

Afşin Özdemir

Report

Sea Surface Height reconstructions using physics based variational approaches

September 19, 2023

Supervisors

Nelly Pustelnik, Antoine Venaille
Ens de Lyon, Physics Laboratory

Contents

1	Introduction	3
2	Image, Inpainting, Minimization Tools	5
2.1	Image Restoration	5
2.2	Variational Approaches	6
2.2.1	Total Variation Regularization	7
2.3	Proximal Methods	7
2.3.1	Moreau Decomposition	8
2.4	Algorithms	9
2.4.1	Proximal Gradient Descent/Forward Backward	9
2.4.2	Chambolle-Pock	10
2.4.3	Condat-Vu	10
2.4.4	Plug and Play	10
3	Oceanographic Equations	13
3.1	Main concepts	13
3.1.1	Coriolis Force	13
3.1.2	Primitive Equations	14
3.1.3	f-plane, beta-plane approximations	15
3.2	Small aspect ratio limit	15
3.3	Hydro-static Balance	16
3.4	Geostrophic Balance	17
3.5	Vorticity	17
3.6	Shallow Water Equations	17
3.6.1	Reduced Gravity Equations for Free Upper surface	18
3.7	Quasi-Geostrophic Equations	19
3.7.1	Rotating Fluid	19
3.7.2	Shallow Water Quasi-Geostrophic Potential Vorticity	20
4	Image Processing Tools for SSH Recovery	21
4.1	Data-set	21
4.1.1	Case-study region, dates and NATL60 data	21
4.1.2	Degraded Data Set	22
4.1.3	Preprocess of the data set	23
4.1.4	Mask	24
4.1.5	Measurements	24
4.1.6	Baseline Method: Optimal Interpolation	25

4.2	Proposed Methods	26
4.2.1	Total Variation	26
4.2.2	Total Variation on Potential Vorticity	27
4.2.3	Hybrid Method	28
4.3	Results and Discussion	28
5	One step more	35
5.1	Data Assimilation methods considering Dynamics	35
5.1.1	4D-Var: 4-dimensional Variational Assimilation	36
5.1.2	Kalman Filtering	37
5.2	Conclusion, Future Work	38
References		39

List of symbols and abbreviations

Image Processing

- \mathcal{P} , probability.
- \mathcal{H} , Hilbert space.
- \mathbf{H} , linear observation operator.
- $\Gamma_0(\mathcal{H})$, class of convex, lower-semi continuous, proper functions.
- D , discrete gradient operator.
- s^t , ground truth.
- s , variable for Sea Surface Height.
- o , observations.
- ϵ, ε , Gaussian noise.

Oceanography

- \mathbf{F}_c , Coriolis Force
- f, f_0 , Coriolis parameter.
- p , pressure.
- H , thickness of the fluid.
- L , horizontal length scale of the fluid.
- U , horizontal velocity scale of the fluid.
- W , vertical velocity scale of the fluid.
- R_o , Rossby Number.
- \mathbf{r} , radius of Earth.
- q , Quasi geostrophic potential vorticity.
- Q , potential vorticity.
- \mathbf{U} , inertial Velocity.
- \mathbf{u} , two dimensional velocity, (u, v) .
- \mathbf{v} , three dimensional velocity, (u, v, w) .
- β , rate of change of f with altitude.
- η , perturbation height.
- Ω, Ω , rotation of the Earth and associated vector.
- θ , latitude.
- ψ , stream function.

- ρ , density.
- ∇^2 , Laplacian operator.
- ζ , vorticity.
- $\frac{\mathbf{D}}{\mathbf{D}t}$, advection operator.

Experiments

- SSH, Sea Surface Height.
- RMS, Root Mean Square.
- RMSE, Root Mean Square Error.
- PSNR, Peak Signal-to-Noise Ratio.
- s^f , prior model, forecast.
- \mathcal{L} , discrete Laplacian operator.

Chapter 1

Introduction

Satellite altimeters have been collecting data from the sea surface since the 1970s, and in the last 20 years, with technological advancements, they have become capable of collecting data with higher resolution. It is challenging to use and interpolate these data sets for accurate predictions.

In this report, our concern is to take into account the Navier-Stokes equation in a geo-physical context and utilize this model to interpolate the sea surface height (SSH) field using image reconstruction techniques. We can establish the relationship as follows:

$$o = \mathbf{H}s^t + \varepsilon, \quad (1.1)$$

where " o " represents the collected data, " s^t " is the 'ground truth,' and " ε " represents the noise caused by the observations.

To address this problem, we aim to find the best estimate since finding the "exact solution" is nearly an NP-Hard problem. This leads us to a branch of mathematics known as the "Inverse Problem." The reason it is called "inverse" is due to the inversion of the operator. In Chapter 2, we will delve into this area, providing more details.

In Chapter 3, we will derive the shallow water quasi-geostrophic equations, which are a special form of the Navier-Stokes equations, considering scale and the Earth's assumptions. We will make the connection with our variational problem.

In Chapter 4, we will begin by explaining the considered data set and how we process it for our case. We will cover the considered measurements and the methods used for the experiments. At the end of the chapter, we will explain the results in comparison with the baseline method.

Last but not least, Chapter 5.1 will elucidate well-known methods in data assimilation, concluding our work with possible research directions.

Chapter 2

Image, Inpainting, Minimization Tools

2.1 Image Restoration

As we will see in Chapter 4, the field of sea surface height is partially observed and the process creates a noise. The ideas of interpolation in this field are similar to image restoration techniques that will be detailed in this chapter.

Before proceeding into the main approaches, it is important to give an introduction to image processing and its tools. Let us start with the simplest denoising problem:
Suppose that we have an observation denoted as o , and the grand truth s^t

$$o = s^t + \varepsilon, \quad (2.1)$$

where ε represents the additive noise that can be mainly caused by the observation device. The main objective of the inverse problem is to obtain a reconstructed image \hat{s} from observation o . The reason why it is a challenging problem is that, in most cases, we don't know what the noise is, but we only have a piece of information on the distribution. We will assume it is an additive white Gaussian noise¹. Before discussing how to solve this problem, we would like to present, what is an inpainting problem and what are the challenges of an inpainting problem.

From a more general perspective, we can mathematically formalize it in the following form,

$$o = \mathbf{H}s^t + \varepsilon \quad (2.2)$$

only difference between equation (2.2) and equation (2.1) is, now there is a linear operator \mathbf{H} . This operator can be a blur, any measurement, or, like in our case, an inpainting operator. Intuitively, the first approach to solve this problem is to try inverting the operator \mathbf{H} . There are a couple of problems in this approach, first, we don't know if it is invertible or not, second, assume we know the linear operator is invertible, this is a

¹Additive white Gaussian noise (AWGN) is a basic noise model used in information theory to mimic the effect of many random processes that occur in nature. The central limit theorem of probability theory indicates that the summation of many random processes will tend to have a distribution called Gaussian or Normal.

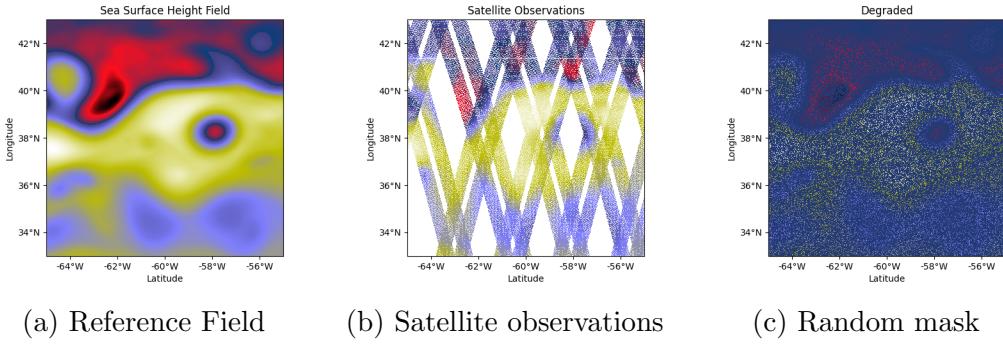


Figure 2.1: Example of different inpainting operations.

costly operation to do, and very popular methods like Gauss-Jordan elimination works in $O(n^3)^2$ [17], and the best possible (theoretically) one "Optimized CW-like³" algorithm work in $O(n^{2.373})$. Therefore, for now, it's not an option to solve this problem.

Especially, it is important to note that in an inpainting problem, there is data with a lot of missing parts, and we want to fill these parts by using information on the general structure of images (or in our case information on the sea surface height) and/or imposing constraints on our main problem. For this purpose, we have introduced the variational approaches.

2.2 Variational Approaches

Consider the equation (2.2), let us rephrase this from a statistical point of view. We want to find s for a given o , so we want to maximize the posterior probability $\mathcal{P}(s|o)$ since we know the distribution of the noise ε . We want to find the best \hat{s}

$$\hat{s} = \arg \max_s \mathcal{P}(s|o). \quad (2.3)$$

This is called "the maximum a posteriori", or simply MAP. By the Bayes' theorem, we have,

$$\hat{s} = \arg \max_s \mathcal{P}(s|o) = \arg \max_s \frac{\mathcal{P}(o|s)\mathcal{P}(s)}{\mathcal{P}(o)}, \quad (2.4)$$

where $\mathcal{P}(o|s)$ is the likelihood. Which can be formulated as,

$$\arg \max_s \log \mathcal{P}(o|s) + \log \mathcal{P}(s) \quad (2.5)$$

where the second term $\mathcal{P}(s)$ is the prior information of a assumption on the image (image prior). In the Gaussian case the objective function can be reformulated as a minimization problem of the form,

$$\arg \min_s \frac{1}{2} \|o - \mathbf{H}s\|_2^2 + \lambda R(s). \quad (2.6)$$

Where $R(s)$ is the regularization term coming from the prior information, and λ is the regularization parameter. Here the challenging part is finding a good regularization term for the considered problem. In this work, we are investigating different effects of different regularization terms on the sea surface height field reconstruction.

² O is the common notation for the time complexity. [17]

³The Coppersmith-Winograd Algorithm, which was published in 90' and revisited in [11], this is the improved version of the first result.

2.2.1 Total Variation Regularization

The intuitive choice would be considering the L2-norm for regularization, however, it has a problem of over-smoothing image details [14]. To solve this problem, anisotropic diffusion-based methods have been used to preserve image details, nevertheless, the edges are still blurred [1],[3].

To overcome this issue, the total variational (TV)-based regularization has been introduced [2]. TV regularization is based on the statistical fact that natural images are locally smooth and the pixel intensity gradually varies in most regions. It is defined as follows [2]:

$$R_{TV}(x) = \|Dx\|_1, \quad (2.7)$$

where D is a discrete gradient operator and $\|x\|_1 = \sum_{i=0}^n |x_i|$ for $x \in \mathbb{R}^n$. The reason why it has used the result of image denoising, it can calculate the optimal solution efficiently but also retain sharp edges. This can cause different problems, textures tend to be over-smoothed, flat areas are approximated by a piecewise constant surface resulting in a stair-casing effect and the image suffers from losses of contrast.

2.3 Proximal Methods

This part is dedicated to the minimization problem 2.6. Let's first write the problem in an abstract sense,

$$\arg \min_x F(x) \quad (2.8)$$

In numerical analysis, there are certain methods to find the minimizer of F , examples can be Newton's method⁴ or gradient descent method. These methods require a strong assumption, a smooth function. What will be presented here is another way to approach the problem in non-smooth cases. For example, in the case that we choose R as total variation (2.7) in the minimization problem (2.6). Before proceeding into the motivation of this operator, let us define it first.

Definition 2.1. *Proximal operator*

Suppose \mathcal{H} is a Hilbert space. Let F be a lower-semi continuous (l.s.c.), convex and proper⁵ function, then the proximal operator of F , $\text{prox}_{\lambda F} : \mathcal{H} \rightarrow \mathcal{H}$ defined as,

$$\text{prox}_{\lambda F}(u) = \arg \min_x F(x) + \frac{1}{2\lambda} \|x - u\|_2^2. \quad (2.9)$$

Note that lower semi-continuity and convexity are necessary assumptions for the existence of minimizer⁶. If F corresponds to the indicator function,

$$i_X(x) = \begin{cases} 0, & \text{if } x \in X \\ \infty, & \text{if } x \notin X \end{cases} \quad (2.10)$$

⁴Newton's method is a root-finding algorithm which produces successively better approximations to the roots (or zeroes) of a real-valued function.

⁵Which not equal to zero always

⁶This is a more specific version of the calculus of variations, which is a branch of mathematics that is interesting in variational problems in abstract settings. For more information check the book of Giusti [4]

Then the proximal map is the projection onto X^7 ,

$$\text{prox}_{\lambda i_X}(u) = \arg \min_{x \in X} |(\lambda i_X(x) + \frac{1}{2}\|x - u\|^2)| \quad (2.11)$$

$$= \arg \min_{x \in X} (\frac{1}{2}\|x - u\|^2) \quad (2.12)$$

$$= \text{Proj}_X(y). \quad (2.13)$$

Until this point, we have stressed that we are developing this method to deal with the minimization problem of non-smooth functions. However, We can motivate it by using the idea of the smooth case, for this, we need to have the following definition.

