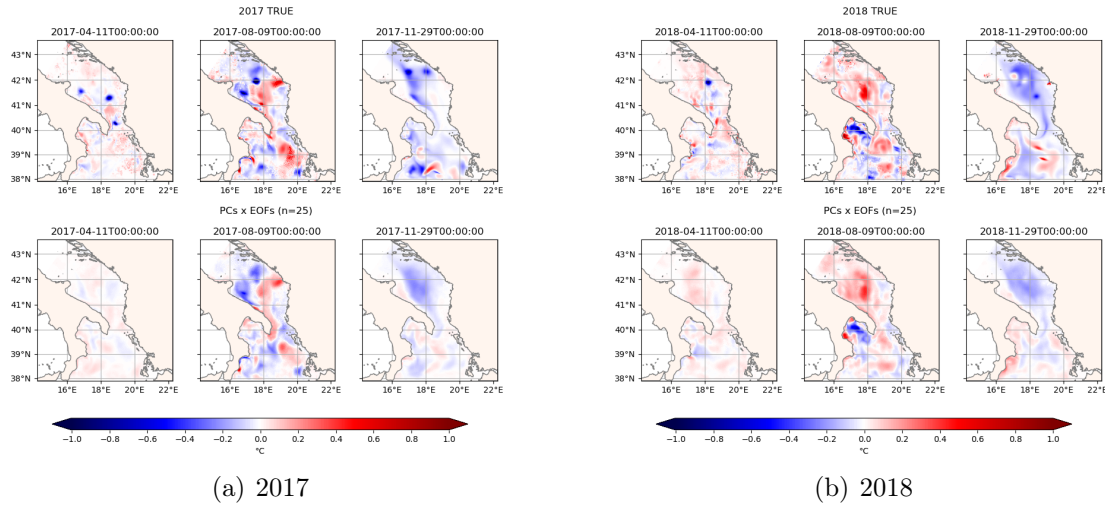

Figure 6.8: True (first row) and reduced space using $n = 25$ PCs/EOFs (second row) SST increment field over 3 different timesteps of 2017 (a) and 3 different timesteps of 2018 (b)

are not well represented and we can only represent the average field shape. Increasing n, we also increase the details that PCs/EOF can represent. For example, looking at the timestep referred to 2017-04-11T00:00:00, we observe in Fig. 6.8 - Fig. 6.11 how increasing n we are increasing the SST increments representation in the observation location. This happens because a higher number of modes capture more field variability. So, to train an LSTM attention-based model that can predict the SST increments field, it is more appropriate to use $n = 500$ PCs/EOFs. Each PCs mode will represent a different time series input for the LSTM model. However, the drawback of using $n = 500$ is that it increases the training process's complexity. Indeed, since the higher modes represent
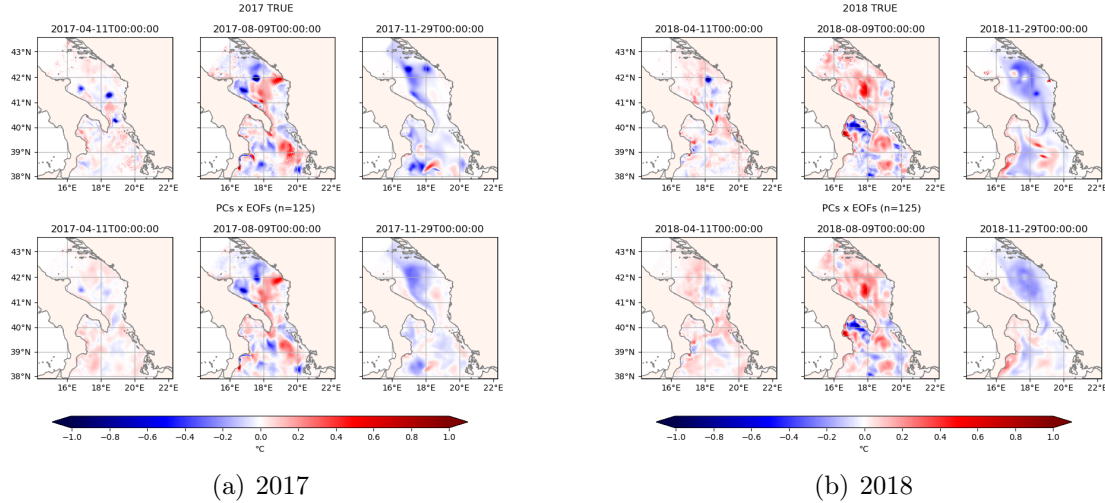
(a) 2017   (b) 2018

Figure 6.9: True (first row) and reduced space using $n = 125$ PCs/EOFs (second row) SST increment field over 3 different timesteps of 2017 (a) and 3 different timesteps of 2018 (b)
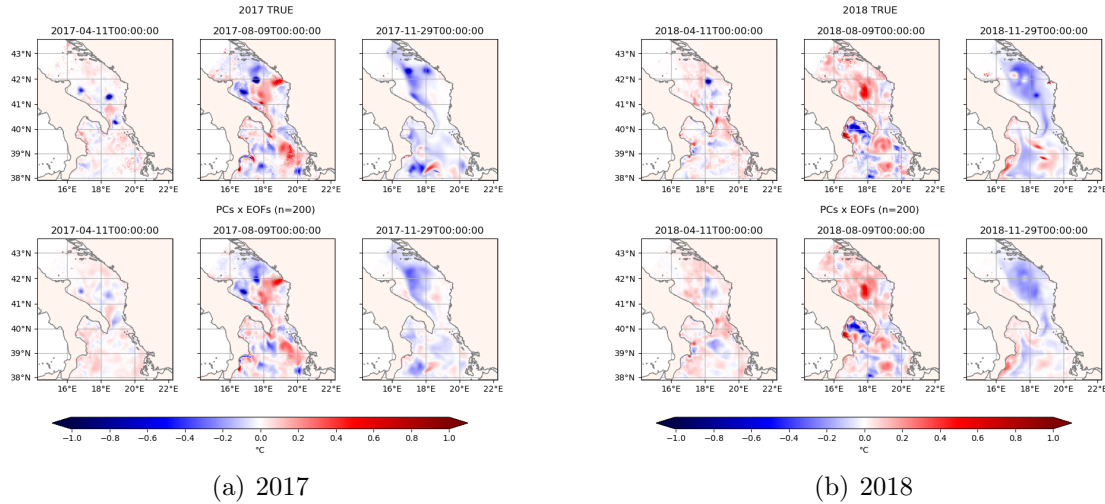

(a) 2017   (b) 2018

Figure 6.10: True (first row) and reduced space using $n = 200$ PCs/EOFs (second row) SST increment field over 3 different timesteps of 2017 (a) and 3 different timesteps of 2018 (b)

lower variability, they introduce noise in the training process. This makes the model training harder concerning the use of lower n. To show that, we will discuss prediction results from an LSTM attention-based model using $n = 25$ and $n = 500$.

