Variational data assimilation for the shallow water equations with applications to tsunami wave prediction

VARIATIONAL DATA ASSIMILATION FOR THE SHALLOW WATER EQUATIONS WITH APPLICATIONS TO TSUNAMI WAVE PREDICTION

By Ramsha KHAN,

A Thesis Submitted to the School of Graduate Studies in the Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

McMaster University
Doctor of Philosophy (2020)
Hamilton, Ontario (Comp. Science & Eng.)

TITLE: Variational data assimilation for the shallow water equations with applications to tsunami wave prediction
AUTHOR: Ramsha KHAN (McMaster University)
SUPERVISOR: Dr. Nicholas KEVLAHAN
NUMBER OF PAGES: xx, 194

Abstract

Accurate prediction of tsunami waves requires complete boundary and initial condition data, coupled with the appropriate mathematical model. However, necessary data is often missing or inaccurate, and may not have sufficient resolution to capture the dynamics of such nonlinear waves accurately. In this thesis we demonstrate that variational data assimilation for the continuous shallow water equations (SWE) is a feasible approach for recovering both initial conditions and bathymetry data from sparse observations. Using a Sadourny finite-difference finite volume discretisation for our numerical implementation, we show that convergence to true initial conditions can be achieved for sparse observations arranged in multiple configurations, for both isotropic and anisotropic initial conditions, and with realistic bathymetry data in two dimensions. We demonstrate that for the 1-D SWE, convergence to exact bathymetry is improved by including a low-pass filter in the data assimilation algorithm designed to remove scale-scale noise, and with a larger number of observations. A necessary condition for a relative L^2 error less than 10% in bathymetry reconstruction is that the amplitude of the initial conditions be less than 1% of the bathymetry height. We perform Second Order Adjoint Sensitivity Analysis and Global Sensitivity Analysis to comprehensively assess the sensitivity of the surface wave to errors in the bathymetry and perturbations in the observations. By demonstrating low sensitivity of the surface wave to the reconstruction error, we found that reconstructing the bathymetry with a relative error of about 10% is sufficiently accurate for surface wave modelling in most cases. These idealised results with simplified 2-D and 1-D geometry are intended to be a first step towards more physically realistic settings, and can be used in tsunami modelling to (i) maximise accuracy of tsunami prediction through sufficiently accurate reconstruction of the necessary data, (ii) attain *a priori* knowledge of how different bathymetry and initial conditions can affect the surface wave error, and (iii) provide insight on how these can be mitigated through optimal configuration of the observations.

Acknowledgements

This thesis could not have been completed without the help and collaboration of my fellow graduate students and friends in the Maths & Stats department, especially the AIMS lab. I especially want to thank Tyler Meadows, Szymon Sobieszek, Samantha Caetano, John Heisey, Graham Keiper, Fletcher Gates, and Lee Van Brussels, for not only being there as rubber ducks I explain my maths woes to, but also for making the office environment something I looked forward to every day. And for taking multiple coffee breaks a day with me. Pritpal Matharu gets a special mention all for himself, because I don't think the thesis in its current form, or my understanding of the field we're both privileged to to be in, would be where it is without all the feedback, advice and animated discussions we've had, whether we've been in Hamilton Hall or the Mojave desert!

I want to thank Dr. Bartosz Protas, Dr. Dmitry Pelinovsky, and the late Dr. Walter Craig, for all their guidance in the six years I've spent attempting to wrangle first a Masters, and then this PhD, and for the foundations they have helped me develop. A big thank you also to Dr. Marilyn Lightstone, for taking the time to be on my supervisory committee and for all her valuable input.

To my supervisor Dr. Nicholas Kevlahan, I am so very grateful for the constant support he has given me in shaping this project, and for seeing my potential six years ago. I could not have asked for more insightful direction and feedback, and I am proud to say that his influence is over the entire thesis. Thank you for taking the time to painstakingly guide me through the theory, troubleshooting Matlab code with me, and for the opportunity to broaden my horizons at the Université Grenoble Alpes.

To my partner-in-(PhD-and)-crime Adrian, it's hard to believe we're approaching the finish line, and I can't tangibly quantify (even though that's what I do) my appreciation for all the support you've given me. Instead I'll just say, here's to the next chapter!

And finally, I dedicate this thesis to the two people who are most responsible for where I am today: my parents. Thank you for always letting me chase my

goals, and for the utter conviction I have that I'll always have their support, no matter where I end up. I couldn't have asked for more.

Contents

List of Figures

List of Tables

Declaration of Authorship

I, Ramsha KHAN, declare that this thesis titled, "Variational data assimilation for the shallow water equations with applications to tsunami wave prediction" and the work presented in it are my own. I confirm that:

- Chapter 1: The introduction is my own work, and was edited based on suggestions by Dr. Nicholas Kevlahan.
- Chapter 2: The derivation of the 2-D data assimilation algorithm in this chapter is extended from the work conducted during my Masters, and published in [27]. All the numerical implementation and analyses were completed by myself under the guidance of Dr. Nicholas Kevlahan.
- Chapter 3: For the 1-D bathymetry data assimilation, all derivations, numerical implementation and analyses were completed by myself under the guidance of Dr. Nicholas Kevlahan. The suggestion for the low pass filter and its derivation is attributed to Dr. Bartosz Protas.
- Chapter 4: The derivation of the second order adjoint sensitivity algorithm is based on the general methods provided in [53] and [52]. The implementation and all subsequent numerical results and analyses were conducted by myself, under the guidance of Dr. Nicholas Kevlahan.
- Chapter 5: The Global Sensitivity Analyses conducted in this chapter are based on the algorithms and methods outlined in [56], [48], and [42]. The numerical implementation was adapted from the the SAFE GSA toolbox [41]. The model design, choice of sampling strategy, and all subsequent analyses were conducted by myself under the guidance of Dr. Nicholas Kevlahan.
- Chapter 6: The conclusion is my own work, and was edited based on suggestions by Dr. Nicholas Kevlahan.

Our intention is to submit these results to peer reviewed journals in the near future.

Chapter 1

Introduction

The purpose of tsunami modelling is to accurately predict the wave height and velocity of the tsunami at coastal areas, given observations of the wave height at discrete locations, and the available bathymetry data. Tsunamis are series of waves caused by large-scale disturbances in the ocean, such as seismic activity. They are characterised by very long wavelengths relative to the ocean depth (sometimes hundreds of kilometres), categorising them as shallow-water waves. However, the perturbation of the free surface in the deep ocean can be less than a metre, rendering it virtually imperceptible away from the shore. The wave speed c is directly proportional to the square root of the water depth, and with an average deep-ocean depth of 4000m, tsunamis can travel faster than 700km/h, and are capable of generating atmospheric gravity waves that can travel into the upper atmosphere [24]. Consequently, variations in the sea floor topography (i.e. bathymetry) can significantly alter the wave speed, direction, and shape, with wide-reaching effects. As the wave approaches shallower regions where the depth decreases, the speed decreases as well. The energy flux of a tsunami remains relatively constant, therefore a decrease in speed results in an increase in the wave height at coastlines. Coupled with the momentum of the wave, this can have devastating effects.

The 2004 Indian ocean tsunami and the 2011 Japanese tsunami have highlighted the need for more effective forecasting models that can be used to create and implement evacuation and emergency protocols effectively in a limited amount of time. However, existing methods are challenged by the limitations of the necessary sets of data [35]. Tsunami forecasts rely on data collected via Deep-Ocean

Assessment & Reporting of Tsunami (DART) buoys [18], initially developed by the National Oceanic and Atmospheric Administration (NOAA), and now used globally to measure height changes in oceanic water columns. Due to their high operational investment, the DART buoy network is limited to approximately 60 systems, spread over the entire Pacific Ocean and located mostly near coastlines. Other observation sources like ocean bottom seismometer (OBS) pressure gauges, are sparsely distributed in the Indian ocean [62]. Consequently, information on the initial conditions of tsunamis is likely to be incomplete or inaccurate. Similarly, available data of ocean floor topography (i.e. bathymetry) is often of uneven quality and low-resolution. Databases like the ETOPO2 global relief model [13] give a resolution of 2 arc-minutes of the earth's surface, integrating land topography and ocean bathymetry from multiple sources, collected using various methods. These include measurements of variations in the gravitational field, sonar, satellite data, data from shipping vessels, and using video data to interpolate the shape of the sea floor, all of which are subject to inaccuracies and can be inconsistent.

An alternative theoretical approach uses surface wave propagation to extrapolate bathymetry [44, 21, 37]. The latter is classically known as the "inverse problem", and is the subject of this thesis. Research in this area has focused predominantly on the refinement of algorithms and methods used to find an inversion formula for the bathymetry. These include Monte Carlo methods, discrete least-squares problems, and least-squares problems for functionals [59]. The general consensus in existing literature is that the inversion problem for finding bathymetry using surface wave data is ill-posed [37], and small amounts of noise can yield erroneous results. However, regularisation methods are used to a successful degree, with relative errors between resultant and exact bathymetry reaching lower orders of magnitude than previously achieved.

The primary objective of this study is to develop and test variational data assimilation techniques to improve estimates of the initial conditions and bathymetry. In this section we provide a concise overview of the mathematical model and underlying assumptions, and the principles of data assimilation. We describe the challenges involved in accurate tsunami modelling, and summarise the analyses and results provided in each chapter to address them in an idealised model.

Our model for tsunami wave propagation is based on the one- and two-dimensional non-dispersive shallow water equations (SWE). These are a coupled system of partial differential equations (PDEs) for travelling waves, governing fluid flow in oceans and coastal regions. Tsunami waves can also be modelled using the Boussinesq water wave approximation. While both shallow water and Boussinesq approximations are widely used for analysing solitary wave propagation, studies such as Dongfang et al [10] have compared the two processes, finding that in certain run up processes the two approximations are identical, and that Boussinesq approximation is most often used to model near shore hydrodynamic behaviour. However, our analysis addresses the optimal configuration of deep-ocean observations required to accurately reconstruct initial conditions and bathymetry and the necessary and/or sufficient conditions for convergence, before coastal dynamics are observed. Therefore, we do not take near-shore behaviour into account in our analysis.

Additionally, we neglect Coriolis effects, bottom friction, and kinematic viscosity, and for the bathymetry assimilation we only consider the 1-D shallow water equations. This idealised configuration with simplified 2-D and 1-D geometry is intended to be a first step for more complex analyses. We aim to validate the basic approach by demonstrating the feasibility of variationally data assimilation for tsunami wave prediction, and investigate fundamental questions for the simpler case first before considering more physically realistic settings.

The general characteristic of shallow water flow is that the length scale of horizontal motion is much larger than the fluid depth ($\lambda \gg H$), and the wave perturbation from the free surface is much less than the average fluid depth ($\eta \ll H$). A representation of the fluid column is given in figure 1.1. The SWE are derived from the three dimensional Navier-Stokes equations for fluid flow, which are themselves derived from the equations for conservation of mass and momentum [9].

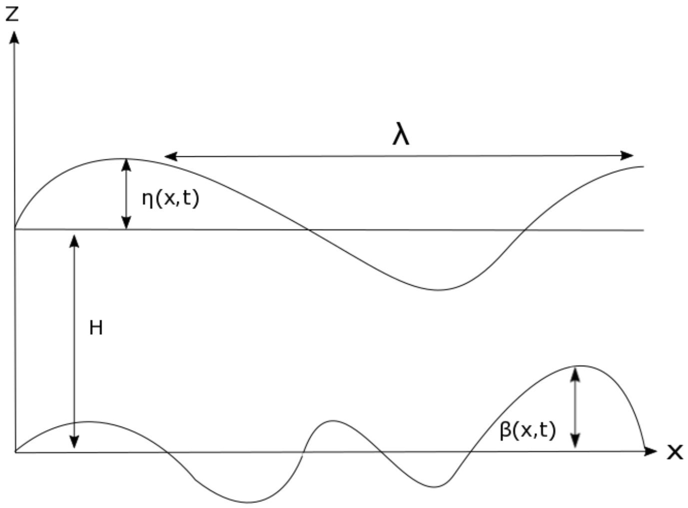


FIGURE 1.1: A single fluid layer, where the total height of the water column is $h = H + \eta - \beta$. H is the average depth, η is the perturbation of the free surface, and β is the bathymetry. Bathymetry can be time-dependent (i.e. sudden shifts of the sea floor due to seismic activity), however in this study we consider static bathymetry. Shallow water approximations require that the surface wave length is much larger than the depth ($\lambda >> H$), and that the free surface perturbation is much less than the depth ($\eta << H$).

We assume hydrostatic balance ($\partial p/\partial z = \rho g$) to integrate the Navier-Stokes equations over depth. Eliminating the vertical dimension gives us the incompressible, nonlinear 2-D shallow water equations (where we assume that the $x - z$ and $y - z$ velocities are irrotational),

$$\begin{cases} \dfrac{\partial \eta}{\partial t} + \dfrac{\partial}{\partial x}\Big((\eta + H - \beta)u\Big) + \dfrac{\partial}{\partial y}\Big((\eta + H - \beta)v\Big) = 0, & \text{(1.0.1a)} \\ \dfrac{\partial u}{\partial t} + u\dfrac{\partial u}{\partial x} + v\dfrac{\partial u}{\partial y} + g\dfrac{\partial \eta}{\partial x} = 0, & \text{(1.0.1b)} \\ \dfrac{\partial v}{\partial t} + u\dfrac{\partial v}{\partial x} + v\dfrac{\partial v}{\partial y} + g\dfrac{\partial \eta}{\partial y} = 0, & \text{(1.0.1c)} \\ \eta(x, y, 0) = \phi(x, y), & \text{(1.0.1d)} \\ u(x, y, 0) = v(x, y, 0) = 0, & \text{(1.0.1e)} \end{cases}$$

where:

- u is the velocity in the x-direction (zonal velocity).
- v is the velocity in the y-direction (meridional velocity).
- $\eta(x, y, t) + H - \beta(x, y)$ is the fluid height, where $\eta(x, t)$ is the perturbation of the free surface, H is the average depth, and $\beta(x, y)$ is the time-independent bathymetry. Note that in the SWE the bathymetry enters as a parameter, and there is no explicit bottom boundary condition.
- g is the gravitational constant.
- $\phi(x, y)$ is the initial condition for the perturbation of the free surface η.
- We assume periodic boundary conditions for u, v, η and β in our rectangular domain $\Omega = [-L\ L] \times [-L\ L]$.

The one-dimensional SWE (as considered for the bathymetry assimilation) are

$$\begin{cases} \dfrac{\partial \eta}{\partial t} + \dfrac{\partial}{\partial x}\Big((H + \eta - \beta)u\Big) = 0, & (1.0.2a) \\ \dfrac{\partial u}{\partial t} + \dfrac{\partial}{\partial x}\left(\dfrac{1}{2}u^2 + g\eta\right) = 0, & (1.0.2b) \\ \eta(x, 0) = \ \phi(x), & (1.0.2c) \\ u(x, 0) = \ 0, & (1.0.2d) \end{cases}$$

where once again we have assumed periodicity at the domain boundaries $x = -L$ and $x = L$.

In the past, numerical simulation of tsunami waves was often done independently of the data used to validate the model prediction. There are shortcomings in this approach. Numerical simulations are subject to inaccuracies and approximation errors, and the observational data often suffers from technical, budgetary and physical restrictions [35, 62]. These include noisy measurements subject to errors, or sparse data at low resolution. Consequently, the logical solution is to combine these processes, using observed data as a correction for the state variables. Data Assimilation is the study of finding the optimal way to integrate observed data into a computational model in order to achieve the most accurate representation of a dynamical system possible. However, optimal reconstruction of missing or inaccurate data can be operationally challenging. Physical systems can be modelled by systems of PDEs, therefore if we have enough initial and boundary data, we can solve this system and accurately forecast the state at a given time. In reality the information available is often orders of magnitude smaller than the complete boundary or initial data required to fully solve the system. Natural systems like ocean currents, waves or climate can have numerous parameters and variables that cannot always be quantified at all needed points. Additionally, they are often nonlinear, and the processing power needed to solve a system of high dimensionality is often not available.

Methodologies for assimilating observations into a model such that the observed data remains consistent include variational methods, or statistical techniques like Kalman filtering. Kalman filters recursively estimate the joint probability distribution of a model and a set of measurements, while variational methods iteratively minimise a cost function representing errors between the model and observations. The variational approach is routinely used in Numerical Weather Prediction (NWP), in an algorithm called 4D-VAR (four dimensional variational data assimilation), that estimates parameters and initial conditions in weather prediction models [2].

Kalman filtering techniques have also been used extensively for both parameter estimation and initial condition reconstruction in ocean models. Mayo et al. [34] use variants of an ensemble Kalman filter EnKF (where error statistics of the model are represented by an ensemble of forecasted model states) to estimate bottom stress terms in the Advanced Circulation (ADCIRC) coastal model using observations of sea surface elevations. They demonstrate accurate estimation of friction parameters in lagoons and estuaries, and highlight the influence of the bottom surface roughness, motivating the need for high-resolution bathymetry. Ghorbanidehno et al. [17] also use Kalman filtering to estimate near shore bathymetry, using a novel compressed-state Kalman Filter to recover both constant and temporally evolving bathymetry profiles, and demonstrate superior accuracy compared to ensemble-based methods with comparable computational costs. Statistical data assimilation has also been used in tsunami modelling efforts. Yang et al. [65] use Optimal Interpolation (OI), a variant of the EnKF to reconstruct real-time tsunami wave fields using measurements of pressure changes on the ocean floor. However, accurate prediction in real time using Kalman filtering techniques can be challenging as forecast accuracy depends on the choice of initial error covariance matrices, and often the error statistics for true state variables are not available, nor easily estimated. Additionally, the ability to resolve details of covariance structure is proportional to the ensemble size, and larger ensembles may be too computationally expensive for predictions in real time.

Nevertheless, comparison between Kalman filtering and variational data assimilation (given a perfect model and same observations and domain), determined that performance was comparatively equivalent for both [14, 30]. In this study we implement a variational data assimilation scheme, using optimal control theory to minimise the error between the state variables and observations. The novelty of our variational approach is that we consider the infinite dimensional case, unlike previous works on 4D-VAR for the SWE like Zou et al. [66] and Maeda et al. [33], and Kalman filtering techniques. Consequently, our data assimilation algorithm is independent of the discretisation used in its numerical implementation.

Additionally, a requirement for variational assimilation is that the system (1.0.1) can be solved in forward time from some initial time t_0 till a final time t_T, given an initial 'guess' from the set of admissible values of the initial condition, or bathymetry. This should yield solutions of the state at measurement points, that can be compared with observations. For nonlinear problems without analytical solutions, the efficacy of this depends on the choice of numerical scheme used to solve the system, introducing a further consideration for successful implementation. Such schemes are subject to stability and order of accuracy issues, which need to be accounted for.

The problem can be summarised thus: what is the optimal way to integrate the information we have access to? In this study, we attempt to find a solution, by (i) finding the minimum information in the form of observations necessary to bring our model output as close to observed values as possible, without exacerbating or amplifying small errors in measurement, in a realistic amount of time, and

(ii) quantifying the effect of the reconstruction error on the free surface wave, and understanding how variations in the algorithm parameters and observations impact accuracy of free surface wave prediction.

The four main topics and contributions of the thesis are summarised briefly below.

1. **Chapter 2:** We implement a variational data assimilation scheme for the 2-D shallow water equations (SWE), where observations of sea surface height are used to reconstruct initial conditions. Our approach is focused on a qualitative investigation of the applicability to 2-D, of results from the 1-D analysis in Kevlahan et al. [27], regarding sufficient conditions for convergence. We present a comparison of observation configurations arranged in straight lines, in a grid array, or along concentric circles, where the domain is a rectangular grid with initial conditions located approximately at the origin. The analyses are divided into the following three topics.
    (a) Assess the optimal number and configuration of observation points such that convergence to the true initial conditions is achieved in the data assimilation for the 2-D SWE.
    (b) Investigate whether the sufficient conditions for convergence in Kevlahan et al. [27] for the 1-D case extends to the 2-D data assimilation.
    (c) Implement the data assimilation algorithm for a relatively realistic tsunami forecasting model, using non-flat bathymetry. Our final analysis uses realistic bathymetry data from the ETOPO2 global topographical relief database and the optimal observation configurations identified in this study.
2. **Chapter 3:** We implement variational data assimilation for the 1-D SWE for the reconstruction of bathymetry for an idealised case, addressing two complementary questions:
    (a) How accurately can bathymetry data be reconstructed from surface wave measurements, and what determines the accuracy?

   (b) How accurate does the bathymetry data need to be to model sea surface waves to given accuracy?

   We quantify key relationships between the initial condition and bathymetry amplitudes relative to the average fluid depth. We also analyse the effect of number of observations on assimilation convergence. We implement a Sobolev gradient smoothing technique (effectively a low-pass filter) within our optimization scheme and illustrate its ability to reduce small-scale noise present in the bathymetry reconstruction. We then investigate the consequences of error in the bathymetry data on the resulting surface wave by observing trends in the surface wave propagation error, as the amplitude of the initial condition, amplitude of the bathymetry, and the number of observation points is varied.

3. **Chapter 4:** We use second order adjoint methods outlined in Shutyaev et al. [53] and Shutyaev et al. [52], to derive the sensitivity of the surface wave error produced by the reconstructed bathymetry, to perturbations in observations for (i) the data assimilation scheme for initial conditions outlined in [27], and (ii) the data assimilation scheme for bathymetry given in Chapter 3. We do this by analytically deriving the Hessian of the cost function minimised in the data assimilation, and take advantage of its properties to derive expressions for the sensitivity of the data assimilation to perturbations in the observations. We present a numerical implementation of the algorithm for the bathymetry data assimilation. The focus is on the latter instead of the initial condition reconstruction, as our main objective is to use the present analysis to further investigate the qualitative results observed in Chapter 3. Specifically, these are (i) the link between low sensitivity and the accuracy of the data assimilation scheme, and (ii) the effect of perturbations in the model parameters and the observations on the sensitivity. Ultimately we wish to see whether the conclusions from the bathymetry reconstruction in Chapter 3 can be better understood using the results of the sensitivity analysis in this chapter.

4. **Chapter 5:** While we discuss the impact of number of observations and

spacing in detail in Chapter 4, we wish to gain more insight on how the position of the bathymetry relative to observation points effects the bathymetry reconstruction error and the surface wave response. We also wish to determine the sensitivity of these to model parameters. In this chapter we use Global Sensitivity Analysis (GSA) to derive sensitivity indices quantifying the influence of (i) bathymetry position relative to the observations, and (ii) the amplitudes of the initial condition and bathymetry, on the error in bathymetry reconstruction and the surface wave. We focus our analysis on a localised surface wave propagating over a compact bathymetry. This work is motivated by the observation in Chapter 3 that there is low sensitivity of surface waves to noise in bathymetry reconstruction. Using variance-based and density-based methods, we quantify and subsequently rank the influence of these input factors on the bathymetry and surface wave errors respectively.

We conclude in Chapter 6 with a summary of the main results of the thesis, and discuss their contribution to current research, and considerations for future work.

The key contributions of this thesis include the derivation and implementation of a variational data assimilation algorithm for the infinite-dimensional SWE, where we demonstrate that variational assimilation is a feasible approach for recovering both initial condition and bathymetry data from sparse observations. We show that convergence can be achieved for sparse observations arranged in multiple configurations, for both isotropic and anisotropic initial conditions, and with realistic bathymetry data in two dimensions. We demonstrate that convergence is improved by including a low-pass filter in the data assimilation algorithm, designed to remove scale-scale noise. We highlight necessary and/or sufficient conditions on the observation operator and model parameters for convergence to the true bathymetry, and comprehensively assess the sensitivity of the surface wave to errors in the bathymetry and perturbations to the observations. These results are a first step towards determining criteria for sufficient convergence of optimally reconstructed bathymetry such that we can model tsunami waves to given accuracy, and highlight the optimal configurations of sparse observation facilitating convergence. These insights can subsequently be used to improve future forecasts of tsunamis.

Chapter 2

Data Assimilation on the 2-D SWE for the reconstruction of initial conditions

2.1 Introduction

In this study we formulate a variational data assimilation scheme for the 2-D shallow water equations (SWE), where observations of sea surface height are used to reconstruct missing or noisy initial conditions data. Compared to similar analyses for the 1-D SWE in [27], the 2-D SWE system has additional features that increase the complexity of the analyses, such as characteristic curves of the PDE that are non-linear, or 2-D bathymetry features. Additionally, there is an increase in the degrees of freedom for the observation operator, which makes finding optimal configurations of observations more difficult. The 2-D SWE system includes an additional conservation of momentum equation in the horizontal y-direction. The state vector is subsequently $(\eta\ u\ v)^T$, and the full system is given in (2.2.1).

Our goal is to determine the necessary and/or sufficient requirements for the observation network such that the reconstruction error for the initial conditions is minimised. In practice, Zou et al. [66] have demonstrated that the discretised 2-D SWE are observable even with measurements of only one of the three variables η, u, v. However, existing data assimilation methods have been applied mostly for tsunami forecasts in North America and Japan where relatively large observation networks exist [62]. Primary sources of observations include the Deep-ocean Assessment and Reporting of Tsunamis (DART) buoy system, consisting of a bottom pressure recorder residing on the ocean floor which transmits data to a surface buoy. The data is then relayed to shore via NOAA's Geostationary Operational Environmental Satellite (GOES) [18]. The large investment required for such apparatus limits the feasibility of a dense network, and currently the global network consists of approximately 60 systems. Other observations used for tsunami detection and reporting include ocean bottom seismometer (OBS) pressure gauges, however these are sparsely distributed in the Indian ocean [62].

Existing data assimilation schemes are able to utilise multiple sources and techniques for assimilating off-shore observations. Maeda et al. [33] assimilated real-time data from an ocean bottom network of tsunameters to simulate the wave field directly in real time instead of approximating initial conditions, and thus mitigating the uncertainties of modelling the seismic source. Wang et al. [62] used interpolation of observed waveforms to create virtual observational data, and demonstrated the success of the assimilation scheme when applied to forecasting simulations of the 2004 Indian ocean tsunami.

The purpose of the current study is not to offer alternative methodologies for simulating observational data or modified assimilation techniques. Instead, we present an analysis of the observation operator aimed at finding conditions on the observation operator necessary to achieve accurate reconstruction of the initial data. We have already proven sufficient conditions for optimal reconstruction of the true initial conditions using sparse observations for the 1-D SWE in [27]. Our objective was to determine the optimal number and locations of wave height measurements, such that that the optimally reconstructed initial conditions obtained via data assimilation converged to the true form. This was a first step in understanding observability conditions for the SWE. In order to extend the analysis to more realistic dynamics in tsunami models, here we implement an analogous scheme on the 2-D SWE system, and investigate whether the key results of our previous study extend to the more complex case with full dimensionality.

Due to the sparsity of observations in tsunami models, our aim for the 2-D case is to find the minimum information, in the form of observations of surface wave height, required for convergence to the true initial conditions. The focus on observation configuration differentiates the current work from existing literature on data assimilation for the 2-D SWE like Zou et al. [66], as they do not analyse the choice of the observation configuration, and focus instead on the minimum number of observational fields. Additionally, while they derive a variational scheme for the discrete 2-D SWE system, a novelty of our approach is that we consider the infinite dimensional case, unlike previous works on 4D-VAR. Our purpose in this chapter is to extend the 1-D data assimilation scheme derived in [27] to 2-D. Subsequently, we investigate whether the conditions for convergence observed for the 1-D case, also hold for the more complex 2-D system. We recall that for the 1-D assimilation, optimal convergence to the true initial conditions occurs when at least one pair of observation points are spaced more closely than half the effective minimum wavelength of the energy spectrum of the initial conditions.

In the 1-D case we were able to exploit the fact that the linear 1-D SWE system can be formulated as the 1-D wave equation, and subsequently we used its analytic properties to derive an exact solution for the adjoint system solved in the variational scheme. This solution was used to prove sufficient conditions for convergence to the true initial conditions. However, this method fails to extend to the 2-D case, as the analytical solution to the adjoint system is not easily found. Existing works, such as Iacono [25] on analytical solutions of the 2-D SWE, involve drastic simplifications of the equilibrium problem. Such simplifications would not help us find configurations of observation points that are effective in tsunami models.

Therefore, our approach is focused on a qualitative investigation of the applicability of results from the 1-D analysis for the 2-D data assimilation, and on a comparison of different observation configurations. We begin in section 2.2 by extending the data assimilation to 2-D, and implement this algorithm numerically in section 2.3. The results and complementary analyses are divided into the following three topics.

1. Section 2.4: Assess the optimal number and configuration of observation points such that convergence to the true initial conditions is achieved in the

data assimilation for the 2-D SWE.