Definition 2.2. (*Subdifferential*) Suppose that $F \in \Gamma_0(\mathcal{H})$, then the subdifferential of F defined as,

$$\partial F : x \rightarrow \{u \in \mathcal{H} | (\forall y \in \mathcal{H}) \langle y - x | u \rangle + F(x) \leq F(y)\} \quad (2.14)$$

The subdifferential takes into account all possible limits contrary to the differential definition. Note that, if F is smooth at the point x , then $\partial F(x) = \{\nabla F(x)\}$. This definition is also related to our minimization problem by the following theorem,

Theorem 2.1. *Fermat's Theorem*

$$0 \in \partial F(\hat{x}) \Leftrightarrow \hat{x} \in \operatorname{Arg} \min_x F(x) \quad (2.15)$$

This implies that instead of finding the minimizer of (2.8), we can estimate it in the subdifferential set. Now it's possible to mimic the gradient descent by just taking a point in the subdifferential. Then the explicit subgradient step would be,

$$x_{k+1} = x_k - u_k \quad \text{where } u_k \in \gamma \partial F(x_k), \quad (2.16)$$

where $\gamma > 0$ is the step size. Where the implicit form will be,

$$\begin{aligned} x_{k+1} &= x_k - u_k \quad \text{where } u_k \in \gamma \partial F(x_k), \\ &= \text{prox}_F(x_k). \end{aligned}$$

Proximal algorithms have a lot of applications. For the properties and different algorithms, we suggest to the reader to check the corresponding chapter of the book of N.Parikh and S.Boyd [9].

Note: Before proceeding into the algorithms of our concern, we will denote the set of convex, lower semi-continuous, and proper functions on a Hilbert space \mathcal{H} with $\Gamma_0(\mathcal{H})$.

2.3.1 Moreau Decomposition

In some cases, it is not possible to write explicitly the proximal operator of a function. So instead of trying to find it, we can look for different solutions. For example, we don't have an exact form for a composition of functions (There are some cases that we know, [9], chapter 2). This is where we the conjugate of a function useful in our setting. First, let us define what the convex conjugate is,

⁷ $\text{Proj}_X(y) = \arg \min_{x \in X} \frac{1}{2}\|x - y\|^2$

Definition 2.3. Let $F : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ be a convex function. Its convex conjugate (Legendre transform) $F^* : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by

$$F^*(y) = \sup_{x \in \mathbb{R}^n} \{\langle y, x \rangle - F(x)\}, \quad (2.17)$$

for all $y \in \mathbb{R}^n$.

A basic example would be the case when $F(x) = \|x\|_2^2$. With just applying the definition we obtain $F^*(y) = \|y\|_2^2$.

We have a relation between the proximal operator of the main function and its convex conjugate, which is,

$$v = \text{prox}_F(v) + \text{prox}_{F^*}(v). \quad (2.18)$$

This property, known as *Moreau decomposition*, is the main relation between the proximal operator and its conjugate.

In the next chapter, we will investigate possible estimation algorithms that we can use. Depending on the problem, one of the previous forms of our problem will be used.

2.4 Algorithms

Now we are changing our focus to find a solution for the problem with multiple functions. Recall that at the beginning of this chapter our objective function (2.6) does not have just one but two terms that should be minimized simultaneously. For this case let us define the general problem as in the following form. Suppose, \mathcal{H}, \mathcal{G} are Hilbert spaces. For functions $f \in \Gamma_0(\mathcal{H}), g \in \Gamma_0(\mathcal{G})$,

$$\arg \min_x f(x) + g \circ L(x), \quad (2.19)$$

where f is smooth function and $L : \mathcal{H} \rightarrow \mathcal{G}$ is a linear operator. Now, we are converging to the problem of the reconstruction (2.6), such as $f(x) = \frac{1}{2}\|Ax - y\|^2$, $g(x) = \|x\|_1$ and $Dx = Lx$. Note that our data fidelity term f will always be the L2-norm as a result of the Gaussian assumption.

Next algorithms are useful to find a minimizer in the case of (2.19). We will present 3 different algorithms here.

2.4.1 Proximal Gradient Descent/Forward Backward

In this approach, we decompose the objective function into two components, one of which is differentiable, following the principles of implicit sub-gradient descent.

Note that different splitting leads to different algorithms of the proximal gradient descent algorithm [9].

The *proximal gradient method* is

$$x_{k+1} := \text{prox}_{\gamma g \circ L}(x^k - \gamma \nabla f(x^k)), \quad (2.20)$$

where $\gamma > 0$ is a fixed step size. There many form of this algorithm, we will use the one presented in [5],

Algorithm 1: aimed to solve $\arg \min f(x) + g \circ L(x)$

Input : For $\gamma > 0$
Initialization: x_0
for $k = 1, \dots, K$ **do**

$$\quad \quad \quad \boxed{x_{k+1} := \text{prox}_{\gamma g \circ L}(x_k - \gamma \nabla f(x_k))};$$

2.4.2 Chambolle-Pock

In the previous algorithm problem emerges when we need to compute the proximal operator of $g \circ L$, which is not always possible. To overcome this we will use the *Chambolle-Pock* algorithm which has been presented in [6]. We can write the algorithm as in the following,

Algorithm 2: aimed to solve $\arg \min_{\hat{x}} f(x) + g(L(x))$

Set : $\tau > 0, \sigma > 0, \sigma \tau \|L\|^2 < 1$
Initialization: (x_0, y_0)
for $k = 1, \dots, K$ **do**

$$\quad \quad \quad x_{k+1} = \text{prox}_{\tau f}(x_k - \tau L^* y_k);$$

$$\quad \quad \quad y_{k+1} = \text{prox}_{\sigma g^*}(y_k + \sigma L(2x_{k+1} - x_k));$$

2.4.3 Condat-Vu

To deal with additional term, we need to introduce a new algorithm. For the minimization problem in this form,

$$\arg \min_x f(x) + g(x) + h(Lx). \quad (2.21)$$

we have used the Condat-Vu algorithm introduced in [8].

Algorithm 3: aimed to solve $\arg \min_x f(x) + g(x) + h(Lx)$

Set : $\tau > 0, \sigma > 0, \beta > 0, \frac{1}{\tau} - \sigma \|L\|^2 \geq \frac{\beta}{2}$
Initialization: (x_0, y_0)
for $k = 1, \dots, K$ **do**

$$\quad \quad \quad \tilde{x}_{n+1} := \text{prox}_{\tau g}(x_n - \tau \nabla f(x_n) - \tau L^* y_n);$$

$$\quad \quad \quad \tilde{y}_{n+1} := \text{prox}_{\sigma h^*}(y_n + \sigma L(2\tilde{x}_{n+1} - x_n));$$

$$\quad \quad \quad (x_{n+1}, y_{n+1}) := \rho_n(\tilde{x}_{n+1}, \tilde{y}_{n+1}) + (1 - \rho_n)(x_n, y_n)$$

2.4.4 Plug and Play

Consider the proximal step in the algorithm 1, which is,

$$\text{prox}_{\gamma g}(y) = \arg \min_x \gamma g(x) + \frac{1}{2} \|x - y\|_2^2. \quad (2.22)$$

If we go backward in the idea of MAP, we can see this problem as a denoising problem which would be equivalent to

$$y = x + \varepsilon, \quad (2.23)$$

where ε is Gaussian noise. By comparing with the result of MAP 2.5, we obtain

$$g(x) = -\log(\mathcal{P}(x)) \quad (2.24)$$

The method *Plug and Play* ([10]) (PnP) is emerging at this point. The main idea is to put a pre-trained denoiser⁸ in the iteration instead of the proximal algorithm because the pre-trained denoiser gives an estimation of the x , which is the same as the proximal operator of g . It is a deep learning method with a wide variety of applications. It has different types of advantages, we can construct a model with various noise levels by using the state-of-art architectures independently of the degradation operator.

What we will present here is the proximal gradient descent version of it, however, it can also be used in different iterative schemes.

Once we have a denoiser D_θ , the plug and play algorithm is,

Algorithm 4: Plug and Play Proximal Gradient Descent algorithm

Input : $y \in \mathbb{R}^M$, step size $\gamma > 0$

Initialization: $x^0 = A^*y$

for $n = 0, 1, 2, \dots$ **do**

| $\tilde{x}^n = x_k - \gamma \nabla f(x_k);$
| $x^{n+1} = D_\theta(\tilde{x}^n);$

⁸Deep Neural Network

Chapter 3

Oceanographic Equations

The main purpose of this chapter is to derive the shallow water quasi-geostrophic, which will make the connection between the potential vorticity field q and the sea surface height value s . To obtain this equation we have to consider rotation and stratification effects on larger-scale fluid on a planet. Our concern will be on the mother planet, Earth. The effects of rotation and stratification¹ will be detailed during this chapter. For whom are curious for the effects of rotation and stratification, which we will give an introduction here, we suggest chapter III from the book of Benoit Cushman-Roisin [7] or the chapter 2 from the book of Geoffrey K. Vallis [13].

On the other hand, there is also another factor introduced by geophysical fluid dynamics, which is called *scale of motion*. These are dimensional quantities that express their significance to the interested constraints, they are more like an estimate than exact numbers and are understood solely as *order of magnitude* of physical variables [7].

In general, key scales are *time*, *length* and *velocity*. For example, Hurricane France during her course over the southeastern United States in early September 2004. By observations, it had a nearly circular feature spanning approximately 7.5° of latitude (830 km). The wind speed of a category-4 hurricane such as Frances ranges from 59 to 69 m/s. In general hurricane tracks change (significantly) in direction and speed of propagation over 2-days. So, with this information, we can choose the scale for a hurricane as, the length scale $L = 800$ km, velocity $U = 60$ m/s and for the time $T = 2 \times 10^5$ s (= 55.6 h). For more examples, we suggest readers check chapter 1.3 from the book of Benoit Cushman-Roisin [7].

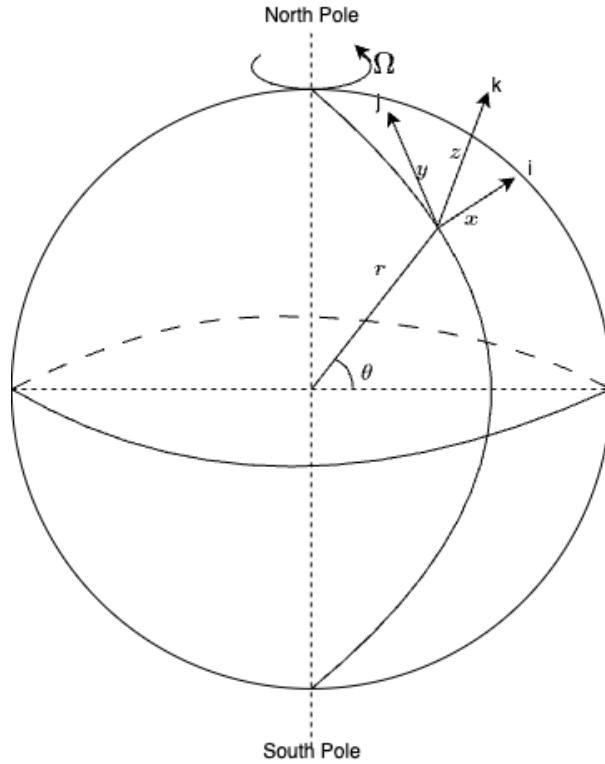
In this chapter we will explain the effects of rotation on geophysical fluids (on oceans more specifically), we will introduce, the primitive equations, what is Rossby scale of radius, Sea Surface height, and the 1.5-layer model that we use.

3.1 Main concepts

3.1.1 Coriolis Force

When we observe the geophysical fluid such as atmosphere, and oceans, we are observing them on a rotating system. From a theoretical point of view working on an inertial frame-

¹Stratification is the interaction between fluids with different densities.



work of reference can be acceptable, it's not the case when we are dealing with numerical approximation. The governing equations should be adapted to a rotating frame of reference, this is emerging an effect called *Coriolis force*, which is a fictitious force. Also, because of rotation, there is an outward force, called *Centrifugal force*. In a rotating frame of reference centrifugal force is accounted for in the gravity constant. So, only the *Coriolis force* will be added the in the forces balance, which is defined as,

Definition 3.1. *The Coriolis force per unit mass is defined as,*

$$\mathbf{F}_c = -2\Omega \times \mathbf{u} \quad (3.1)$$

where relative velocity is $\mathbf{u} = u\mathbf{i} + v\mathbf{j}$ and Ω^2 is defined as,

$$\Omega = \frac{2\pi \text{ radians}}{\text{time of one revolution}}. \quad (3.2)$$

3.1.2 Primitive Equations

In the previous part, we talked about changing the reference frame, but we haven't given the exact results of changing the reference frame, without giving further details, we will just present it.

These will be our main equations of concern, which are represented in a rotating frame. Now, we will define a tangent plane on the surface of the Earth at the latitude θ_0 and then we will use the Cartesian coordinate system (x, y, z) to define the motion. Also, define

²For the Earth, it is $2\pi/24$ hours $\approx 7.2921 \times 10^{-5} \text{ s}^{-1}$.

the velocity $\mathbf{v} = (u, v, w)$, where u, v, w are the components of velocity on the tangent plane.

$$\text{x-momentum: } \frac{\mathbf{D}u}{\mathbf{D}t} - fv = -\frac{1}{\rho} \frac{\partial p}{\partial x} \quad (3.3)$$

$$\text{y-momentum: } \frac{\mathbf{D}v}{\mathbf{D}t} + fu = -\frac{1}{\rho} \frac{\partial p}{\partial y} \quad (3.4)$$

$$\text{z-momentum: } \frac{\mathbf{D}w}{\mathbf{D}t} = -\frac{1}{\rho} \frac{\partial p}{\partial z} - g \quad (3.5)$$

$$\text{Continuity: } \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{x}) = 0 \quad (3.6)$$

where $\mathbf{D}\mathbf{u}/\mathbf{D}t = \partial \mathbf{u}/\partial t + \mathbf{v} \cdot \nabla \mathbf{u}$ is the advection operator. There is also one more equation in this system called, *equation of state*, which describes the thermodynamic properties of the considered fluid. However, it is not in the scope of this work.

3.1.3 f-plane, beta-plane approximations

There are two kinds of approximation methods that can be considered in this case, the first one is called *f-plane* approximation, and the second one is called β -plane approximation.