The LSTM model is trained using 85% of the PCs time series length and 15% for prediction validation. The LSTM model implemented is a sequence-to-sequence that predicts the next 24 hourly SST increment field from the previous 24 timesteps. The hidden layers are 100 for the model trained using $n = 25$ and 24 for $n = 500$ PCs/EOFs. We use fewer hidden layers with $n = 500$ to reduce the network complexity, which is a drawback when the input dataset has a high level of complexity. The activation function is tanh (Fig. 6.3) for each LSTM state. Finally, the model has been trained for 2000 epochs.

Fig. 6.12 shows the prediction chain. It is such that the last available 24 hours of SST increments from the analysis are the input for the LSTM model to predict the first 24
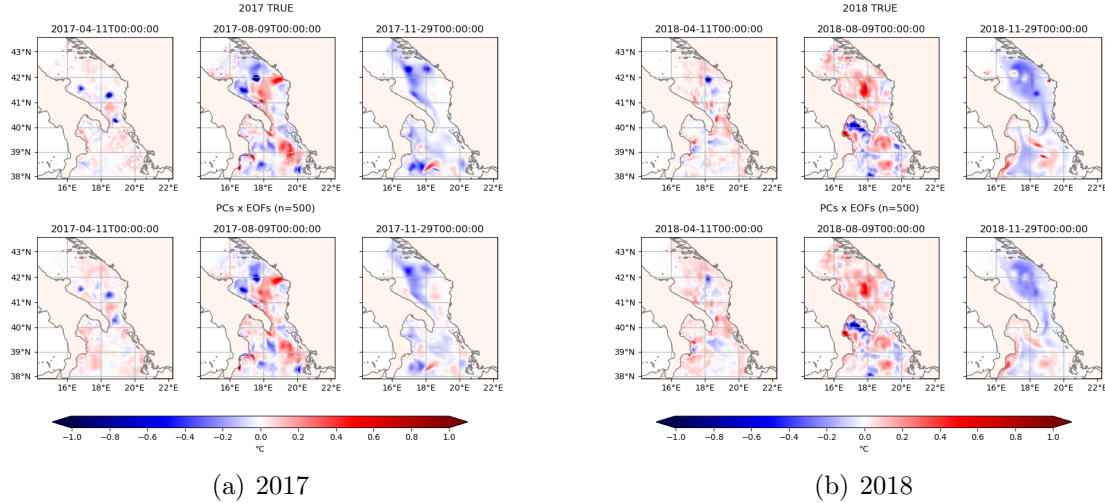
Figure 6.11: True (first row) and reduced space using $n = 500$ PCs/EOFs (second row) SST increment field over 3 different timesteps of 2017 (a) and 3 different timesteps of 2018 (b)

timesteps. The predicted timesteps are the input for the next 24 hours. This process will continue up to the decided number of timesteps after the last assimilation time.

In this thesis, we predict 72 hourly timesteps from the last analysis SST increments
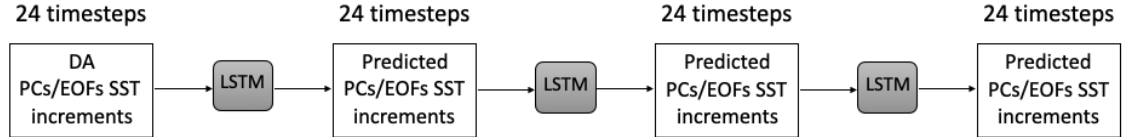

Figure 6.12: Prediction chain implemented using LSTM.

considering two initial days, 2018-09-19 and 2018-10-13. Each of those days represents the input LSTM sequence of 24 timesteps. Results for the midnight (which is also the analysis time in the STSLA experiment) of each predicted sequence for $n = 25$ are shown in Fig. 6.13 and Fig. 6.15. For $n = 500$ in Fig. 6.14 and Fig. 6.16. We observe in the first column of Fig. 6.13 - Fig. 6.16 the difference in the reduced true SST increments field features using $n = 25$ and $n = 500$ as we discussed for Fig. 6.8 - Fig. 6.11. However, we can see how, in the case of Fig. 6.15 and Fig. 6.16, the two representations are much more similar with respect to Fig. 6.13 and Fig. 6.14. This is happening because in the case of Fig. 6.13 and Fig. 6.14, the SST increment field complexity is captured by higher modes than in the case of Fig. 6.15 and Fig. 6.16. For both the examples days, we observe from the predicted field (second column in Fig. 6.13 - Fig. 6.16) from the absolute error (third column in Fig. 6.13 - Fig. 6.16) and the scatter plot of predicted increments against the reduced true (fourth column in Fig. 6.13 - Fig. 6.16 in which the red line represents the

reduced true) that the model trained using $n = 25$ has better prediction abilities with respect to $n = 500$. Furthermore, the same figures show how the prediction error increases as we increase the number of prediction steps.

The problems related to the model trained with $n = 500$ are for sure linked to the higher complexity of the input dataset with respect to $n = 25$ and the presence of higher modes that add noise to the learning process. Here the cases of $n = 125$ and $n = 200$ are not shown. However, experiments using them showed the same results as $n = 500$.