2. Section 2.4.1: Investigate whether the sufficient conditions for convergence in [27] for the 1-D case extends to the 2-D data assimilation.
3. Section 2.5: Implement the data assimilation algorithm for a relatively realistic tsunami forecasting model, using non-flat bathymetry. Our final analysis uses realistic bathymetry data from the ETOPO2 global topographical relief database and the optimal observation configurations identified in this study.

We conclude with a summary of the main results and future considerations in section 2.6.

2.2 Derivation

The 2-D shallow water equations (SWE) are

$$\frac{\partial \eta}{\partial t} + \frac{\partial}{\partial x}\Big((\eta + H - \beta)u\Big) + \frac{\partial}{\partial y}\Big((\eta + H - \beta)v\Big) = 0, \tag{2.2.1a}$$

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + \frac{\partial \eta}{\partial x} = 0, \tag{2.2.1b}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + \frac{\partial \eta}{\partial y} = 0, \tag{2.2.1c}$$

$$\eta(x, y, 0) = \phi(x, y), \tag{2.2.1d}$$

$$u(x, y, 0) = 0, \tag{2.2.1e}$$

$$v(x, y, 0) = 0, \tag{2.2.1f}$$

where H is the average depth of the fluid and the system has been normalised such that $\sqrt{gH} = 1$. η and β are the perturbation of the free surface and sea floor respectively, and u and v are the velocities in the x and y direction. We assume the initial conditions $\phi \in L^2(\Omega)$ is compactly supported on the spatial domain $\Omega = \{(x, y); x \in [-L, L],\ y \in [-L, L]\}$ and is periodic at the boundaries. The notation used in the data assimilation algorithm is summarised in table 2.1.

Our objective is to minimise the least squares error between the observations $m_j^{(o)}(t)$ and the forecast solution of the wave height $\eta^{(f)}(x, y, t)$ given some initial

Symbol	Definition
$\eta(x, y, t)$	General solution for the height perturbation
$\phi^{(t)}(x, y)$	True initial conditions
$\beta(x, y)$	Time-independent bathymetry
$\phi^{(g)}(x, y)$	Starting guess for initial conditions
$\phi^{(n)}(x, y)$	Approximate bathymetry at iteration n of the assimilation algorithm
$\phi^{(b)}(x, y)$	Best approximation to the bathymetry (e.g., fixed point of iterations)
$m^{(o)}(t)$	Observations of the true height perturbation at positions $\{x_j, y_j\}, j = 1, ..., N_{obs}$
$\eta^{(f)}(x, y, t)$	Approximate ("forecast") solution generated by approximate bathymetry
$\mathcal{J}^{(n)}$	Cost function at iteration n
$(\cdot)^*$	Adjoint

TABLE 2.1: Notation used in the derivation of data assimilation scheme of the 2-D SWE to find the optimal initial conditions.

conditions ϕ. We express this as a cost functional $\mathcal{J} : L^2(\Omega) \to \mathbb{R}$, constrained by the system (2.2.1),

$$\mathcal{J}(\phi) = \frac{1}{2} \int_0^T \sum_{i=1}^{M} \left[\eta^{(f)}(x_j, y_j, t; \phi) - m_j^{(o)}(t) \right]^2 dt. \tag{2.2.2}$$

Then the optimal initial conditions $\phi^{(b)}$ is the minimiser defined as

$$\phi^{(b)} = \operatorname{argmin}_{\phi \in L^2(\Omega)} \mathcal{J}(\phi). \tag{2.2.3}$$

Since the minimum of (2.2.2) is achieved when

$$\nabla^{L^2} \mathcal{J}(\phi^{(b)}) = 0, \tag{2.2.4}$$

we formulate a dual adjoint system in terms of some appropriately chosen adjoint variables, such that $\nabla^{L^2} \mathcal{J}(\phi^{(b)})$ can be derived more efficiently than direct computation of the gradient of (2.2.2), given ϕ.

The first variation of $\mathcal{J}$, given some arbitrary perturbation ϕ' of scale ε is given by the Gateaux derivative,

$$\mathcal{J}'(\phi;\phi') = \lim_{\varepsilon\to 0} \frac{J(\phi+\varepsilon\eta') - J(\phi)}{\varepsilon}. \tag{2.2.5}$$

Expanding the perturbation to $\mathcal{O}(\varepsilon)$, we can reformulate (2.2.5) as

$$\mathcal{J}'(\phi;\phi') = -\int_0^T \Big(\eta^{(f)}(x_j, y_j, t;\phi) - m^{(o)}(t)\Big)\eta' \; dt, \tag{2.2.6}$$

where (η', u', v') are the solutions of the perturbed system given the perturbation in the initial conditions ϕ', found by linearising about (η, u, v) and extracting the $\mathcal{O}(\varepsilon)$ system,

$$\frac{\partial \eta'}{\partial x} + \frac{\partial}{\partial x}\Big(u\eta' + (H+\eta-\beta)u'\Big) + \frac{\partial}{\partial y}\Big(v\eta' + (H+\eta-\beta)v'\Big) = 0, \tag{2.2.7a}$$

$$\frac{\partial u'}{\partial t} + u\frac{\partial u'}{\partial x} + v\frac{\partial u'}{\partial y} + \frac{\partial \eta'}{\partial x} = 0, \tag{2.2.7b}$$

$$\frac{\partial v'}{\partial t} + u\frac{\partial v'}{\partial x} + v\frac{\partial v'}{\partial y} + \frac{\partial \eta'}{\partial y} = 0, \tag{2.2.7c}$$

$$\eta'(x,y,0) = \phi'(x,y), \tag{2.2.7d}$$

$$u'(x,y,0) = 0, \tag{2.2.7e}$$

$$v'(x,y,0) = 0. \tag{2.2.7f}$$

As the Gateaux derivative is a directional derivative in the direction of the perturbation ϕ', we can express (2.2.6) as the inner product between $\nabla\mathcal{J}$ and ϕ,

$$\mathcal{J}'(\phi;\phi') = \langle \nabla\mathcal{J}, \phi'\rangle_{L^2(\Omega)} = \int_\Omega \nabla^{L_2}\mathcal{J}\;\eta' \; d\Omega. \tag{2.2.8}$$

Then the following forms of $\mathcal{J}(\beta;\beta')$ are equivalent,

$$J'(\phi;\phi') = -\int_0^T \Big(\eta^{(f)}(x_j, y_j, t;\beta) - m^{(o)}(t)\Big)\eta' \; dt = \int_\Omega \nabla^{L^2}\mathcal{J}\;\phi' \; d\Omega. \tag{2.2.9}$$

We form a Lagrangian of our linearised system (2.2) with some arbitrary adjoint variables (η^*, u^*, v^*),

$$\begin{aligned}
&\int_0^T \int_\Omega \eta^*(x,y,t)\Big[\frac{\partial \eta'}{\partial x} + \frac{\partial}{\partial x}\Big(u\eta' + (H+\eta-\beta)u'\Big) + \frac{\partial}{\partial y}\Big(v\eta' + (H+\eta-\beta)v'\Big]\\
&\quad + u^*(x,t)\Big[\frac{\partial u'}{\partial t} + u\frac{\partial u'}{\partial x} + v\frac{\partial u'}{\partial y} + \frac{\partial \eta'}{\partial x}\Big]\\
&\quad + v^*(x,t)\Big[\frac{\partial v'}{\partial t} + u\frac{\partial v'}{\partial x} + v\frac{\partial v'}{\partial y} + \frac{\partial \eta'}{\partial y}\Big]\ d\Omega\ dt = 0.
\end{aligned} \tag{2.2.10}$$

By integrating by parts in time and space, and using the fact that due to periodicity our boundary terms at $\partial\Omega$ vanish, (3.3.9) reduces to

$$\begin{aligned}
0 = -\int_0^T \int_\Omega \Bigg\{ &\eta'\Big[\frac{\partial \eta^*}{\partial t} + u\frac{\partial \eta^*}{\partial x} + v\frac{\partial \eta^*}{\partial y} + \frac{\partial u^*}{\partial x} + \frac{\partial v^*}{\partial y}\Big]\\
&+u'\Big[\frac{\partial u^*}{\partial t} + u\frac{\partial u^*}{\partial x} + v\frac{\partial u^*}{\partial y} + (H+\eta-\beta)\frac{\partial \eta^*}{\partial x}\Big]\\
&+v'\Big[\frac{\partial v^*}{\partial t} + u\frac{\partial v^*}{\partial x} + v\frac{\partial v^*}{\partial y} + (H+\eta-\beta)\frac{\partial \eta^*}{\partial y}\Big]\Bigg\}\ d\Omega\ dt\\
&-\int_\Omega \eta^*\eta'\Big|_{t=T} d\Omega\ - \int_\Omega \eta^*\eta'\Big|_{t=0} d\Omega\\
&+\int_\Omega u^*u'\Big|_{t=T} d\Omega - \int_\Omega u^*u'\Big|_{t=0} d\Omega.
\end{aligned} \tag{2.2.11}$$

If we pick (η^*, u^*, v^*) as the solution to

$$\frac{\partial \eta^*}{\partial t} + u\frac{\partial \eta^*}{\partial x} + v\frac{\partial \eta^*}{\partial y} + \frac{\partial u^*}{\partial x} + \frac{\partial v^*}{\partial y} = m(x_i, y_i, t) - H\eta(x,y,t;\beta), \tag{2.2.12a}$$

$$\frac{\partial u^*}{\partial t} + u\frac{\partial u^*}{\partial x} + v\frac{\partial u^*}{\partial y} + (H+\eta-\beta)\frac{\partial \eta^*}{\partial x} = 0, \tag{2.2.12b}$$

$$\frac{\partial v^*}{\partial t} + u\frac{\partial v^*}{\partial x} + v\frac{\partial v^*}{\partial y} + (H+\eta-\beta)\frac{\partial \eta^*}{\partial y} = 0, \tag{2.2.12c}$$

$$\eta^*(x,y,T) = 0, \tag{2.2.12d}$$

$$u^*(x,y,T) = 0, \tag{2.2.12e}$$

$$v^*(x,y,T) = 0. \tag{2.2.12f}$$

Algorithm 1 Data Assimilation Algorithm for initial conditions Estimation for the 2-D SWE.

```
Pick initial estimate for $\phi^{(g)}$.
Solve the initial value problem for $(\eta, u, v)$ from $t = 0$ to $t = T$.
Solve adjoint problem for $(\eta^*, u^*, v^*)$ backwards in time from $t = T$ to $t = 0$ to find $\eta^*(x, y, 0)$.
Define $\nabla^{L^2}\mathcal{J} = -\eta^*(x, y, 0)$.
Compute the optimal time step $\tau_n$ at the $n$-th iteration through a line minimisation algorithm
    $\tau_n = \text{argmin}_{\tau\in\mathbb{R}}\ \mathcal{J}\Big(\phi^{(n)}(x, y) - \tau\nabla^{\mathcal{L}_2}\mathcal{J}\Big(\phi^{(n)}(x, y)\Big)\Big).$   (2.2.13)
Use a gradient descent algorithmn to compute the guess for $\phi$ at the next time step
    $\phi^{(n+1)}(x, y) = \phi^{(n)}(x, y) - \tau_n\nabla^{L^2}\mathcal{J}\Big(\phi^{(n)}(x, y)\Big).$   (2.2.14)
Repeat until $\parallel \eta^*(x, y, 0) \parallel\approx 0$.
Set $\phi^{(b)}(x, y) := \phi^{(n)}(x, y)$.
```

Then (2.2.11) is reduced to

$$\int_0^T \int_\Omega \Big(\eta^{(f)}(x_j, y_j, t; \phi) - m^{(o)}(t)\Big)\eta' \, d\Omega \, dt = -\int_\Omega \eta^*\eta'\Big|_{t=0} d\Omega, \tag{2.2.15}$$

Combining this result with the equivalence given by (2.2.9), we have

$$-\int_\Omega \eta^*\eta'\Big|_{t=0} d\Omega = -\int_\Omega \eta^*\Big|_{t=0} \phi' \, d\Omega = \int_\Omega \nabla^{L^2}\mathcal{J} \, \phi' \, d\Omega, \tag{2.2.16}$$

and thus since our functional is linear and bounded, and belongs to the space of square-integrable functions, we can use the Riesz representation theorem to extract $\nabla^{L_2}\mathcal{J}$, giving us

$$\nabla^{L^2}\mathcal{J}(\phi) = -\eta^*(x, y, 0). \tag{2.2.17}$$

We utilise an iterative steepest descent algorithm to find our minimiser $\phi^{(b)}$ yielding $\nabla^{L^2}\mathcal{J} = 0$, given some starting guess $\phi^{(g)}$. The steps of the process at each iteration are summarised in algorithm 1.

2.3 Numerical implementation

Our spatial domain is $\Omega = [-L\ L] \times [-L\ L]$, where $L = 3$. The assimilation is carried out in the time interval $[0\ T]$ where the final time $T = 2$ is selected such that there are no boundary effects.

For the spatial discretisation we implement a second-order finite difference-finite volume Sadourny energy conserving scheme on a C-grid. A C-grid arrangement (as shown in figure 2.1), is a regular quadrilateral grid symmetric about the x and y directions, and has been demonstrated to mitigate group velocity errors in the numerical discretisation of the SWE. The Sadourny scheme ensures that this discretisation inherits the conservation properties (such as energy or enstrophy conservation) of the original system [47].

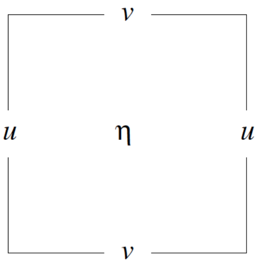


Figure 2.1: Staggered arrangement of variables on a square.

If the SWE are expressed in the following equivalent form,

$$\frac{\partial\eta}{\partial t} + \frac{\partial}{\partial x}\Big((\eta + H - \beta)u\Big) + \frac{\partial}{\partial y}\Big((\eta + H - \beta)v\Big) = 0 \tag{2.3.1}$$

$$\frac{\partial u}{\partial t} - \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)v + \frac{\partial}{\partial x}\left(g\eta + \frac{1}{2}u^2 + \frac{1}{2}v^2\right) = 0 \tag{2.3.2}$$

$$\frac{\partial v}{\partial t} + \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)v + \frac{\partial}{\partial y}\left(g\eta + \frac{1}{2}u^2 + \frac{1}{2}v^2\right) = 0, \tag{2.3.3}$$

Then the discretisation

$$\frac{\partial u}{\partial t} = \frac{1}{\Delta x}\delta_i\Big(g\eta + \frac{1}{2}u^2 + \frac{1}{2}v^2\Big) + \Big(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\Big)\overline{\overline{(\big(\overline{1+\eta-\beta)}^{j}v\big)}^{i}}^{j}, \tag{2.3.4}$$

$$\frac{\partial v}{\partial t} = \frac{1}{\Delta y}\delta_i\Big(g\eta + \frac{1}{2}u^2 + \frac{1}{2}v^2\Big) + \Big(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\Big)\overline{\overline{\big(\overline{(1+\eta-\beta)}^{i}u\big)}^{j}}^{i}, \tag{2.3.5}$$

where

$$\delta_i f = f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}}, \tag{2.3.6}$$

$$\overline{f}^{i} = \frac{1}{2}\Big(f_{i+\frac{1}{2}} + f_{i-\frac{1}{2}}\Big), \tag{2.3.7}$$

ensures that the total energy is conserved, such that

$$\begin{aligned} 0 =& \sum_{ij}\Delta x\Delta y\Big[g\eta\frac{\partial\eta}{\partial t} + \Big(\frac{1}{2}u^2 + \frac{1}{2}v^2\Big)\frac{\partial}{\partial t}(1+\eta-\beta)\Big] \\ &+ \sum_{ij}\Delta x\Delta y\Big[(1+\eta-\beta)u\frac{\partial u}{\partial t}\Big] \\ &+ \sum_{ij}\Delta y\Delta x\Big[(1+\eta-\beta)v\frac{\partial v}{\partial t}\Big] \end{aligned} \tag{2.3.8}$$

where the staggered centre difference operator (2.3.6) and centre linear interpolation operator (2.3.7) ensure that the velocities u and v and their time derivatives are discretised at the centre of the cell boundaries, and η and β and their derivatives are located at cell centres, as in figure 2.1. The spatial derivations are discretised using a second order finite difference scheme with periodic boundary conditions. The system is then integrated using a four stage third order Runge-Kutta scheme [57]. The resolution is 128 × 128 due to computational cost considerations, as the memory requirements for the solver are quite high. The time step $\Delta t = \Delta x/3$ is chosen to satisfy the Courant-Friedrichs-Lewy (CFL) condition, and subsequently the solutions for η, u, v are $128 \times 128 \times 193$ matrices, which need to be stored for all time steps. The line minimisation for (2.2.13) is carried out using the Matlab function `fminunc`.

We also consider the following anisotropic Gaussian initial conditions,

$$\phi^{(t)}(x,y) = \frac{1}{20}\exp\Big(-\frac{x^2+y^2}{0.1^2}\Big). \tag{2.3.9}$$

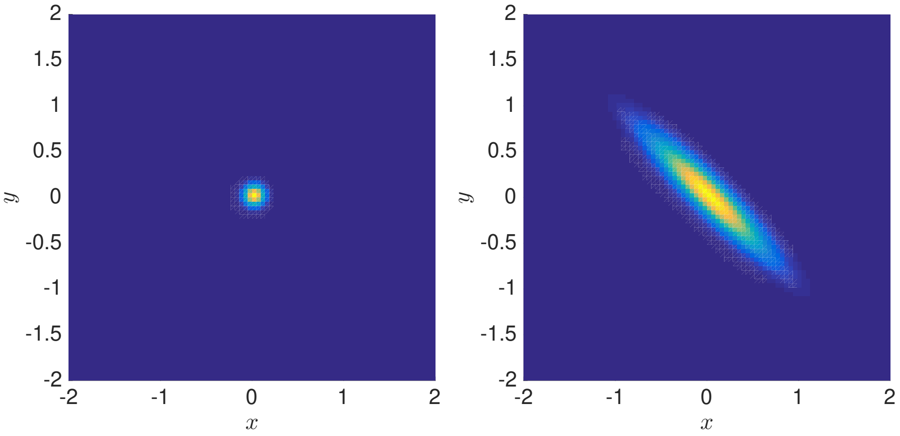
For cases where a non-isotropic initial conditions is used,

$$\phi^{(t)}(x,y) = \frac{1}{20}\exp\Big(-\frac{x_r^2+(y_r/5)^2}{0.15^2}\Big), \tag{2.3.10}$$

$$x_r = x\cos\Big(-\frac{\pi}{4}\Big) - y\sin\Big(-\frac{\pi}{4}\Big), \tag{2.3.11}$$

$$y_r = x\cos\Big(-\frac{\pi}{4}\Big) + y\sin\Big(-\frac{\pi}{4}\Big). \tag{2.3.12}$$

Figure 2.2 shows the isotropic and anisotropic initial conditions. In the results, we indicate an approximation of this support to highlight the position of the the observation points relative to the initial conditions. While Gaussian initial conditions are not technically compactly supported, they are in practise for numerical simulations, for example when the exponential is smaller than machine precision.



(a) Isotropic (b) Non-isotropic

FIGURE 2.2: Planar view of the truel initial conditions $\phi^{(t)}$. The highlighted regions indicate the support of the initial conditions.

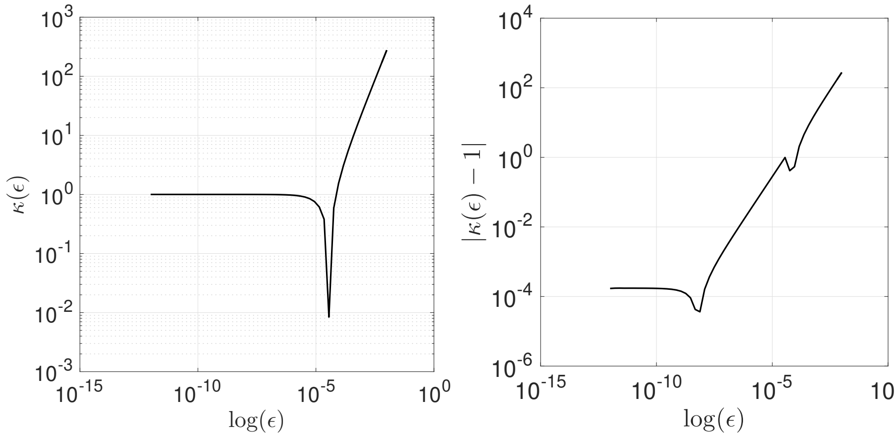
The results in section 2.4 and 2.4 are for assimilation with flat bathymetry $\beta(x,y) = 0$ as in [27]. However, in section 2.5 we implement the data assimilation algorithm with (i) a Gaussian form bathymetry, and (ii) realistic bathymetry data taken from the ETOPO2 database for global topography and bathymetry.

Verification test

To verify our numerical implementation, we conduct a verification test using the equivalent forms for $\mathcal{J}'(\phi, \phi')$ used in our derivation, the Gateaux derivative and the Riesz gradient. If the implementation is correct, the parameter κ should converge to 1 as $\varepsilon \to 0$,

$$\kappa(\varepsilon) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \frac{\mathcal{J}(\phi + \varepsilon\phi') - \mathcal{J}(\phi)}{\int_\Omega -\eta^*(x, y, 0)\phi' \, d\Omega}. \tag{2.3.13}$$

The results of the verification test are shown in figure 2.3, with a non-zero bathymetry.



(a) Isotropic (b) Non-isotropic

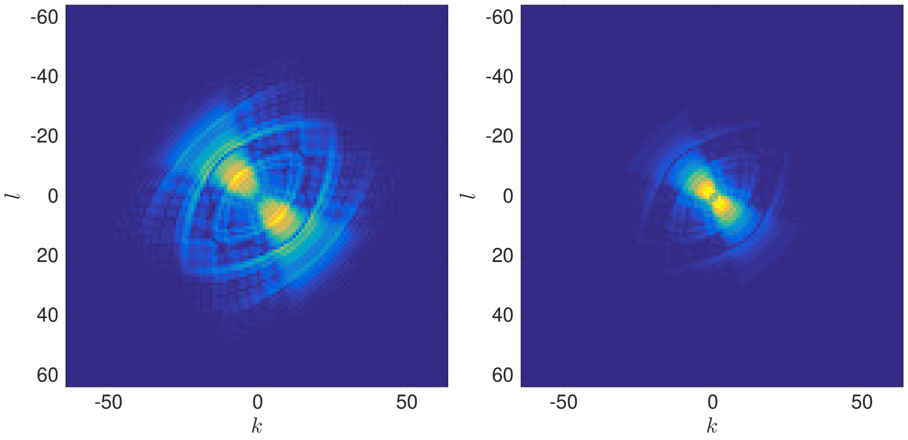
Figure 2.3: Verification test for numerical implementation. (a) is the loglog graph of κ as $\varepsilon \to 0$, and (b) is the difference $|\kappa(\varepsilon) - 1|$ as $\varepsilon \to 0$.

Smoothing of $\nabla^{L^2}\mathcal{J}(\phi)$

To mitigate noise in the gradient $\nabla^{L^2}\mathcal{J}$ which may accumulate within the optimisation, we can implement the following low-pass filter in Fourier space that smooths $\nabla\mathcal{J}$ such that it exists in $H^1(\mathbb{R}^2)$ and subsequently has a smooth first derivative as well. The smoothing can be described by

$$(\widehat{\nabla^{H^1}\mathcal{J}})_{(k,l)} = \frac{1}{1+m^2(k^2+l^2)}(\widehat{\nabla^{L^2}\mathcal{J}})_{k,l}, \tag{2.3.14}$$

where k and l are the wavenumbers for the x and y directions respectively, and the filter co-efficient m is chosen appropriately to adjust regularity. All modes larger than the cut-off frequency will effectively be damped. An illustration of the filter applied to $\nabla\mathcal{J}$ after the first iteration in case 2.7(a) is given in figure 2.4, where all but the lowest 10% frequencies are damped to zero.



(a) $|\widehat{\nabla^{L^2}\mathcal{J}(\phi)}_{k,l}|$

(b) $|\nabla\widehat{\mathcal{J}^{H^1}(\phi)}_{k,l}|$

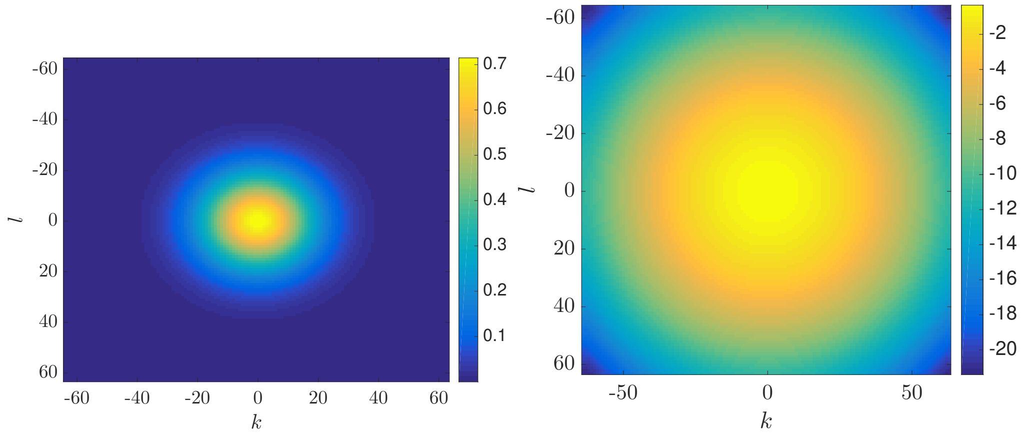
FIGURE 2.4: Energy spectrum of gradient $\nabla\mathcal{J}(\phi)$ after the first iteration in 2.7(a),before (a) and after (b) smoothing using (2.3.14).

The filter was tested in most cases where convergence was sub-optimal in order to gauge whether errors were due to noise in the reconstruction, however generally it was found that there was little change in the overall result, suggesting the source of error is not primarily due to noise in the gradient $\nabla\mathcal{J}(\phi)$.

2.4 Numerical results

The analysis for the 2-D data assimilation is more complex than the 1-D case. The configuration of observations has more degrees of freedom than simply the number and spacing of points on the real line $\mathbb{R}$. With the increased dimensionality from 1-D to 2-D, there is far more freedom and geometric complexity in the placement of the observation points. With this in mind, in this section we explore the effect of configuration and number of observations on the overall convergence of the reconstructed initial conditions. In [27], all observation points were located to one side of the initial condition support. In the present analysis, we consider observations in all quadrants of our spatial domain (where the initial conditions is centred at $(0, 0)$). We note that though in most cases our initial conditions is isotropic and the subsequent wave propagation is radially symmetric given a flat bathymetry, this azimuthal symmetry property cannot be observed by observation points along characteristics in a single quadrant alone. This suggests the isotropic nature of the initial conditions is only captured by observations along part of the circle $x^2 + y^2 = R(t)^2$, where $R(t)$ is the radial position of the propagating wave $\eta(x, y, t)$ at each time $t \in [0\ T]$.

We analyse results for three configurations of observation placement. (i) In a square grid centred at $(0, 0)$, (ii) along the characteristics $x = y$ and $x = -y$, and (iii) on arcs along the two circles $x^2 + y^2 = r_i^2$ for $i = 1, 2$. In section 2.4.1 we extend the analysis for observation spacing from the $1 - D$ case as in [27] for each configuration. For all cases, we assume the initial guess $\phi^{(g)} = 0$. In section 2.4.2 we build on the optimal results for observation spacing, and investigate whether convergence improves when increasing the number of observation points. In reality, it is unlikely there will be very many observations within the support of the initial conditions of a tsunami. Therefore, we also investigate the effect of removing all observation points located within the support of both isotropic and anisotropic initial conditions on convergence.



(a) (b)

FIGURE 2.5: scaled colour map of (a) $|\widehat{(\phi)}_{k,l}|$, and (b) $\log|\widehat{(\phi)}_{k,l}|$

2.4.1 Observation spacing

In order to investigate the effect of observation spacing for the 2-D initial conditions assimilation, we use the results of [27] as a benchmark. We recall that for the 1-D assimilation, optimal convergence to the true initial conditions $\phi^{(t)}$ occurs when at least one pair of observation points are spaced more closely than half the effective minimum wavelength of the energy spectrum of the initial conditions. Subsequently given a pair of observation points (m_1, m_2) we require

$$|m_1 - m_2| \leq \frac{\pi}{k_{max}}, \tag{2.4.1}$$

where k_{max} is the largest effective wavenumber of $\phi^{(t)}$. In the absence of a rigorous equivalent result for the 2-D assimilation, we qualitatively investigate whether convergence improves when (2.4.1) is satisfied for the 2-D wavenumbers k and l. To find k_{max} and l_{max}, we examine the energy spectrum of the true isotropic 2-D initial conditions given in figure 2.5. We observe that the modes for $k, l \geq 40$ have relatively negligible magnitudes, and thus $k_{max} = l_{max} = 40$ is a good approximation. Thus the distance between at least two observation points should be less than 0.0785.