If the flow field extends over a meridional span that is not too wide, the variation of the Coriolis parameter with latitude is negligible, and f can be taken as a constant [7]. The frame of reference is then called the *f-plane*. On the *f-plane*, we will take $f = f_0 = 2\Omega \sin \theta_0$.

The reason why we have the second approximation is the magnitude of the vertical components of rotation varies with latitude, and it has important dynamical consequences. It is possible to approximate these small variations,

$$f = 2\Omega \sin \theta \approx 2\Omega \sin \theta_0 + 2\Omega(\theta - \theta_0) \cos \theta_0. \quad (3.7)$$

It is possible to mimic this effect by having a small change the Coriolis parameter,

$$f = f_0 + \beta y, \quad (3.8)$$

where $\beta = \partial f / \partial y = (2\Omega \cos \theta_0) / r$. This is what we call β -plane approximation, and during the text, we will consider this approach.

3.2 Small aspect ratio limit

In general cases, geophysical fluids have a very big length scale (denoted by L) if you compare it with its thickness (denoted by H). Like, in the oceans, the length scale L is around 10^4 meters, whereas the thickness of the fluid is around 10^2 meters. This gives us an opportunity to neglect some values, *hydrostatic balance* is the result of this, which we will give the definition above. But before going to that, let us check why it's happening.

For incompressible fluids density is constant, so the continuity equation 3.6 becomes,

$$\frac{\partial u}{\partial t} + \frac{\partial v}{\partial t} + \frac{\partial w}{\partial t} = 0 \quad (3.9)$$

Variable	Scale	Unit	Atmospheric value	Oceanic value
x, y	L	m	$100 \text{ km} = 10^5 \text{ m}$	$10 \text{ km} = 10^4 \text{ m}$
z	H	m	$1 \text{ km} = 10^3 \text{ m}$	$100 \text{ m} = 10^2 \text{ m}$
t	T	s	$\leq \frac{1}{2} \text{ days} \approx 4 \times 10^4 \text{ s}$	$\leq \frac{1}{2} \text{ days} \approx 4 \times 10^4 \text{ s}$
u, v	U	m/s	10 m/s	0.1 m/s
p	P	$\text{kg m}^{-1} \text{s}^{-2}$	variable	variable
ρ	$\nabla \rho$	kg/m^3	variable	variable

Table 3.1: Typical scales of Atmospheric and Oceanic flows [7]

Let us assume, for the x, y axis we have the same length and velocity scale L and U respectively, and for the z axis, let us denote them with W and H (thickness of the fluid). This means, in the above continuity equation 3.9, we have three terms, of respective order of magnitude:

$$\frac{U}{L}, \frac{U}{L}, \frac{W}{H}$$

This leads us to three different cases: $\frac{W}{H}$ is much less than, on the order of, or much greater than $\frac{U}{L}$. The third option is not possible, ([7], chapter 4.3). This leads us to the point where the vertical-velocity scale is constrained by the horizontal-velocity scale;

$$W \lesssim \frac{H}{L} U \quad (3.10)$$

by the light of the assumption that we made at the beginning ($H \ll L$), we have,

$$W \ll U \quad (3.11)$$

Therefore, the large-scale geophysical fluids are shallow $H \ll L$ and nearly two-dimensional. This scaling idea is an important aspect when you consider the geophysical fluid dynamics, one can say that this is just a numerical perspective, but that would be a correct guess. The assumptions and scaling are the reason for simplifying the equations of motion, as a result not all the assumptions that we will make are close to reality but with respect to our needs during the studies.

3.3 Hydro-static Balance

When we continue the argument of scaling, we will see that in the equation 3.5 the term $\frac{Dw}{Dt}$ is sufficiently small compared to the gravity, so it's negligible. This what we call *hydro-static balance*.

Definition 3.2. *Hydro-static balance is the state where the gravitational acceleration balances the vertical component of the pressure gradient force. The hydrostatic pressure gradient only has a vertical component, which is equal to,*

$$\frac{\partial p}{\partial z} = -\rho g \quad (3.12)$$

In both oceanographic and atmospheric cases, the equation 3.12 is a good way to approximate to so-called *z-momentum equation*.

3.4 Geostrophic Balance

There is a way to compare the effect of Coriolis acceleration and relative acceleration, the ratio is defined as *Rossby number* is essentially characterizing the effects of rotation and is defined as;

$$R_o \equiv \frac{U}{fL} \quad (3.13)$$

where $f = 2\Omega \sin \theta$, is the Coriolis parameter. If the Rossby number is small then rotation effects are important and possible dominant ([13], chapter 3.4). This will be our assumption during these studies, we assume that we have a small Rossby number, so the rotation term dominates the non linear advection term in the horizontal momentum equations, so only term that can balance the rotation term is the pressure term, so the equations 3.3 and 3.4, turn into,

$$fu \approx -\frac{1}{\rho} \frac{\partial p}{\partial y}, \quad fv \approx \frac{1}{\rho} \frac{\partial p}{\partial x} \quad (3.14)$$

This balance is known as *geostrophic balance*. Note that it is an important assumption to simplify the Navier-Stokes equations so that we can use it.

3.5 Vorticity

To understand the motion of fluid, there is a concept called *vorticity* in continuum mechanics, which is a vector field that models the circulation of a fluid. This is an important aspect of fluid dynamics to explain the complex nature of the flow. Mathematically, for the velocity field $\mathbf{u} = (u, v)$ it is defined as,

$$\zeta = \nabla \times \mathbf{u} = \partial_x v - \partial_y u \quad (3.15)$$

However, in our case, it is different from all the previous quantities that we worked on. This new vorticity field will be affected by the assumptions that we made and will do.

3.6 Shallow Water Equations

As section 3.2 suggests, our system is shallow ($H \ll L$). In this system, momentum equations turn into,

$$\frac{\mathbf{D}\mathbf{u}}{\mathbf{D}t} + f \times \mathbf{u} = -g \nabla \eta, \quad (3.16)$$

where f may be constant or change with latitude like in the previous cases. Also, the adapted version of the mass continuity equation will be,

$$\frac{\mathbf{D}h}{\mathbf{D}t} + h \nabla \cdot \mathbf{u} = 0. \quad (3.17)$$

where h is total fluid thickness.

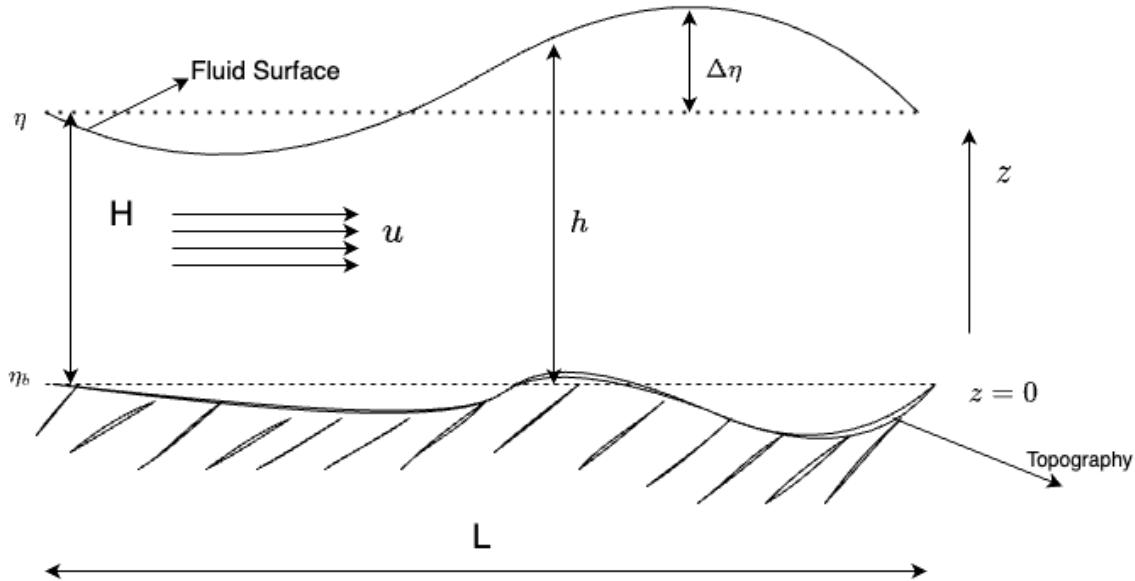


Figure 3.1: One layer Shallow Water

3.6.1 Reduced Gravity Equations for Free Upper surface

Now consider single shallow layer is on top of a quiescent³ fluid (Fig 3.2). This model is often used to model the upper ocean, which is our consideration, also, the result of section 3.2 gives better approximations than the one-layer model. Moreover, this part will give a clear idea of why we use sea surface height field in our studies which will be explained in the chapter 4.

Here we will derive the equations for the oceanic case (active layer top). The pressure in the upper layer is given by integrating the hydrostatic equation down from the upper surface. Thus, at a height \$z\$ in the upper layer

$$p_1(z) = g\rho_1(\eta_0 - z), \quad (3.18)$$

where the \$\eta_0\$ is the height of the upper surface. Hence we obtain,

$$\frac{1}{\rho_1} \nabla p_1 = g \nabla \eta_0, \quad (3.19)$$

and the momentum equation is

$$\frac{\mathbf{D}\mathbf{u}}{\mathbf{D}t} + f \times \mathbf{u} = -g \nabla \eta_0. \quad (3.20)$$

At the lower layer pressure is the weight of the first layer plus it own pressure, which is,

$$p_2(z) = g\rho_1(\eta_0 - \eta_1) + g\rho_2(\eta_1 - \eta_z). \quad (3.21)$$

However, velocity of this layer is zero, the horizontal pressure gradient in it is zero, therefore,

$$\rho_1 g \eta_0 = -\rho_1 g' \eta_1 + c, \quad (3.22)$$

³Fluid with zero velocity.

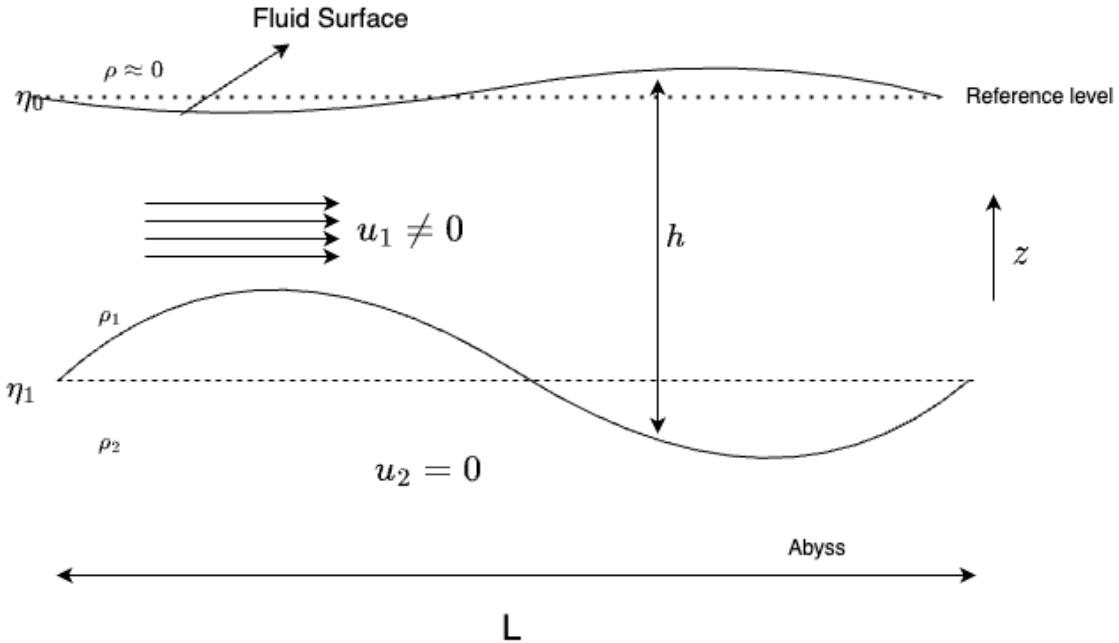


Figure 3.2: 1.5-layer Shallow water system

where c is a constant and $g' = g(\rho_2 - \rho_1)/\rho_1$, and normally $(\rho_2 - \rho_1)/\rho_1 \ll 1$ and $g' \ll g$. The momentum equation becomes,

$$\frac{\mathbf{D}\mathbf{u}}{\mathbf{D}t} + f \times \mathbf{u} = g\nabla\eta_1, \quad (3.23)$$

and the mass continuity equation becomes,

$$\frac{\mathbf{D}h}{\mathbf{D}t} + h\nabla \cdot \mathbf{u} = 0. \quad (3.24)$$

where $h = \eta_0 - \eta_1$. Because $g \gg g'$, 3.22 shows that surface displacement is *much smaller* than the displacements at the interior interface. Hereby, we can make the relation between surface fluctuation and interior changes, where we will use the surface level change, we can ignore it in the main equations by relation 3.22.

3.7 Quasi-Geostrophic Equations

3.7.1 Rotating Fluid

Rotation has a big effect on the shallow water system because the solution can exist with non-zero gradients in the height field (the associated pressure gradients being balanced by the Coriolis force) and the potential vorticity conservation provides a powerful constraint on the fluid evolution. In a rotating shallow fluid that conservation is represented by

$$\frac{\partial Q}{\partial t} + \mathbf{u} \cdot \nabla Q = 0, \quad (3.25)$$

where $Q = (\zeta + f)/h$. This quantity is called *shallow water potential vorticity*. Before deriving the considered potential vorticity equation let us summarize what we have done so far.