Considering that building an LSTM model was not the primary purpose of this thesis and that the prediction of the SST increment field has been just a first approach to this topic, we can conclude that: using an LSTM attention-based network, it is possible to have a reasonable prediction of the increments to be added to the model forecast. Nevertheless, the high complexity of the input dataset is a drawback, and PCs/EOFs are needed to reduce the space dimensionality. If the domain considered in the DA process has sparse observations, as in the study case of this thesis and in general in ocean applications, the model training becomes difficult. Indeed, since the increments field has peaks in correspondence of the location of the observations, a sparse field of observations introduces high variability to be reproduced in the reduced space. A high number of PCs/EOFs is necessary to capture the essential increment features. However, the drawback of increasing the number of modes is that it also increases the training process's complexity. The future perspective of this application is to solve this problem to predict the essential increment features. Furthermore, considering that the increment field features are similar on different variables (see chapter 5 Fig. 5.12 and Fig. 5.13), this approach could be applied and tested also on salinity or sea level anomaly increments field. This will allow us to introduce a suitable correction also in the forecast. In combination with the correction introduced by DA, the LSTM increment prediction could be a further added value to improve the model skills.
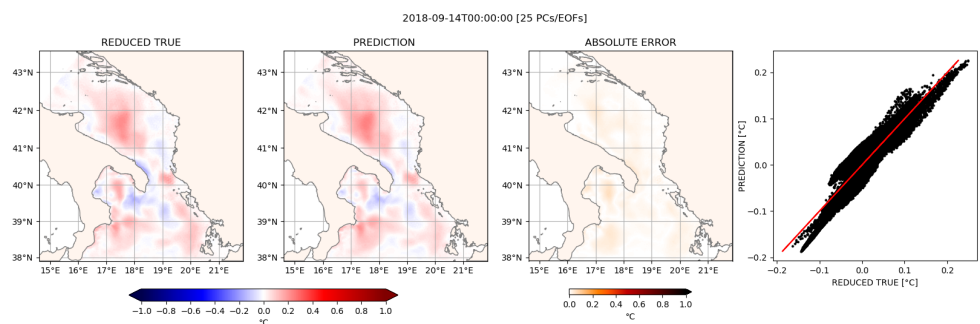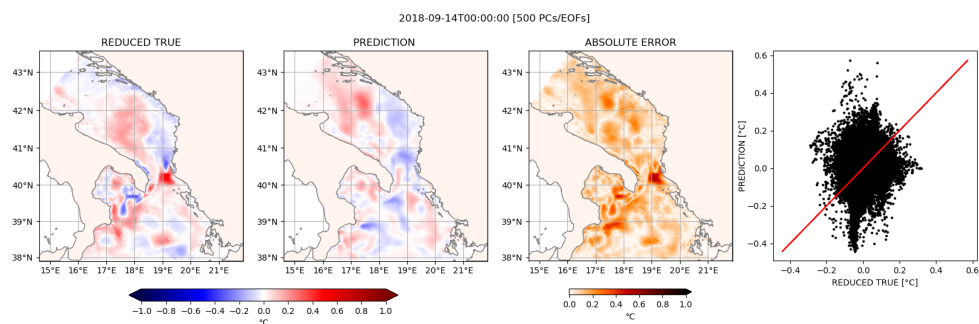
(a) 2018-09-14T00:00:00
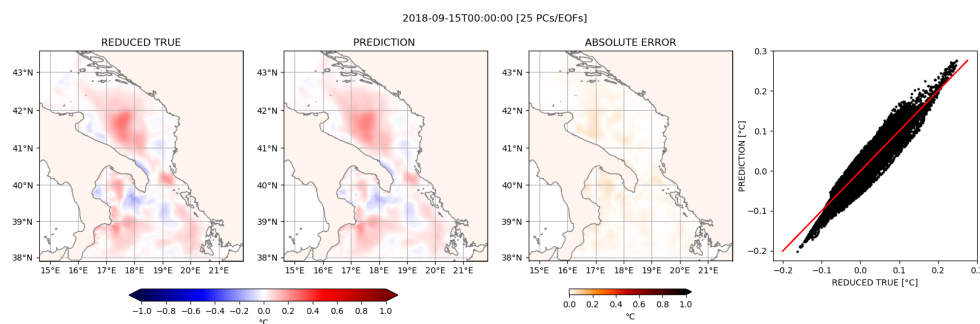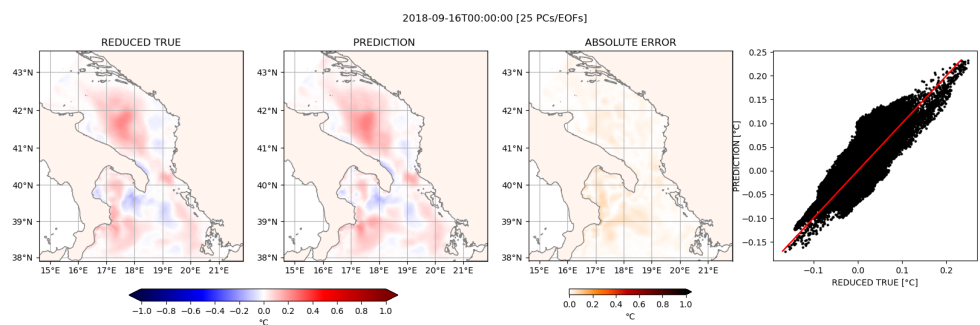
(b) 2018-09-15T00:00:00

(c) 2018-09-16T00:00:00

Figure 6.13: Prediction of three timesteps of SST increments field using 25 PCs/EOFs starting from 24 timesteps of 2018-09-13.

(a) 2018-09-14T00:00:00

(b) 2018-09-15T00:00:00

(c) 2018-09-16T00:00:00

Figure 6.14: Prediction of three timesteps of SST increments field using 500 PCs/EOFs starting from 24 timesteps of 2018-09-13.