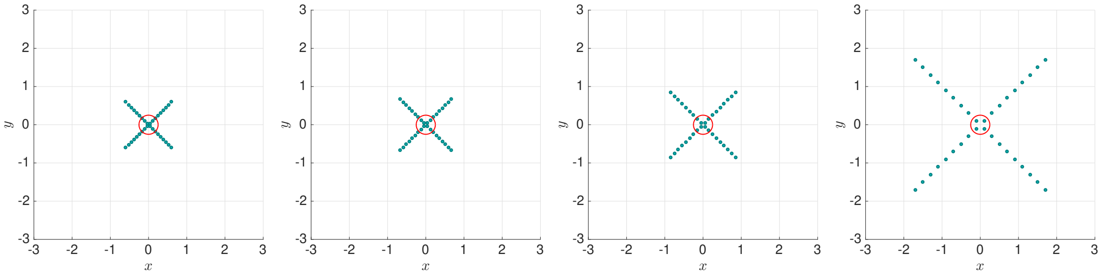
In the 1-D case, at least one pair of points was required to satisfy (2.4.1).

However it may be that a greater number of points satisfying this condition is correlated with improved convergence in the 2-D case, due to the need to resolve the two-dimensional shape of the initial conditions (e.g. azimuthal symmetry). To investigate this, we compare results with different spacing of observations $\Delta x = \Delta y$ for (Δr for the arc configurations) in a grid, along characteristics $x = y$ and $x = -y$, and in radial arcs centred on the initial conditions. We also investigate convergence with different numbers of pairwise observations with a Euclidean distance fulfilling (2.4.1). We begin with observations along characteristic lines $x = \pm y$.

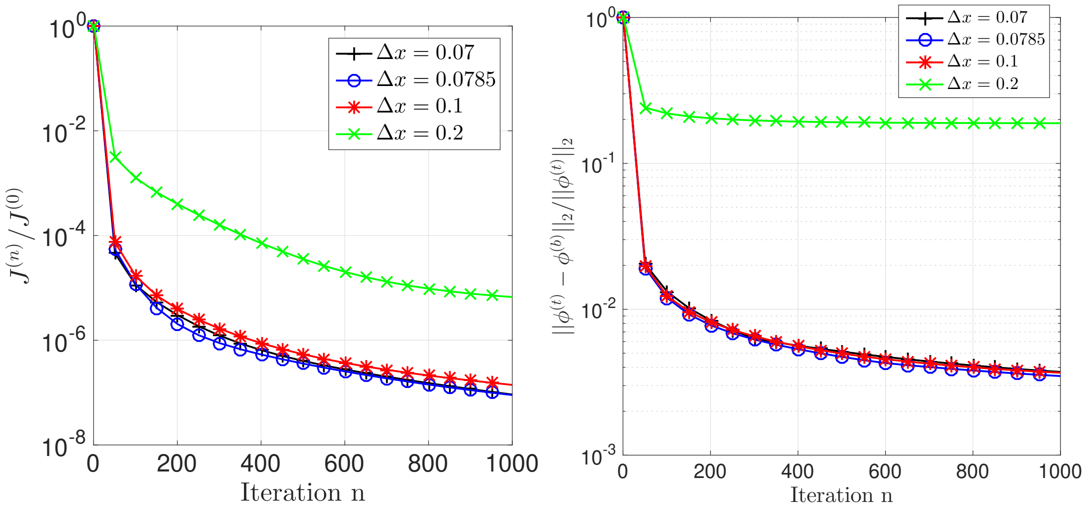
Observations along characteristic $x = \pm y$

For this case, we initially choose $N_{obs} = 36$. This initial choice of N_{obs} is larger than those considered in [27] where $N_{obs} \leq 5$ was considered. However, given the higher dimensionality of the 2-D problem, we scale up our initial choice of N_{obs}. We evaluate results for the $\Delta x = \Delta y = 0.07,\ 0.0785,\ 0.1$, and 0.2. These values were chosen to include cases both when (2.4.1) is satisfied, and when it is violated. Since observations are placed along $x = \pm y$, the Euclidean distance between points is $\sqrt{(\Delta x)^2 + (\Delta y)^2}$. Comparing pairwise distances between any two observation points in the configuration, we observe that at least 8 out of a total of $36(36-1)$ pairwise observations have a Euclidean distance less than 0.0785 for $\Delta x = 0.07,\ 0.0785$, and 0.1. There are zero pairs satisfying (2.4.1) for $\Delta x = 0.2$
.

The characteristic configurations can be seen in figure 2.6(a), (b), (c), and (d). The red circle represents the boundary of the initial condition. It is an approximate representation intended for reference only, and the actual region where $\phi^{(t)} \neq 0$ may be slightly larger, and corresponds to figure 2.2. The configurations in 2.6(a), 2.6(b), and 2.6(c) satisfy (2.4.1), and those in 2.6(d) do not. We wish to determine whether convergence is worse where spacing is larger than 0.0785.



(a) $\Delta x = 0.07$ (b) $\Delta x = 0.0785$ (c) $\Delta x = 0.1$ (d) $\Delta x = 0.2$



(e) Cost Function (f) Relative L^2 Error

Figure 2.6: Observations along characteristics $x = \pm y$ with varying spacing such that $\Delta x = \Delta y$, and $N_{obs} = 36$. The green circles represent the observation points, and the area inside the red circle approximates the support of the initial conditions. The assimilation time is $t \in [0\ 2]$. (d) shows the convergence of the cost function after 1000 iterations and (e) represents the relative L^2 error in the initial conditions reconstruction. We note that configuration (d) with $\Delta x = 0.2$ fails to converge.

The convergence results for both cases are given in 2.6(e) and 2.6(f), where the former is the relative decrease in the cost function after 1000 iterations, and the latter is the relative L^2 error in the reconstruction. We observe that the assimilation converges for for the three cases with $\Delta x \leq 0.1$. As 0.0785 is just an estimate of the minimum spacing, it is not surprising that the slightly larger 0.1

case also converges. There is no significant difference between the three cases and the error is reduced to $\mathcal{O}(10^{-3})$. However, the configuration with $\Delta x = 0.2$ fails to converge. These results confirm condition (2.4.1) (and therefore are consistent with the minimum spacing theorem in [27]), as $\Delta x = 0.2$ was the only case where there were no pairs of observations with a Euclidean distance less than 0.0785. To analyse the smoothness of the convergence, we considered the relative L^2 error for an extended range of Δx, and observed that the error with spacing of observations such that $0.1 < \Delta x \leq 0.2$ still converged despite having a higher error compared to $\Delta x \leq 0.1$, and no configuration with $\Delta x \geq 0.2$ achieved convergence of the L^2 error less than 10%.

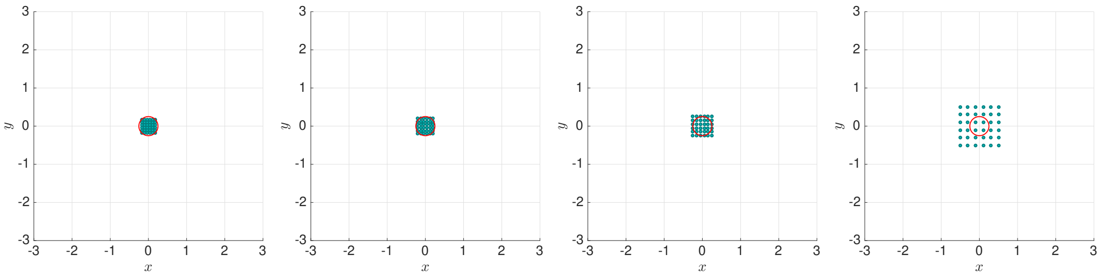
We observe that the placement along characteristics in all four quadrants appears to sufficiently capture the radial symmetry of the propagating wave and reconstruct the initial conditions accurately. We verify whether these results by considering observations placed in a grid, and along concentric circles centred at the initial conditions.

Observations in a grid formation

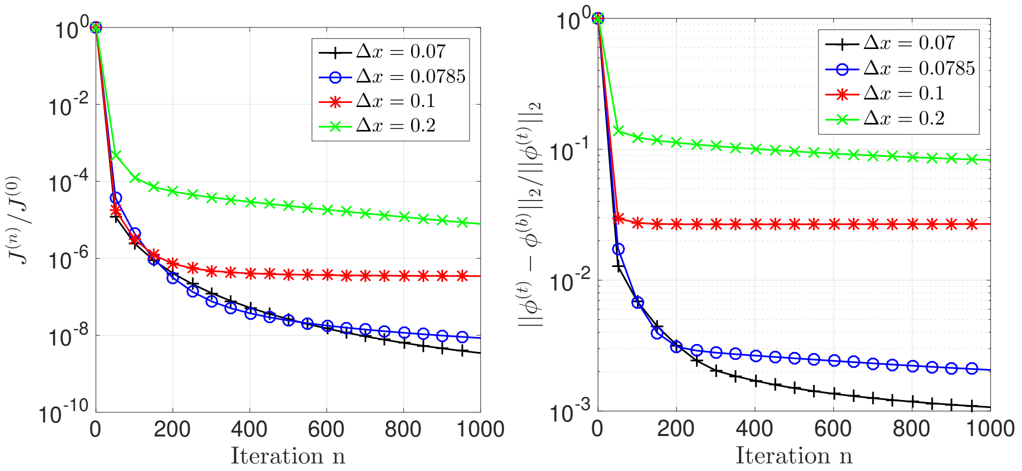
The second configuration of observation points we consider is a grid layout in the xy plane. To maintain comparison with the characteristic configuration, we initially choose $N_{obs} = 6^2$, and observations are arranged in a six-by-six square grid centred at $(0, 0)$. These configurations are presented in figure 2.7(a), (b), (c), and (d) with the same values of Δx as for the characteristic case. The trend in the cost function and relative error decrease over 1000 iterations are presented in figures 2.7(c) and (d).

We can see that like the characteristic configuration, the results for observations in a grid configuration show increasing convergence to zero as the spacing between adjacent observation points in the x and y directions decrease. Even the worst performing configuration 2.7(d) has a relative error at the final iteration of $\mathcal{O}(10^{-2})$. This suggests that if we classify a result as convergent if the relative error is lower than 10%, then each spacing considered in figure 2.7 convergences, though it is clear the lower values $\Delta x = 0.07$ and $\Delta x = 0.0785$ perform better by an order of magnitude. Once again, we note that these configurations both had

64 pairwise observations satisfying (2.4.1), whereas the ones with $\Delta x = 0.1$ and $\Delta x = 0.2$ had none, confirming the hypothesis that (2.4.1) is a sufficient condition for convergence with a grid configuration, for the 2-D case. While convergence with $\Delta x = 0.07$ is marginally better for the grid configuration than the characteristic configuration, they are both $\mathcal{O}(10^{-3})$ and we may consider them equivalent in performance thus far. We note that these results are also consistent with the 1-D theorem in [27].



(a) $\Delta x = 0.07$ (b) $\Delta x = 0.0785$ (c) $\Delta x = 0.1$ (d) $\Delta x = 0.2$

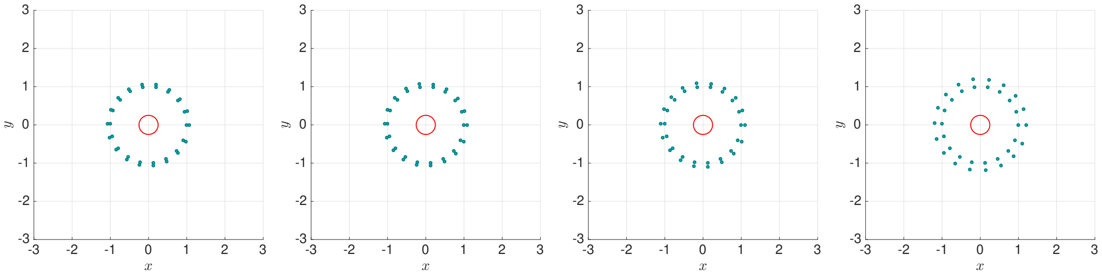


(e) Cost Function (f) Relative L^2 Error

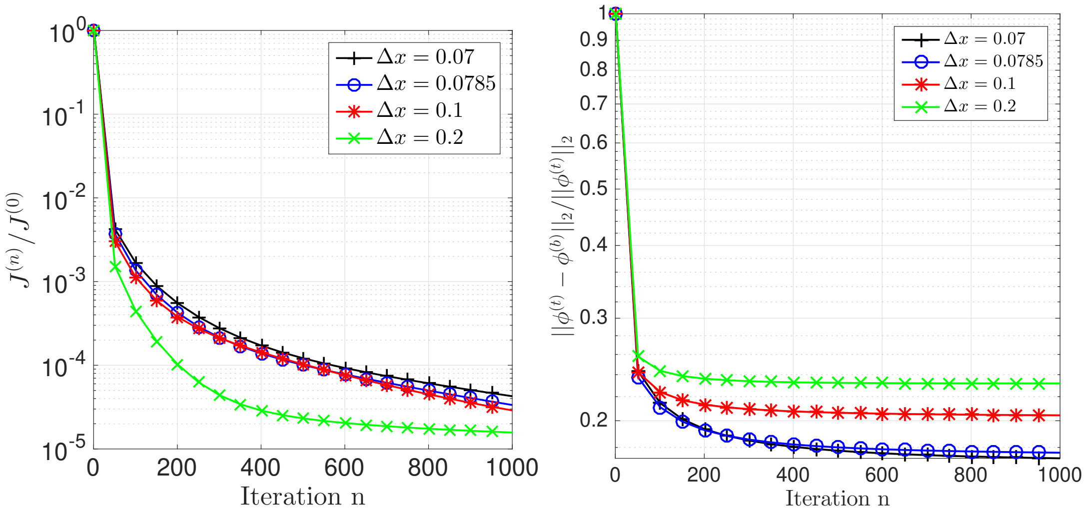
FIGURE 2.7: Observation configuraton in a grid format with varying spacing such that $\Delta x = \Delta y$, and $N_{obs} = 36$. (d) shows the convergence of the cost function after 1000 iterations and(e) represents the relative L^2 error in the initial conditions reconstruction. All configurations converge (L^2 error less than 10%), however it is marginal for (d).

It should be noted however that in 2.7, most of the observations are placed within the support of the initial conditions. This is a consequence of having a grid centred at the initial conditions support with only $N_{obs} = 36$ and small values of Δx. Consequently in section 2.4.2 we investigate the effect of increasing the number of observation points, and subsequently in section 2.4.3 the effect of removing all points that lie in the initial conditions support.

Observations along arcs



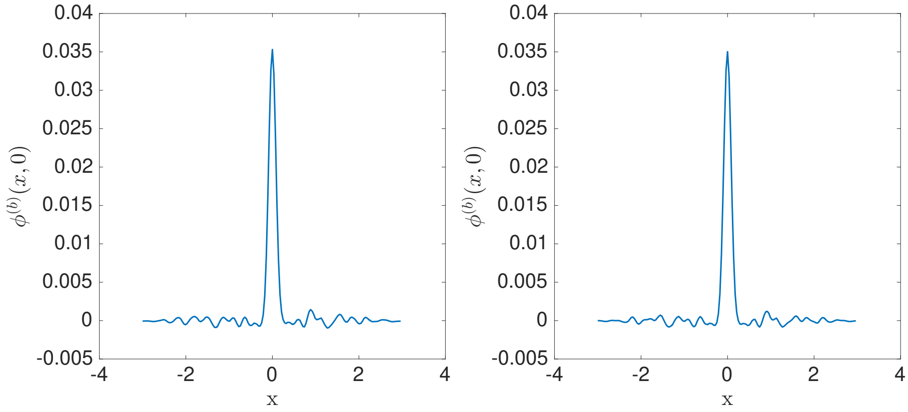
(a) $\Delta r = 0.07$ (b) $\Delta r = 0.0785$ (c) $\Delta r = 0.1$ (d) $\Delta x = 0.2$



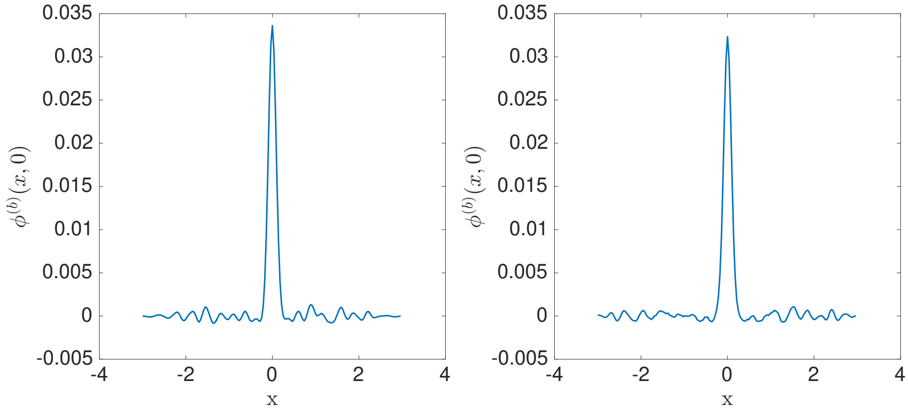
(e) Cost Function (f) Relative L^2 Error

FIGURE 2.8: Observations along the two concentric circles $x^2+y^2 = r_i^2$ for $i = 1, 2$ and with spacing between circles $\Delta r = |r_1 - r_2|$ and $N_{obs} = 36$. (d) shows the convergence of the cost function after 1000 iterations and(e) represents the relative L^2 error in the initial conditions reconstruction.

Current observations used for detection and forecasting of tsunamis such as DART buoys are usually arrayed in an arc-like formation along coastlines [18]. Therefore, we consider a circular arc configuration for observations, as shown in figure 2.8(a)-(d). The observation points are placed along two concentric circles $x^2 + y^2 = r_i^2$, for $i = 1, 2$. We vary the spacing $\Delta r = |r_1 - r_2|$ such that 2.8(a) and 2.8(b) satisfy (2.4.1) and 2.8(c) and 2.8(d) don't, similar to the characteristic and grid configurations.



(a) $\Delta r = 0.07$ (b) $\Delta r = 0.0785$



(c) $\Delta r = 0.1$ (d) $\Delta r = 0.2$

FIGURE 2.9: Cross-sections of the reconstructed initial conditions at $y = 0$ for cases in figure 2.8 with spacing $\Delta r = |r_1 - r_2|$ and $N_{obs} = 36$. We observe that the peak of the Gaussian is not fully resolved (true amplitude is 0.05) and there is small-scale noise in the reconstruction for all cases.

We observe in 2.8(d) that the error does not converge for any of the values of Δr considered, and is $\mathcal{O}(10^{-1})$. This is despite the fact that there are 14 pairwise observations satisfying (2.4.1) for $\Delta r = 0.07$ and 8 pairs for $\Delta r = 0.0785$. We present the reconstructed initial conditions for each Δr in figure 2.9. We observe that in each case, the Gaussian peak is not fully resolved and the amplitude is smaller than the exact initial conditions (0.05). Additionally in each case there is small-scale noise in the reconstruction.

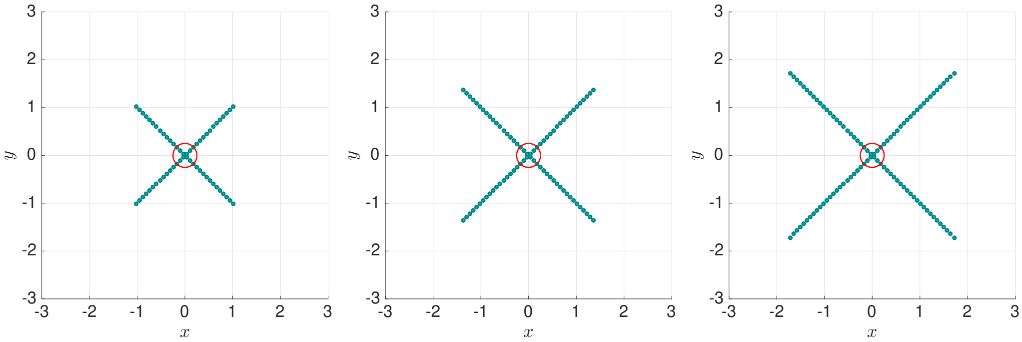
In theory, observations placed in concentric circles around the true initial conditions should be able to capture the radial propagation of the free surface wave in all directions, however these results suggest that either small spacing between pairwise observations is not sufficient for convergence with an arc configuration or that a larger number of pairwise observations satisying (2.4.1) is necessary. In the following section we investigate whether convergence improves as we increase the number of observations in each of the three configurations.

2.4.2 Results with large or small number of observations

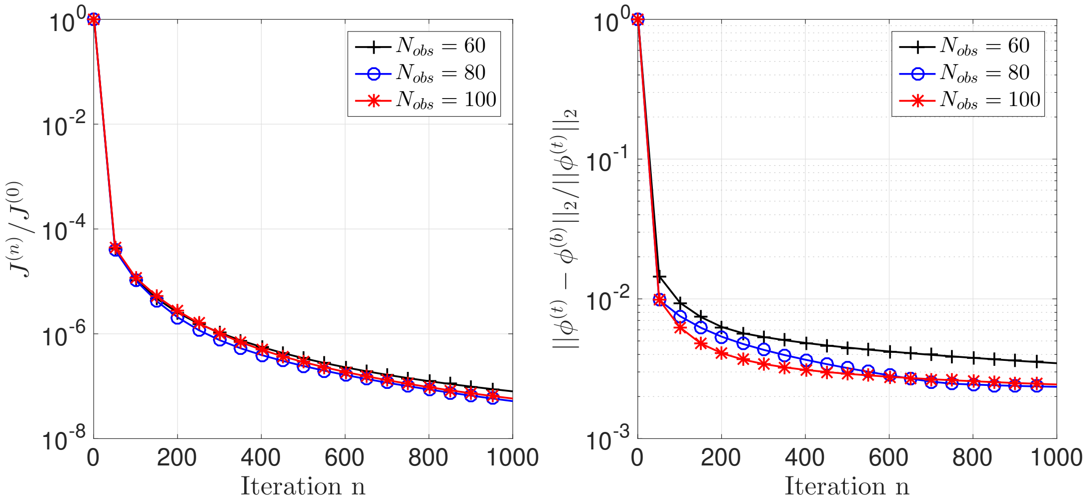
In this section observe the convergence of the assimilation when the number of observation points is either large or small for the characteristic, grid, and arc configurations respectively. The spacing Δx and Δr is fixed at the best-performing case considered in section 2.4.1 ($\Delta x = 0.07$ or $\Delta x = 0.0785$). We first show convergence for large N_{obs}, and then subsequently investigate the minimum number of observations required for convergence. For the grid and arc configurations, we initially consider results for $N_{obs} = 10^2, 12^2, 14^2$ respectively, with $\Delta x = \Delta r = 0.07$. For the characteristic configuration we chose $N_{obs} = 60, 80, 100$ and $\Delta x = 0.0785$. The comparatively smaller values for the characteristic configuration are due to the fact that we required all observation points to interact with the propagating wave in the assimilation time T, and this would not have been possible for larger values of N_{obs} placed along characteristics.

Characteristic configuration

The results for the characteristic configuration are presented in figure 2.10, where (a), (b) and (c) show the configurations for each value of N_{obs}. We observe that the relative reconstruction error presented in figure 2.10(e) does not change significantly with the number of observation points, and is $\mathcal{O}(10^{-3})$. We note that the same level of convergence was achieved with only 36 observations, as seen in figure 2.6(e). This is encouraging from a tsunami modelling perspective, as it would suggest that a smaller observation network placed along characteristics is sufficient to recreate the initial conditions effectively.



(a) $N_{obs} = 60$ (b) $N_{obs} = 80$ (c) $N_{obs} = 100$

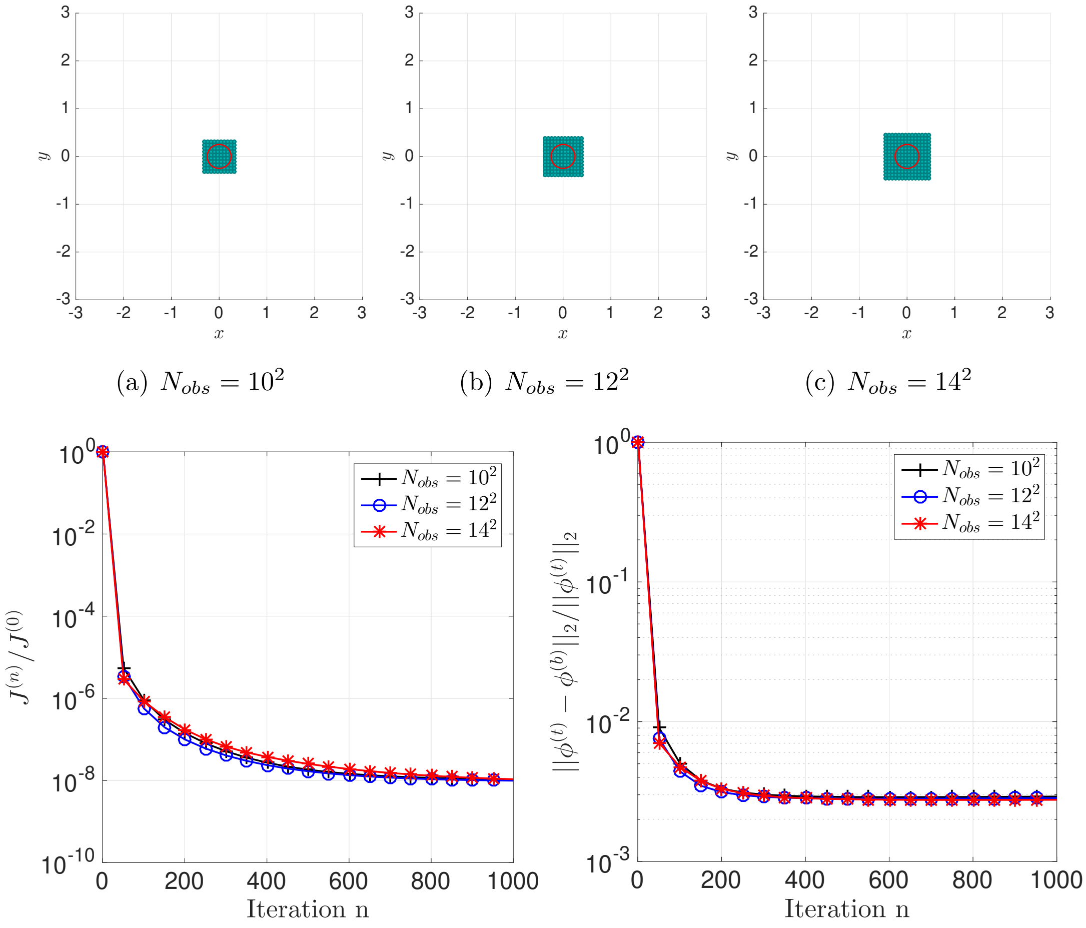


(d) Cost (e) Error

Figure 2.10: Characteristic formation with increased number of points and $\Delta x = 0.0785$.

Grid Configuration

The configurations of observations placed in a grid format with varying number of observation points is given in figures 2.11(a), (b), and (c). The convergence results are presented in 2.11(d) and 2.11(e).



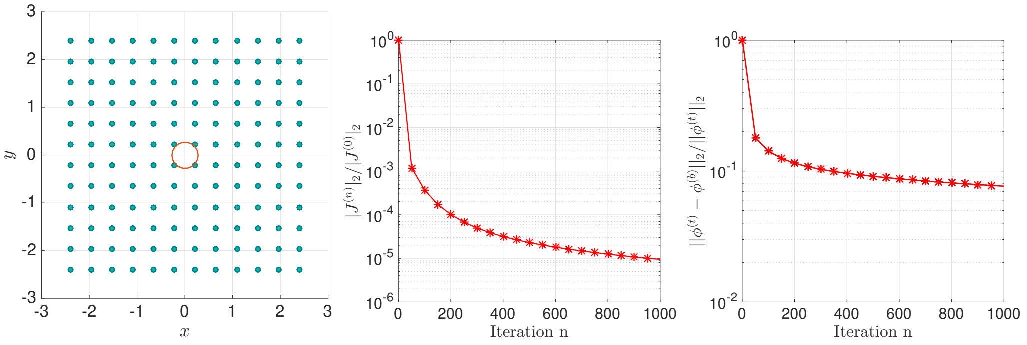
(a) $N_{obs} = 10^2$ (b) $N_{obs} = 12^2$ (c) $N_{obs} = 14^2$

(d) Cost (e) Error

Figure 2.11: Grid formation with increased number of points and $\Delta x = 0.07$.

As with the characteristic configuration, we observe that the convergence of the L^2 relative error does not significantly change for the large values of N_{obs}, and is $\mathcal{O}(10^{-3})$. To determine whether having a large number of observations is a sufficient condition for convergence, we consider results where $N_{obs} = 12^2$, but

$\Delta x = 0.44$. We see from figure 2.12(c) that the assimilation does not converge, and therefore large N_{obs} is not a sufficient condition for convergence. To verify whether it is a necessary condition for convergence, we explore results when N_{obs} is small, in section 2.4.2.