3.7.2 Shallow Water Quasi-Geostrophic Potential Vorticity

In every scale and assumption, there are different cases. The *quasi geostrophic* equations are appropriate for the scales of the same order as the deformation radius L_d ($= \frac{\sqrt{g'H}}{f_0}$ [13], chapter 5.2.2). We assume the geophysical fluid is almost in geostrophic balance, which means,

$$R_o \ll 1, \quad \frac{L}{L_d} = \mathcal{O}(1) \quad (3.26)$$

Also, suppose that the Coriolis parameter is almost constant. As in the limit $R_o \rightarrow 0$, the geostrophic balance 3.14 turns into,

$$f_0 u = -g \frac{\partial \eta}{\partial y}, \quad -f_0 v = g \frac{\partial \eta}{\partial x} \quad (3.27)$$

Where we assume that we have a divergence free fluid. We made the following assumptions.

- The Rossby number is small, $R_o \ll 1$.
- Length scales are order of the same order as the deformation radius, $\frac{L}{L_d} = \mathcal{O}(1)$
- Variations in the Coriolis parameter is small, which is $|\beta y| \sim R_o f_0$

Now we are ready to give the definition of *stream function* ψ ,

$$v = \partial_x \psi, \quad u = -\partial_y \psi, \quad \text{and} \quad \eta = \frac{f_0 \psi}{g}. \quad (3.28)$$

Last part is the result of the assumption on geostrophic balance. Hence we have,

$$\zeta = \partial_x v - \partial_y u = \nabla^2 \psi \quad (3.29)$$

Now lets consider the shallow water potential vorticity Q ,

$$Q = \frac{f + \zeta}{h},$$

Now let $h = H(1 + \eta'/H)$, where η' is the perturbation of the free-surface height, and assume that η'/H is small to obtain,

$$Q = \frac{f + \zeta}{h} \approx \frac{1}{H}(f + \zeta)(1 - \frac{\eta'}{H}) \approx \frac{1}{H}(f_0 + \beta y + \zeta - f_0 \frac{\eta'}{H}). \quad (3.30)$$

f_0/H is constant and so it has no effect in the evolution equation, and the quantity given by

$$q = \beta y + \zeta - f_0 \frac{\eta'}{H}, \quad (3.31)$$

is materially conserved. By the assumption on the geostrophic balance we have $\zeta = \nabla^2 \psi$ and $\eta' = f_0 \psi/g$. The equation, 3.25 turns into,

$$\frac{\partial q}{\partial t} + \vec{u} \cdot \nabla q = 0, \quad (3.32)$$

where $\vec{u} = (-\partial_y \psi, \partial_x \psi)$. Also, we obtain,

$$q = \nabla^2 \psi - \frac{1}{L_d^2} \psi + \beta y. \quad (3.33)$$

This quantity is called the *shallow water quasi-potential vorticity*.

Chapter 4

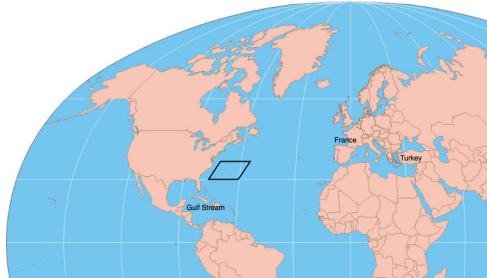
Image Processing Tools for SSH Recovery

4.1 Data-set

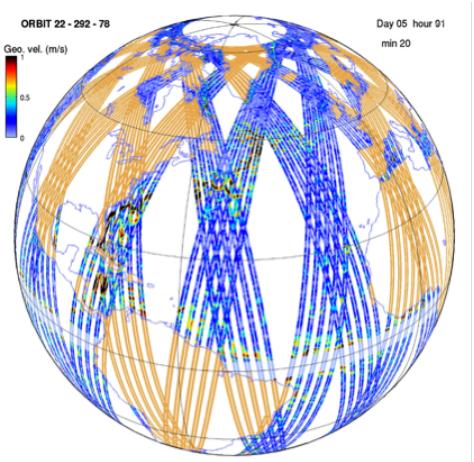
4.1.1 Case-study region, dates and NATL60 data

The main objective is to reconstruct the sea surface height (SSH) value for the partially observed sea surface region. For this purpose the case region that has been chosen is $10^\circ \times 10^\circ$ subpart of GULFSTREAM domain [33° N, 43° N ; -65° W, -55° W].

However, the data we consider is not original satellite data but the pseudo data extracted from the reference simulation. The reference simulation is based on a nature run of NATL60 configuration [15] of the NEMO (Nucleus for European Modeling of the Ocean) model [18], and extraction of SWOT-like data is briefly explained in the article [12], the track of the SWOT-like data can be seen the figure 4.1b. This data is hosted on the website AVISO+, and in this document format of the data is explained in [16]. The observation data is available from '2012-10-02' to '2013-09-30', and for each day there are different numbers of points, which makes it hard to see as discrete data per day. As you can see in the histogram 4.6. One of the main reasons that we consider observation-like



(a) World Map



(b) Swot Track simulations Credit: MITGCM

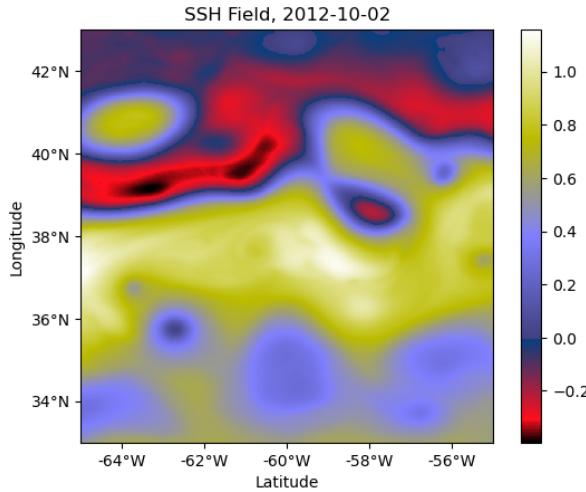


Figure 4.2: First entry of the numerical simulations, 2012-10-02

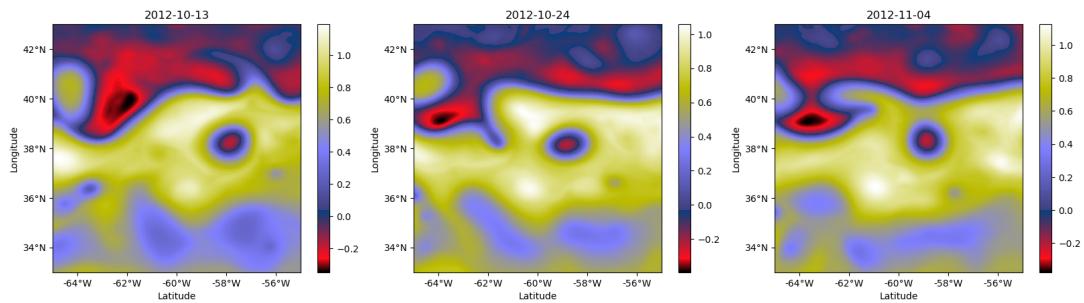


Figure 4.3: Average SSH field of subjected days

data is in real life is already hard to deal with big data sets, so we want to deal with reasonable settings. It also has clear divergent features associated with quasi-geostrophic flows, which makes it suitable for this study.

We worked on three dates, '2012-10-13', '2012-10-24', and '2012-11-04'. Which has different aspects, dynamics etc.

Also, even though the subject is deeply related to dynamics, we consider it a static image. In the experiments, we haven't included the time dynamical part. Other methods regarding dynamics are explained in the next chapter.

The swot-like data has an interval of 11 days, which is the least amount of time window in which we have enough information to reconstruct the desired field, which we can clearly see in the figure 4.7. Because of this reason, we will consider the ground truth data of an average of 11 days, instead of just using a snapshot reference value.

4.1.2 Degraded Data Set

To measure the performance level of the algorithms, we also tested them with a randomly generated Gaussian mask, as in the figure 4.5. We also increase the percentage of the mask data in the observations, in the degraded data set it is around 84%. In the satellite observations it is around 48%, the challenge is that we don't have any information from the big connected parts of the images.

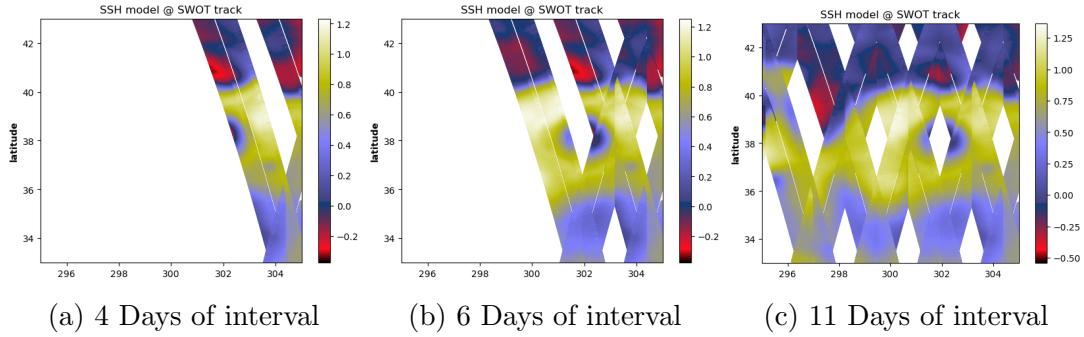


Figure 4.4: Different number of day intervals, with central date '2012-10-13'

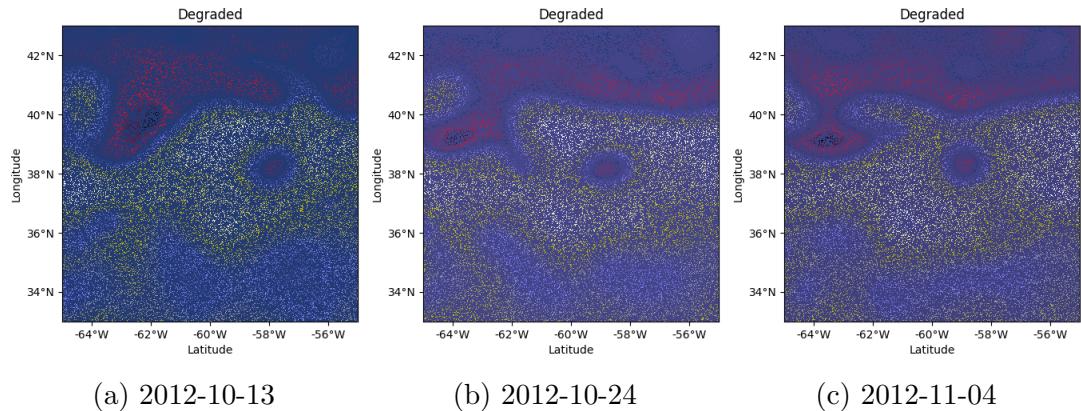


Figure 4.5: Random Gaussian mask

4.1.3 Preprocess of the data set

The number of data points per day changes on the subjected area due to the fact that the SWOT is observing different parts of the ocean during the time period as shown in the figure 4.1b. As can be seen in the example histogram 4.6, it varies a lot from day to day. Also, there are certain regions that have data from different times. This makes the problem harder in the sense that we don't have just one value for one point in the region. Also, time series are almost continuous, we have a lot of close data points. To overcome this problem/s, we have processed the data by taking the mean value for certain grid structures. The purpose is to have a data set in which we can apply non-smooth optimization methods which are developed in discrete settings.

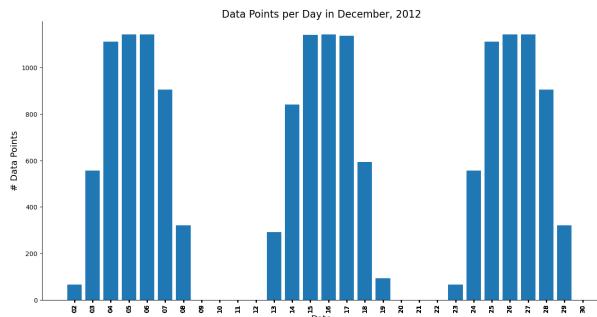


Figure 4.6: Histogram

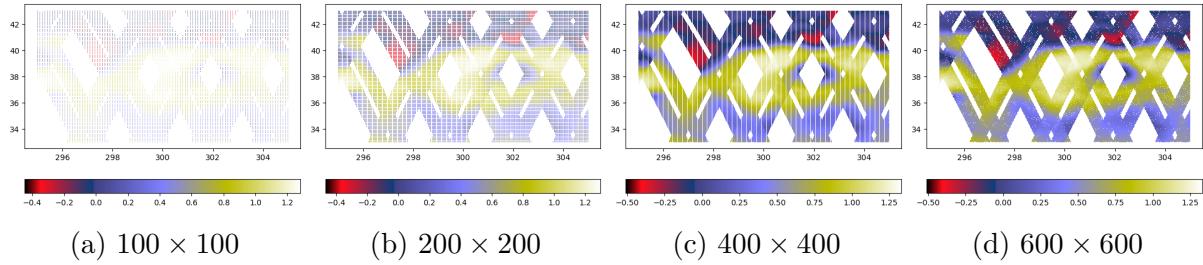


Figure 4.7: Effects of Discretization

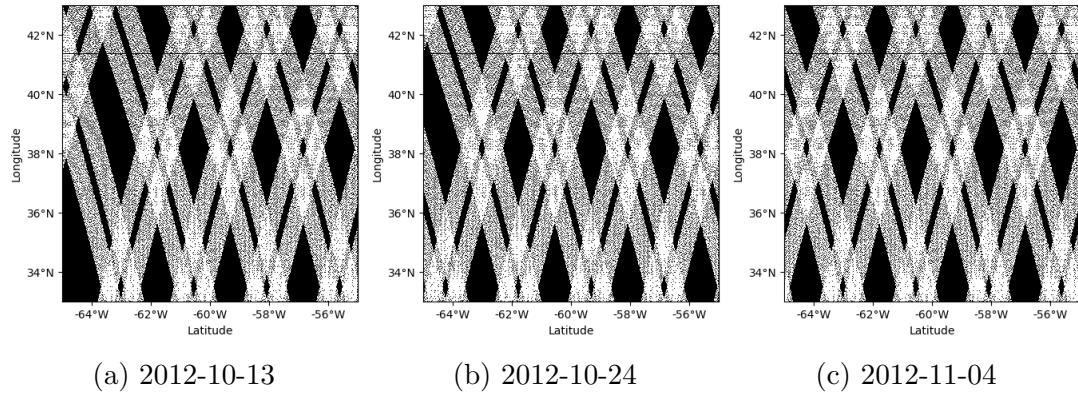


Figure 4.8: Mask

4.1.4 Mask

Observation Operator

In all the iterative algorithms we consider the inpainting operator \mathbf{H} as a 600×600 matrix, where $(\mathbf{H}_{ij}) \in \{0, 1\}$, which we can see them clearly in the figure 4.8.