(a) 2018-10-14T00:00:00
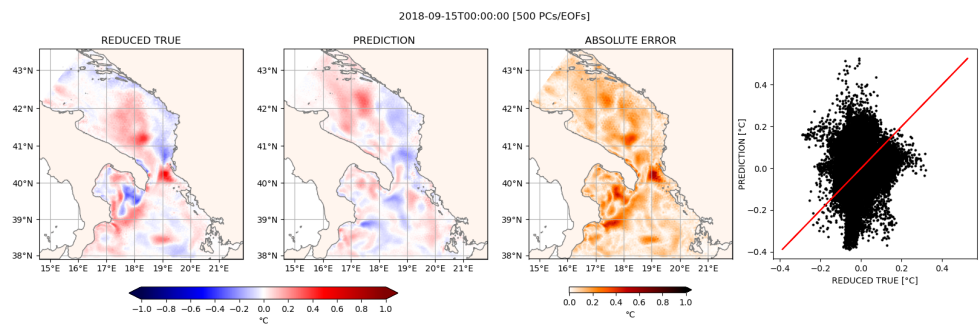

(b) 2018-10-15T00:00:00
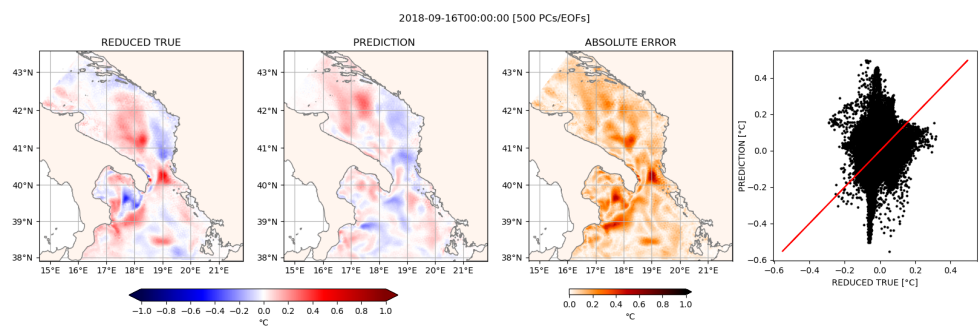

(c) 2018-10-16T00:00:00

Figure 6.15: Prediction of three timesteps of SST increments field using 25 PCs/EOFs starting from 24 timesteps of 2018-10-13.
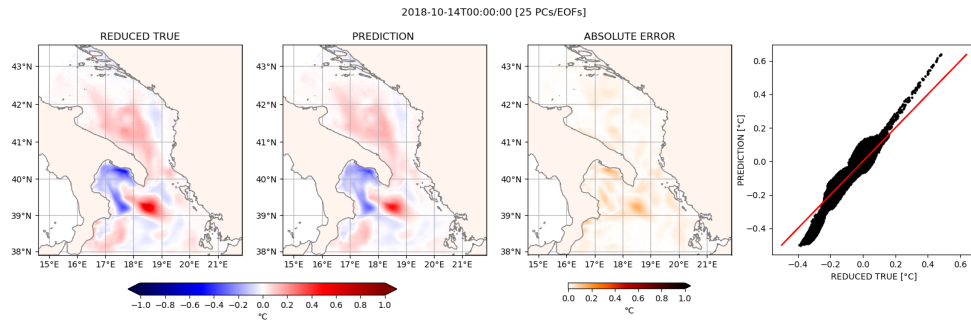
(a) 2018-10-14T00:00:00
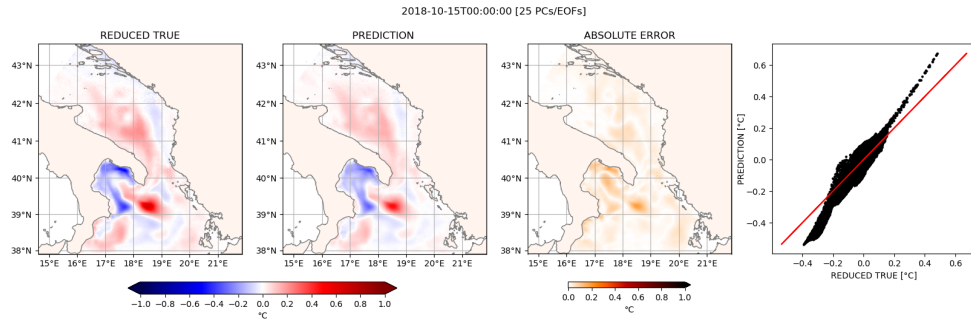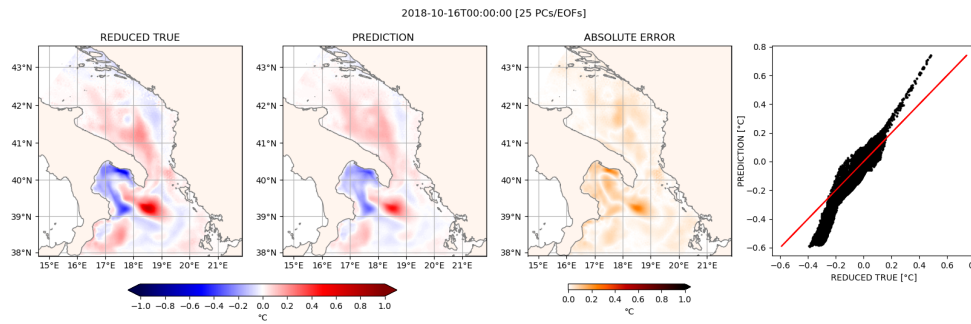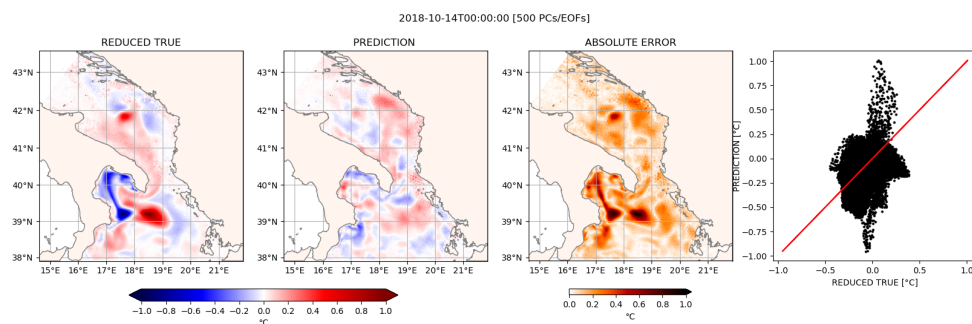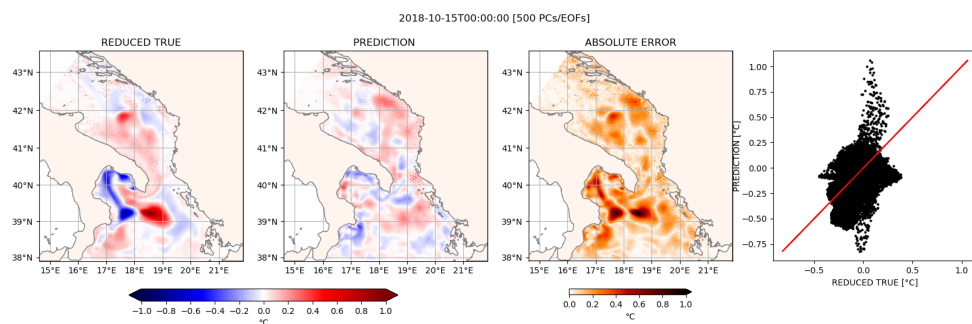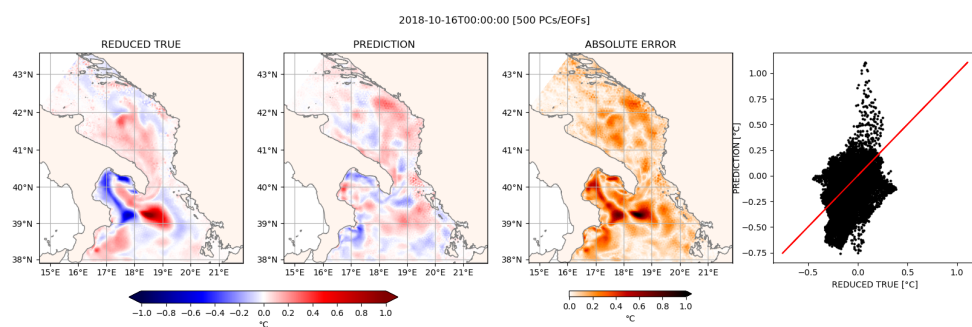