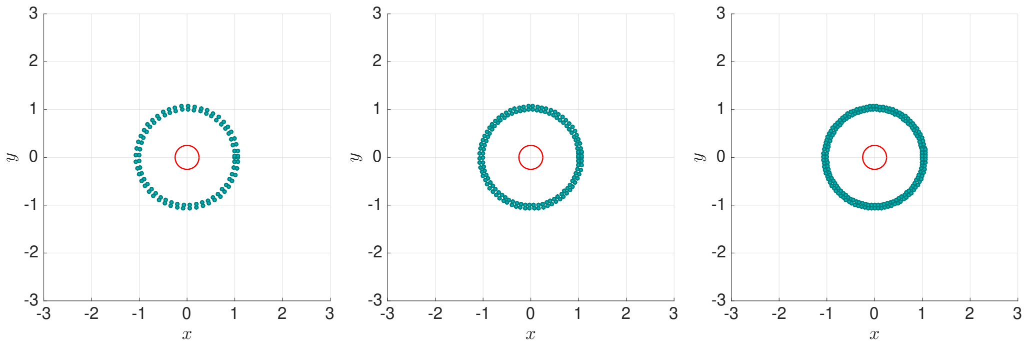
(a) $N_{obs} = 12^2$ (b) Cost (c) Error

FIGURE 2.12: Grid formation with increased number of points and $\Delta x = 0.44$

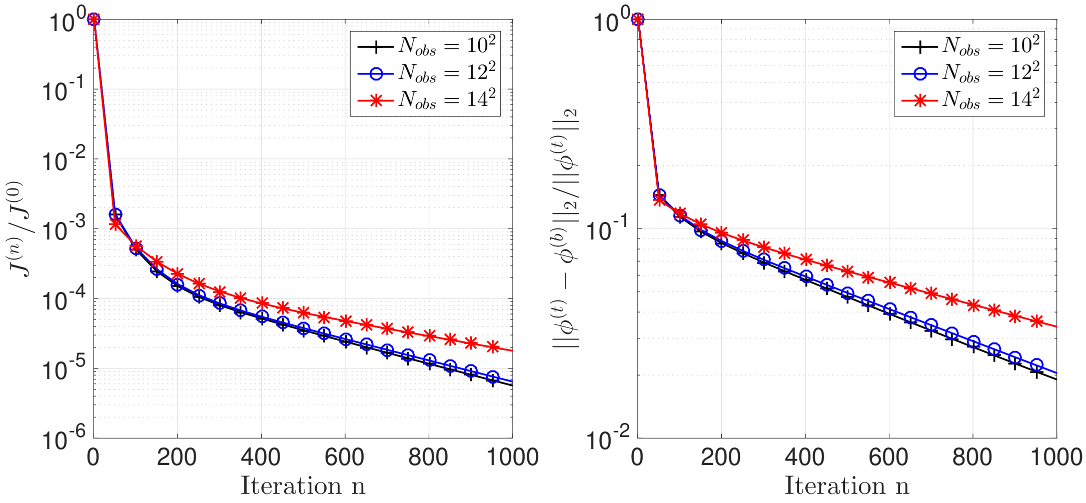
Arc configuration

We observed in figure 2.8 that unlike the characteristic or grid configurations, the results for the arc configuration failed to converge for all spacings Δr. Consequently it would be of significant interest to see whether these results can be improved by increasing N_{obs}. We present the configurations for $N_{obs} = 10^2, 12^2$, and 14^2 in figures 2.13(a), (b), and (c) respectively.

The relative L^2 error shown in 2.13(e) converges to $\mathcal{O}(10^{-2})$ for each values of N_{obs} and consequently all three cases are convergent. This result suggests that a relatively larger network of observations is required to fully resolve the initial conditions for arc configurations. A rigorous verification would be to analyse results with a larger values of N_{obs} and $\Delta r > 0.2$. However this is not possible with our current problem setup as a case where both N_{obs} and Δr (or $|m_i - m_j|$ with a single circular arc) are large enough such that that no points satisfy (2.4.1), would result in some observations not interacting with the free surface wave in the assimilation time $T = 2$. As the latter was chosen to prevent boundary effects from effecting the reconstruction, it cannot be altered without impacting the results.



(a) $N_{obs} = 10^2$ (b) $N_{obs} = 12^2$ (c) $N_{obs} = 14^2$



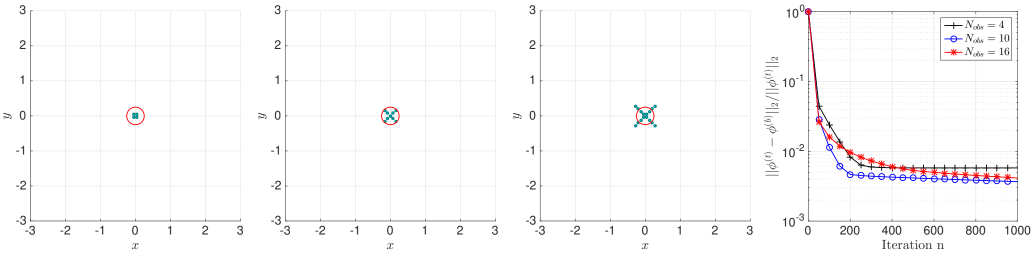
(d) Cost (e) Error

Figure 2.13: Arc formation with increased number of points and $\Delta r = 0.07$

Results with small N_{obs}

In this section we investigate the minimum number of observation points required for convergence. We have shown that when N_{obs} is large (greater than 60), increasing the number of points does not have a significant impact on the convergence of the relative minimum error for the characteristic and grid configurations. We now consider results when $4 \leq N_{obs} \leq 16$, to determine if there is a significant increase in the relative L^2 reconstruction error when N_{obs} is small. We do not consider $N_{obs} = 1$, 2, or 3 as the smallest number of points we can have while still having points positioned in each quadrant (to capture the azimuthal symmetry of the initial conditions) is 4.

Results for observations along a characteristic with $N_{obs} = 4$, 10, and 16 (with spacing $\Delta x = 0.0785$ are presented in figure 2.14.

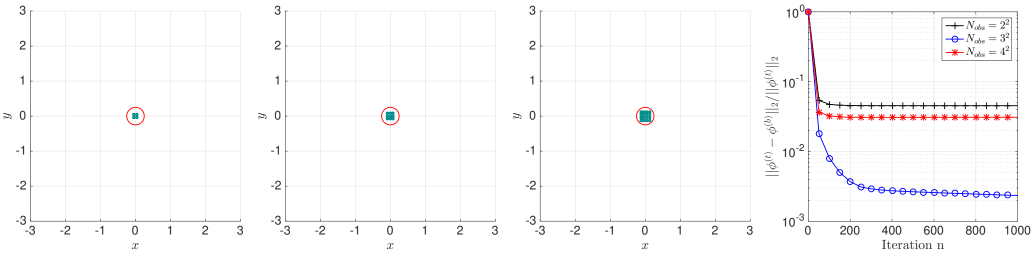


(a) $N_{obs} = 4$ (b) $N_{obs} = 10$ (c) $N_{obs} = 16$ (d) Cost

Figure 2.14: Characteristic formation with a small number of points and $\Delta x = 0.0785$.

We observe that the minimum relative L^2 error is less than 1% for each case. Therefore we conclude that for the characteristic configuration, the minimum number of observation points necessary to achieve convergence of an isotropic initial condition is $N_{obs} = 4$.

Results for observations in a grid configuration for $N_{obs} = 2^2$, 3^2, and 4^2 are presented in figure 2.15. We observe that the relative L^2 error for all three configurations converges. Configurations with 2^2 and 4^2 have a minimum relative L^2 error of 5% and 3% respectively. The best convergence is for $N_{obs} = 3^2$, with a minimum error of 0.2%, suggesting that convergence does not necessarily improve with an increase in N_{obs}, even with a small number of observations.



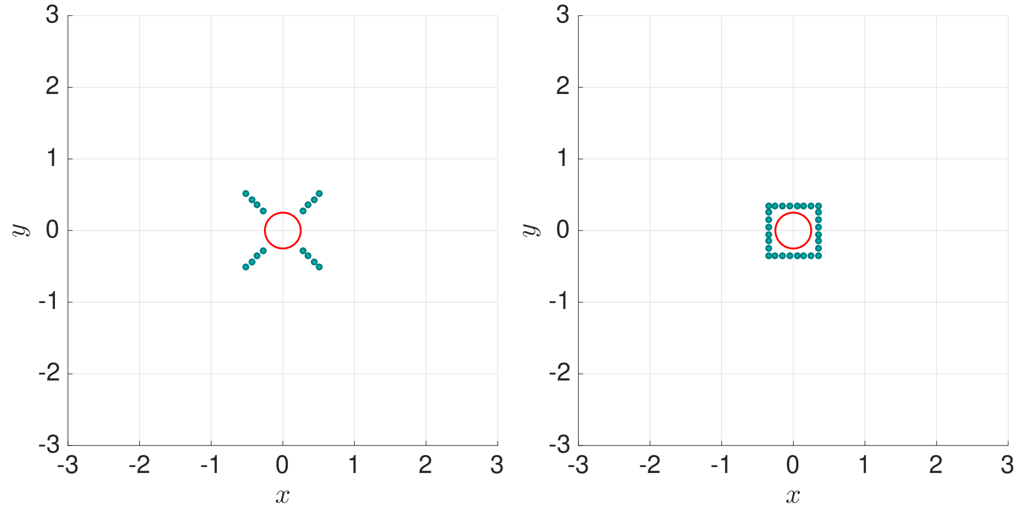
(a) $N_{obs} = 4$ (b) $N_{obs} = 9$ (c) $N_{obs} = 16$ (d) Cost

Figure 2.15: Characteristic formation with a small number of points and $\Delta x = 0.0785$.

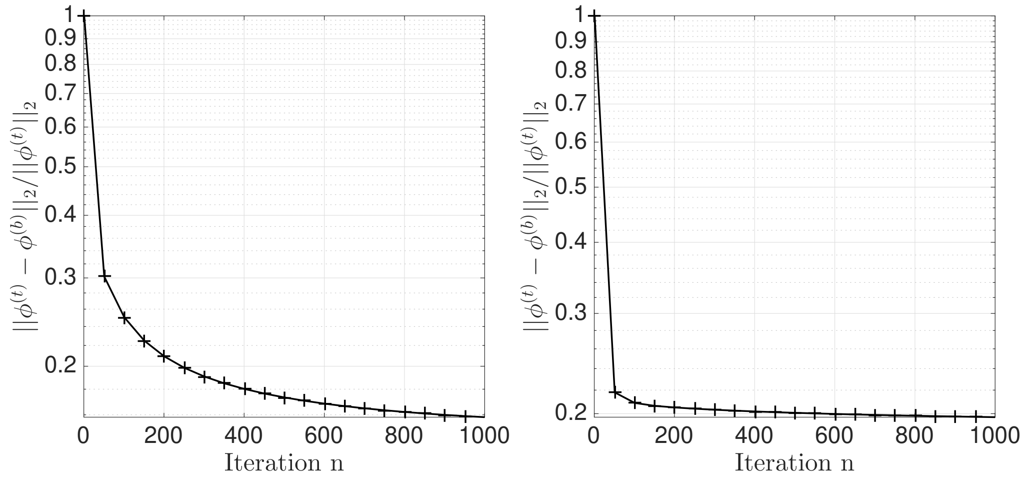
To summarise, we have demonstrated that convergence can be achieved for the characteristic and grid configurations with a relatively small number of points, and does not necessarily improve as N_{obs} is increased. And so we conclude that, based on the results so far, having a large number of observation points *is neither necessary nor sufficient* for convergence with observations in a grid or along two characteristics.

In the result observed thus far, we note that for all cases that converge (as in figures 2.11 and 2.10), there were multiple observations inside the support of the initial conditions. However, there were no points within the initial conditions support for the arc configuration. Consequently, we must consider whether convergence can be achieved for the characteristics and grid configurations when there are no observation points inside the support of the initial conditions. In tsunami models, it is unrealistic to assume such observations are available, and thus a logistically applicable forecasting scheme should avoid such assumptions. The results of this analysis are presented in section 2.4.3.

2.4.3 Results with no observations within the support of the initial conditions



(a) $N_{obs} = 16$ (b) $N_{obs} = 28$



(c) L^2 Error for characteristics (d) L^2 error for grid

FIGURE 2.16: Convergence of the relative L^2 error for observations in a (a) characteristic configuration, and (b) grid configuration, with small N_{obs} and $\Delta x = 0.07$, and points within the support of the initial conditions removed.

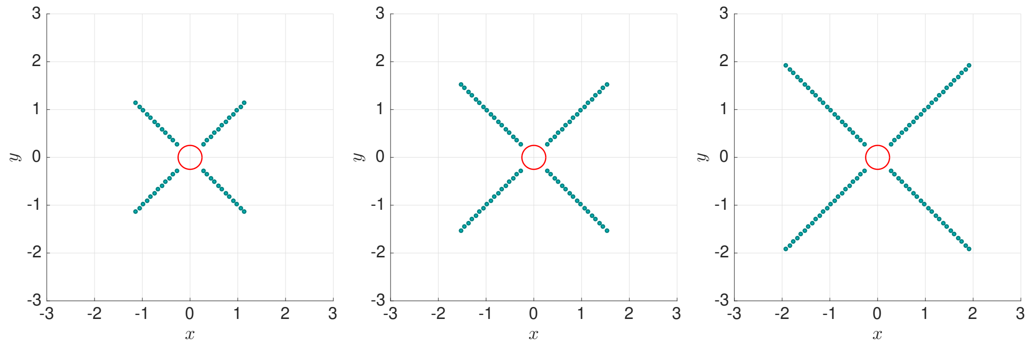
In this section we present results for the characteristic and grid configurations, with points inside the initial conditions support removed. The spacing Δx is fixed at the sufficient value for convergence of 0.07. As removal of these points results in smaller values of N_{obs}, the initial number of N_{obs} before removal is increased slightly so that the resulting number of observation points are comparable to the results in section 2.4.2. First, we consider results for small numbers of observation points, such that $\Delta x = 0.07$, and there are no points within the support of the initial condition, in figure 2.16.

We observe that even with $N_{obs} = 28$, the configuration in figure 2.16(b) (with observations in a grid) fails to converge when there are no points in the support of the initial conditions, despite having converged when $N_{obs} = 4$ in figure 2.15. Similarly, a configuration of 16 observations along the characteristics $x = \pm y$ fails to converge when there are no points within the support of the initial conditions, as in figure 2.16(a). Consequently, we investigate whether convergence can be achieved when N_{obs} is relatively large.

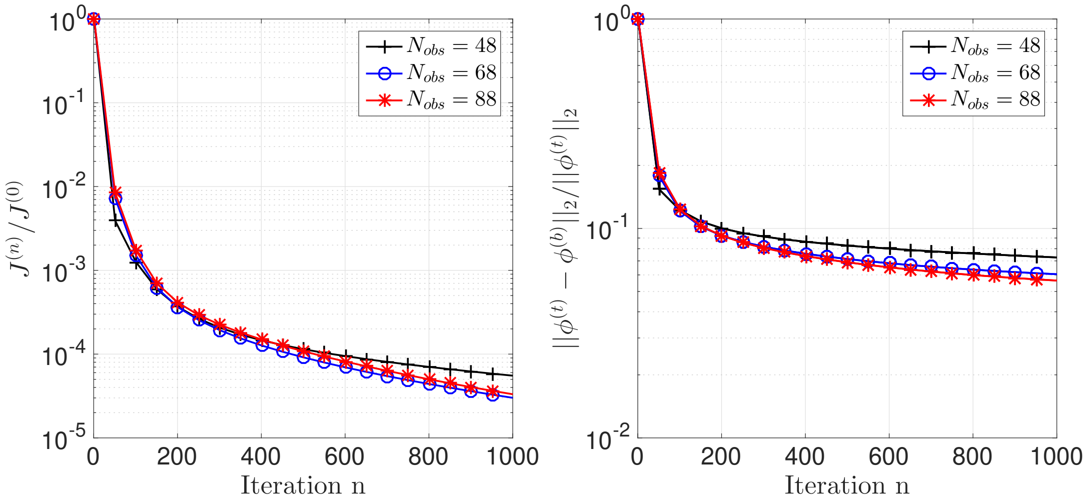
Figure 2.17 shows the configurations and results for observations along characteristics, with $N_{obs} = 48,\ 68$, and 88 (after removal). The relative L^2 error at the final iteration for each case is approximately 7%, as shown in 2.17(e). We note that the error is almost two orders of magnitude higher than it was in figure 2.10. In figure 2.18 we show a comparison of cross-sections of the optimally reconstructed initial conditions $\phi^{(b)}(x, 0)$, both with and without observations in the support of the initial conditions. We observe that there is some small-scale noise in figure 2.18(c) when there are no points within the initial condition support.

These results are not surprising, as having observations that are able to measure the true initial conditions may significantly improve the reconstructed initial conditions. It is therefore to be expected that convergence is slightly worse when such observations are no longer included in the assimilation. However based on the tolerance that the relative L^2 error be less than 10%, these results are still convergent. Indeed, we can see that while the reconstruction in figure 2.18(c) has some small-scale noise, the shape of the Gaussian is accurately resolved.

The configurations and result for observations in a grid format are presented in figure 2.19. We note some significant differences between the results here and those in figure 2.11. First, it is clear from figure 2.19(e) that there is improved



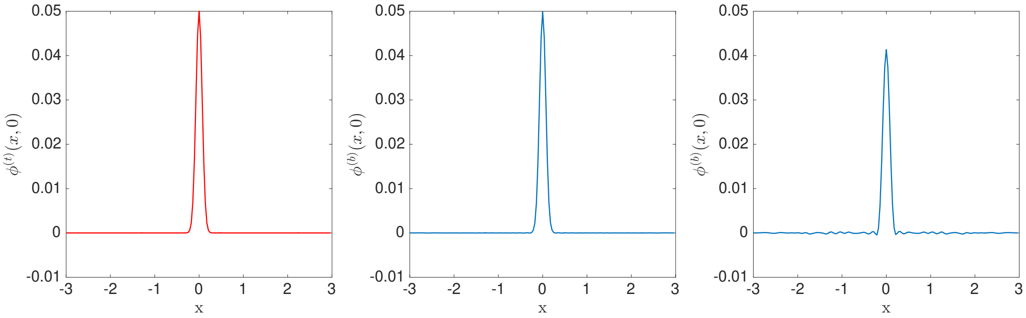
(a) $N_{obs} = 48$ (b) $N_{obs} = 68$ (c) $N_{obs} = 88$



(d) Cost (e) Error

FIGURE 2.17: Characteristic configuration with points within the support of the initial conditions removed.

convergence for cases with $N_{obs} > 80$, suggesting that the number of observations is an important factor when we do not have observations in the support of the initial conditions. In contrast, the results in figure 2.11(e) did not indicate difference in convergence for the different values of N_{obs}. Secondly, the best convergence in figure 2.19(e) is a relative L^2 error of approximately 0.1%, indicating better convergence than observed in figure 2.11(e), despite the overall number of observations being comparable. It is interesting to note that in figure 2.11(e) the error stagnated after approximately 400 iterations. This suggests that perhaps

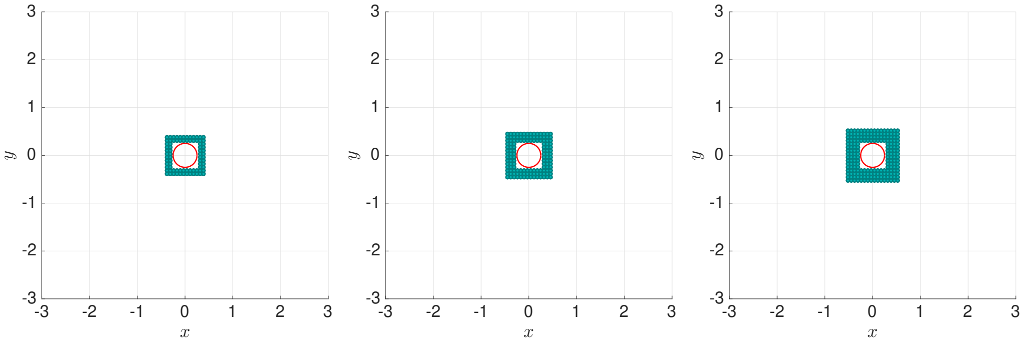


(a) Exact initial conditions (b) $N_{obs} = 80$ (c) $N_{obs} = 88$

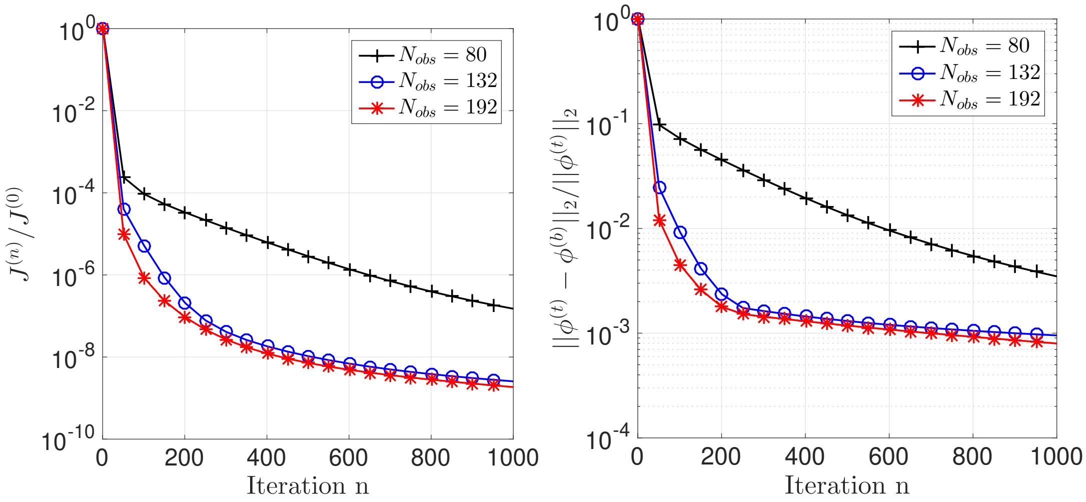
FIGURE 2.18: Cross section of the (a) exact initial conditions $\phi^{(t)}(x,0)$, and reconstructed initial condition $\phi^{(b)}(x,0)$ for observations along characteristics with (b) observations within the support of the initial condition, and (c) no observations within the support of the initial conditions.

there is some overfitting that results in small-scale noise in the reconstructed initial conditions, when there are too many observations of the true initial conditions. This would impact the convergence of the gradient descent algorithm used in the data assimilation. In comparison, in figure 2.19(e) we observe that the error is still maintaining a negative slope even at the final iteration, suggesting improved convergence. To verify this hypothesis, figure 2.20 shows the energy spectra of the relative absolute error of the reconstructed initial conditions for configurations 2.11(b) and 2.19(b) with $N_{obs} = 144$ and $N_{obs} = 132$ respectively .

We can see that the energy of larger wavenumbers in the reconstruction error $|\phi^{(t)} - \phi^{(b)}|/ \parallel \phi^{(t)} \parallel_{L^2}$ is higher for configuration 2.11(b), verifying our initial hypothesis regarding small-scale noise present in the reconstruction that impacted convergence. On the other hand, the energy of the error spectrum for configuration 2.19(b) is mostly restricted to smaller wave-numbers. This may explain why the latter showed convergence of the L^2 error to 0.1% whereas the former converged to 0.3% .

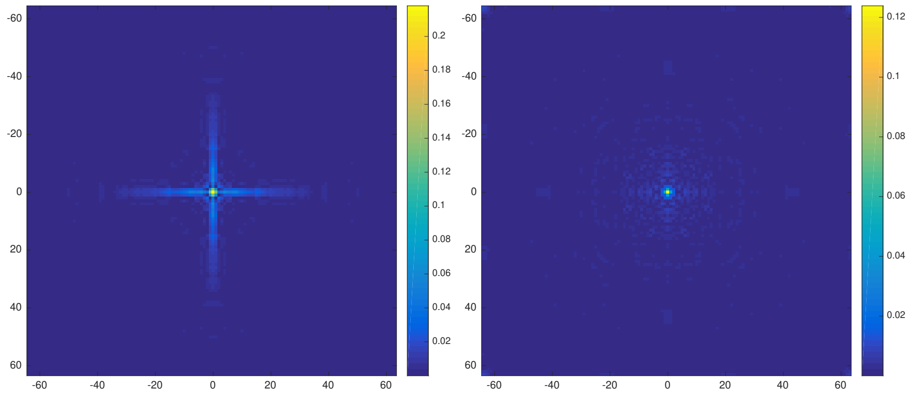


(a) $N_{obs} = 80$ (b) $N_{obs} = 132$ (c) $N_{obs} = 192$



(d) Cost (e) Error

Figure 2.19: Grid configuration with points within the support of the initial conditions removed.



(a) 2.11(b) (b) 2.19(b)

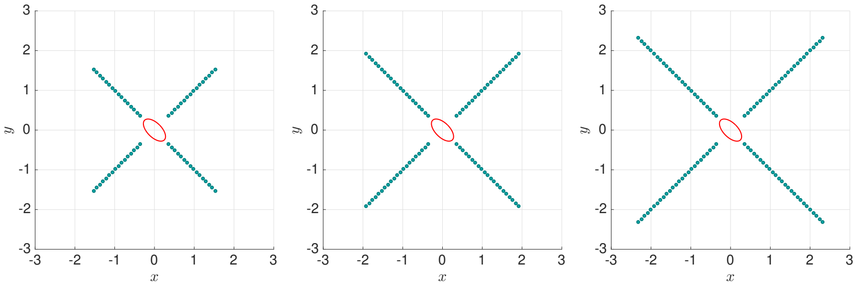
Figure 2.20: Energy spectrum of the error $|\phi^{(t)} - \phi^{(b)}| / \parallel \phi^{(t)} \parallel_{L^2}$ for the grid configurations in 2.11(b) (points inside the support of $\phi^{(b)}$))and 2.19(b) (no points inside the support of $\phi^{(b)}$)).

In summary, we conclude that a necessary (but not sufficient for all configurations) condition for convergence is that Δx be small enough that the Euclidean distance between pairwise observations satisfies the condition (2.4.1). Additionally, the convergent configurations for the grid configuration in figure 2.7 had 64 pairwise observations satisfying the spacing condition (2.4.1), while the non-convergent cases with spacing $\Delta x \geq 0.1$ had none. Similarly, the non-convergent arc configurations in figure 2.8 had at most 14 pairwise observations with a Euclidean distance less than 0.0785. This suggests that there may be a minimum number of pairwise observations required that satisfy the spacing condition 2.4.1. We observed that much larger numbers of observations were required for convergence in the arc configuration, and this was also true for the grid configuration when observations inside the support of the initial conditions were removed. In fact, we found that the assimilation error was lower in the grid configuration when there were no observations within the support of the initial conditions. This is likely due to overfitting generating small-scale noise in the reconstructed initial conditions, as shown in figure 2.20.

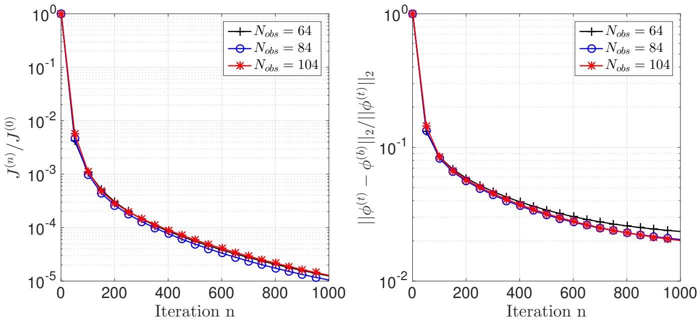
To extend these results for more realistic conditions, we conducted equivalent analyses with the anisotropic initial conditions given in (2.3.10).

2.4.4 Results for anistropic initial conditions

We have observed convergence for each of the observation point configurations considered in this study (characteristic, grid, and arc) by varying the spacing between observations and the number of observation points. We have also demonstrated convergence even when there are no observations within the support of the initial conditions. We now verify whether convergence is still achieved with anisotropic initial conditions. Our goal is to investigate the observability at measurement points when the surface wave is no longer azimuthally symmetric, and constrast with results for isotropic Gaussian initial conditions. Our goal is to provide results that are more comparable with realistic scenarios for tsunami observations, where the both the initial conditions and bathymetry produces highly anisotropic surface waves.