4.1.5 Measurements

PSNR

The Peak Signal-to-Noise Ratio (PSNR) is a commonly used metric to measure the quality of a reconstructed or processed signal, such as an image or a video, in comparison to the original or reference signal. It quantifies the distortion or loss of information in the signal due to compression or other forms of processing.

The PSNR is calculated using the mean squared error (MSE) between the original signal I and the reconstructed signal \hat{I} . The formula for PSNR in decibels (dB) is given by:

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{\text{MAX}^2}{\text{MSE}} \right) \quad (4.1)$$

Where:

PSNR : Peak Signal-to-Noise Ratio (in dB)

MAX : Maximum possible pixel value (e.g., 255 for an 8-bit image)

MSE : Mean Squared Error between I and \hat{I}

Root Mean Square Error

One of the measurement that we are using is described in the data challenge [16], the statistical comparison is based on *Root mean square error based* score, $RMSE_S$ defines as:

$$RMSE_S(t) = 1 - \frac{RMSE(t)}{RMS(SSH_{true})} \quad (4.2)$$

where the root mean square function with,

$$RMSE(t) = \sqrt{\frac{1}{N}(SSH_{est}(t, i) - SSH_{true}(t, i))^2}. \quad (4.3)$$

N is the number of pixels included in the study domain.

4.1.6 Baseline Method: Optimal Interpolation

The optimal interpolation method is the baseline method of the Ocean data challenge [16]. Which is using the MAP and additionally defining a covariance matrix for the distributions. During this section we will denote the prior model with s^f which is correlated with the s via matrix \mathbf{P}^f and \mathbf{R}^{-1} will denote the covariance matrix of the observations. Then for Gaussian distribution, we have the following quantities that we will use to solve the MAP,

Prior:

$$\mathcal{P}(s) \propto \exp\left\{-\frac{1}{2}(s - s^f)^T \mathbf{P}^{f-1} (s - s^f)\right\}$$

Likelihood:

$$\mathcal{P}(o|s) \propto \exp\left\{-\frac{1}{2}(o - \mathbf{H}s)^T \mathbf{R}^{-1} (o - \mathbf{H}s)\right\}$$

Posterior:

$$\mathcal{P}(s|o) \propto \exp\left\{-\frac{1}{2}(s - s^f)^T \mathbf{P}^{f-1} (s - s^f) + -\frac{1}{2}(o - \mathbf{H}s)^T \mathbf{R}^{-1} (o - \mathbf{H}s)\right\}$$

We want to find a maximum a posterior probability $\mathcal{P}(s|o)$, so to minimize this, we can minimize,

$$J(s) = -\frac{1}{2}(o - \mathbf{H}s)^T \mathbf{R}^{-1} (o - \mathbf{H}s) - \frac{1}{2}(s - s^f)^T \mathbf{P}^{f-1} (s - s^f) \quad (4.4)$$

For the notation we can write it in the following form,

$$J(s) = \frac{1}{2}\{\|o - \mathbf{H}s\|_{\mathbf{R}^{-1}}^2 + \|s - s^f\|_{\mathbf{P}^{f-1}}^2\}, \quad (4.5)$$

where $\|s\|_A = \langle As, s \rangle$. As the usual method, (since the norms are differentiable and \mathbf{H} is linear), we are taking the derivative of $J(s)$,

$$\begin{aligned} J(s)' &= \left(\frac{1}{2}\{\|o - \mathbf{H}s\|_{\mathbf{R}^{-1}}^2 + \|s - s^f\|_{\mathbf{P}^{f-1}}^2\}\right)' \\ &= \left(\frac{1}{2}(\langle -\mathbf{R}^{-1}\mathbf{H}\mathcal{I}_n, o - \mathbf{H}s \rangle + \langle -\mathbf{R}^{-1}(o - \mathbf{H}s), \mathbf{H}\mathcal{I} \rangle + \langle \mathbf{P}^{f-1}\mathcal{I}, s - s^f \rangle + \langle \mathbf{P}^{f-1}(s - s^f), \mathcal{I} \rangle\right)' \\ &= \langle -(\mathbf{R}^{-1}\mathbf{H})^T(o - \mathbf{H}s), \mathcal{I} \rangle + \langle \mathbf{P}^{f-1}(s - s^f), \mathcal{I} \rangle \\ &= 0. \end{aligned}$$

So, we have,

$$\begin{aligned}
\mathbf{P}^{f^{-1}}(s - s^f) &= (\mathbf{R}^{-1}\mathbf{H})^T(o - \mathbf{H}s) \\
\mathbf{P}^{f^{-1}}s - \mathbf{P}^{f^{-1}}s^f &= \mathbf{H}^T\mathbf{R}^{-1}o - \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}s \\
(\mathbf{H}^T)^{-1}\mathbf{P}^{f^{-1}}(s - s^f) &= \mathbf{R}^{-1}o - \mathbf{R}^{-1}\mathbf{H}s \\
\mathbf{R}(\mathbf{P}^f\mathbf{H}^T)^{-1}(s - s^f) &= o - \mathbf{H}s \\
\mathbf{R}(\mathbf{P}^f\mathbf{H}^T)^{-1}(s - s^f) &= o - \mathbf{H}s + \mathbf{H}s^f - \mathbf{H}s^f \\
\mathbf{R}(\mathbf{P}^f\mathbf{H}^T)^{-1}(s - s^f) + \mathbf{H}(s - s^f) &= o - \mathbf{H}s^f \\
\mathbf{R}(\mathbf{P}^f\mathbf{H}^T)^{-1}(s - s^f) + \mathbf{H}(\mathbf{P}^f\mathbf{H}^T)(\mathbf{P}^f\mathbf{H}^T)^{-1}(s - s^f) &= o - \mathbf{H}s^f \\
(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R})(\mathbf{P}^f\mathbf{H}^T)^{-1}(s - s^f) &= o - \mathbf{H}s^f \\
(\mathbf{P}^f\mathbf{H}^T)^{-1}(s - s^f) &= (\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R})^{-1}(o - \mathbf{H}s^f).
\end{aligned}$$

Therefore we have the solution as,

$$\hat{s} = s^f + \mathbf{P}^f\mathbf{H}^T(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R})^{-1}(o - \mathbf{H}s^f) \quad (4.6)$$

Minimizing the function in the equation 4.5 is also known as 3-dimensional variational assimilation (abbreviated as "3Dvar"), in the following chapter other methods in data assimilation will be investigated such as 4Dvar, and Kalman filtering.

4.2 Proposed Methods

The methods that have been considered in this section depend on different variational problems and we have used two main algorithms that we introduced in the chapter 2. Which are *Chambolle Pock* and *Condat Vu*.

4.2.1 Total Variation

As a first approach, we used the total variation penalization term that we mentioned in the chapter 2, recall this variational problem as,

$$\arg \min \frac{1}{2} \|o - \mathbf{H}s\|_2^2 + \lambda \|Ds\|_1, \quad (4.7)$$

where $\lambda > 0$ is the regularization parameter.

During this section, we will consider the initialization of the algorithms as s_0 and $v_0 = D^*(s_0)$. $s_0 = H(s)$ where the 'NaN' values of discretized SSH filled are filled with 0. The reason why we are choosing is because it is the reference value for the sea surface height (figure 3.2). D^* represents the conjugate of the discrete gradient operator D and corresponds to the discrete divergence operator.

We have chosen the step size $\sigma = 1$, and to make the convergence guarantee $\tau = \frac{1}{\sigma\|D\|^2}$.

The proximal operator of the data fidelity term turns to

$$\text{prox}_{\sigma f}(v) = (\sigma I + \mathbf{H}^*\mathbf{H})^{-1}(\sigma\mathbf{H}^*o + v), \quad (4.8)$$

where $\sigma > 0$ is the step size. The distinction between the proximal gradient descent and others happens here, we don't have a closed form expression for the regularization term $g \circ D$, so instead of this we will use the proximal operator of g^* where $g(\cdot) = \|\cdot\|_1$, which is called soft-thresholding,

$$\left(\text{soft}_{[-\gamma, \gamma]}(x_k) \right)_{1 \leq n \leq N} \quad (4.9)$$

where,

$$\text{soft}_{[-\gamma, \gamma]}(x_k) = \begin{cases} x + \gamma & \text{if } x < -\gamma \\ 0 & \text{if } -\gamma \leq x \leq \gamma \\ x - \gamma & \text{if } \gamma < x \end{cases} \quad (4.10)$$

By Moreau decomposition 2.3.1, we have the convex conjugate as,

$$\text{prox}_{\sigma g^*}(v) = v - \text{prox}_{\sigma g}(v). \quad (4.11)$$

Algorithm 5: Chambolle Pock Algorithm with the minimization problem 4.7

Set : $\sigma = 1, \tau = \frac{1}{\sigma \|D\|^2}$
Initialization: $s_0 = \mathbf{H}(s^t), v_0 = D(s_0)$
for $k = 1, \dots, K$ **do**
 $t_k = s_n - \tau D^* v_n$;
 $s_{k+1} = (\sigma I + \mathbf{H}^* \mathbf{H})^{-1}(\sigma \mathbf{H}^* o + t_k)$;
 $v_{k+1} = \text{prox}_{\sigma g^*}(v_k + \sigma D(2s_{k+1} - s_k))$;

4.2.2 Total Variation on Potential Vorticity

This method is inspired by the Total Variation method, where we replace the $D(s)$ with $D(q(s))$. To have a smooth field that respects to underlying physical system, in our case, it is the potential vorticity. Then the minimization problem turns into,

$$\arg \min_s \|o - \mathbf{H}s\|_2^2 + \lambda \|D(\mathcal{L} \frac{g}{f_0} s - \frac{1}{L_d^2 f_0} s + \beta y)\|_1 \quad (4.12)$$

where $\lambda > 0$ is the regularization parameter and the quantity $\frac{g}{f_0} \mathcal{L} s - \frac{1}{L_d^2 f_0} s + \beta y (= q(s))$ is the potential vorticity 3.33 that we have derived in the chapter 3. Note that here \mathcal{L} denotes the discrete Laplacian operator which is denoted as ∇^2 in continuous settings before.

We can calculate the regularization term,

$$\begin{aligned} \|D(\mathcal{L} \frac{g}{f_0} s - \frac{1}{L_d^2 f_0} s + \beta y)\|_1 &= \|D(\mathcal{L} \frac{g}{f_0} s - \frac{1}{L_d^2 f_0} s) + c_1\|_1 \\ &= \frac{g}{f_0} \|D(\mathcal{L} - \frac{1}{L_d^2} I)s + c_2\|_1 \\ &= \frac{g}{f_0} \|Bs + c_2\|_1, \end{aligned}$$

where we define the $B = D(\mathcal{L} - \frac{1}{L_d^2} I)$ and then its conjugate $B^* = (\mathcal{L})^*(D^*) - \frac{1}{L_d^2} D^*$.

Now for the convergence guarantee, we need to consider the operator norm of B . Also, for the proximal step we need to consider the translation of g^* , which is equal to

$$\text{prox}_{\sigma g_c^*}(v) = v - c + \text{prox}_{\sigma g}(v - c). \quad (4.13)$$

Therefore, the algorithm of our consideration turn into,

Algorithm 6: Chambolle Pock Algorithm with the minimization problem 4.12

Set : $\sigma = 0.9, \tau = \frac{1}{\sigma \|D\|^2}$
Initialization: $s_0 = \mathbf{H}(s^t), v_0 = B(s_0)$
for $k = 1, \dots, K$ **do**
 $t_k = s_n - \tau B^* v_n$;
 $s_{k+1} = (\sigma I + \mathbf{H}^* \mathbf{H})^{-1}(\sigma \mathbf{H}^* o + t_k)$;
 $v_{k+1} = \text{prox}_{\sigma g_c^*}(v_k + \sigma B(2s_{k+1} - s_k))$;

4.2.3 Hybrid Method

In this method the considered minimization problem is,

$$\arg \min_s \|o - \mathbf{H}s\|_2^2 + \frac{\chi}{2} \|Ds\|_2^2 + \lambda \|D(\mathcal{L} \frac{g}{f_0} s - \frac{1}{L_d^2} f_0 s + \beta y)\|_1. \quad (4.14)$$

where $\chi, \lambda > 0$. In this method, we have considered the Condat-Vu algorithm to deal with the extra term we have. As before we are dealing with the first and last term with the same computations, the gradient of the $\frac{\chi}{2} \|Ds\|_2^2$ has been added, which is,

$$\nabla(\frac{\chi}{2} \|Ds\|_2^2) = \chi D^*(Ds). \quad (4.15)$$

To make the convergence guarantees, there is an extra term κ which is the Lipschitz constant of the discrete gradient D . We have chosen $\sigma = 100$, therefore $\tau = 1/(\frac{\chi}{2} \|D\|^2 + \sigma \|B\|^2)$. Hereby, we have used the following algorithm to minimize 4.14