(b) 2018-10-15T00:00:00

(c) 2018-10-16T00:00:00

Figure 6.16: Prediction of three timesteps of SST increments field using 500 PCs/EOFs starting from 24 timesteps of 2018-10-13.

7. Conclusions and Future Perspectives

In this thesis, we presented for the first time the implementation of an assimilation system using SANIFS (a SHYFEM unstructured grid finite element ocean model implementation) and OceanVar based on a 3Dvar assimilation scheme. A complete adaptation of OceanVar to an unstructured grid was necessary to reach this goal. Indeed, OceanVar and all its components have been adapted to be coupled with ocean models with regular grids, e.g., NEMO. The adaptation involves mainly the vertical and horizontal components of the background error covariance matrix. The vertical component is modeled using 25 trivariate (temperature, salinity and sea level anomaly) EOFs. Those EOFs are computed using singular value decomposition methodology from 4 years of SANIFS integration and the cross-covariance matrix has been derived. A novel first order recursive filter algorithm for the unstructured grid (uRF) has been introduced to adapt the horizontal component. This algorithm is based on the triangle's edges ordering with respect to latitude and longitude. An ordering algorithm for backward and forward passes in the x and y direction has been developed using this principle. The uRF algorithm has been tested on idealized test case grids. In particular, we showed that it is equivalent to the regular grid algorithm and produce the same results on a regular grid. Then we moved on to the unstructured meshes and highlighted the sensitivity to the mesh tessellations here. In particular, on an unstructured Delaunay Frontal grid, we showed that the uRF algorithm exhibits a gaussian shape elongated with respect to a particular direction which depends on the grid symmetry. Finally, on the SANIFS grid, we showed that uRF produces a gaussian shape over all the domains except the open sea in the Southern Ionian region, which has a homogeneous triangles tessellation.

Then we designed two assimilation experiments: STSLA, in which we assimilate temperature and salinity from Argo profiling floats and along-track sea level anomaly, and ST, in which we assimilate only temperature and salinity. From these experiments, we showed that the introduction of data assimilation improves the model skills in modeling temperature ad salinity fields. Indeed, looking at the profiles, we showed how the analysis reproduces surface temperature and thermocline better than the model. The maximum gain is found for the temperature up to 10% with respect to the control run without data assimilation.

Concerning the sea level anomaly, we showed a positive improvement in the STSLA exper-

iment and a negative one in the ST experiment. However, one source of these contrasting results could be due to the tidal inputs included in SANIFS at the boundaries. These results on sea level could be improved in the near future by introducing an FGAT interface in SHYFEM. In conclusion, the assimilation system SHYFEM-OceanVar improves the model skills. This makes this system suitable for a more detailed cross-scale forecast. Finally, using the sea surface temperature increment field, we trained a deep learning model using an LSTM attention-based network for SST increment field prediction. This application aimed to predict the increments field to be applied as a correction to the model forecast, which cannot benefit from data assimilation. Here we showed that the high dimensionality of the model space requires a space dimensionality reduction using PCs and EOFs. Also, we showed that the number of PCs/EOFs chosen is an important parameter to be defined. We were able to train a suitable prediction model using 25 PCs/EOFs. However, the drawback is that more than 25 PCs/EOFs are needed to represent all the essential features of the sea surface temperature increment field. The future perspectives of this work are related to a better investigation of the novel first order recursive filter algorithm introduced for the unstructured triangular grid, reducing the sensitivity to the mesh tessellations. Future investigations could go toward consolidating the LSTM algorithm to complement the data assimilation approach.

At this stage, the assimilation cycle includes open sea observations, which could help improve the propagation of these more accurate fields in coastal zones. The next step is introducing coastal observations (e.g., tide gauges, coastal radar) in the assimilation cycle to improve the local coastal dynamics and foster the mutual exchange between the scales (e.g., from coastal to open ocean).

Bibliography

ARTICLES

Accarino, Gabriele et al. "A multi-model architecture based on Long Short-Term Memory neural networks for multi-step sea level forecasting". In: *Future Generation Computer Systems* 124 (2021), pp. 1–9.

Agrawal, JD and MC Deo. "On-line wave prediction". In: *Marine structures* 15.1 (2002), pp. 57–74.

Amendola, Maddalena et al. "Data assimilation in the latent space of a neural network". In: *arXiv preprint arXiv:2012.12056* (2020).

Aydoğdu, Ali et al. "OSSE for a sustainable marine observing network in the Sea of Marmara". In: *Nonlinear Processes in Geophysics* 25.3 (2018), pp. 537–551.

Ayyad, Mahmoud, Muhammad R Hajj, and Reza Marsooli. "Machine learning-based assessment of storm surge in the New York metropolitan area". In: *Scientific Reports* 12.1 (2022), pp. 1–12.

Bai, Xue, Changbo Wang, and Chenhui Li. "A streampath-based RCNN approach to ocean eddy detection". In: *IEEE Access* 7 (2019), pp. 106336–106345.

Bajo, Marco. "Improving storm surge forecast in Venice with a unidimensional Kalman filter". In: *Estuarine, Coastal and Shelf Science* 239 (2020), p. 106773.

Barker, DM et al. "A three-dimensional variational (3DVAR) data assimilation system for use with MM5". In: *NCAR Tech Note* 68 (2003).

Barletta, Ivano. "Seamless Modelling of the Southern European Seas with an optimized MPI-based Unstructured-Grid Model". In: (2021).