(a) $N_{obs} = 60$ (b) $N_{obs} = 84$ (c) $N_{obs} = 104$



(d) Cost (e) Error

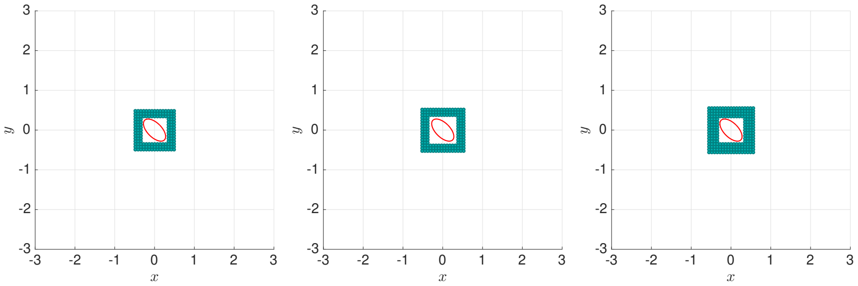
Figure 2.21: Characteristic formation with anisotropic initial conditions with points within the support of the initial conditions removed.

Briefly, we found that there was no significant impact on convergence with anisotropic initial conditions (as given in (2.3.10)), compared to results in section 2.4.3. The results for a characteristic configuration are presented in figure 2.21, the results for a grid format in figure 2.22, and the results for an arc configuration in figure 2.23. We observe that the convergence of the relative L^2 error for the characteristic and arc configurations is $\mathcal{O}(10^{-2})$, and there is little difference between results with different values of N_{obs}. The convergence for the grid configuration is $\mathcal{O}(10^{-4})$ with $N_{obs} = 208$, and $\mathcal{O}(10^{-3})$ for the other values of N_{obs}.

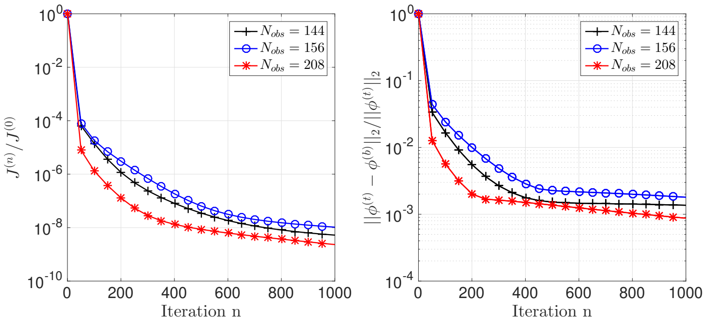
Summary of main results

To summarise, the main results we have observed are as follows:

- Convergence was achieved (a relative L^2 error less than 10%) for each of the three configuration with both isotropic and anisotropic initial conditions



(a) $N_{obs} = 144$ (b) $N_{obs} = 156$ (c) $N_{obs} = 208$



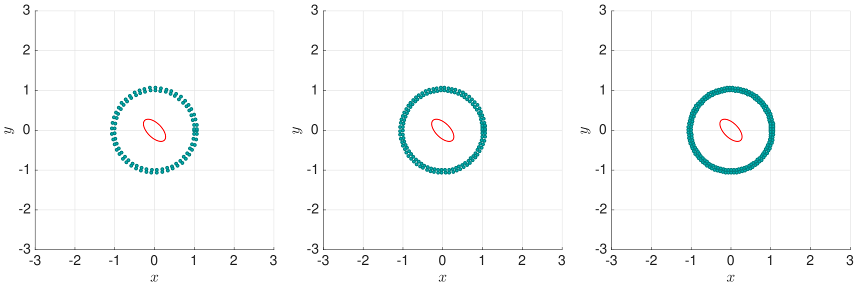
(d) Cost (e) Error

FIGURE 2.22: Grid formation with anisotropic initial conditions with points within the support of the initial conditions removed.

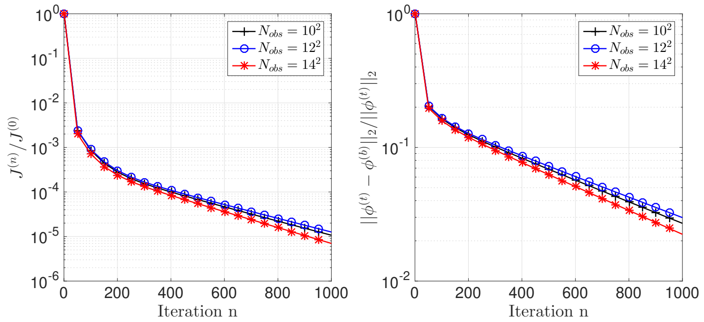
respectively.

- The best convergence in the present analysis was achieved with observations placed in a grid formation. Even with no points inside the initial conditions support, the relative L^2 error in the reconstruction error was 0.1% with both isotropic and anisotropic initial conditions. Convergence was achieved for the characteristic and grid configurations with as few as 4 observation points, the minimum number required to capture the azimuthal symmetry of the initial conditions.

- A necessary condition for convergence was that the spacing Δx for grid and characteristic configurations be chosen such that at least some pairwise observations satisfy (2.4.1) . However convergence was not achieved for the arc configuration even with small Δr (the difference between radii of observations placed in two concentric circles), suggesting this is not a sufficient condition



(a) $N_{obs} = 10^2$ (b) $N_{obs} = 12^2$ (c) $N_{obs} = 14^2$



(d) Cost (e) Error

Figure 2.23: Arc formation with anisotropic initial conditions with points inse the support of the initial conditions removed and $\Delta r = 0.07$.

for convergence, or there is a minimum number of pairwise observations with a Euclidean distance satisfying (2.4.1), which the arc configuration did not satisfy (having only 14 such points, whereas the convergent results for the grid configurations in figure 2.7 had 64 such pairwise observations).

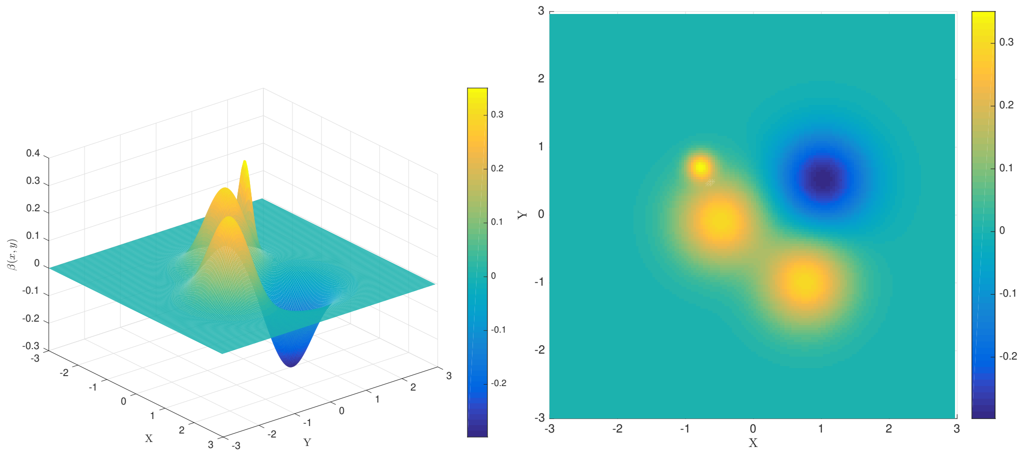
- Increasing the number of observation points improved convergence of the arc configuration, but had little effect on the grid and characteristic configurations when there were points placed inside the initial conditions support. However, without any observations in the support of the initial conditions, larger numbers of observations N_{obs} resulted in improved convergence with the grid configuration, and even surpassed convergence achieved in the former case (points inside initial conditions support). This suggests that observations of the true initial conditions can lead to overfitting and small-scale noise in the reconstruction, as demonstrated in figure 2.20.

2.5 Results with non-zero bathymetry

We now consider results for simulations where the bathymetry is no longer flat. Bathymetry can have a significant impact on propagating shallow water waves, where wave speed $c = \sqrt{gh}$ changes as depth h varies. A tsunami's energy flux remains relatively constant, and so as the tsunami's speed varies, so does wave height (shoaling effects).

The purpose of this section is to take the best results observed thus far, and integrate them into a relatively more realistic model which takes us closer to conditions necessary for an active tsunami forecast model. As such, we limit our results to $N_{obs} = 88$. Even though we observed higher convergence with a greater number of N_{obs} (for example in figure 2.19), there are logistical difficulties in creating dense observation networks. Therefore we impose an upper limit. The current number of observation buoys in the DART network is approximately 60, therefore accounting for additional observation sources and methods (as described in section 2.1, we maintain that $N_{obs} = 88$ is a reasonable choice. Additionally, we note that convergence of the relative L^2 error to $\mathcal{O}(10^{-2})$ in the initial conditions reconstruction was achieved with ≤ 88 points in each of the three configurations considered in this study.

We begin with a simple Gaussian bathymetry model simulating peaks and basins, and then subsequently extend our analysis to a subsample of ETOPO2, a digital database of land elevation and sea floor topography, where the datum represents the vertical deviation in metres from the mean sea level.



(a) (b)

FIGURE 2.24: Gaussian Bathymetry with 3 peaks and a basin as described in (2.5.1). The y-axis represents the amplitude of the bathymetry relative to the mean depth H, which has been normalised to $H = 1$.

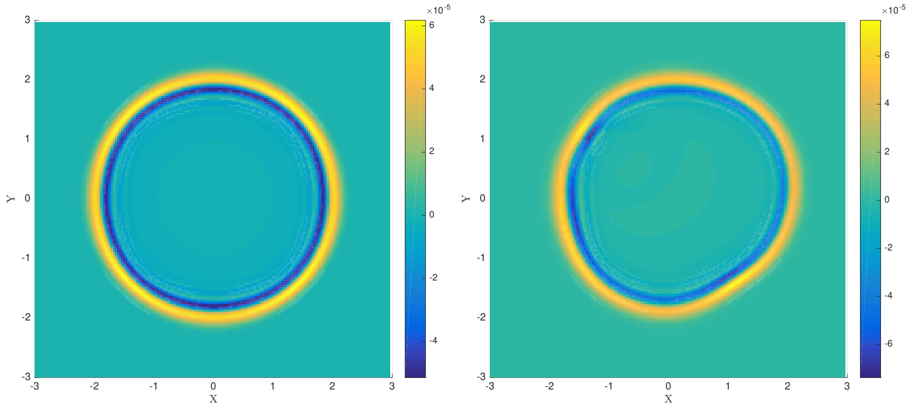
2.5.1 Gaussian bathymetry

In this section we analyse results with the Gaussian bathymetry described by

$$\begin{aligned}\beta(x,y) = \frac{3}{10}\exp\left[-\frac{\Big((x+0.5)^2+(y+0.1)^2\Big)}{0.65^2}\right] - \frac{3}{10}\exp\left[-\frac{\Big((x-1)^2+(y-0.5)^2\Big)}{0.7^2}\right] \\ + \frac{3}{10}\exp\left[-\frac{\Big((x-0.75)^2+(y+1)^2\Big)}{0.65^2}\right] + \frac{3}{10}\exp\left[-\frac{\Big((x+0.8)^2+(y-0.7)^2\Big)}{0.2^2}\right]\end{aligned} \tag{2.5.1}$$

where a representation is given in figure 2.24. This shape was chosen as it contains Gaussian peaks that are 30% of the average depth H and with varying widths, simulating underwater mountains. Additionally, a negative Gaussian represents a basin, and all features are placed such that there are relatively quick changes in the depth, within the assimilation time T. The amplitude of the true initial conditions is set to $0.001H$, in order to accurately simulate tsunami conditions where amplitude of the surface wave in the deep ocean can be just a few metres (depending on the generating mechanism).

We demonstrate the effect of the bathymetry on the free surface propagation in figure 2.25, where figure 2.25(a) is the free surface at $t = T$ with a flat bottom bathymetry, and figure 2.25(b) is the free surface at $t = T$ with the bathymetry given in (2.5.1).

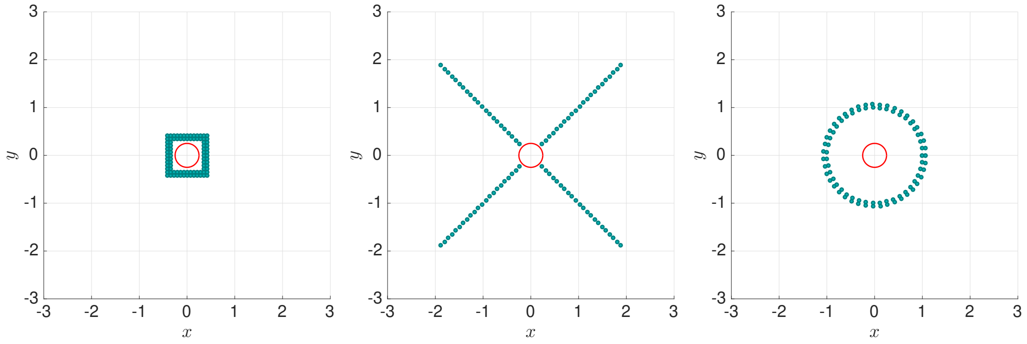


(a) 2.11(a)

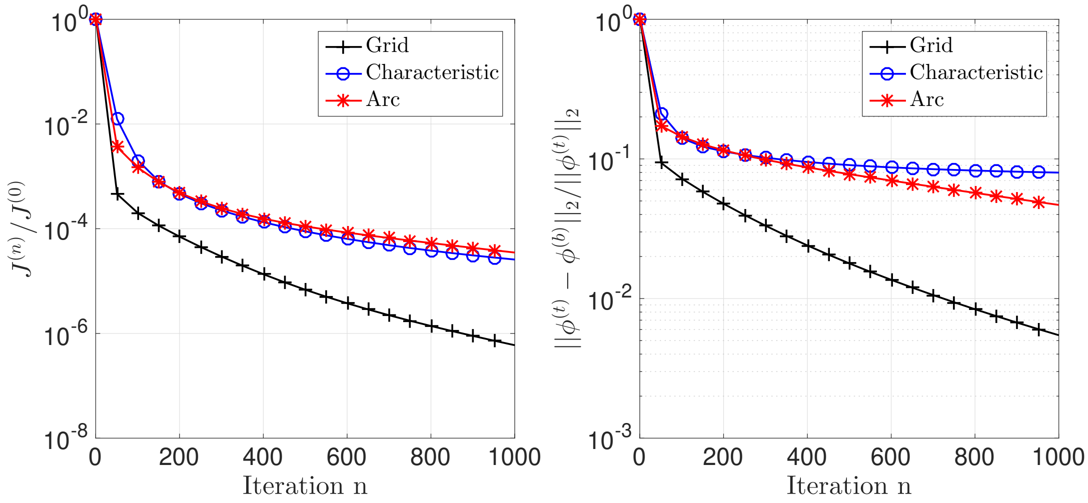
(b) 2.19(a)

FIGURE 2.25: Planar view of the free surface wave without bathymetry (a), and with bathymetry given in (2.5.1) (b), at the final time T, showing the effect of the bathymetry in distorting the surface wave.

We observe in figure 2.25(b) that due to bathymetry effects we no longer have azimuthal symmetry in the free surface wave propagation, and the wave amplitude is higher at positions close to $(-1, 1)$ and $(-\frac{3}{2}, -\frac{3}{2})$. This suggests the influence of the peaks in 2.24. Similarly the amplitude of the wave is slightly damped in the first quadrant, indicating the effect of the basin.



(a) $N_{obs} = 88$ (b) $N_{obs} = 88$ (c) $N_{obs} = 88$



(d) Cost (e) Error

FIGURE 2.26: Results with Gaussian Bathymetry. Note that the configuration in (b) can no longer be referred to as observations placed along characteristics, as these are no longer characteristics of the distorted wave.

The results with isotropic initial conditions (and no points inside the initial conditions support) are show in figure 2.26. We note that observations placed along the lines $x = \pm y$ can no longer be referred to as observations placed along characteristics, as these are no longer characteristics of the distorted wave, due to bathymetry effects. However, we include the configuration to demonstrate convergence when observations are placed along lines.

In each case we have $N_{obs} = 88$ and $\Delta x = \Delta r = 0.07$. We observe that the grid configuration has the best convergence of the relative L^2 error, at 0.5%. The results for the line and arc configuration also converge, however with a larger error of 5%. There is a relative decrease in convergence compared to results for with a flat bottom bathymetry as shown in section 2.4.3. This is most likely due to the fact that there is no longer azimuthal symmetry in the surface wave propagation, and not all observation points are able to capture the interaction of the free surface wave with localised bathymetry. Despite this, we note that results for each case converge. This is encouraging, considering our results included effects of localised bathymetry features with amplitudes as much as 30% of the fluid depth. Another possible reason for convergence is that the amplitude of the initial conditions is relatively small, which would mitigate wave breaking effects that could hinder convergence.

We now investigate whether similar results can be observed when the bathymetry is an actual representation of ocean floor topography in section 2.5.2.

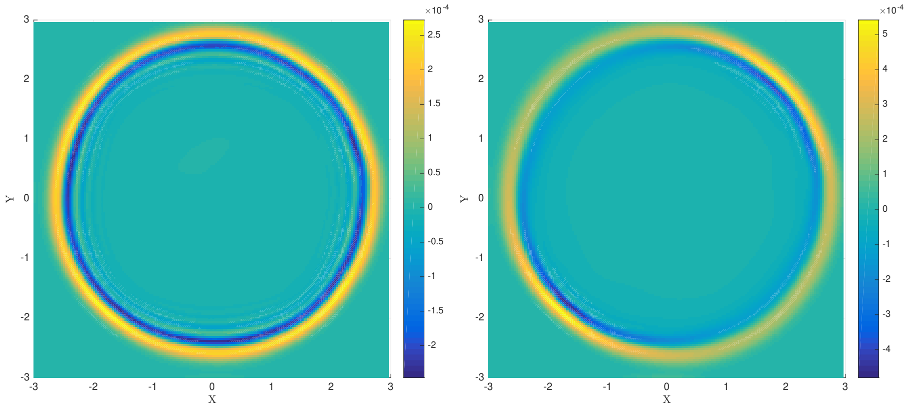
2.5.2 ETOPO2 bathymetry

ETOPO2 is a database of two-minute global relief data hosted by the National Geophysical Data Centre at NOAA (2006). This is an amalgamation of data collected via multiple sources including satellite altimetry observations and shipboard echo-sounding measurements [13]. The grid values represent elevation at the cell edges, averaged over the cell area. The horizontal grid spacing is 2-minutes of latitude and longitude, where 1 minute of latitude represents 1.853 km at the Equator, and the vertical precision is 1 metre, where $z = 0$ represents the mean sea level.

For the current analysis, a square sub-interval of the ETOPO2V2c database was chosen within the specified latitude and longitude limits $[15,\ 25]$ and $[90,\ 100]$ respectively. This is equavalent to a rectangular grid of approximately 1100km ×1100km (accounting for slight differences between degrees of longitude due to the equatorial bulge, where the maximum variation is approximately 60km). This bathymetry section represents the Wharton basin, a topographical feature in the ocean floor located off the western coast of Australia. It was chosen because it includes both basin and ridge features, and because of the relative ease with which periodic boundary conditions could be implemented. Additionally, it has been a documented source for seismic activity and strike-slip events such as the 2012 and 2016 events in the region. As most earthquakes are of strike-slip mechanisms, such events are potential sources of tsunamis in the Wharton basin [23]. A visual representation of the region and a relief map as taken directly from the ETOPO2v2c database are given in figure 2.28(a) and 2.28(b).

In order to find classical solutions of the shallow water equations, we require the bathymetry $\beta(x, y)$ and its first derivative to be smoother than the raw ETOPO2 data. Additionally, we require periodic boundaries. Thus the boundaries are artificially padded with zeros, and the low pass filter (2.3.14) is used to damp all frequencies after the lowest 5% modes to zero. A moving average filter was implemented to remove sharp curves at the boundaries, using the Matlab `smooth` function. The control time T was adjusted so that the free surface wave does not actually reach the padded boundary region.

The average depth $H = -5000$m was normalised to 1 and the length scales adjusted to $[-L,\ L]$ in both x and y directions. The amplitude of the initial conditions was set to $0.001H$ (equivalent to 5m). The smoothed and scaled relief map, and 3-D plot of the bathymetry can be seen in figure 2.28(c) and 2.28(d).



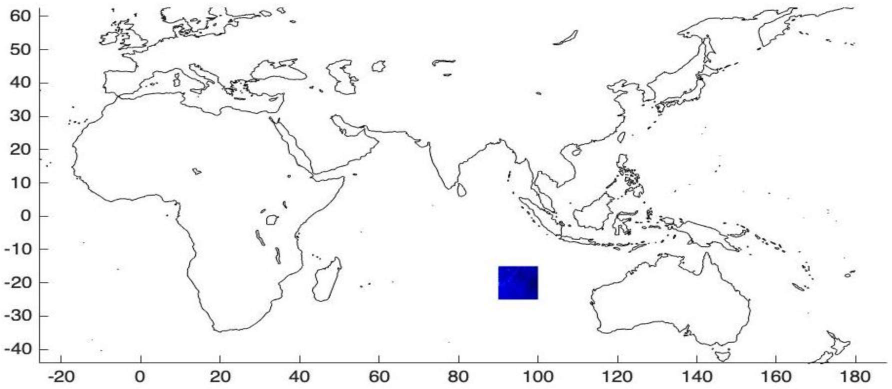
(a) Isotropic initial conditions (2.3.9) (b) Anisotropic initial conditions (2.3.10)

Figure 2.27: Planar view of the free surface wave with ETOPO2 bathymetry and (a) Isotropic initial conditions (2.3.9), and (b) Anisotropic initial conditions (2.5.1), at the final time T.

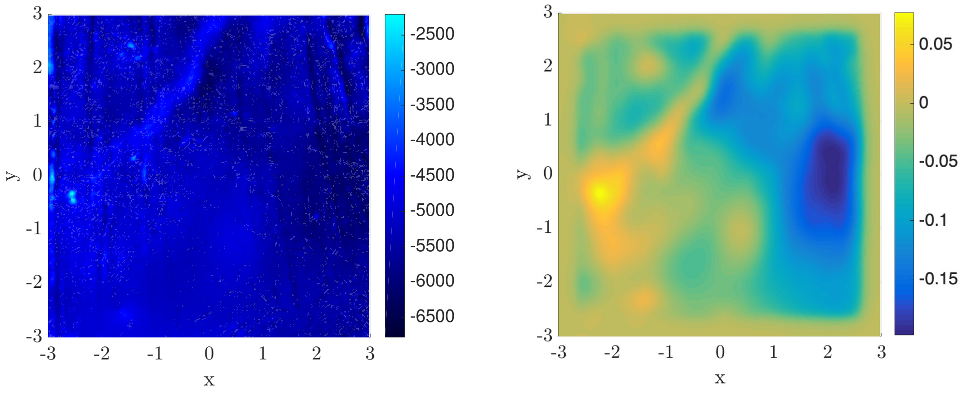
We implement the data assimilation for observations along the lines $x = \pm y$, a grid, and an arc, with isotropic initial conditions and anisotropic initial conditions respectively. Figures 2.27(a) and 2.27(b) show the free surface wave at the final time T. The effect of the bathymetry on the free surface wave is relatively small compared to the flat bottom case in figure 2.25(a), than for the Gaussian bathymetry. However, we note that the highest amplitude in the ETOPO2 bathymetry (as shown in figure 2.28 is only approximately 0.06, which is significantly less than the bathymetry in (2.5.1). Nevertheless, we observe slightly deeper troughs in the free surface wave indicated by the dark blue regions in 2.27(a), which are indicative of bathymetry effects. We conclude that the wave speed affected by the bathymetry structure may result in different coastal communities having varied arrival time and wave energy.

The observation configurations and results with isotropic initial conditions and anisotropic initial conditions are presents in figures 2.29 and 2.30 respectively.

For the isotropic initial conditions, we see in figure 2.29(e) that the lowest relative error is for the grid configuration, with a minimum error of 0.3%. The straight line configuration has minimum a relative error of 5% and thus is convergent. However, as in the idealised case with Gaussian bathymetry, it is again the worst performing configuration. The results for the arc configuration are a bit better than those observed in figure 2.26 for the Gaussian bathymetry, with a minimum error of 5%. Subsequently we conclude that the main results of this study extend to realistic bathymetry, such as the smoothed topography of the Wharton basin.

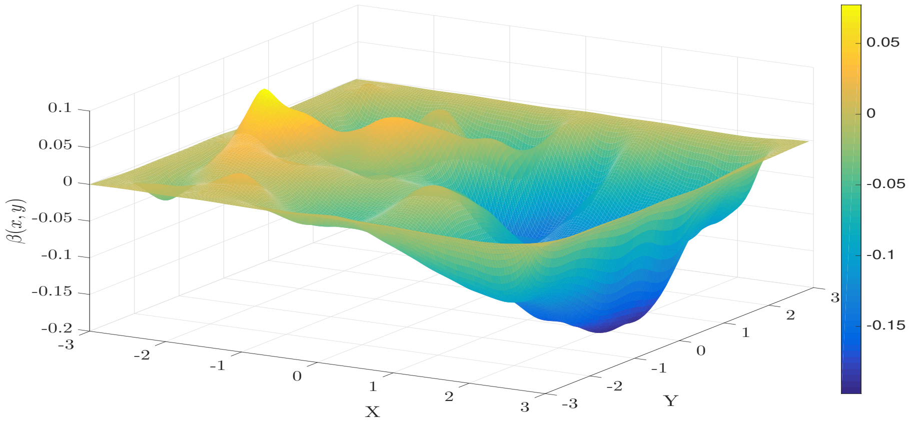


(a) The selected bathymetry subsample on a world map.



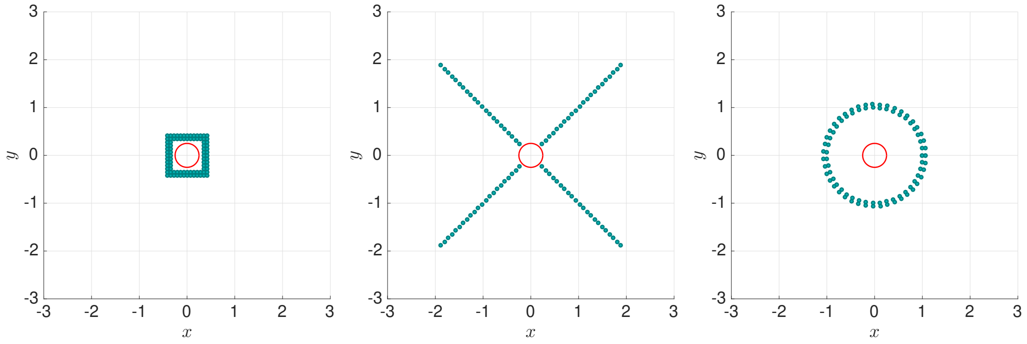
(b) ETOPO2V2c relief of the Wharton Basin.

(c) Smoothed version of (b).

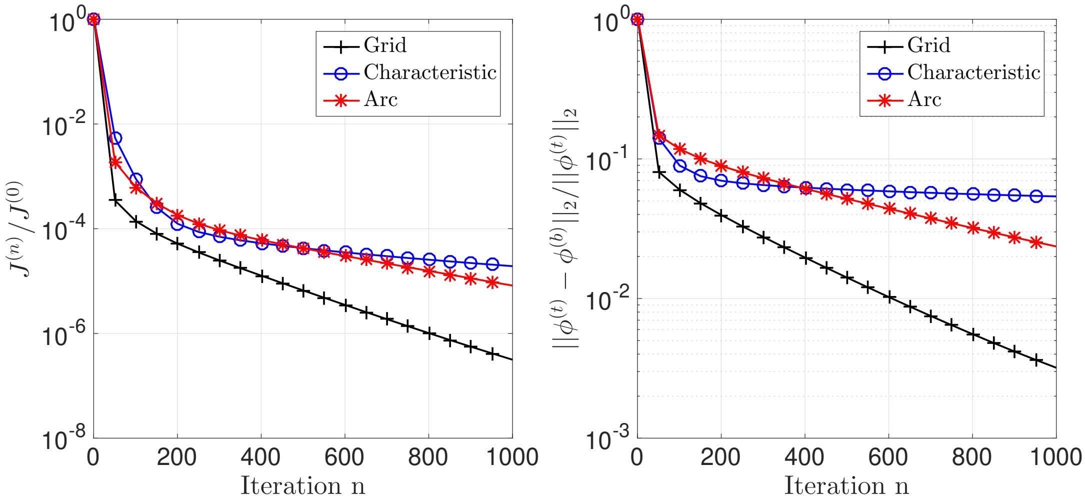


(d) 3-D version of (c).

Figure 2.28: Smoothed Etopo2v2c bathymetry relief within [15, 25] and [90, 100].