Algorithm 7: Condat-Vu algorithm aim to solve 4.14

Set : $\sigma = 100, \tau = \frac{1}{\frac{\chi}{2} \|D\|^2 + \sigma \|B\|^2}, \rho \in]0, 2[$
Initialization: $s_0 = \mathbf{H}(s^t), v_0 = D(s_0)$
for $k = 1, \dots, K$ **do**
 $t_n := s_n - \tau D^*(Ds_n) - \tau B^* v_n$;
 $\tilde{s}_{n+1} := (\sigma I + \mathbf{H}^* \mathbf{H})^{-1}(\sigma \mathbf{H}^* o + t_n)$;
 $\tilde{v}_{n+1} := \text{prox}_{\sigma g_c^*}(v_n + \sigma B(2\tilde{s}_{n+1} - s_n))$;
 $(s_{n+1}, v_{n+1}) := \rho(\tilde{s}_{n+1}, \tilde{v}_{n+1}) + (1 - \rho)(s_n, v_n)$

4.3 Results and Discussion

In the figures 4.9, 4.10, we can see the effects of the regularization parameters both on the observational data and the degraded data in the algorithms 5, where we use the total variation as the regularization term. Even though the performances of the iteration have increased up to certain points, a high regularization parameter causes a high smoothing effect on the estimations. We can see in the tables 4.1 and 4.2 that the performance

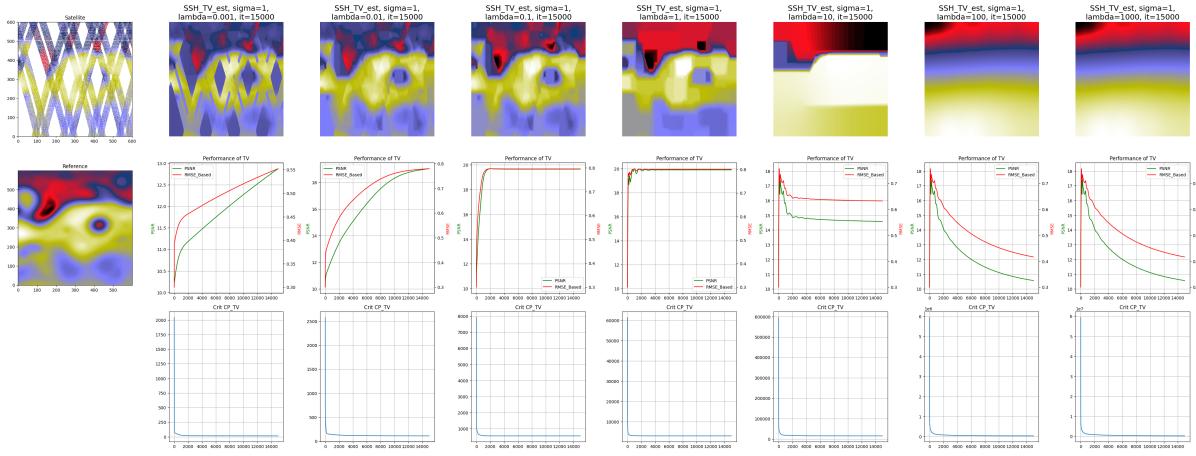


Figure 4.9: Total Variation Penalty with different regularization parameters. Real mask

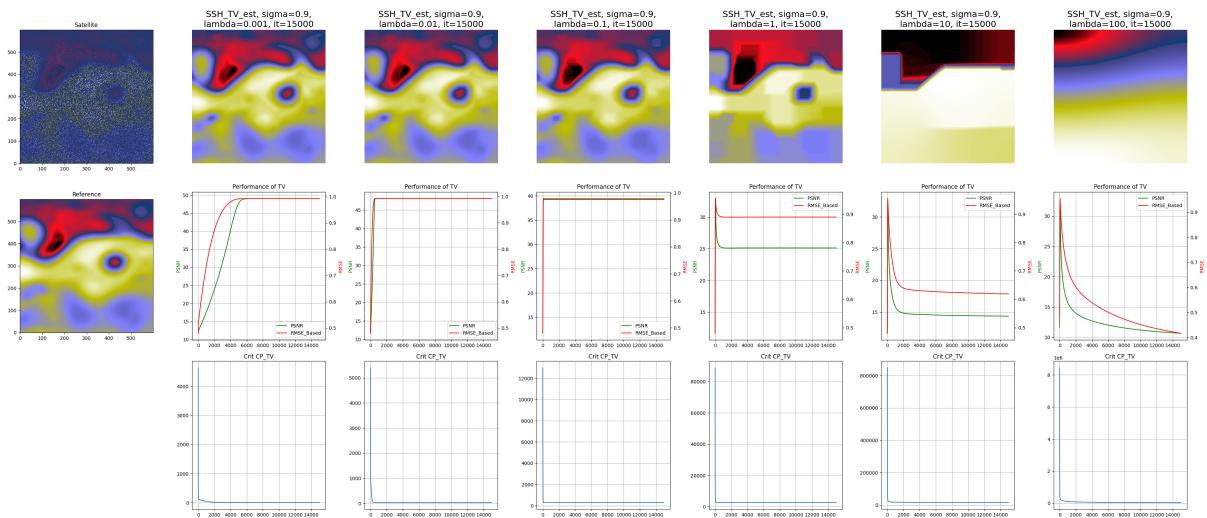


Figure 4.10: Total Variation Penalty with different regularization parameters. Random Mask

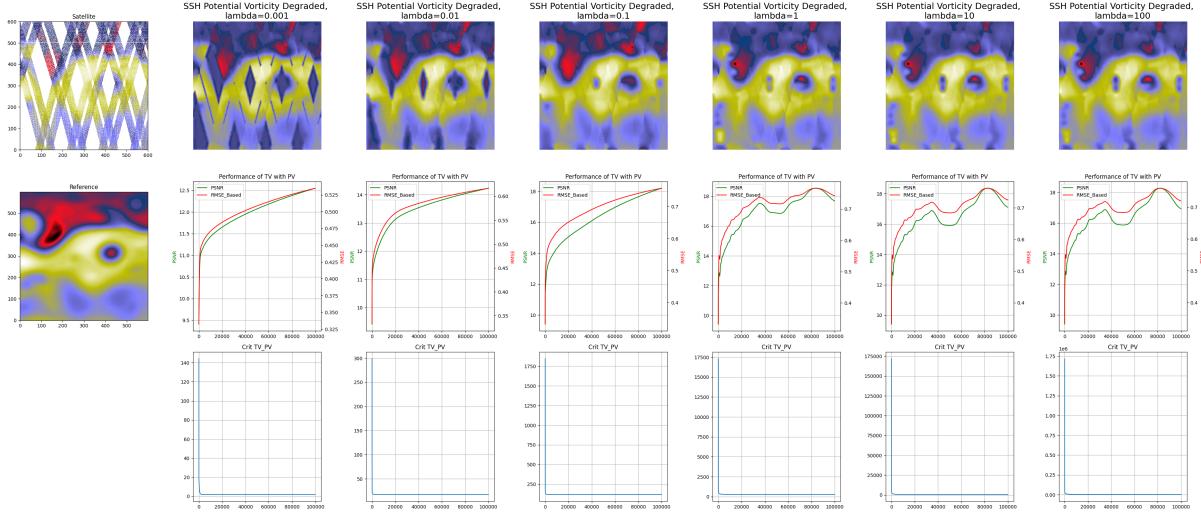


Figure 4.11: Total Variation combined with Potential Vorticity for different regularization parameters. Real mask.

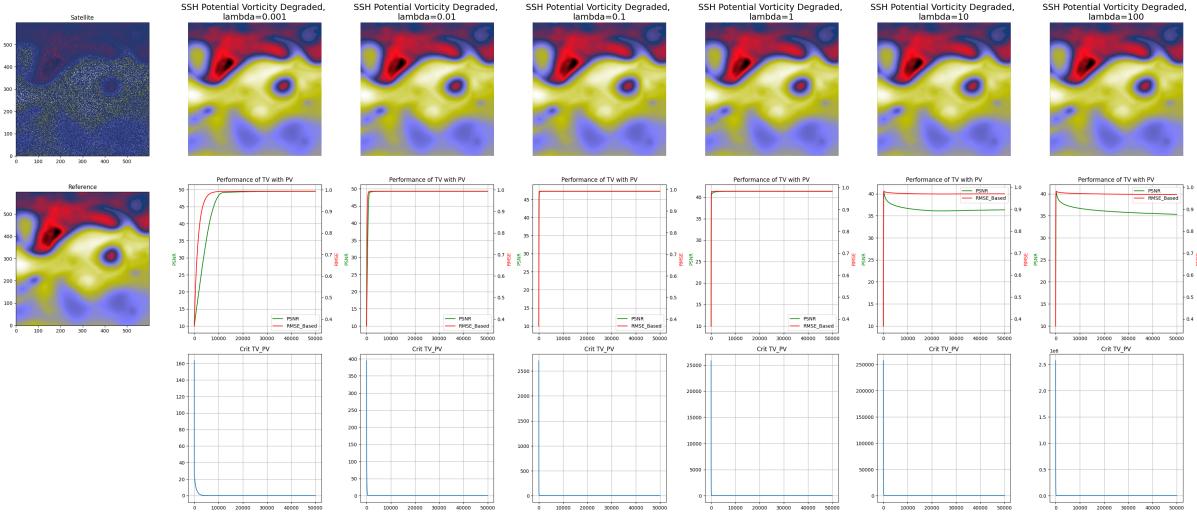
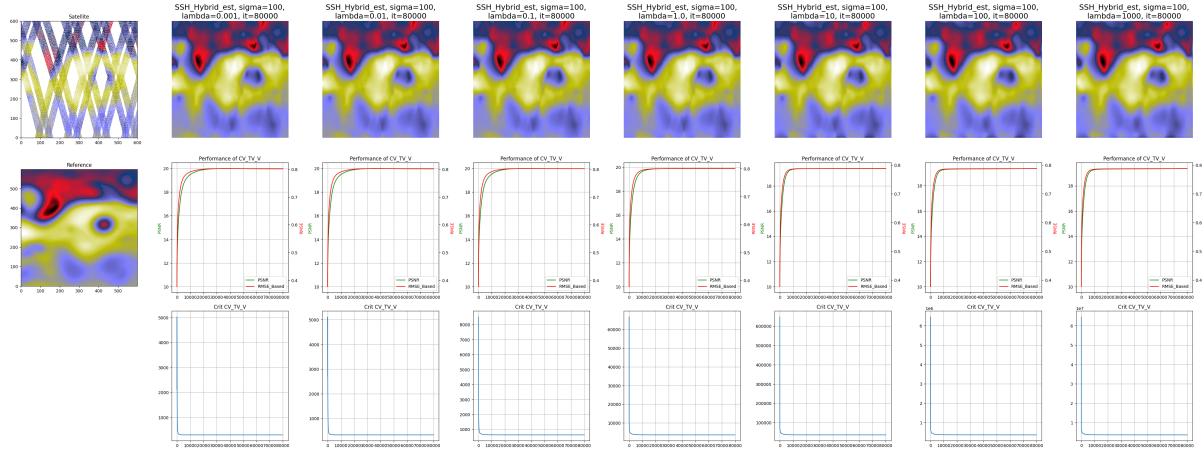
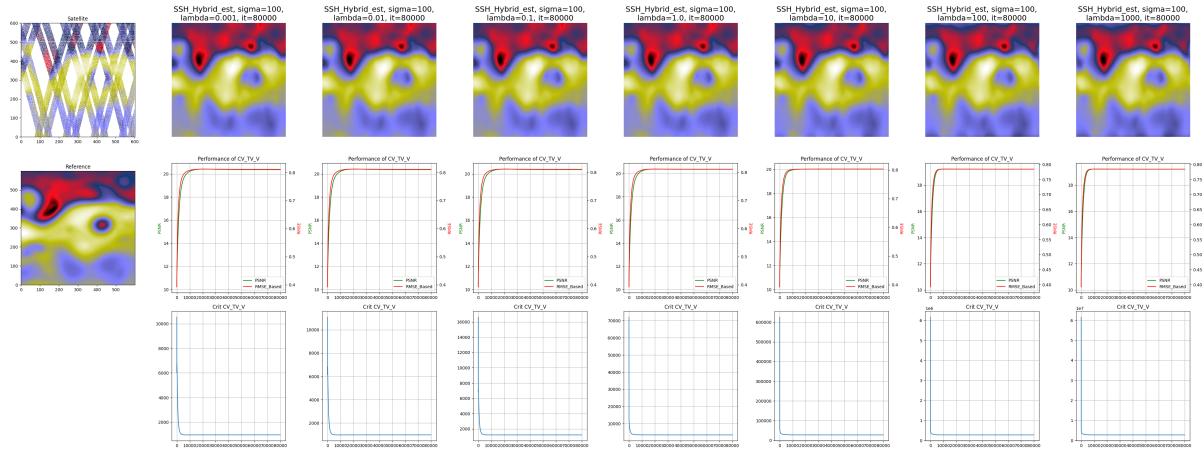


Figure 4.12: Total Variation combined with Potential Vorticity for different regularization parameters. Random mask.

drastically decreases after $\lambda = 1$. Also, we can see the expected over-smoothing effects of the TV, especially the effects of stair-casing effect is more obvious when we look into the potential vorticity field in figure 4.17. For the satellite observations, lack of information causes artificial sea surface predictions whereas it works well for the degraded data.

In the figures 4.11 and 4.12, we see the results of algorithm 6, where we impose the underlying physical constraint known as the QG equation. We have expected to recover the underlying potential vorticity field, however, because of the lack of data, it created non-existent vortexes in the empty areas. In this method, the PSNR and $RMSE_b$ increase up to 18.19 and 0.75 for the date '2012-10-13', respectively, and this has one of the lowest scores among the proposed methods. On the other hand, the algorithm works well with the degraded images. However, the main difference between this and the previous method would be the time cost. We had an average of 3 iterations per second, whereas in the TV method, this is up to 60 iterations per second on the same CPU.