Barnes, Stanley L. "A technique for maximizing details in numerical weather map analysis". In: *Journal of Applied Meteorology and Climatology* 3.4 (1964), pp. 396–409.

Bellafiore, Debora and Georg Umgiesser. "Hydrodynamic coastal processes in the North Adriatic investigated with a 3D finite element model". In: *Ocean Dynamics* 60.2 (2010), pp. 255–273.

Bengio, Yoshua, Patrice Simard, and Paolo Frasconi. "Learning long-term dependencies with gradient descent is difficult". In: *IEEE transactions on neural networks* 5.2 (1994), pp. 157–166.

Blumberg, Alan F and George L Mellor. "A description of a three-dimensional coastal ocean circulation model". In: *Three-dimensional coastal ocean models* 4 (1987), pp. 1–16.

Buizza, Caterina et al. "Data learning: Integrating data assimilation and machine learning". In: *Journal of Computational Science* 58 (2022), p. 101525.

Burchard, Hans and Ole Petersen. "Models of turbulence in the marine environment,ÄîA comparative study of two-equation turbulence models". In: *Journal of Marine Systems* 21.1-4 (1999), pp. 29–53.

Byrd, Richard H et al. "A limited memory algorithm for bound constrained optimization". In: *SIAM Journal on scientific computing* 16.5 (1995), pp. 1190–1208.

Carrassi, Alberto et al. "Data assimilation in the geosciences: An overview of methods, issues, and perspectives". In: *Wiley Interdisciplinary Reviews: Climate Change* 9.5 (2018), e535.

Chen, Changsheng, Hedong Liu, and Robert C Beardsley. "An unstructured grid, finite-volume, three-dimensional, primitive equations ocean model: application to coastal ocean and estuaries". In: *Journal of atmospheric and oceanic technology* 20.1 (2003), pp. 159–186.

Chiggiato, Jacopo et al. "Dynamics of the circulation in the Sea of Marmara: numerical modeling experiments and observations from the Turkish straits system experiment". In: *Ocean Dynamics* 62.1 (2012), pp. 139–159.

Cooper, Mike and Keith Haines. "Altimetric assimilation with water property conservation". In: *Journal of Geophysical Research: Oceans* 101.C1 (1996), pp. 1059–1077.

Courtier, Philippe, J-N Thépaut, and Anthony Hollingsworth. "A strategy for operational implementation of 4D-Var, using an incremental approach". In: *Quarterly Journal of the Royal Meteorological Society* 120.519 (1994), pp. 1367–1387.

Cressman, George P. "An operational objective analysis system". In: *Monthly Weather Review* 87.10 (1959), pp. 367–374.

Cucco, Andrea and Georg Umgiesser. "Modeling the Venice Lagoon residence time". In: *Ecological modelling* 193.1-2 (2006), pp. 34–51.

Danilov, Sergey. "Ocean modeling on unstructured meshes". In: *Ocean Modelling* 69 (2013), pp. 195–210.

Danilov, Sergey, Gennady Kivman, and Jens Schröter. "A finite-element ocean model: principles and evaluation". In: *Ocean Modelling* 6.2 (2004), pp. 125–150.

Danilov, Sergey et al. "The finite-volume sea ice–ocean model (fesom2)". In: *Geoscientific Model Development* 10.2 (2017), pp. 765–789.

Derber, J. "The National Meteorological Center's Spectral Statistical-Interpolation Analysis System". In: *Mon. Weath. Rev* 120 (1992), pp. 1747–63.

Derber, John and Anthony Rosati. "A global oceanic data assimilation system". In: *Journal of physical oceanography* 19.9 (1989), pp. 1333–1347.

Dinther, Ylona van, Hans R Künsch, and Andreas Fichtner. "Ensemble data assimilation for earthquake sequences: probabilistic estimation and forecasting of fault stresses". In: *Geophysical Journal International* 217.3 (2019), pp. 1453–1478.

Dobricic, Srdjan and Nadia Pinardi. "An oceanographic three-dimensional variational data assimilation scheme". In: *Ocean modelling* 22.3-4 (2008), pp. 89–105.

Dobricic, Srdjan et al. "Mediterranean Forecasting System: An improved assimilation scheme for sea-level anomaly and its validation". In: *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography* 131.613 (2005), pp. 3627–3642.

Dobricic, Srdjan et al. "Daily oceanographic analyses by the Mediterranean basin scale assimilation system". In: *Ocean Science Discussions* 3.6 (2006), pp. 1977–1998.

Dueben, Peter D and Peter Bauer. "Challenges and design choices for global weather and climate models based on machine learning". In: *Geoscientific Model Development* 11.10 (2018), pp. 3999–4009.

Egbert, Gary D and Svetlana Y Erofeeva. "Efficient inverse modeling of barotropic ocean tides". In: *Journal of Atmospheric and Oceanic technology* 19.2 (2002), pp. 183–204.

Evensen, Geir. "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics". In: *Journal of Geophysical Research: Oceans* 99.C5 (1994), pp. 10143–10162.

Farchi, Alban et al. "Using machine learning to correct model error in data assimilation and forecast applications". In: *Quarterly Journal of the Royal Meteorological Society* 147.739 (2021), pp. 3067–3084.

Farina, Raffaele et al. "A revised scheme to compute horizontal covariances in an oceanographic 3D-VAR assimilation system". In: *Journal of Computational Physics* 284 (2015), pp. 631–647.

Federico, Ivan et al. "Coastal ocean forecasting with an unstructured grid model in the southern Adriatic and northern Ionian seas". In: *Natural Hazards and Earth System Sciences* 17.1 (2017), pp. 45–59.

Ferrarin, Christian, Marco Bajo, and Georg Umgiesser. "Model-driven optimization of coastal sea observatories through data assimilation in a finite element hydrodynamic model (SHYFEM v. 7_5_65)". In: *Geoscientific Model Development* 14.1 (2021), pp. 645–659.

Ferrarin, Christian et al. "Tide-surge-wave modelling and forecasting in the Mediterranean Sea with focus on the Italian coast". In: *Ocean Modelling* 61 (2013), pp. 38–48.

Fletcher, SJ. "Mixed Gaussian-lognormal four-dimensional data assimilation". In: *Tellus A: Dynamic Meteorology and Oceanography* 62.3 (2010), pp. 266–287.

Fofonoff, NP. "Physical properties of seawater: A new salinity scale and equation of state for seawater". In: *Journal of Geophysical Research: Oceans* 90.C2 (1985), pp. 3332–3342.