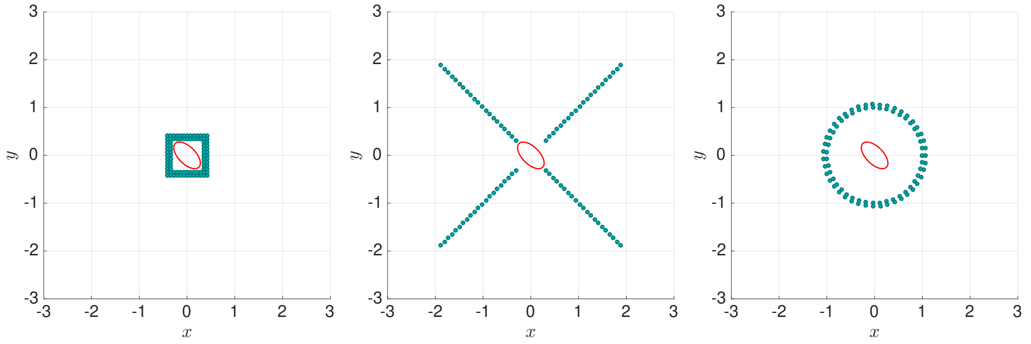
(a) $N_{obs} = 88$ (b) $N_{obs} = 88$ (c) $N_{obs} = 88$



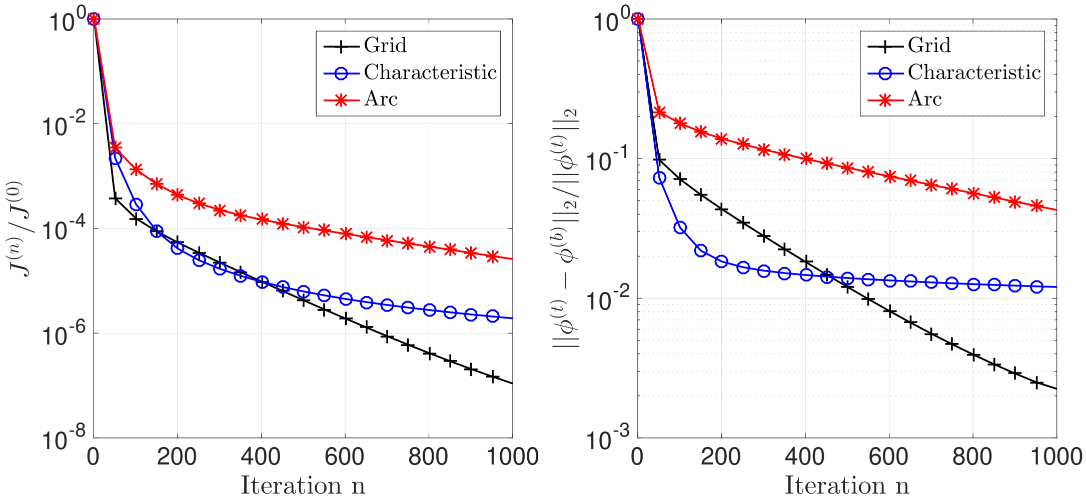
(d) Cost (e) Error

FIGURE 2.29: Results with ETOPO2 and isotropic initial conditions. We observe better convergence with an arc configuration of observations than for a straight-line configuration, despite the latter being in closer proximity to the initial condition region.

We consider analogous results with a anisotropic initial conditions in figure 2.30. The main trends in the relative L^2 error shown in 2.30(e) are similar to 2.29(e) in that the grid formation has the lowest minimum relative error, 0.2%. However, in this case the straight line configuration performs better than observations along the arc, although both are convergent. This is not surprising as we have already observed in section 2.4.4 that the arc and straight line configurations show relatively worse convergence with a radially asymmetric surface wave and bathymetry



(a) $N_{obs} = 88$ (b) $N_{obs} = 88$ (c) $N_{obs} = 88$



(d) Cost (e) Error

Figure 2.30: Results with ETOPO2 and anisotropic initial conditions.

features. The main difference across the two cases 2.29 and 2.30 is the relatively better performance of the straight line formation in 2.30(e). Nevertheless, overall convergence for both isotropic and anisotropic initial conditions is relatively equivalent.

2.6 Conclusion and further considerations

In this study, we implemented a data assimilation scheme for the 2-D SWE to reconstruct the initial conditions, using observations of free surface wave height.

Our objective was to find the optimal configuration of observations such that reconstruction error is minimised. We compared observations placed along straight lines (i.e. characteristics of the surface wave with isotropic initial conditions and flat bathymetry), in a grid array, and along concentric arcs. We compared results where observations satisfied the minimum distance criterion between pairwise observations (2.4.1) from the 1-D results in [27], and when they did not. We analysed the effect of a small or large number of observations, and the effect on convergence when observations inside the initial conditions support were removed. The algorithm was implemented for both isotropic and anisotropic true initial conditions, both with both flat and non-flat bathymetry. Additionally, computational considerations such as the low-pass filter (2.3.14) were utilised to investigate whether they improved convergence.

Summarising the main results, we observed that a necessary (but not sufficient) condition for convergence was that some of the pairwise observations are sufficiently closely spaced to observe the minimum lengthscales of the initial condition, i.e. satisfy equation (2.4.1). While this confirms results from the 1-D case in part, the main difference is that for the 1-D assimilation, the spacing condition (2.4.1) was sufficient for convergence. We have demonstrated that this is not the case (as for results in figure 2.8) in the 2-D assimilation, and we may require a minimum number of points satisfying (2.4.1) in order to achieve convergence. This could be due to the need to capture the shape of true initial conditions with higher dimensionality. We recall that equation (2.4.1) was not analytically derived for the 2-D case, and therefore is not expected to account for its complexities. Nonetheless, it is a useful reference tool for the qualitative analysis in this study.

For the arc configuration, we observed that a higher number of observations was necessary for convergence, whereas convergence was achieved for the straight line and grid configurations with as few as 4 observation points. However, small N_{obs} was insufficient for convergence when there were no observations within the support of the initial conditions. Increasing the number of observation points improved convergence for the grid and characteristic configurations when observation points were removed from the support of the initial conditions. The grid configuration showed the best convergence of the relative L^2 error in the reconstructed initial conditions. We concluded that observations are not required within the support of the IC, and that this can actually degrade the results due to over-fitting. The error was $\mathcal{O}(10^{-4})$ with a flat bottom bathymetry, and $\mathcal{O}(10^{-3})$ with a Gaussian bathymetry, and bathymetry from the ETOPO2v2 database. Convergence was slightly worse for the characteristic configuration with anistropic initial conditions and when bathymetry features were included, but overall all configurations showed convergence (a relative error in the reconstructed initial conditions less than 10%) with 88 observations for both isotropic and anisotropic initial conditions, and with bathymetry. The main results for the analyses conducted in this study are summarised in table 2.2.

Analysis	**Lines** $\boldsymbol{x = \pm y}$	**Grid**	**Arc**
$N_{obs} = 36$, minimum spacing $\Delta x \leq 0.1$	0.3%	0.1%	18%
Large N_{obs}, $\Delta x = 0.07$	0.2%	0.3%	2%
Small N_{obs}, $\Delta x = 0.07$	0.4%	0.3%	None
No observations within support of $\phi^{(t)}$, small N_{obs}, $\Delta x = 0.07$	18%	20%	None
No observations within support of $\phi^{(t)}$, large N_{obs}, $\Delta x = 0.07$	7%	0.1%	2%

Table 2.2: Summary of the minimum relative L^2 error for the different analyses conducted in this study for observations along (i) straight lines, (ii) in a grid array, and (iii) along arcs. Entries highlighted in red indicate non-convergent results. Convergence was achieved for all configurations with no points within the support of the true initial conditions $\phi^{(t)}$, with large N_{obs} and $\Delta x = 0.07$.

We now consider the feasibility of integrating these findings with a realistic tsunami model. The primary questions are:

1. Is it realistic to assume we have a sufficient number of observations for convergence, for a grid array of observations centred on the initial conditions?
2. Is it logistically possible to have a sufficient number of observations of surface wave heights for all possible initial conditions (i.e. associated with known seismically active regions)?

Addressing the former question, we note that the most destructive tunamis are those generated by shallow earthquakes, with epicentres along fault lines. Of these, tectonic subduction at the plate boundaries are the most likely causes of tsunamis, and subsequently their position coincides with the support of the initial conditions. Schellart et al. [51] provide a comprehensive overview of global subduction sites as well as the velocities of the respective plates. Concentrating observations around subduction zones may sufficiently capture the propagation of waves triggered by seismic events originating here. However, this is not necessarily exhaustive; As we have indicated, the Wharton basin (which does not lie on a subduction zone) is a potential zone for small tsunamis due to strike-slip earthquakes [23], and there are many such regions. However, the most destructive tsunamis in recent history such as the 2004 Indian ocean tsunami, have been due to earthquakes with epicentres falling on subduction zones, and hence merit primary focus.

The scope of positioning observations at all possible such locations, brings us to the second question, concerning the logistical feasibility of such configurations. Requiring observation configurations around subduction zones with the highest probability of tsunamis occurring requires a large number of observations. In the present analysis we have demonstrated convergence with $N_{obs} = 88$ placed around a single source of potential tsunamis, and in reality there are several such regions. However as we previously noted, innovative methods have already been introduced to generate observations even when existing networks are relatively sparse. For example, introducing virtual observation data interpolated from neighbouring real observations as demonstrated by Wang et al. [62]. These methodologies can be implemented to pad the existing observational network, such that requirements for N_{obs} are satisfied. Additionally, there have been significant advancements in altimetry observations taken from geospatial satellite data. Since the Geodynamics Experimental Ocean Satellite 3 (GEOS-3) satellite was first launched in 1975, remote sensing of ocean wave height from space was predicted to provide a quantum increase in forecasting capabilities [5]. Recent projects such as the ESA's GLOBWAVE project (2010-2013) have expanded upon this, by providing access to consolidated datasets on ocean waves to the scientific community for the purpose of modelling and forecasting. In 2017 the Copernicus Marine Environment Monitoring Service (CMEMS) released the first real-time global wave product, containing wave height data collected via multiple satellites, available within three hours of acquisition. Additionally, these observations have been collected and made accessible for forecasting models, as in Ribal and Young [46]. With a global spatial resolution of 7km × 7km, the observation network is very similar to the optimal grid configuration observed in this study, and thus has the potential to generate similar results.

While tsunami waves are characteristic of relatively small amplitudes, Smith et al. [54] demonstrated that it is still possible to measure tsunamis using altimetry data, and restrictions for forecasting in real time were due to the delay in access to data. With innovated products such as the CMEMS data-sets accessible in under three hours, these restrictions can be mitigated and real-time forecasting capabilities for tsunamis using altimetry data are on the horizon.

In conclusion, we have developed a 2-D variational data assimilation algorithm for reconstruction of initial conditions of surface waves, with the primary aim of extending the results of the 1-D variational assimilation outlined in [27]. We have confirmed the feasibility of variationally data assimilation for tsunami waves in idealised 2D configurations, and have qualitatively demonstrated the necessary and/or sufficient conditions for convergence of the reconstructed initial conditions to the true shape. We have analysed different configurations of observations, their spacing and their number, with a variety of bathymetry and initial conditions. Importantly, we have extended the 1-D results regarding necessary conditions for the maximum distance between pairwise observations as given by equation (2.4.1), and have additionally shown that for the 2-D case this is not a sufficient condition. We have demonstrated that when there are no observations within the support of the initial conditions, a sufficiently large number of observations is a necessary condition for convergence. Based on these criteria, we have shown that convergence can be achieved for observations arranged in straight lines, grids, and along concentric circular arcs, for both isotropic and anisotropic initial conditions, and with realistic bathymetry data. With advances in altimetry observation data availability and accuracy, there is scope to use these results for more accurate tsunami models, with real observation data.

Chapter 3

A variational data assimilation scheme for bathymetry detection from surface wave observations

3.1 Introduction

The process of observing ocean bathymetry using measurements of surface waves is ill-posed, often exhibiting sensitivity to small amounts of noise in the system, and susceptible to instability inherent in the inversion process [39]. Data assimilation is one such inversion process, where observations of a true state are combined with a mathematical model in order to recover missing data governing the system evolution. The verification of a variational adjoint-based scheme as presented in this study is challenging, especially without analytical solutions for the nonlinear shallow water system with non-zero bathymetry. Despite this, the effect of bathymetry is essential when predicting tsunamis, since bathymetry variation modifies speed, direction, and stability of the propagating wave [7]. Bathymetry information can be either static or dynamic, where the latter accounts for shifts in the ocean floor due to seismic activity.

The high number of degrees of freedom in this problem makes it difficult to determine the criteria for optimal bathymetry reconstruction. We must consider the shallow water system governing tsunami propagation, as well as parameters such as amplitude and shape of the bathymetry, and initial conditions. This is in addition to calibration of the optimization scheme, and deriving optimal configurations of the observation network. In this study we address two complementary questions:

1. How accurately can bathymetry data be reconstructed from surface wave measurements, and what determines the accuracy?
2. How accurate does the bathymetry data need to be model sea surface waves to given accuracy?

We attempt to address these questions for an idealised 1-D case, as a first step for more complex analyses. We quantify some key relationships between the initial conditions and bathymetry amplitudes relative to the average fluid depth. We also analyse the effect of number of observations on assimilation convergence. We implement a Sobolev gradient smoothing technique (effectively a low-pass filter) within our optimization scheme and illustrate its ability to reduce small-scale noise present in the bathymetry reconstruction. We then investigate the consequences of error in the bathymetry data on the resulting surface wave by observing trends in the surface wave propagation error, as the amplitude of the initial condition, amplitude of the bathymetry, and the number of observation points is varied.

Section 3.2 provides a review of the effect of bathymetry on surface waves, and efforts to date to map ocean bathymetry, highlighting empirical, numerical and theoretical approaches. In section 3.3, we provide a concise overview of the shallow water system and derivation of the first order adjoint data assimilation scheme using principles of optimal control theory, and a summary of the algorithm. Section 3.4 gives preliminary results for different choices of initial conditions and exact bathymetry for the data assimilation scheme. This investigation reveals small-scale noise in the optimal reconstruction. To reduce this noise, section 3.5 proposes a low pass filter, which effectively removes higher frequencies in our reconstruction by increasing the regularity of our estimate at each iteration, taking it from the space $\mathrm{L}^2(\mathbb{R})$ to $\mathrm{H}^2(\mathbb{R})$. We discuss results of the smoothed optimisation scheme and illustrate the removal of noise in the reconstructed bathymetry in multiple cases. Section 3.6 analyses the relationships between the amplitudes of the initial conditions and bathymetry relative to the average depth, and attempts to formulate a relationship summarising certain necessary conditions for convergence. We also analyse the effect of number of observation points on the optimal reconstruction. Finally, we provide a sensitivity analysis of the surface wave to errors in the bathymetry reconstruction.

We observe that the low pass filter effectively reduces small-scale noise in the bathymetry reconstruction for multiple bathymetry shapes. Additionally, a necessary condition for convergence is that the amplitude of the initial conditions be at least two orders of magnitudes smaller than the amplitude of the bathymetry. Convergence is significantly improved by increasing the number of observation points, however this is not a sufficient condition for convergence. Finally, we show qualitatively that the surface wave exhibits low sensitivity to bathymetry reconstruction error, motivating second order adjoint sensitivity analysis (in Chapter 4) and global sensitivity analysis (GSA) (in Chapter 5), that investigate sensitivity of the surface wave to observations and the amplitudes of the initial conditions and bathymetry.

Conclusions are summarised in section 3.7, where we provide insights that motivate analyses in chapters two and three, with more rigorous sensitivity analysis to quantify effects of observations, and bathymetry shape and amplitudes on surface wave predictions.

3.2 Review of bathymetry effects and previous work

Bathymetry can have a significant impact on propagating shallow water waves, by altering the depth-dependent wave-speed $c = \sqrt{gH}$, where h is the total fluid depth. A demonstration of bathymetry effects is given in figure 3.1, where we observe the surface wave with and without bathymetry effects. We observe that both the speed and height of the surface wave are perturbed by bathymetry effects. This is because a tsunami's energy flux remains relatively constant, and so as the tsunami's speed varies, so does its height (shoaling). Thus bathymetry can significantly impact the arrival time of tsunami waves, and differing coastal communities can receive differing forms of tsunami energy based on local bathymetry effects. A detailed sea floor topography is an essential component of accurate tsunami models.

Attempts to create an accurate map of oceanic bathymetry have been made by direct measurements, or using information from propagating surface waves. Direct measurement includes platforms like ship-based high frequency radars. However many of these methods are either too costly or have poor spatial resolution. Often it is easier to measure waves propagating on the free surface, and use this information to create a map of the bottom topography from classical wave theory. This is an *inverse problem*. In practice, this can be difficult as bathymetry maps are extremely variable in quality and resolution, and often use inconsistent grids.

Wunsch [64] provides a detailed review of inverse methods for ocean circulation models detailing methods for both deep and shallow water. In coastal regions, depth-inversion methods have been refined to account for observational data in water areas with large interference from human activities and muddy water [15]. Additionally, inverse methods are also routinely used in open channel flow modelling, where bed topography is approximated using surface measurements [16].

A longstanding approach to solving this inverse problem uses the dispersion relation of surface waves. Earlier works such Lubard et al. [32] used measurements of the frequency-wavenumber spectrum made via optical images, obtained using cameras mounted on an oceanographic research tower. Since then, various methods using dispersion relations have been investigated, where bathymetry is measured by fitting the theoretical dispersion relation for gravity waves (where depth is a system parameter), and derived using inversion formulas.

More recent works include Dugan [11], extended by Piotrowski and Dugan [43], where image sequences of shoaling ocean waves taken from an aircraft are used to retrieve maps of water depth via the linear dispersion relation. The accuracy of this method were found to within 5% if the waves are reasonably linear. However, Grilli [22] builds on the research conducted by Dugan [11], arguing that the latter is limited due to neglect of amplitude dispersion effects, which accumulate through increasing nonlinearity as waves approach breaking in shallow water. He compares the linear frequency dispersion to a third order polynomial relationship between wave speed c as a function of wavenumber k and depth h, showing that due to amplitude dispersion effects, linear wave theory may greatly under estimate c, and lead to poorer estimates of bathymetry inversion formulas based on a linear dispersion relation.

These inversion algorithms are calibrated based on results of simulated periodic waves over mild slopes in a two-dimensional 'numerical wave tank'. This FNPF (Fully Nonlinear based on Potential Flow) numerical wave tank methodology was developed by Grilli and Subramanya [20] with wave generation and absorption methods, to calculate speed and height variation for a number of shoaling waves over slopes ranging from $1 : 35$ to $1 : 70$, as shown in Grilli [22]. It was demonstrated in Grilli [22] to have higher accuracy in coastal simulations than the nonlinear shallow water equations.

Tsai and Yue [60] also demonstrated how FNPF numerical tanks allow calculation of "numerically exact" properties of shoaling wave up to breaking point, and can provide accurate representation of surface waves independent of nonlinearity parameters.

While these methods are based on an empirical formula for the nonlinear inversion problem, Nicholls and Taber [36] derive an inversion formula for bathymetry analytically, using the nonlinearity of the governing Euler equations for ideal fluid flow to detect bathymetry information. The governing equations for the surface wave are expressed as a Hamiltonian system, and a Dirichlet-to-Neumann Operator (DNO) is applied to the system in order to remove some implicit dependencies. The result is a single equation of the wave height at the surface in terms of the bathymetry, and subsequently an inversion is derived. However, because their inversion formula is linear, they are required to assume a small amplitude for the bathymetry and their numerical results only consider normalised bathymetry amplitudes of 0.07 and 0.025.

Additionally, surface waves can be used to characterize a rapid dynamic change in bathymetry, rather than approximate static bathymetry. Jang et al. [26] take a similar approach to Grilli [22] to the related problem of measuring a sudden shift in the sea floor, e.g. due to seismic activity, using measurements of the surface waves. Their inversion formula however, is based on the same approach as Dugan [11]; using the linear dispersion relation with bathymetry as a parameter, and using transforms to show that the problem becomes one of solving an integral equation involving the known surface wave data. The uniqueness of this solution is demonstrated, and analysis concludes that there is a lack of stability in the measurement of the bottom displacement, and a question as to whether it depends continuously on the wave elevation. They overcome this using regularisation methods iteratively as a stabilisation technique, and show numerical convergence to the integral solution.

Each of the aforementioned methodologies has its strengths and weaknesses. Piotrowski and Dugan [43] and Grilli [22] both discuss practical measurement techniques of surface waves, whereas theoretical approaches such as Nicholls and Taber [36], and Jang et al. [26] show high degree of convergence, but assume full knowledge of the surface wave, and do not address the complexities involved in obtaining accurate measurements in a real-world scenario, such as noise or incomplete measurements. Nicholls and Taber [36] state that their future objectives aim to find an effective way of extracting wave fields from full observational data.

In summary, while the theoretical results from such models are promising, their applicability to real-world measurements has not been established. In the present study we attempt to combine theoretical results with realistic assumptions. We consider 1-D geometry as a first step to validate the basic approach and investigate fundamental questions. Assuming a finite set of observations, we analyse the effects of different bathymetry features on convergence, as well as the effect of different amplitudes and shape of the initial condition. These effects have not been considered in detail in the reviewed works.

Another issue is the difference between the inversion formulas derived for linear and nonlinear systems. Grilli [22], and Nicholls and Taber [36] give results accounting for nonlinearity and the resulting dispersion effects, whereas Piotrowski and Dugan [43], and Jang et al. [26] are restricted to linear dispersion relations for gravity waves. However, there remains the question of whether the empirical formulation of the inversion as derived by Grilli [22] is as rigorous as the analytically derived solution of Nicholls and Taber [36]. Both assume periodic waves, but the practical limitations on the accuracy of the free surface data make it difficult to assess the relative efficacy of these two methods. Ultimately, a complete evaluation of these approaches will depend on research which systematically compares inversion formulas based on the nonlinear governing equations for free surface wave propagation, as well as addresses the practical issues of collecting realistic field wave data.

In the present study, we approach the bathymetry estimation problem from a variational data assimilation perspective, with the goal of formulating an algorithm for the 1-D nonlinear shallow water system that predicts bathymetry from a small set of observation. We neglect rotational Coriolis effects and assume a static bathymetry with lateral boundary conditions. Our approach does not involve inversion of the dispersion relation. Instead we formulate an optimisation problem, seeking to minimise the error between observations and forecasts of the free surface wave. The analytical derivation of the variational algorithm differs from the empirical techniques introduced by Grilli [22] and is derived for the infinite-dimensional case. The use of sparse observations aims to provide a relatively more realistic set of assumptions than Nicholls and Taber [36], who require complete observations of surface wave fields.

Our aim is to use conclusions of this idealised case, to better understand the role of model parameters and the observation operator on bathymetry reconstruction. Our goal is to take the qualitative results observed here, and use them as a benchmark for more rigorous sensitivity analyses based on second order adjoint and Global Sensitivity Analysis approaches in Chapters 4 and 5. Consequently, our objective is to highlight results and techniques than can improve existing methodologies for accurate tsunami prediction .

Symbol	Definition
$\eta(x,t)$	General solution for the height perturbation
$\phi(x)$	General initial condition, i.e., $\phi(x) := \eta(x,0)$
$\hat{\eta}$	Amplitude of the initial conditions $\phi(x)$
$\eta^{(t)}(x,t)$	True solution for the height perturbation $\eta(x,t)$
$\beta^{(t)}(x)$	True bathymetry
$\hat{\beta}$	Amplitude of the true bathymetry $\beta^{(t)}(x)$
$\beta^{(g)}(x)$	Starting guess for bathymetry
$\beta^{(n)}(x)$	Approximate bathymetry at iteration n of the assimilation algorithm
$\beta^{(b)}(x)$	Best approximation to the bathymetry (e.g., fixed point of iterations)
$y^{(o)}(t)$	Observations of the true height perturbation at positions $\{x_j\}, j = 1, ..., N_{obs}$
$\eta^{(f)}(x,t)$	Approximate ("forecast") solution generated by approximate bathymetry
$\mathcal{J}^{(n)}$	Cost function at iteration n
$(\cdot)^*$	Adjoint

Table 3.1: Notation used in the derivation of data assimilation scheme of the SWE to find the optimal bathymetry, using same format as given in Kevlahan et al. [27].

3.3 Derivation of adjoint based data assimilation scheme

The nonlinear shallow water equations are a coupled system of equations for travelling free surface waves. They are derived from the two-dimensional Euler equations, under the assumption that the wavelength λ of free surface waves is much larger than the total ocean depth h, allowing us to average the Euler equations over the vertical dimension. The fluid column height becomes $h = H + \eta(x,t) - \beta(x)$, where H is the average depth, η is the perturbation of the free surface, and β is time-independent sea floor perturbation from zero, i.e. the "bathymetry". Appropriate vertical averaging gives the irrotational, constant density incompressible one-dimensional nonlinear shallow water equations,

$$\frac{\partial \eta}{\partial t} + \frac{\partial}{\partial x}\Big((H + \eta - \beta)u\Big) = 0, \tag{3.3.1a}$$

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\Big(\frac{1}{2}u^2 + g\eta\Big) = 0, \tag{3.3.1b}$$

$$\eta(x,0) = \ \phi(x), \tag{3.3.1c}$$

$$u(x,0) = \ 0. \tag{3.3.1d}$$

We assume that the initial conditions $\phi(x)$ is compactly supported, and that we have periodic boundary conditions on some domain $\Omega = \{x; x \in [-L, L]\}$. Our objective is to implement a variational data assimilation scheme constrained by (3.3.1) in order to estimate the bathymetry $\beta(x)$. We wish to derive an optimal estimate of the bottom topography using a finite number of observations of the free surface perturbation, for all times t in our temporal domain $[0, T]$. To simplify further, we normalise the system (3.3.1) by the average height H and gravitational acceleration g such that the wave speed $c = \sqrt{gH} = 1$.

We can quantify our objective as the partial differential equation (PDE) constrained minimisation of some cost function $\mathcal{J}$,

$$\mathcal{J}(\beta) = \frac{1}{2}\int_0^T \sum_{i=1}^{M} \Big[\eta^{(f)}(x_j, t; \beta) - y_j^{(o)}(t)\Big]^2 \, dt, \tag{3.3.2}$$

where $y_j^{(o)}(t)$ are the observations of the true free surface perturbations taken at positions x_j, $j = 1, ..., N_{obs}$, and $\eta^{(f)}(x_j, t; \beta)$ is the solution of our system at x_j generated by the bathymetry β. We define the optimal bathymetry $\beta^{(b)}$ by

$$\beta^{(b)} = \operatorname{argmin}_{\beta \in L^2(\Omega)} \mathcal{J}(\beta). \tag{3.3.3}$$

This is equivalent to solving

$$\nabla^{L^2} \mathcal{J}(\beta^{(b)}) = 0. \tag{3.3.4}$$

As direct computation of this optimization problem is too computationally expensive, we formulate a dual adjoint system in terms of Lagrange multipliers (also called adjoint variables), that allow us to find $\beta^{(b)}$ efficiently [19].