Figure 4.13: Hybrid Method $\chi = 10$ Figure 4.14: Hybrid Method $\chi = 100$

A drastic change happened in the Hybrid algorithm, in which we used both the L2-norm of the discrete gradient and the total variation of potential vorticity. In this method, we have two different regularization terms for our penalization terms, χ and λ , respectively. As we see in this method, the changes of λ do not affect the performance and the algorithm even for the extreme points. However, the changes of the parameter χ affected the algorithm and increased the performance up to 20.36 PSNR and 0.81 RMSE_b (Figures 4.13,4.14). It has repeated the stair-casing effect as a common behavior of the TV term, but it has better results for the potential vorticity field (figure 4.15).

As a last application, we have used the Plug-and-Play method. Even though it made a proper estimation in certain parts, it has not been successful for the recovery of the vortexes hereby the potential vorticity field.

All the methods have a good performance for the degraded data, even with the high percentage mask they reached out the almost the original image. However, among all the methods, only the results of the Hybrid method are close to the results of the baseline method. OI has a 0.835 RMSE_b score in an average of 3 days, whereas the Hybrid method scored 0.828.

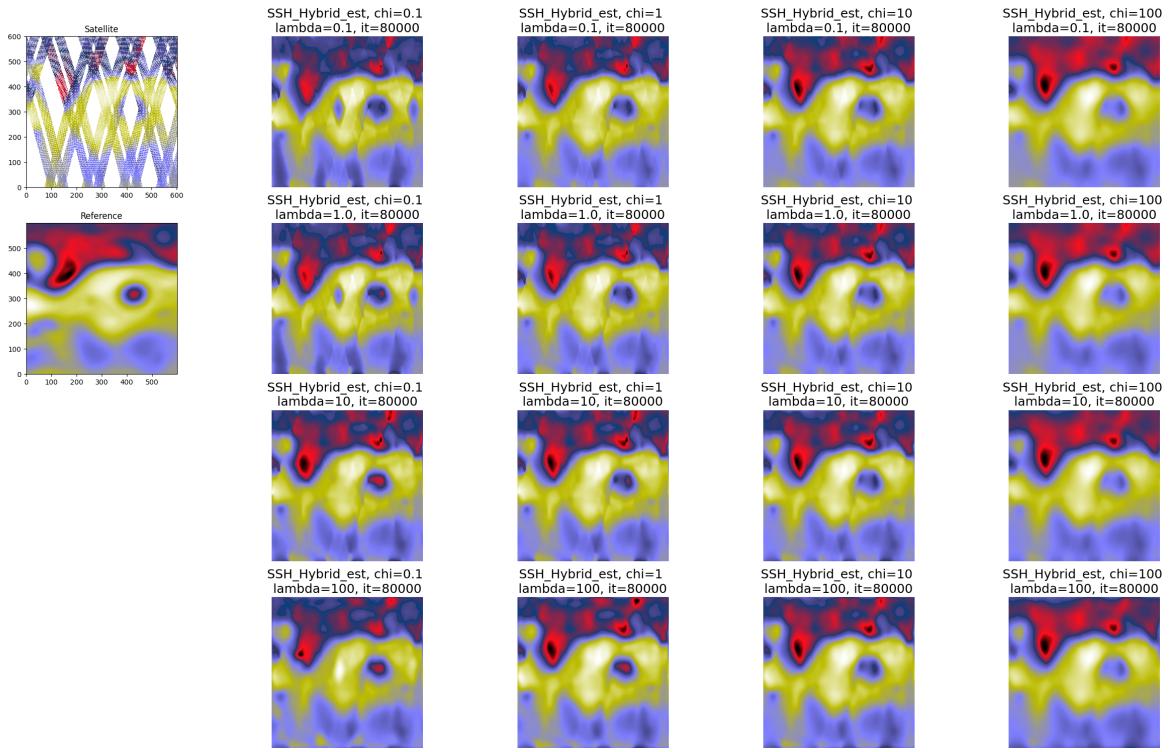


Figure 4.15: Hybrid Method: Comparison of Parameters. Horizontal axis: χ variations, Vertical axis: λ variations.

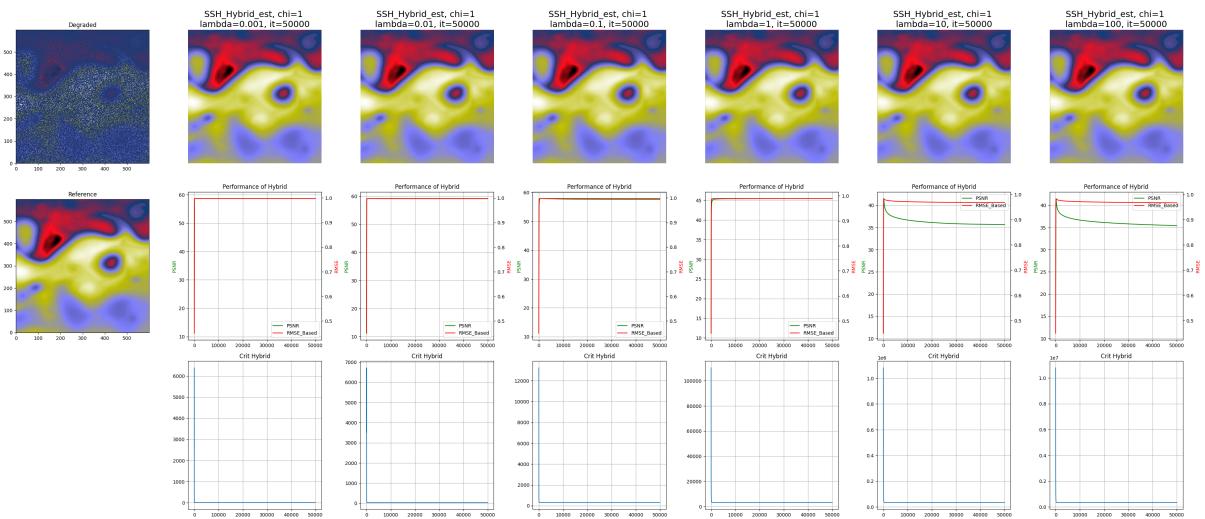


Figure 4.16: Hybrid Method $\chi = 1$, degraded image.

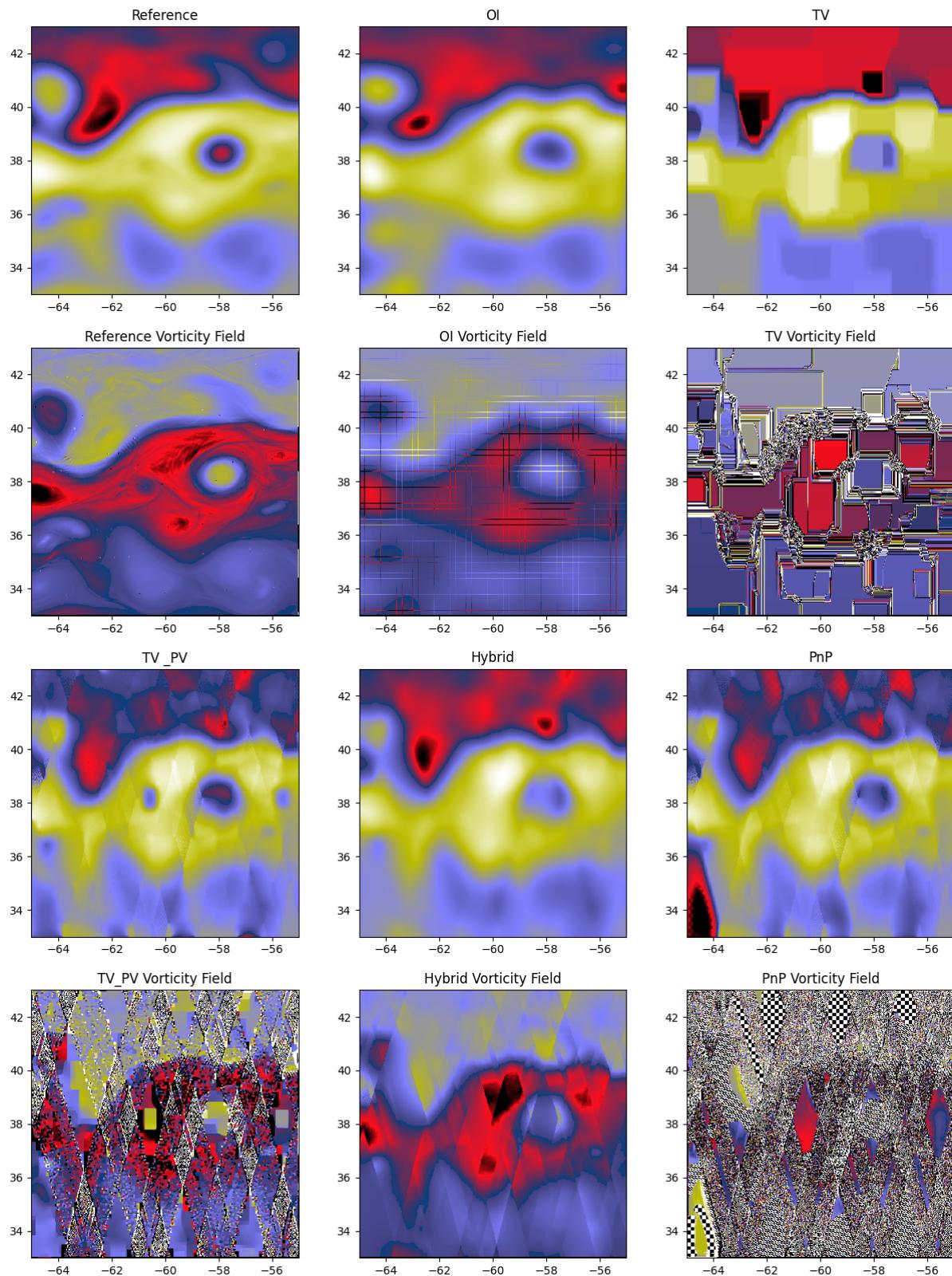


Figure 4.17: Estimations, and their potential vorticity field on the day '2012-10-13'

Methods \lambda	0.001	0.01	0.1	1	10	100
Total Variation	12.8991	19.1098	19.7588	19.9206	14.5077	10.5564
Total Variation of PV	12.5487	14.233	18.1927	17.684	17.1096	16.9282
Hybrid $\chi = 0.1$	16.4262	16.5075	16.8682	17.4905	19.0896	18.91
Hybrid $\chi = 1$	19.4289	19.4514	19.5156	19.1617	19.2164	19.275
Hybrid $\chi = 10$	19.9517	19.9533	19.9623	19.8736	19.3598	19.3582
Hybrid $\chi = 100$	20.365	20.3653	20.3678	20.3753	20.0169	19.2166
Hybrid $\chi = 1000$	16.9927	16.9927	16.9926	16.9911	16.9452	15.8143

Table 4.1: Table of PSNR, 2012-10-13

Methods \lambda	0.001	0.01	0.1	1	10	100
Total Variation	0.553329	0.7815	0.79723	0.800972	0.628845	0.415049
Total Variation of PV	0.534944	0.616918	0.757167	0.742522	0.724917	0.719114
Hybrid $\chi = 0.1$	0.7024	0.705174	0.717165	0.73672	0.78099	0.776415
Hybrid $\chi = 1$	0.78938	0.789927	0.791472	0.782802	0.784166	0.785616
Hybrid $\chi = 10$	0.801685	0.80172	0.801925	0.799894	0.7877	0.78766
Hybrid $\chi = 100$	0.8109	0.810906	0.810961	0.811124	0.803166	0.784171
Hybrid $\chi = 1000$	0.721191	0.721191	0.721188	0.721139	0.719663	0.680681

Table 4.2: Table of $RMSE_b$, 2012-10-13

Chapter 5

One step more

We have constructed our methods on static images and have not included the time factor of the dynamical system. It is possible to add the time factor to the variational problem. In this chapter, we will give the two most common interpolation methods, which have considered the time dimension in the variational problem in different ways.

In general, these methods are considered in the context of weather forecasting and/or oceanography and are called Data assimilation techniques. The optimal interpolation that has been introduced in Chapter 4 is the introduction to these methods, where it does not require additional time in the variational problem.

This chapter begins with the data assimilation methods, then we will conclude our work with possible directions for future research.

5.1 Data Assimilation methods considering Dynamics

Data assimilation refers to the use of both model output and observations. Before starting this part let us fix the notation that we will use in the rest of this section. Note that we will use a similar notation to the book of Benoit Cushman-Roisin and Jean-Marie Beckers. [7] Our analysis \mathbf{s}^a will be a linear combination of the forecast \mathbf{s}^f and the observation \mathbf{o} , we have,

$$\mathbf{s}^a = \mathbf{s}^f + \mathbf{K}(\mathbf{o} - \mathbf{H}\mathbf{s}^f), \quad (5.1)$$

where \mathbf{H} is a linear observation operator as before.

The matrix \mathbf{K} of size $M \times P$, where M is the number of model variables and P will be the number of observations for "an optimal blending of data into the model". The forecast error will be

$$\boldsymbol{\epsilon} = \mathbf{s} - \mathbf{s}^t. \quad (5.2)$$

During this section, we will assume that we have Gaussian distribution for our errors, so we have $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{s}^f, \mathbf{P}^f)$, where we denoted the covariance of the forecast by \mathbf{P}^f . Also, we have an observational error which is,

$$\boldsymbol{\epsilon}^o = \mathbf{o} - \mathbf{o}^t, \quad (5.3)$$

where the distribution is $\boldsymbol{\epsilon}^o \sim \mathcal{N}(\mathbf{o}^t, \mathbf{R})$.