Ford, R et al. "A nonhydrostatic finite-element model for three-dimensional stratified oceanic flows. Part I: Model formulation". In: *Monthly Weather Review* 132.12 (2004), pp. 2816–2831.

Gao, Hang et al. "A hybrid method for fine-scale wind field retrieval based on machine learning and data assimilation". In: *IEEE Transactions on Geoscience and Remote Sensing* 60 (2022), pp. 1–12.

Gunduz, Murat, Emin Özsoy, and Robinson Hordoir. "A model of Black Sea circulation with strait exchange (2008–2018)". In: *Geoscientific Model Development* 13.1 (2020), pp. 121–138.

Haben, Stephen A, Amos S Lawless, and Nancy K Nichols. "Conditioning and preconditioning of the variational data assimilation problem". In: *Computers & Fluids* 46.1 (2011), pp. 252–256.

Hanert, Emmanuel, Vincent Legat, and Éric Deleersnijder. "A comparison of three finite elements to solve the linear shallow water equations". In: *Ocean Modelling* 5.1 (2003), pp. 17–35.

Hannachi, A and A O'Neill. "Atmospheric multiple equilibria and non-Gaussian behaviour in model simulations". In: *Quarterly Journal of the Royal Meteorological Society* 127.573 (2001), pp. 939–958.

Hannachi, Abdel, Ian T Jolliffe, and David B Stephenson. "Empirical orthogonal functions and related techniques in atmospheric science: A review". In: *International Journal of Climatology: A Journal of the Royal Meteorological Society* 27.9 (2007), pp. 1119–1152.

Hayden, Christopher M and R James Purser. "Recursive filter objective analysis of meteorological fields: Applications to NESDIS operational processing". In: *Journal of Applied Meteorology and Climatology* 34.1 (1995), pp. 3–15.

Hellerman, Sol and Mel Rosenstein. "Normal monthly wind stress over the world ocean with error estimates". In: *Journal of Physical Oceanography* 13.7 (1983), pp. 1093–1104.

Hochreiter, Sepp and Jürgen Schmidhuber. "Long short-term memory". In: *Neural computation* 9.8 (1997), pp. 1735–1780.

Jain, Pooja and MC Deo. "Artificial intelligence tools to forecast ocean waves in real time". In: *The Open Ocean Engineering Journal* 1.1 (2008).

James, Scott C, Yushan Zhang, and Fearghal O'Donncha. "A machine learning framework to forecast wave conditions". In: *Coastal Engineering* 137 (2018), pp. 1–10.

Janjić, Tijana et al. "On the representation error in data assimilation". In: *Quarterly Journal of the Royal Meteorological Society* 144.713 (2018), pp. 1257–1278.

Jofre, Lluís et al. "Conservation properties of unstructured finite-volume mesh schemes for the Navier-Stokes equations". In: *Numerical Heat Transfer, Part B: Fundamentals* 65.1 (2014), pp. 53–79.

Kalman, Rudolph Emil. "A new approach to linear filtering and prediction problems". In: (1960).

Kara, A Birol, Alan J Wallcraft, and Harley E Hurlbut. "A correction for land contamination of atmospheric variables near land–sea boundaries". In: *Journal of physical oceanography* 37.4 (2007), pp. 803–818.

Katavouta, Anna and Keith R Thompson. "Downscaling ocean conditions with application to the Gulf of Maine, Scotian Shelf and adjacent deep ocean". In: *Ocean Modelling* 104 (2016), pp. 54–72.

Le Roux, Daniel Y, Andrew Staniforth, and Charles A Lin. "Finite elements for shallow-water equation ocean models". In: *Monthly Weather Review* 126.7 (1998), pp. 1931–1951.

Le Roux, DY et al. "On some spurious mode issues in shallow-water models using a linear algebra approach". In: *Ocean Modelling* 10.1-2 (2005), pp. 83–94.

Lee, Der-Tsai and Bruce J Schachter. "Two algorithms for constructing a Delaunay triangulation". In: *International Journal of Computer & Information Sciences* 9.3 (1980), pp. 219–242.

Lee, Seunghee et al. "Air quality forecasts improved by combining data assimilation and machine learning with satellite AOD". In: *Geophysical Research Letters* 49.1 (2022), e2021GL096066.

Lermusiaux, PFJ. "Evolving the subspace of the three-dimensional multiscale ocean variability: Massachusetts Bay". In: *Journal of Marine Systems* 29.1-4 (2001), pp. 385–422.

Lins, Isis Didier et al. "Prediction of sea surface temperature in the tropical Atlantic by support vector machines". In: *Computational Statistics & Data Analysis* 61 (2013), pp. 187–198.

Lorenc, AC et al. "The Met. Office global three-dimensional variational data assimilation scheme". In: *Quarterly Journal of the Royal Meteorological Society* 126.570 (2000), pp. 2991–3012.

Lorenc, Andrew. "Iterative analysis using covariance functions and filters". In: *Quarterly Journal of the Royal Meteorological Society* 118.505 (1992), pp. 569–591.

Lorenc, Andrew C. "Analysis methods for numerical weather prediction". In: *Quarterly Journal of the Royal Meteorological Society* 112.474 (1986), pp. 1177–1194.

Lorenz, Edward N. "Deterministic nonperiodic flow". In: *Journal of atmospheric sciences* 20.2 (1963), pp. 130–141.

Luong, Minh-Thang, Hieu Pham, and Christopher D Manning. "Effective approaches to attention-based neural machine translation". In: *arXiv preprint arXiv:1508.04025* (2015).

Maicu, Francesco et al. "Downscaling with an unstructured coastal-ocean model to the Goro Lagoon and the Po River Delta branches". In: *Frontiers in Marine Science* 8 (2021), p. 647781.

Makarynskyy, O. "Improving wave predictions with artificial neural networks". In: *Ocean Engineering* 31.5-6 (2004), pp. 709–724.

Malanotte-Rizzoli, Paola and Eli Tziperman. "The oceanographic data assimilation problem: overview, motivation and purposes". In: *Elsevier Oceanography Series* 61 (1996), pp. 3–17.

Mavriplis, DJ. "Unstructured grid techniques". In: *Annual Review of Fluid Mechanics* 29.1 (1997), pp. 473–514.

Pain, CC et al. "Three-dimensional unstructured mesh ocean modelling". In: *Ocean Modelling* 10.1-2 (2005), pp. 5–33.