The first variation of $\mathcal{J}$, given some arbitrary perturbation β' of size ε is given by the Gâteaux derivative,

$$\mathcal{J}'(\beta;\beta') = \lim_{\varepsilon\to 0} \frac{J(\beta+\varepsilon\beta') - J(\beta)}{\varepsilon}. \tag{3.3.5}$$

Expanding the perturbation to $\mathcal{O}(\varepsilon)$, we can reformulate (3.3.5) as

$$\mathcal{J}'(\beta;\beta') = -\int_0^T \Big(\eta^{(f)}(x_j,t;\beta) - y^{(o)}(t)\Big)\eta'\ dt, \tag{3.3.6}$$

where (η',u') are the solutions of the perturbed system of (3.3.1) given β', found by linearising about (η,u) and extracting the $\mathcal{O}(\varepsilon)$ system. As the Gâteaux derivative is a directional derivative in the direction of the perturbation β', we can express (3.3.6) as the inner product between $\nabla\mathcal{J}$ and β',

$$\mathcal{J}'(\beta;\beta') \ = \ \langle\nabla\mathcal{J},\beta'\rangle_{L^2(\Omega)} \ = \ \int_{-L}^{L} \nabla^{L_2}\mathcal{J}\ \beta'\ dx. \tag{3.3.7}$$

Then the following forms of $\mathcal{J}(\beta;\beta')$ are equivalent,

$$J'(\beta;\beta') \ = -\int_0^T \Big(\eta^{(f)}(x_j,t;\beta) - y^{(o)}(t)\Big)\eta'\ dt = \int_{-L}^{L} \nabla^{L^2}\mathcal{J}\ \beta'\ dt. \tag{3.3.8}$$

We form the Lagrangian associated to linearised system for (η',u') with some arbitrary adjoint variables (η^*,u^*),

$$\int_0^T\int_{-L}^{L} \eta^*(x,t)\Big[\frac{\partial\eta'}{\partial t}+\frac{\partial}{\partial x}\Big((\eta'-\beta')u+(\eta+1-\beta)u'\Big)\Big]+u^*(x,t)\Big[\frac{\partial u'}{\partial t}+\frac{\partial}{\partial x}\Big(\eta'+uu'\Big)\Big]dx\,dt = \ 0. \tag{3.3.9}$$

Integrating by parts in time and space reduces (3.3.9) to

$$\begin{aligned}
0 = &- \int_0^T \int_0^L \eta' \left\{ \frac{\partial \eta^*}{\partial t} + u\frac{\partial \eta^*}{\partial x} + \frac{\partial u^*}{\partial x} \right\} + u' \left\{ \frac{\partial u^*}{\partial t} + (\eta + 1 - \beta)\frac{\partial \eta^*}{\partial x} + u\frac{\partial u^*}{\partial x} \right\} - \beta' u \frac{\partial \eta^*}{\partial x} \; dxdt \\
&+ \int_0^T \eta^* \Big[(\eta' - \beta')u + (\eta + 1 - \beta)u' \Big] \Big|_{-L}^{L} dt \; + \int_0^T u^* \Big[\eta' + uu' \Big] \Big|_{-L}^{L} dt \\
&+ \int_{-L}^{L} \eta^* \eta' \Big|_{t=T} dx \; - \int_{-L}^{L} \eta^* \eta' \Big|_{t=0} dx \\
&+ \int_{-L}^{L} u^* u' \Big|_{t=T} dx - \int_{-L}^{L} u^* u' \Big|_{t=0} dx. \qquad (3.3.10)
\end{aligned}$$

Due to periodicity the boundary terms vanish. If we choose (η^*, u^*) as the solution to

$$\begin{aligned}
\frac{\partial \eta^*}{\partial t} + u\frac{\partial \eta^*}{\partial x} + \frac{\partial u^*}{\partial x} =& \; \Big(\eta^{(f)}(x, t; \beta) - y^{(o)}(t) \Big) \delta(x - x_j), && (3.3.11a) \\
\frac{\partial u^*}{\partial t} + (1 + \eta - \beta)\frac{\partial \eta^*}{\partial x} + u\frac{\partial u^*}{\partial x} =& \; 0, && (3.3.11b) \\
\eta^*(x, T) =& \; 0, && (3.3.11c) \\
u^*(x, T) =& \; 0, && (3.3.11d)
\end{aligned}$$

then (3.3.10) is reduced to

$$\int_0^T \int_0^L \Big(\eta^{(f)}(x_j, t; \beta) - y^{(o)}(t) \Big) \eta' \; dxdt = \int_0^T \int_0^L \beta' u \frac{\partial \eta^*}{\partial x} \; dxdt, \qquad (3.3.12)$$

and (3.3.11) are called the adjoint equations. We note that the normalised mean depth $H = 1$ is preserved in the adjoint system. Combining this result with the equivalence given by (3.3.8), we have

$$\int_{-L}^{L} \int_0^T u \frac{\partial \eta^*}{\partial x} \beta' \; dt \; dx = \int_{-L}^{L} \nabla^{L^2} \mathcal{J} \; \beta' \; dx, \qquad (3.3.13)$$

and thus, since our functional is linear and bounded and belongs to the space of square-integrable functions, we can use the Riesz representation theorem to extract

$\nabla^{L_2}\mathcal{J}$, giving

$$\nabla^{L^2}\mathcal{J} = \int_0^T u\frac{\partial\eta^*}{\partial x}\,dt. \tag{3.3.14}$$

Losch and Wunsch [31] utilise a similar adjoint based minimisation for their bathymetry detection analysis, however they do not consider the infinite-dimensional case as we have here. The benefits of our approach is that it is independent of the discretisation used in its numerical implementation.

To verify that our formulation for $\nabla^{L^2}\mathcal{J}$ is correct, we define the Kappa test

$$\kappa(\varepsilon) = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\frac{J(\beta+\varepsilon\beta') - J(\beta)}{\langle\nabla^{L^2}\mathcal{J},\beta'\rangle_{L^2(\Omega)}}, \tag{3.3.15}$$

where $\kappa(\varepsilon)$ is the quotient of the two equivalent forms for the variation $\mathcal{J}(\beta;\beta')$ we used in the above derivation. Given some perturbation β', if we have correctly defined $\nabla^{L^2}\mathcal{J}$, then as $\varepsilon \to 0$, we should see $\kappa(\varepsilon) \to 1$. The results of the kappa test for different cases are presented in sections 3.4 and 3.5.

The minimiser $\beta^{(b)}$ yielding $\nabla^{L^2}\mathcal{J} = 0$, is computed using an steepest descent algorithm given some starting guess $\beta^{(g)}$. Using a line minimisation algorithm to find the optimal step size at each iteration, this can be summarised as

$$\beta^{(n+1)} = \beta^{(n)} - \tau_n\nabla^{\mathcal{L}_2}\mathcal{J}\Big(\beta^{(n)}\Big) \tag{3.3.16}$$

where

$$\tau_n = \operatorname{argmin}_{\tau\in\mathbb{R}}\ \mathcal{J}\Big(\beta^{(n)}(x) - \tau\nabla^{\mathcal{L}_2}\mathcal{J}\Big(\beta^{(n)}(x)\Big)\Big). \tag{3.3.17}$$

The optimal bathymetry reconstruction $\beta^{(b)}$ is the fixed point of this iterative scheme. The steps for the process are outlined in algorithm 2.

Algorithm 2 Data Assimilation Algorithm for Bathymetry Estimation

```
1: Pick initial estimate for $\beta^{(g)}$.
2: Solve the initial value problem for $(u, \eta)$ from $t = 0$ to $t = T$.
3: Solve adjoint problem for $(u^*, \eta^*)$ backwards in time from $t = T$ to $t = 0$ to find $\eta^*(x, t)$.
4: Approximate $\int_0^T u \frac{\partial \eta^*}{\partial x}\, dt$ at every point in spatial domain $\Omega$.
5: Define $\nabla^{L^2} \mathcal{J} = \int_0^T u \frac{\partial \eta^*}{\partial x}\, dt$.
6: Compute the optimal time step $\tau_n$ through a line minimisation algorithm.
7: Use a gradient descent algorithmn to compute $\beta^{(n+1)}(x) = \beta^{(n)}(x) - \tau_n \nabla^{L^2} \mathcal{J}\Big(\beta^{(n)}(x)\Big)$.
8: Repeat until $\parallel \nabla^{L^2} \mathcal{J} \parallel < \varepsilon$ for some small $\varepsilon$ ($\parallel \int_0^T u \frac{\partial \eta^*}{\partial x}\, dt \parallel \approx 0$).
9: Set $\beta^{(b)}(x) := \beta^{(n)}(x)$.
```

3.4 Initial results using L^2 gradients

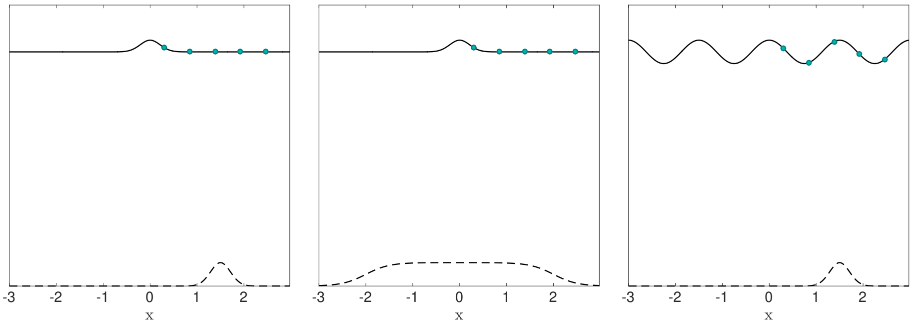
Case	Bathymetry	Initial conditions
I	Gaussian	Gaussian
II	Sandbar	Gaussian
III	Gaussian	Sinusoidal

Table 3.2: Cases considered for data assimilation algorithm 2.

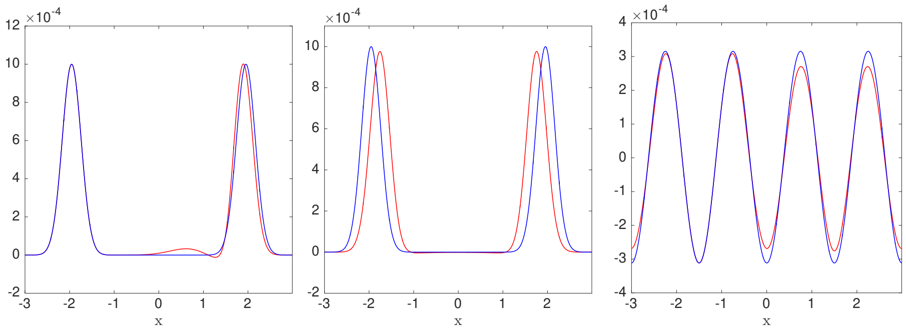
To verify the numerical implementation of algorithm 2, we consider three cases, characterised by different initial conditions $\phi(x)$ and the true bathymetries $\beta^{(t)}(x)$. These cases are represented in Table 3.2, and shown in fig 3.1. These cases were chosen to analyse convergence in scenarios where the support of $\phi(x)$ and the support of $\beta^{(t)}(x)$ overlap or are disjoint. Additionally, we want to evaluate the effect of a surface wave with compactly supported initial conditions (cases I and II) or periodic initial conditions (case III). We consider Gaussian and sandbar profiles for the bathymetry, similar to Nicholls and Taber [36], as a 1-D approximation for peaks and ridges characterising ocean bathymetry. The primary application of this study is tsunami modelling given some optimal reconstruction of missing bathymetry data, hence we are primarily interested in a non-periodic propagating surface wave, as in case I and case II. However, including the periodic initial conditions case III in our analysis helps understand the effects of the observation operator and model parameters, on the optimal reconstruction.

We implement these schemes using a second order finite difference approximation in space, and a four stage third order Runge-Kutta scheme [57] in time. The resolution of our spatial grid is $N = 512$, and our spatial domain is $\Omega = \{x \in \mathbb{R}; -L \leq x \leq L\}$.

The system is integrated for $t \in [0, T]$. In the variational data assimilation for the initial conditions reconstruction in Kevlahan et al. [27], the system is integrated in time for $t \in [0\ T]$, where $T = 2$ is chosen such that the free surface wave does not reach the boundary. This was appropriate when our objective was to reconstruct the initial conditions of a tsunami wave in real time. For the bathymetry assimilation, as we have no constraints on the assimilation time, the final time is chosen to be $T = 2L$, where boundary effects were present. Periodic boundary conditions are artificial in the sense that they are not a substitute for realistic ocean conditions and are not intended to simulate coastlines. However they are appropriate in our idealised 1-D case as they simplify the dynamics, and do not violate underlying principles. In our case we restrict the choice of free surface wave to (i) a travelling Gaussian, and (ii) a sinusoidal wave. The latter is periodic and is not significantly altered by the boundary conditions. For the Gaussian initial conditions, the wave is effectively reflected at the boundary, and consequently is a wave travelling in the opposite direction. We observed that there were no errors generated at the boundaries that propagated inside the domain and impacted results. The main conclusions of this study were not altered by increasing the assimilation time from $T = 2$ to $T = 2L$, however convergence of the reconstructed bathymetry improved with the longer assimilation time.



(a) Case I (b) Case II (c) Case III



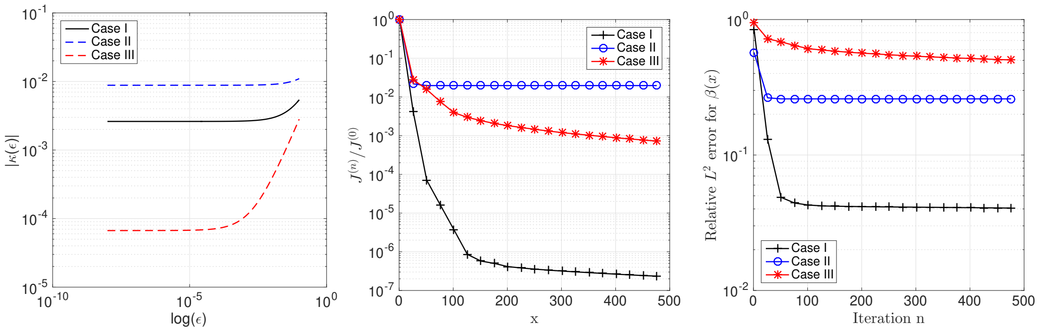
(d) Case I surface wave (e) Case II surface wave (f) Case III surface wave

FIGURE 3.1: (a), (b), & (c) show the three test cases for bathymetry $\beta(x)$ and free surface perturbation initial conditions $\eta(x,0)$ for the data assimilation scheme. The green circles represent the observations, with $N_{obs} = 5$ and $y_1^{(o)} = 0.1L$. Note that while the spatial distribution is correct, amplitude of the initial conditions $\hat{\eta}$, amplitude of the bathymetry $\hat{\beta}$, and average depth H are not to scale in these diagrams, as $\hat{\eta}$ was restricted to 1% of $\hat{\beta}$ across most of the numerical tests. Plots (d), (e) & (f) show the propagating free surface wave at $t = 1.95$ with flat bottom (blue) and bathymetry (red) for each case I, II and III respectively, to highlight the effect of bathymetry on surface wave propagation.

We integrate the system (3.3.1) on a staggered grid where $u(x,t)$ values are located at grid edges and $\beta(x)$ and $\eta(x,t)$ are located at grid centres. Periodic boundary conditions are imposed at $x = L$ and $x = -L$ where $L = 3$. We assume we have no background information for bathymetry a priori, and set $\beta^{(g)}(x) = 0$.

The results in fig 3.2 illustrate the convergence of the data assimilation scheme for each case outlined in table 3.2, as well as the convergence of the kappa test (3.3.15). In all results we take a relative L^2 error less than 10% as the threshold for "converged" bathymetry. Let us first consider the convergence of the cost function (3.3.2) to zero. Ultimately the purpose of the optimization scheme is to minimise the error between the observations of the true height perturbation $y^{(o)}(t)$ and the approximated solution, $\eta^{(f)}$ given the optimal bathymetry $\beta^{(b)}$. Accurate reconstruction of the bathymetry is predicted to be a consequence of minimising the error. However, due to the ill-posed nature of the problem, convergence of the cost function may not guarantee accurate bathymetry reconstruction. Indeed, the results highlight the difference between these two objectives. We see in figure 3.2(b), that the relative decrease in the cost function over 500 iterations of algorithm 2 is greatest for case I at $\mathcal{O}(10^{-6})$, and thus the algorithm converges successfully for this case. The relative decrease in $\mathcal{J}^{(n)}$ for cases II and III is not less than $\mathcal{O}(10^{-2})$ and $\mathcal{O}(10^{-4})$ respectively. We show the errors in the reconstructed bathymetry $\beta^{(b)}$ corresponding to these cost functions in figure 3.2(c).

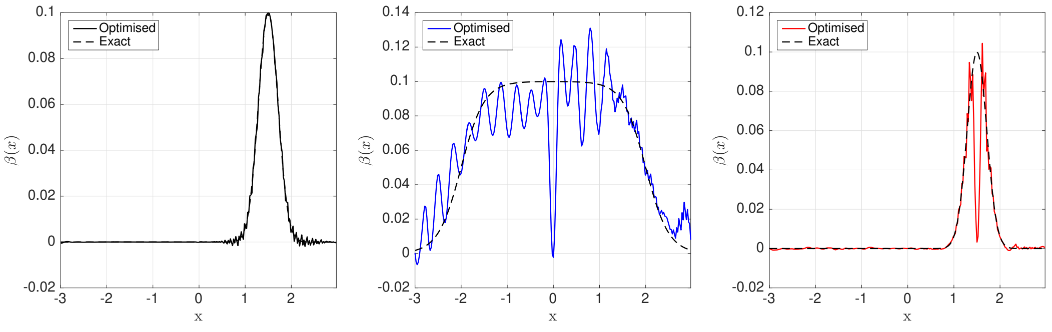
It is clear that for each case, convergence of the cost function does not necessarily correspond to the true bathymetry $\beta^{(t)}$. Only in Case I does the reconstructed bathymetry converge to the true bathymetry, although even in this case the error is relatively large, 0.04. Figures 3.2(d), (e) and (f) give the reconstructed bathymetry for each case. It is immediately clear that the primary source of error is small-scale noise in the reconstruction. For case I, we see that although the Gaussian bathymetry is well-resolved, there is still small-scale noise present in the tails. However, for case III where we assume periodic initial conditions $\phi(x)$, the error in the reconstructed bathymetry is large, and we conclude that the observability of the bathymetry by sensors measuring a sinusoidal propagating surface wave is significantly lower than that of a travelling Gaussian wavefront, as in case I.



(a) $|\kappa(\varepsilon) - 1|$

(b) Relative reduction in $\mathcal{J}^{(n)}$

(c) $\|\ \beta^{(t)} - \beta^{(n)}\ \|^2_{L^2} / \|\ \beta^{(t)}\ \|^2_{L^2}$



(d) $\beta^{(b)}$ for Case I

(e) $\beta^{(b)}$ for Case II

(f) $\beta^{(b)}$ for Case III

FIGURE 3.2: Results for iterative data assimilation scheme outlined in algorithm 2, with $\nabla\mathcal{J} \in L^2(\Omega)$. Only in case I do we consider that the assimilation has reconstructed the bathymetry with sufficient accuracy ($<$ 10% relative error). (a) shows convergence of the kappa test for the three cases. (b) shows the relative reduction in the cost function after 500 iterations. (c) shows the relative error in the reconstructed bathymetry. (d), (e) & (f) show the optimal reconstructed bathymetry for each case. We observe noise in the reconstruction for each case, especially in case II.

For case II, it is interesting to note that the noise is larger scale on the front slope of the bathymetry ($x < 0$) than the back slope of the bathymetry ($x < 0$). Since the observation points are placed on the right hand side of the Gaussian initial conditions ϕ (as shown in figure 3.1), and as such there are no observations of the left propagating wave. Additionally, unlike cases I and III, the observations are not positioned before the support of the bathymetry. This could suggest that observations for $x > 0$ may not be able to sufficiently capture the bathymetry effects on the left-propagating wave. We give a deeper analysis of sensitivity to the placement of the observations and the resulting effect on reconstruction error in chapters 2 and 3.

We expect that if the kappa test converges, the error of the associated variational algorithm should also converge (even though this is not necessary). However, we see in figure 3.2(a) that the best convergence for the kappa test is for case III, whereas figure 3.2(c) shows that case III actually has the least accurate bathymetry reconstruction.

In the following section we show that an optimisation scheme where the gradient of the cost function (3.3.14) exists in $L^2(\Omega)$ is not smooth enough to obtain classical solutions to (3.3.1), and thus in section 3.5 we analytically derive a low pass filter by smoothing our gradient such that $\nabla\mathcal{J} \in H^2(\Omega)$ and provide results of the numerical implementation. Using smoother H^2 gradients largely eliminates the small scale errors in the reconstructed bathymetry we have seen here using L^2 gradient.

3.5 Smoothing using Sobolev gradients

In system (3.3.1), the bathymetry is incorporated via the $(\beta u)_x$ term. Thus, a classical solution to this system requires smoothness not just on β, but also its derivative. Because of this, we require the gradient (3.3.14) to be in the Sobolev space $H^2(\Omega)$, which imposes smoothness conditions on β_x as well as β. The following derivation of the corresponding Sobolev gradient is adapted from Protas [45].

$H^2(\Omega)$ is a Sobolev space equipped with the inner product

$$\begin{aligned}\langle v_1, v_2\rangle_{H^2(\Omega)} &= \langle v_1, v_2\rangle_{L^2(\Omega)} + l_1^2\left\langle \frac{\partial v_1}{\partial s}, \frac{\partial v_2}{\partial s}\right\rangle_{L^2(\Omega)} + l_2^4\left\langle \frac{\partial^2 v_1}{\partial s^2}, \frac{\partial^2 v_2}{\partial s^2}\right\rangle_{L^2(\Omega)} \\ &= \int_{s=-L}^{L}\left[v_1 v_2 + l_1^2\frac{\partial v_1}{\partial s}\frac{\partial v_2}{\partial s} + l_2^4\frac{\partial^2 v_1}{\partial s^2}\frac{\partial^2 v_2}{\partial s^2}\right] ds, \end{aligned} \tag{3.5.1}$$

where $v_1, v_2 \in H^2(\Omega)$, and $l_1, l_2 \in \mathbb{R}$ are the length scale parameters used to adjust the regularity. As long as l_1, l_2 are finite, by the Riesz representation theorem,

$$\begin{aligned}\mathcal{J}'(\beta;\beta') &= \langle \nabla^{L^2}\mathcal{J},\beta'\rangle_{L^2(\Omega)}\\ &= \langle \nabla^{H^2}\mathcal{J},\beta'\rangle_{H^2(\Omega)}\\ &= \langle \nabla^{H^2}\mathcal{J},\beta'\rangle_{L^2(\Omega)} + l_1^2\Big\langle \frac{\partial \nabla^{H^2}\mathcal{J}}{\partial s},\frac{\partial \beta'}{\partial s}\Big\rangle_{L^2(\Omega)} + l_2^4\Big\langle \frac{\partial^2 \nabla^{H^2}\mathcal{J}}{\partial s^2},\frac{\partial^2 \beta'}{\partial s^2}\Big\rangle_{L^2(\Omega)}.\end{aligned} \tag{3.5.2}$$

In order to extract the gradient as we did in (3.3.13), we need to isolate the β' term, and define an equivalent expression for $\nabla^{L^2}\mathcal{J}$. And so we integrate by parts as before, on the second and third term in (3.5.2). We impose periodic conditions in space to eliminate the resulting boundary terms, and subsequently we have

$$\begin{aligned}\langle \nabla^{L^2}\mathcal{J},\beta'\rangle_{L^2(\Omega)} &= \langle \nabla^{H^2}\mathcal{J},\beta'\rangle_{H^2(\Omega)}\\ &= \int_{s=-L}^{L}\Big[\nabla^{H^2}\mathcal{J} - l_1^2\frac{\partial^2\nabla^{H^2}\mathcal{J}}{\partial s^2} + l_2^4\frac{\partial^4\nabla^{H^2}\mathcal{J}}{\partial s^4}\Big]\beta'\, ds.\end{aligned} \tag{3.5.3}$$

Since this holds for every arbitrary perturbation β', the process of smoothing the gradient from $L^2(\Omega)$ to $H^2(\Omega)$ is equivalent to solving the inhomogeneous boundary value problem

$$\nabla^{L^2}\mathcal{J}(s) = \nabla^{H^2}\mathcal{J}(s) - l_1^2\frac{\partial^2\nabla^{H^2}\mathcal{J}(s)}{\partial s^2} + l_2^4\frac{\partial^4\nabla^{H^2}\mathcal{J}(s)}{\partial s^4}, \tag{3.5.4a}$$

$$\frac{\partial^{(2i+1)}\nabla^{H^2}\mathcal{J}(s)}{\partial s^{(2i+1)}}\Big|_{s=L} = \frac{\partial^{(i)}\nabla^{H^2}\mathcal{J}(s)}{\partial s^{(i)}}\Big|_{s=L}, \quad i = 0,1. \tag{3.5.4b}$$

In Fourier space, solving (3.5.4) is simplified to

$$(\widehat{\nabla^{H^2}\mathcal{J}})_k = \frac{1}{1+l_1^2k^2+l_2^4k^4}(\widehat{\nabla^{L^2}\mathcal{J}})_k. \tag{3.5.5}$$

This is effectively a low pass filter applied to the L^2 gradient. We can make this filter as aggressive as needed by “tuning” l_1 and l_2, where the case $l_1 = l_2 = 0$ is equivalent to the L^2 gradient. For our smoothed data assimilation algorithm, we set $l_1 = 0$ and calibrate l_2, as this gives us the desired regularity and reduces the number of degrees of freedom in the problem. We now consider the numerical

Algorithm 3 Data Assimilation algorithm with low pass filter for bathymetry estimation

```
Pick initial estimate for $\beta^{(g)}$.
Solve the initial value problem for $(u, \eta)$ from $t = 0$ to $t = T$.
Solve adjoint problem for $(u^*, \eta^*)$ backwards in time from $t = T$ to $t = 0$ to find $\eta^*(x, t)$.
Approximate $\int_0^T u \frac{\partial \eta^*}{\partial x}\, dt$ at every point in spatial domain $\Omega$.
Define $\nabla^{L^2} \mathcal{J} = \int_0^T u \frac{\partial \eta^*}{\partial x}\, dt$.
Apply low pass filter (3.5.5) to $\nabla^{L^2} \mathcal{J}$ to get $\nabla^{H^2} \mathcal{J}$
Compute the optimal time step $\tau_n$ through a line minimisation algorithm.
Use a gradient descent algorithmn to compute $\beta^{(n+1)}(x) = \beta^{(n)}(x) - \tau_n \nabla^{H^2} \mathcal{J}\big(\beta^{(n)}(x)\big)$.
Repeat until $\| \nabla^{H^2} \mathcal{J} \| < \varepsilon$ for some small $\varepsilon$ ($\| \int_0^T u \frac{\partial \eta^*}{\partial x}\, dt \| \approx 0$).
Set $\beta^{(b)}(x) := \beta^{(n)}(x)$.
```

implementation of this updated optimization scheme, summarised in algorithm 3.

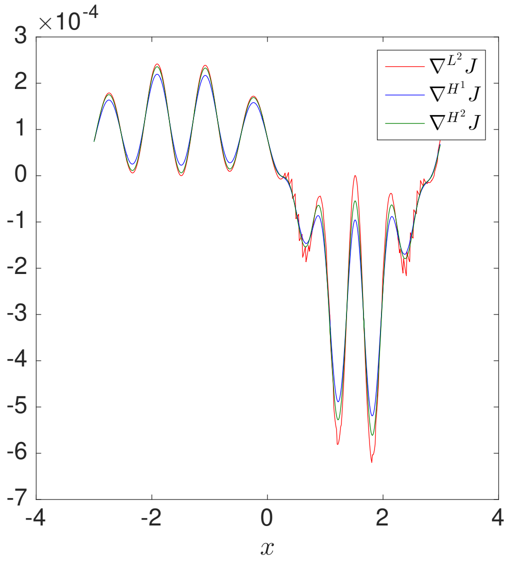
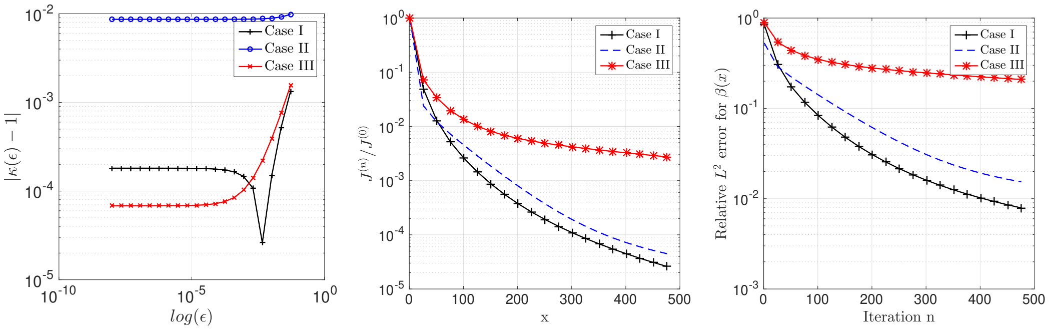


FIGURE 3.3: The gradient of $\mathcal{J}(\beta)$, obtained after one iteration for case III, for H^1 and H^2 Sobolev smoothing compared to the (unsmoothed) L^2 gradient.

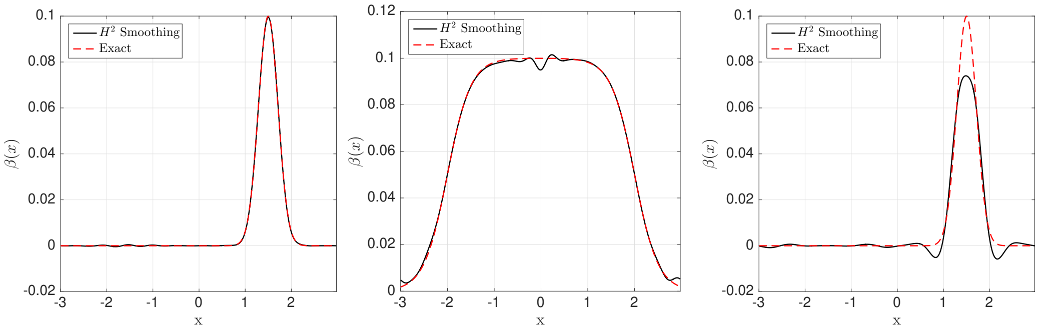
Before we present the results of the data assimilation scheme, we illustrate the efficacy of the filtering in fig 3.3. We compare the gradient in L^2 with the gradient in H^1 and H^2 obtained after smoothing for case III. The purpose of this comparison is to illustrate that requiring $\nabla\mathcal{J} \in H^2$ is more effective at reducing noise than H^1 smoothing, where the H^1 inner product is equivalent to (3.5.1) with $l_2 = 0$. In both cases, we try to choose optimal values of l_1 (for H^1) and l_2 (for H^2) to filter out higher frequencies that contribute to the noise in the bathymetry reconstruction, without also getting rid of necessary information for observing the bathymetry propagated by the lower wavenumbers. We observe in figure 3.3 that the H^1 gradient filters out the noise, but also reduces the amplitude of the signal peaks, whereas the H^2 gradient filters the noise and is closer to the original signal shape. We therefore conclude that H^2 Sobolev smoothing is optimal.