5.1.1 4D-Var: 4-dimensional Variational Assimilation

Let us assume that we have the observational data for successive times $k = 0, \dots, K$,

$$\mathbf{o}_k = \mathbf{H}_k \mathbf{s}_k^t + \boldsymbol{\epsilon}_k^o, \quad (5.4)$$

where \mathbf{s}_k^t is the true state of the flow at the time k , \mathbf{H}_k is the linear observation operator, and $\boldsymbol{\epsilon}$ is an observational error with the covariance matrix \mathbf{R}_k . Note that the observational errors are assumed to be uncorrelated in time.

Also, we want to add the time evolution of the dynamical system, which is,

$$\mathbf{s}_{k+1}^t = \mathcal{M}(\mathbf{s}_k^t) + \boldsymbol{\eta}_k. \quad (5.5)$$

To facilitate matters, despite practical challenges, we will consider \mathcal{M} 's linearized version, which we assumed is already known for each k , then we have,

$$\mathbf{s}_{k+1}^t = \mathbf{M}_k \mathbf{s}_k^t + \boldsymbol{\eta}_k. \quad (5.6)$$

Assume that, like in the optimal interpolation, a background x_0^f (it can be notated as \mathbf{s}_0^b , in our setting this notation is more convenient), with error covariance \mathbf{P}_0^f and this error is uncorrelated with the observational error in 5.4, and available for time $k = 0$.

Let's first assume that there is no model error, so, any initial condition \mathbf{s}_0 at time $k = 0$ defines a model solution,

$$\mathbf{s}_{k+1}^t = \mathbf{M}_k \mathbf{s}_k^t, \quad k = 0, \dots, K - 1. \quad (5.7)$$

Then our new objective function will be,

$$J_{\text{4Dvar}}(\mathbf{s}_0) = \frac{1}{2} (\mathbf{s}_0 - \mathbf{s}_0^f)^T \mathbf{P}_0^{f-1} (\mathbf{s}_0 - \mathbf{s}_0^f) + \sum_{k=0}^K \frac{1}{2} (\mathbf{o}_k - \mathbf{H}_k \mathbf{s}_k)^T \mathbf{R}_k^{-1} (\mathbf{o}_k - \mathbf{H}_k \mathbf{s}_k). \quad (5.8)$$

This is called *strong-constraint four-dimensional variational assimilation*, the word strong-constraint emphasizes the model 5.7 must be exactly satisfied by the sequence of estimated vector space.

Now, we assume the model error $\boldsymbol{\eta}_k$ is different than zero, with a covariance matrix \mathbf{Q}_k , so the previous equation turns into,

$$\begin{aligned} J(\mathbf{s}_0, \mathbf{s}_1, \dots, \mathbf{s}_K)_{\text{4Dvar}} &= \frac{1}{2} (\mathbf{s}_0 - \mathbf{s}_0^f)^T \mathbf{P}_0^{f-1} (\mathbf{s}_0 - \mathbf{s}_0^f) \\ &\quad + \frac{1}{2} \sum_{k=0}^K (\mathbf{o}_k - \mathbf{H}_k \mathbf{s}_k)^T \mathbf{R}_k^{-1} (\mathbf{o}_k - \mathbf{H}_k \mathbf{s}_k) \\ &\quad + \frac{1}{2} \sum_{k=0}^{K-1} (\mathbf{s}_{k+1} - \mathbf{M}_k \mathbf{s}_k)^T \mathbf{Q}_k^{-1} ((\mathbf{s}_{k+1} - \mathbf{M}_k \mathbf{s}_k)), \end{aligned} \quad (5.9)$$

This is called *weak-constraint four-dimensional variational assimilation*. Therefore the minimization problem that can be considered is,

$$\operatorname{Arg min} J_{\text{4Dvar}} \quad (5.10)$$

in both cases.

5.1.2 Kalman Filtering

One of the ways of adding the information of physical models is using the Kalman filtering, but before that let us examine what is the approach. First, let us explain what we mean by the physical model, we take into account the fact between the assimilation cycle n and $n + 1$, the dynamical evolution of the state is,

$$\mathbf{s}_{n+1} = \mathbf{M}(\mathbf{s}_n) + \boldsymbol{\eta}_n$$

where $\boldsymbol{\eta}_n$ is the error of the model \mathbf{M} is the model. We are assuming the simulation cycle between the two analyses started with the analyses of the previous assimilation and assuming a linearized model, we have

$$\mathbf{s}_{n+1}^f = \mathbf{M}\mathbf{s}_n^a + \boldsymbol{\eta}_n \quad (5.11)$$

where the forecast error $\boldsymbol{\epsilon}^f = \mathbf{s}^f - \mathbf{s}^t$ satisfies,

$$\boldsymbol{\epsilon}_{n+1}^f = \mathbf{M}\boldsymbol{\epsilon}_n^a + \boldsymbol{\eta}_n \quad (5.12)$$

where we can obtain,

$$\mathbf{P}_{n+1}^f = \mathbf{M}\mathbf{P}_n^a\mathbf{M}^T + \mathbf{Q}_n = \mathbf{M}(\mathbf{M}\mathbf{P}_n^a)^T + \mathbf{Q}_n, \quad (5.13)$$

where \mathbf{Q}_n , model-error variance defined as,

$$\mathbf{Q}_n = \langle \boldsymbol{\eta}_n, \boldsymbol{\eta}_n^T \rangle \quad (5.14)$$

Now we can combine this method with OI, for this, we will use analyses state of the step n , to estimate $n + 1$, here is the iterative algorithm that we can use,

Algorithm 8: Kalman Filter Algorithm

- 1: **Initialization:**
 - 2: $\mathbf{s}_0^a \leftarrow \mathbf{s}$,
 - 3: $\mathbf{P}_0^a \leftarrow \mathbf{P}$
 - 4:
 - 5: **Forecast:**
 - 6: $\mathbf{s}_{n+1}^f \leftarrow \mathbf{M}(\mathbf{s}_n^a)$,
 - 7: $\mathbf{P}_{n+1}^f \leftarrow \mathbf{M}_n\mathbf{P}_n^a\mathbf{M}_n^T + \mathbf{Q}_n$
 - 8:
 - 9: **Analysis:**
 - 10: $\mathbf{s}_{n+1}^a \leftarrow \arg \min_{\mathbf{s}} \|\mathbf{o}_{n+1} - \mathbf{H}_{n+1}\mathbf{s}\|_{\mathbf{R}_{n+1}^{-1}}^2 + \|\mathbf{s} - \mathbf{s}_{n+1}^f\|_{(\mathbf{P}_{n+1}^f)^{-1}}^2$,
 - 11: $\mathbf{P}_{n+1}^a \leftarrow \mathbf{P}_{n+1}^f - \mathbf{K}_{n+1}\mathbf{H}_{n+1}\mathbf{P}_{n+1}^f$,
 - 12: where,
 - 13: $\mathbf{K}_{n+1} \leftarrow \mathbf{P}_{n+1}^f\mathbf{H}_{n+1}^T(\mathbf{H}_{n+1}\mathbf{P}_{n+1}^f\mathbf{H}_{n+1}^T + \mathbf{R}_{n+1})^{-1}$
-

5.2 Conclusion, Future Work

In conclusion, this study has addressed the variational problem in the context of non-smooth optimization, employing proximal methods and exploring various algorithmic choices. We introduced the shallow water quasi-geostrophic equations to establish a connection between the variational problem and the sea surface height field. Notably, these equations yield different results compared to the shallow water vorticity equation due to factors such as Earth's rotation, fluid stratification, and the selection of appropriate scaling for our specific focus.

We also introduced the case study region, the NATL60 data set, and the specific duration of observation. Our methodology for processing this data set was thoroughly explained. To assess the performance of the methods employed, we presented images generated using a random mask. Moreover, the Optimal Interpolation has been introduced to make comparisons with our results.

In this study, we have introduced four different methods and their performances in both real, and random masks. The first method that has been introduced is the well-known method for image reconstruction called total variation. The results of this method are highly affected by the stair-casing effect, which is a common behavior of total variation as a penalization term. In the second method, we have used the potential vorticity equation in the total variation to introduce more information to the underlying physical part, however, the results have artifacts on the regions we have no information. However, the 'Hybrid' method outperformed previous approaches in terms of performance. In the end, we presented the result of the Plug and Play method, which has been successful for certain local parts, but creates artifacts and was not able to reconstruct the potential vorticity field. But neither of these methods could outperform the baseline method, optimal interpolation.

One promising method for the improvements of our results could have been choosing an optimized first guess, which means, instead of putting reference level 0 on the points of 'NaN' in the whole region, an educated guess that could have been taken into account by looking at the closest local regions or using another model as in the first steps in the reconstruction algorithms.

Another potential area of exploration is taking into account the dynamics, by using consecutive selected data sets and connecting them by using primitive equations and QG equations, potentially using 4Dvar and/or Kalman filtering in the context of image processing. Using these aspects would contribute to a more comprehensive understanding and could help overcome the limitations observed in our methods.

References

- [1] P. Perona and J. Malik. “Scale-Space and Edge Detection Using Anisotropic Diffusion”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12.7 (July 1990), pp. 629–639. ISSN: 01628828. DOI: 10.1109/34.56205.
- [2] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. “Nonlinear Total Variation Based Noise Removal Algorithms”. In: *Physica D: Nonlinear Phenomena* 60.1-4 (Nov. 1992), pp. 259–268. ISSN: 01672789. DOI: 10.1016/0167-2789(92)90242-F.
- [3] Joachim Weickert. “Anisotropic Diffusion in Image Processing”. In: (1998).
- [4] Enrico Giusti. *Direct Methods in the Calculus of Variations*. WORLD SCIENTIFIC, Jan. 2003. ISBN: 978-981-238-043-2 978-981-279-555-7. DOI: 10.1142/5002. URL: <https://www.worldscientific.com/worldscibooks/10.1142/5002> (visited on 09/09/2023).
- [5] Amir Beck and Marc Teboulle. “Gradient-Based Algorithms with Applications to Signal-Recovery Problems”. In: *Convex Optimization in Signal Processing and Communications*. Ed. by Daniel P. Palomar and Yonina C. Eldar. 1st ed. Cambridge University Press, Dec. 3, 2009, pp. 42–88. ISBN: 978-0-521-76222-9 978-0-511-80445-8. DOI: 10.1017/CBO9780511804458.003.
- [6] Antonin Chambolle and Thomas Pock. “A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging”. In: *Journal of Mathematical Imaging and Vision* 40.1 (May 2011), pp. 120–145. ISSN: 0924-9907, 1573-7683. DOI: 10.1007/s10851-010-0251-1.
- [7] Benoit Cushman-Roisin and Jean-Marie Beckers. *Introduction to Geophysical Fluid Dynamics: Physical and Numerical Aspects*. 2nd ed. International Geophysics Series v. 101. Waltham, MA: Academic Press, 2011, pp. 7–10, 150–162. 828 pp.
- [8] Laurent Condat. “A Primal–Dual Splitting Method for Convex Optimization Involving Lipschitzian, Proximable and Linear Composite Terms”. In: *Journal of Optimization Theory and Applications* 158.2 (Aug. 2013), pp. 460–479. ISSN: 0022-3239, 1573-2878. DOI: 10.1007/s10957-012-0245-9.
- [9] N. Parikh and S. Boyd. *Proximal Algorithms*. Foundations and Trends in Optimization. Now Publishers, 2013. ISBN: 978-1-60198-716-7.

- [10] Singanallur V Venkatakrishnan, Charles A Bouman, and Brendt Wohlberg. “Plug-and-play priors for model based reconstruction”. In: *2013 IEEE global conference on signal and information processing*. IEEE. 2013, pp. 945–948.
- [11] Virginia Vassilevska Williams. “Multiplying matrices in O(n^{2.373}) time”. In: 2014. URL: <https://api.semanticscholar.org/CorpusID:124418960>.
- [12] Lucile Gaultier, Clément Ubelmann, and Lee-Lueng Fu. “The Challenge of Using Future SWOT Data for Oceanic Field Reconstruction”. In: *Journal of Atmospheric and Oceanic Technology* 33.1 (Jan. 2016), pp. 119–126. ISSN: 0739-0572, 1520-0426. DOI: 10.1175/JTECH-D-15-0160.1. (Visited on 06/01/2023).
- [13] Geoffrey K. Vallis. *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale Circulation*. 2nd ed. Cambridge University Press, June 30, 2017. DOI: 10.1017/9781107588417.
- [14] Linwei Fan et al. “Brief Review of Image Denoising Techniques”. In: *Visual Computing for Industry, Biomedicine, and Art* 2.1 (July 8, 2019), p. 7. ISSN: 2524-4442. DOI: 10.1186/s42492-019-0016-7. URL: <https://doi.org/10.1186/s42492-019-0016-7> (visited on 06/08/2023).
- [15] Adekunle Ajayi et al. “Spatial and Temporal Variability of the North Atlantic Eddy Field From Two Kilometric-Resolution Ocean Models”. In: *Journal of Geophysical Research: Oceans* 125.5 (May 2020). DOI: 10.1029/2019JC015827.
- [16] CLS/MEOM. *SWOT Data Challenge NATL60: Version 2020*. In collab. with Maxime Ballarotta, Emmanuel Cosme, and Aurelie Albert. application/x-netcdf. 2020. DOI: 10.24400/527896/A01-2020.002. URL: <https://www.aviso.altimetry.fr/en/data/products/ocean-data-challenges/2020a-ssh-mapping-natl60.html>.
- [17] Thomas H. Cormen et al. *Introduction to Algorithms*. Fourth edition. Cambridge, Massachusetts: The MIT Press, 2022. 1291 pp. ISBN: 978-0-262-04630-5.
- [18] Gurvan Madec et al. “NEMO Ocean Engine Reference Manual”. Version v4.2.1. In: (). ISSN: 1288-1619. DOI: 10.5281/ZENODO.1464816. URL: <https://zenodo.org/record/1464816>.