Paulson, Clayton A and James J Simpson. "Irradiance measurements in the upper ocean". In: *Journal of Physical Oceanography* 7.6 (1977), pp. 952–956.

Petersen, Mark R et al. "An evaluation of the ocean and sea ice climate of E3SM using MPAS and interannual CORE-II forcing". In: *Journal of Advances in Modeling Earth Systems* 11.5 (2019), pp. 1438–1458.

Pettenuzzo, D, WG Large, and N Pinardi. "On the corrections of ERA-40 surface flux products consistent with the Mediterranean heat and water budgets and the connection between basin surface total heat flux and NAO". In: *Journal of Geophysical Research: Oceans* 115.C6 (2010).

Peyron, Mathis et al. "Latent space data assimilation by using deep learning". In: *Quarterly Journal of the Royal Meteorological Society* 147.740 (2021), pp. 3759–3777.

Piggott, Matthew D et al. "Unstructured adaptive meshes for ocean modeling". In: *Ocean modeling in an eddying regime* 177 (2008), pp. 383–408.

Pinardi, N, A Rosati, and RC Pacanowski. "The sea surface pressure formulation of rigid lid models. Implications for altimetric data assimilation studies". In: *Journal of Marine Systems* 6.1-2 (1995), pp. 109–119.

Purser, R James, Miodrag Rancic, and Manuel SFV De Pondeca. "The Multigrid Beta Function Approach for Modeling of Background Error Covariance in the Real-Time Mesoscale Analysis (RTMA)". In: *Monthly Weather Review* 150.4 (2022), pp. 715–732.

Purser, R James et al. "Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances". In: *Monthly Weather Review* 131.8 (2003), pp. 1524–1535.

Raj, Nawin et al. "Assessment and Prediction of Sea Level Trend in the South Pacific Region". In: *Remote Sensing* 14.4 (2022), p. 986.

Rus, Marko et al. "HIDRA2: deep-learning ensemble storm surge forecasting in the presence of seiches–the case of Northern Adriatic". In: *EGUsphere* (2022), pp. 1–26.

Sasaki, Yoshikazu. "An objective analysis based on the variational method". In: *Journal of the Meteorological Society of Japan. Ser. II* 36.3 (1958), pp. 77–88.

Scher, Sebastian and Gabriele Messori. "Weather and climate forecasting with neural networks: using general circulation models (GCMs) with different complexity as a study ground". In: *Geoscientific Model Development* 12.7 (2019), pp. 2797–2809.

Shi, Lei, Liujuan Tang, and Edward Myers. "Variational Data Assimilation of Tides". In: *Journal of Marine Science and Engineering* 8.1 (2020), p. 54.

Smagorinsky, Joseph. "General circulation experiments with the primitive equations: I. The basic experiment". In: *Monthly weather review* 91.3 (1963), pp. 99–164.

Storto, Andrea, 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* 142.695 (2016), pp. 738–758.

Storto, Andrea et al. "Assimilating along-track altimetric observations through local hydrostatic adjustment in a global ocean variational assimilation system". In: *Monthly Weather Review* 139.3 (2011), pp. 738–754.

Umgiesser, Georg et al. "A finite element model for the Venice Lagoon. Development, set up, calibration and validation". In: *Journal of Marine Systems* 51.1-4 (2004), pp. 123–145.

Verri, Giorgia et al. "River runoff influences on the Central Mediterranean overturning circulation". In: *Climate dynamics* 50.5 (2018), pp. 1675–1703.

Vincent, Grace et al. "Hybrid Machine Learning Models for Storm Surge Prediction". In: (2022).

Wang, Guosong et al. "A Hybrid Multivariate Deep Learning Network for Multistep Ahead Sea Level Anomaly Forecasting". In: *Journal of Atmospheric and Oceanic Technology* 39.3 (2022), pp. 285–301.

Wang, Qiang, Sergey Danilov, and Jens Schröter. "Finite element ocean circulation model based on triangular prismatic elements, with application in studying the effect of topography representation". In: *Journal of Geophysical Research: Oceans* 113.C5 (2008).

Weaver, Anthony and Philippe Courtier. "Correlation modelling on the sphere using a generalized diffusion equation". In: *Quarterly Journal of the Royal Meteorological Society* 127.575 (2001), pp. 1815–1846.

Xiao, Changjiang et al. "A spatiotemporal deep learning model for sea surface temperature field prediction using time-series satellite data". In: *Environmental Modelling & Software* 120 (2019), p. 104502.

Xiao, Changjiang et al. "Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach". In: *Remote Sensing of Environment* 233 (2019), p. 111358.

Yang, Mingzhi, Xinchun Li, and Yue Liu. "Sequence to point learning based on an attention neural network for nonintrusive load decomposition". In: *Electronics* 10.14 (2021), p. 1657.

Zhang, Suming et al. "Estimating the grade of storm surge disaster loss in coastal areas of China via machine learning algorithms". In: *Ecological Indicators* 136 (2022), p. 108533.

Zhang, Yinglong and António M Baptista. "SELFE: A semi-implicit Eulerian–Lagrangian finite-element model for cross-scale ocean circulation". In: *Ocean modelling* 21.3-4 (2008), pp. 71–96.

Zheng, Xinye et al. "Detecting comma-shaped clouds for severe weather forecasting using shape and motion". In: *IEEE Transactions on Geoscience and Remote Sensing* 57.6 (2019), pp. 3788–3801.

Zhu, Ze-Nan et al. "Assimilation of coastal acoustic tomography data using an unstructured triangular grid ocean model for water with complex coastlines and islands". In: *Journal of Geophysical Research: Oceans* 122.9 (2017), pp. 7013–7030.

BOOKS

Asch, Mark, Marc Bocquet, and Maëlle Nodet. *Data assimilation: methods, algorithms, and applications*. SIAM, 2016.

Fletcher, Steven J. *Data assimilation for the geosciences: From theory to application*. Elsevier, 2017.

Lahoz, Boris Khattatov William and Richard Menard. *Data assimilation*. Springer, 2010.

Menke, William. *Geophysical data analysis: Discrete inverse theory*. Academic press, 2018.

Navarra, Antonio and Valeria Simoncini. *A guide to empirical orthogonal functions for climate data analysis*. Springer Science & Business Media, 2010.

Wang, Sun-Chong. *Interdisciplinary computing in Java programming*. Vol. 743. Springer Science & Business Media, 2003.