The results of algorithm 3 are given in fig 3.4. The plots (b), (e), and (h) show the relative error in the reconstruction for cases I, II, and III respectively. The first thing we note is that in each case, the error is lower for the results with H^2 smoothing compared to results in figure 3.2. Especially with cases I and II, we observe the error decreases by at least an order of magnitude. The reconstructed bathymetry shown in (c) and (f) illustrates how the noise has been greatly reduced, and for case I is almost negligible. For case II we see some noise remaining on the plateau of the sandbar, however it is a drastic improvement from the unfiltered result. This increase is reflected in the kappa test results given in (a) and (d) for case I and II respectively, though the increased convergence does not scale proportionally with the error, as we saw before in fig 3.2.

However, for case III there is no increase in convergence for the kappa test in figure 3.4(a) , and while the reconstruction of $\beta(x)$ has less small-scale noise, we observe that amplitude of the Guassian in figure 3.4(f) is significantly smaller than the exact bathymetry. These results did not improve with a more restrictive choice of l_2, leading us to consider whether additional factors such as the system parameters and placement of the sensors $y^{(o)}(t)$ affect the observability of the non-localised bathymetry. In section 3.6 we attempt to analyse some of these effects.



(a) Kappa test Convergence (b) Cost Function (c) Relative L^2 Error



(d) Case I (e) Case II (f) Case III

FIGURE 3.4: Results for assimilation scheme with Sobolev H^2 smoothing applied to $\nabla^{L^2} J$. (a) shows the convergence of the kappa test, (b) shows the convergence of the cost function, (c) shows the relative L^2 error $\| \beta^{(t)} - \beta^{(n)} \|^2_{L^2} / \| \beta^{(t)} \|^2_{L^2}$ between the exact and reconstructed bathymetry at each iteration. (d), (e), & (f) shows the reconstructed bathymetry with H^2 smoothing and the exact bathymetry for cases I, II and III respectively. Convergence is improved compared to results without smoothing given in figure 3.2.

3.6 Necessary conditions on model parameters and the observation operator

In the previous section we showed that using smoother H^2 gradients is necessary to avoid small scale errors in the reconstructed bathymetry. We now conduct a qualitative analysis to understand the effect of model parameters and the observation operator on the convergence of the data assimilation. The qualitative results observed here motivate the more rigorous analyses in Chapter 4 and Chapter 5, using second order variation and global sensitivity analyses respectively. All subsequent results are for $\mathcal{J}(\beta) \in H^2$.

3.6.1 Necessary conditions on parameters

For the purpose of this study, we restrict our parameter analysis to understanding the relationship between the amplitude of the initial conditions $\hat{\eta}$, the amplitude of the true bathymetry $\hat{\beta}$, and the average depth H, which we have normalised to 1. The number of observation points N_{obs} is 45, and they are positioned as in figure 3.1 with $y_1^{(o)} = 0.1L$. As this research is focused on improvement in tsunami prediction, our objective is to understand how surface waves propagate over bathymetry, and these factors play an important part when considering the scales involved. Tsunamis are characterised by their long wavelength, often reaching 100-150km in the deep oceans, and their relatively small amplitude can be between $0.1 - 1$m, making them often imperceptible. Even when approaching coastlines the amplitude of the surface wave can be $20 - 50$m whereas the wavelength may still be up to 2km. As the energy flux of the wave speed is dependent on depth-dependent wave speed $c = \sqrt{g(H + \eta(x,t) - \beta(x))}$, the amplitude and smoothness of the bathymetry can have a big impact on tsunami propagation [8]. Additionally, in deeper water when h is larger the effects of bathymetry are smaller. Consequently we also investigate the effectiveness of free surface observations in reconstructing bathymetries with amplitudes much smaller than the average depth H.

We therefore give special consideration to case I, Gaussian initial conditions with Gaussian bathymetry as shown in figure 3.1(a). We wish to highlight how convergence is affected when (i) the ratio $\hat{\eta}/\hat{\beta}$ is increased, and (ii) the normalised bathymetry amplitude $\hat{\beta}/H$ varies. As $H = 1$, for the latter we may consider values of $\hat{\beta}$.

We investigated convergence of the kappa test and relative error in the reconstructed bathymetry when $\hat{\eta}/\hat{\beta}$ is $\mathcal{O}(10^{-1})$ ("large"), $\mathcal{O}(10^{-2})$ ("small"), and $\mathcal{O}(10^{-3})$ ("very small") respectively. The experiments are conducted for $\hat{\beta} = \mathcal{O}(10^{-1})$ and for $\hat{\beta} = \mathcal{O}(10^{-3})$. The results of the six configurations are presented in table 3.3. For each case we provide the error in the kappa test $|\kappa(\varepsilon)-1|$ and the relative L^2 error in the reconstruction. We set a tolerance for the kappa test, such that all values of $|\kappa(\varepsilon)-1| \geq 10^{-2}$ are non-convergent. The entries highlighted in red indicate cases where the optimisation algorithm became unstable and the gradient $\nabla\mathcal{J}$ failed to converge to zero or became unbounded.

	$\frac{\hat{\eta}}{\hat{\beta}} = \mathcal{O}(10^{-1})$	$\frac{\hat{\eta}}{\hat{\beta}} = \mathcal{O}(10^{-2})$	$\frac{\hat{\eta}}{\hat{\beta}} = \mathcal{O}(10^{-3})$
$\hat{\beta} = \mathcal{O}(10^{-1})$	$\lvert\kappa(\varepsilon)-1\rvert = 3\times10^{-2}$ Failed to converge	$\lvert\kappa(\varepsilon)-1\rvert = 2.7\times10^{-2}$ L^2 Error $= 1.3\times10^{-3}$	$\lvert\kappa(\varepsilon)-1\rvert = 4.2\times10^{-4}$ L^2 Error $= 4.6\times10^{-3}$
$\hat{\beta} = \mathcal{O}(10^{-3})$	$\lvert\kappa(\varepsilon)-1\rvert = 2.5\times10^{-3}$ L^2 Error $= 7.4\times10^{-3}$	$\lvert\kappa(\varepsilon)-1\rvert = 3\times10^{-2}$ Failed to converge	$\lvert\kappa(\varepsilon)-1\rvert = 2.7\times10^{-3}$ Failed to converge

Table 3.3: Analysis of six experiments for case I where $\hat{\eta}/\hat{\beta}$ and $\hat{\beta}$ are varying orders of magnitude. The results show the convergence error in the kappa test and the error reconstruction error $\| \beta^{(t)} - \beta^{(b)} \|^2_{L^2(\Omega)} / \| \beta^{(t)} \|^2_{L^2(\Omega)}$. Entries highlighted in red denote non-convergent cases.

We observe there are two cases where convergence was achieved (the only cells not highlighted in red): (i) when $\hat{\eta}/\hat{\beta} = \mathcal{O}(10^{-1})$ ("large"), $\hat{\beta} = \mathcal{O}(10^{-3})$ ("very small"), and (ii) when $\hat{\eta}/\hat{\beta} = \mathcal{O}(10^{-3})$ ("very small"), $\hat{\beta} = \mathcal{O}(10^{-1})$ ("large"). We note that for accurate and stable results we require either $\hat{\eta}/\hat{\beta}$ to be small or $\hat{\beta}$ to be very small ($\mathcal{O}(10^{-3})$), however when both are very small ($\mathcal{O}(10^{-3})$) the results are non-convergent. We also note that convergence fails when both are large ($\mathcal{O}(10^{-1})$). Additionally, we see that when $\hat{\eta}/\hat{\beta} = \mathcal{O}(10^{-2})$, $\hat{\beta} = \mathcal{O}(10^{-1})$ the bathymetry reconstruction error is relatively low, despite the kappa test error being higher than the set tolerance, suggesting that the latter may be relaxed to permit $|\kappa(\varepsilon)-1| = \mathcal{O}(10^{-2})$.

Consequently the only admissible cases where the reconstruction error converges are when (i) $\hat{\eta}/\hat{\beta} \leq \mathcal{O}(10^{-2})$, or (ii) $\hat{\beta}$ is small enough that the ratio between $\hat{\eta}$ and H is $\mathcal{O}(10^{-3})$ or smaller as a consequence.

Having explored the relationship between $\hat{\eta}/\hat{\beta}$ and convergence, we now consider the effects of the amplitude of the bathymetry $\hat{\beta}$ when $\hat{\eta}$ is fixed. We analyse convergence of algorithm 3 for amplitudes ranging from 1% of the average depth to 30%. The results are summarised in figure 3.5, and we have included analysis for cases II and III for more insight. In all cases we fix $\hat{\eta}$ to be 0.001, and $N_{obs} = 45$. We see that for case I, figure 3.5(a) indicates that the error is highest when $\hat{\beta}$ is small, however it shows a steady decrease even when $\hat{\beta}$ is 30% of H. Case III shows a similar trend, even though the error is two orders of magnitude smaller than for case I. For case II, the error remains stable at approximately $\mathcal{O}(10^{-2})$ for all values of $\hat{\beta}$.

As bathymetry effects on the surface wave decrease when the fluid depth h is large, it is reasonable that higher amplitudes of bathymetry lead to a slight decrease in the reconstruction error, as observability of the bathymetry by the surface wave may improve. However, to consider the effects of higher amplitudes of the bathymetry in tsunami models, it is logical to analyse the error in the surface wave error given the reconstructed bathymetry as $\hat{\beta}$ increases. Results for this analysis are given in section 3.6.3.

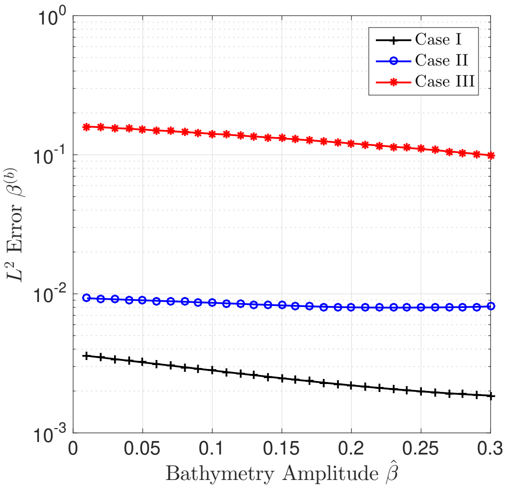


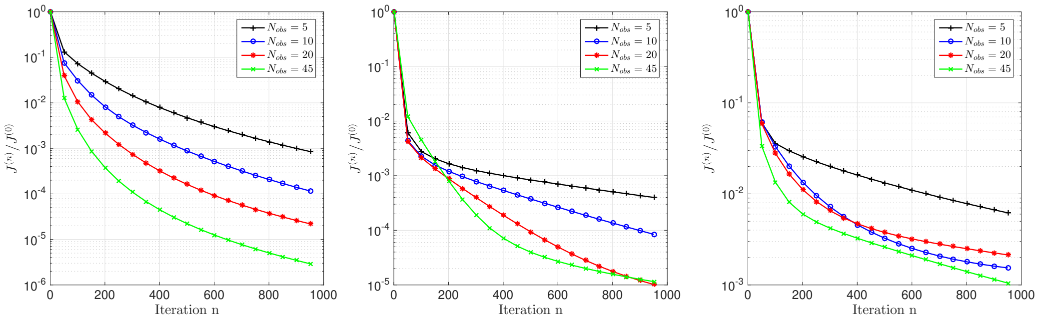
Figure 3.5: The relative error in the bathymetry reconstruction $\| \beta^{(t)} - \beta^{(b)} \|_{L^2(\Omega)} / \| \beta^{(t)} \|_{L^2(\Omega)}$ (where $\Omega = [-L, L]$), shown for different amplitudes $\hat{\beta}$, with amplitude of initial conditions $\hat{\eta} = 0.001$. Note that case III has barely converged for any value of bathymetry amplitude.

Based on the results of table 3.3, we suggest that for situations like case I (i.e. a localized surface wave propagating over a compact bathymetry feature of a similar size) a necessary condition for our data assimilation scheme is that $\hat{\eta}/\hat{\beta}$ be at most $\mathcal{O}(10^{-2})$ when bathymetry amplitudes are large ($\mathcal{O}(10^{-1})$). We observed in figure 3.5 that increasing the amplitude of the bathymetry does not effect the error in the bathymetry, however further analysis is needed on the resulting error in surface wave given the reconstructed bathymetry, in order to quantify the effects of bathymetry in tsunami propagation.

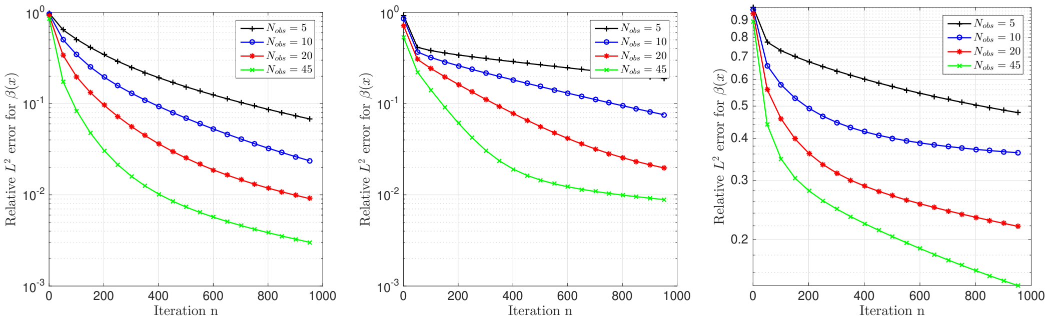
3.6.2 Observation Operator

The observation operator (i.e. the number and location of observations of sea surface height) used in our optimisation scheme is significant, as we have direct control over the number and location of sea surface sensors in real world simulations. We are therefore interested in configuring observation points to optimize convergence to the exact bathymetry.

We consider the effect of the number of observation points on convergence. In fig 3.6 we present the results of algorithm 3 for $5, 10, 20,$ and 25 observation points. Each configuration is a set of equidistant points, with the first point placed at $x = L/10$. For cases I and II the first point is within the support of the Gaussian initial condition. We show the convergence of both the cost function and the relative L^2 error. In all three test cases, the general conclusion is that more observation points produces better convergence. For cases I and II the error decreases by approximately two orders of magnitude as N_{obs} is increased from 5 to 45. However, none of the results for case III converge, and the error remains greater than 10% for all values of N_{obs}.



(a) Case I Cost (b) Case II Cost (c) Case III Cost

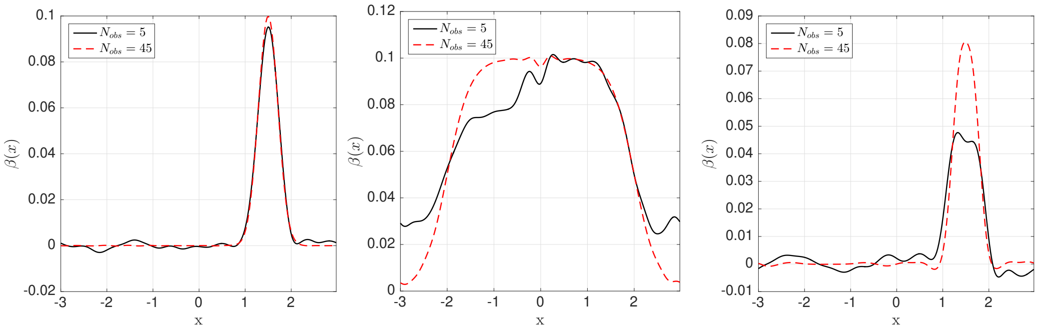


(d) Case I Error (e) Case II Error (f) Case III Error

FIGURE 3.6: Relative cost function and relative L^2 error for different numbers of observation points.

To better understand the effects of N_{obs}, in figure 3.7 we present the reconstructed bathymetry for each case, with $N_{obs} = 5$ and $N_{obs} = 45$ respectively. We observe in figure 3.7(a) that even with $N_{obs} = 5$, the bathymetry is well-resolved, and increasing the number of observations successfully eliminates some small-scale noise at the base of the Gaussian.

In figure 3.7(b) it is clear that $N_{obs} = 5$ is not enough to resolve the bathymetry shape, especially for the bathymetry features in the region $x < 0$. However, increasing N_{obs} to 45 leads to accurately reconstructed bathymetry. The only significant error is a small amount of noise at the plateau of the sandbar profile near $x = 0$. For case III, we note that while neither of the reconstructed bathymetries in figure 3.7(c) converge to the true bathymetry in figure 3.1(c), convergence is much better qualitatively for $N_{obs} = 45$. However the peak of the Gaussian is not fully resolved and the amplitude of the reconstruction with $N_{obs} = 45$ is smaller than the amplitude of the true bathymetry.



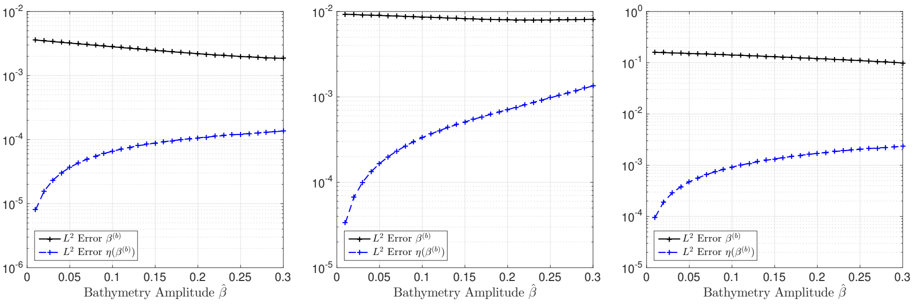
(a) Case I (b) Case II (c) Case III

Figure 3.7: Reconstructed bathymetry for cases I, II and III with $N_{obs} = 5$ and $N_{obs} = 45$ respectively.

Based on these results, we conclude that a sufficiently number of observations N_{obs} is a necessary condition for optimal convergence. However, as indicated by the reconstructions in figure 3.7(c), it is not a sufficient condition for convergence. Despite this, results in figure 3.7(a) demonstrate that observations of a free surface wave with Gaussian initial conditions are able to reconstruct both a Gaussian bathymetry and a Sandbar bathymetry with relatively small error in convergence. The question remains, how does this reconstruction error affect prediction of the free surface wave? In the context of tsunami modelling, our priority is accurate prediction of the free surface wave given the reconstructed bathymetry. Therefore, motivated by results observed thus far, in section 3.6.3 we analyse the L^2 error in the surface wave given the reconstructed bathymetry, as N_{obs} is increased. Additionally, a rigorous sensitivity analysis of the surface wave error to observations is presented in Chapter 4 using second order variational techniques, and in Chapter 5 using Global Sensitivity Analysis (GSA). The objective of this analysis is to explore further not only the effect of the number of observation, but also the position of the observations in the spatial domain.

3.6.3 Sensitivity of propagating surface wave to to bathymetry reconstruction error

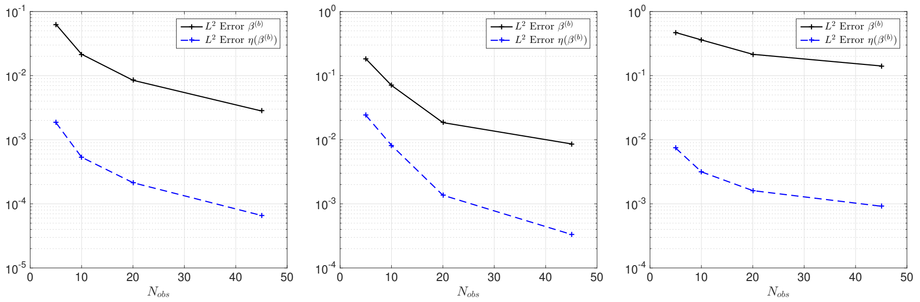
Our final analysis concerns the sensitivity of the surface wave to the bathymetry, in particular, to errors in the bathymetry data. Given some optimal reconstruction $\beta^{(b)}$, we wish to gauge the sensitivity of the propagating surface wave $\eta(x,t)$ to the errors in the reconstruction. Our objective is to address the question: how does the reconstruction error in the bathymetry impact the propagating surface wave? If our goal is to utilise reconstructions of the bathymetry to generate more accurate predictions of tsunami waves, then the main consideration is not the error between the optimal bathymetry reconstruction and the true bathymetry, but the error resulting error in $\eta(x,t)$. In other words, we just need the reconstructed bathymetry to be "good enough" to produce sufficiently accurate surface waves. Thus far we have qualitatively assessed the convergence to the exact bathymetry as we varied the observation operator, and the amplitudes of the initial conditions and bathymetry $\hat{\eta}$ and $\hat{\beta}$. We now address the corresponding error in the surface wave given the reconstructed bathymetry $\eta(x,t;\beta^{(o)})$ as a function of $\hat{\eta}$, $\hat{\beta}$, and N_{obs}.



(a) Case I (b) Case II (c) Case III

Figure 3.8: The relative L^2 error in the bathymetry reconstruction, shown for different amplitudes $\hat{\beta}$, and the corresponding relative L^2 error in the propagating surface wave $\eta(x,t)$. The amplitude of the initial conditions $\hat{\eta}$ is 0.001, and $N_{obs} = 45$.

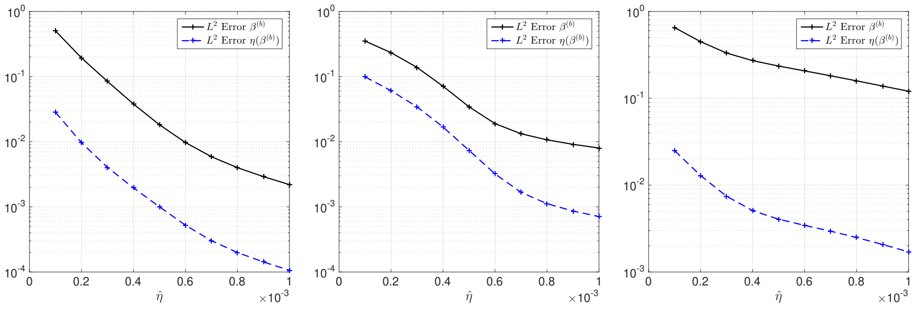
The results for each are summarised in figures 3.8 , 3.9, and 3.10 respectively. The L^2 error in the bathymetry is plotted alongside the resulting L^2 error (in space and time) in the surface wave. Figure 3.8 shows the respective errors as a function of bathymetry amplitude, where $\hat{\eta} = 0.001$. In each of the three cases, we observe the error in the surface wave for $\hat{\beta} \lesssim 0.1$ is almost 2 orders of magnitude lower than the error in the bathymetry estimation. As the relative amplitude of the bathymetry $\hat{\beta}$ increases, so does the surface wave error, even though the bathymetry error remains relatively constant. This is intuitive, as we observe that since the effect of bathymetry decreases as the fluid depth increases (a consequence of smaller amplitudes $\hat{\beta}$), the height $\eta(x, t)$ is affected more by larger bathymetry amplitudes. As such, errors in the reconstruction of larger bathymetry perturbations are more likely to be observed in the resulting surface wave. Therefore we consider $\hat{\beta} \leq 0.1$ a necessary condition for optimal convergence of the surface wave. Overall the results suggest there is low sensitivity of the surface wave to bathymetry reconstruction errors as $\hat{\beta}$ varies.



(a) Case I (b) Case II (c) Case III

FIGURE 3.9: The relative L^2 error in the bathymetry reconstruction, shown for different values of N_{obs}, and the corresponding relative L^2 error in the propagating surface wave $\eta(x, t)$. The amplitude of bathymetry $\hat{\beta}$ is 0.1. Amplitude of the initial conditions $\hat{\eta}$ is fixed to be 1% of $\hat{\beta}$

Figure 3.9 presents the L^2 error in the bathymetry alongside the resulting L^2 error in the surface wave, as a function of the number of observation points. We consider $N_{obs} = 5, \ 10, \ 20,$ and 45 and plot the resulting errors for each of the three cases. $\hat{\beta}$ is fixed at 0.1 and $\hat{\eta} = 0.01\hat{\beta}$. We observe that for cases I and III, the error in the surface wave is almost two orders of magnitude smaller than the bathymetry error, and both decrease proportionally as the number of observation points increases. For case II, the difference in the errors is $\mathcal{O}(10^{-1})$ for smaller values of N_{obs}, however this difference increases as we increase the number of the observation points, suggesting the sensitivity of the surface wave decreases with more observation points for a sandbar profile. A reason for this may be that for case II there are no observation points places before the support of the bathymetry, and consequently a smaller value of N_{obs} may not be sufficient to create a an accurate profile, resulting in increased error in the reconstruction (as observed in figure 3.7). Consequently, the error in the surface wave increases as well. However, as more observation points are added in figure 3.7 the noise in the bathymetry is smoothed, and the smoother profile may explain the relative larger decrease in the surface wave error, up to two orders of magnitude smaller than the bathymetry error.



(a) Case I (b) Case II (c) Case III

Figure 3.10: The relative L^2 error in the bathymetry reconstruction and the corresponding relative L^2 error in the propagating surface wave $\eta(x,t)$,as a function of the initial conditions amplitude $\hat{\eta}$. The amplitude of the bathymetry is fixed to be 0.2.

Finally, figure 3.10 shows the bathymetry reconstruction and surface wave errors as a function of the initial conditions amplitude $\hat{\eta}$. We have already observed a correlation in the amplitudes $\hat{\eta}$ and $\hat{\beta}$, and we concluded in section 3.6.1 for optimal convergence we require $\hat{\eta}/\hat{\beta} = \mathcal{O}(10^{-2})$. Therefore in figure c we present results for small values of $\hat{\eta}$, and $\hat{\beta} = 0.2$. For each case we observe that the error in the surface wave is orders of magnitude lower than the bathymetry error for most values of $\hat{\eta}$. Both errors are relatively high for smaller values of $\hat{\eta}$, but as $\hat{\eta} \to 10^{-3}$, they both decrease, suggesting that the free surface wave is more sensitive to bathymetry reconstruction error when $\hat{\eta}$ is very small.

To summarise, we observed that across variations in (i) the number of observation points, and (ii) model parameters such as $\hat{\eta}$ and $\hat{\eta}$, the error in the surface wave was orders of magnitude lower than the bathymetry reconstruction error, suggesting low sensitivity of the surface wave to bathymetry reconstruction error. The consequences of this low sensitivity are significant. It essential to understand the sensitivity of the surface wave to the bathymetry for two distinct but related reasons. First, the assimilation described in this study is contingent upon the assumption that the surface wave (and observations) are affected by the bathymetry. We have demonstrated this is true in figure 3.1. Additionally, the variation in the surface wave error in figure 3.8 illustrates that changes in bathymetry amplitude and shape does significantly affect surface wave propagation. However, we observe that the sensitivity of the surface wave to the bathymetry reconstruction error is low. This means that the surface wave is not sensitive to the details of the bathymetry: a roughly accurate reconstruction of bathymetry is sufficient in many cases. This suggests that for tsunami modelling we only need the bathymetry to be accurate enough that the surface wave is modelled correctly, and that we can derive criteria for when the bathymetry reconstruction is reconstructed sufficiently well. Additionally, we also do not require our optimal reconstruction $\beta^{(b)}$ to be unique. These implications motivate a more rigorous analysis of surface wave sensitivity, and we explore this further in chapters 2 and 3.

3.7 Conclusion

This study provides a first step in better understanding the role of observations and model parameters in variational bathymetry assimilation. We are limited somewhat by the lack of analytical solution for the shallow water system with non-zero bathymetry, but these computational results provide key insights for further analyses. The 1-D geometry considered in this analysis is intended to be a foundation for extensions to more realistic 2-D cases, and the results observed here should prove an insightful reference for future work.

In summary, we derived a variational data assimilation algorithm to reconstruct the bathymetry from a set of free surface wave observations, by minimising a functional $\mathcal{J}(\beta)$ representing the least squares error between observations and forecast solutions. We observed that $\nabla\mathcal{J} \in L^2(\Omega)$ resulted in small-scale noise in the bathymetry reconstruction and impacted convergence. Consequently, we showed that smoother H^2 gradients are necessary to avoid small-scale errors in the reconstructed bathymetry. We were able to accurately reconstruct the bathymetry for test cases with a Gaussian initial condition, and (i) a Gaussian bathymetry, and (ii) a sandbar profile bathymetry. We analysed the relationship between the normalised bathymetry and initial conditions amplitudes $\hat{\beta}$ and $\hat{\eta}$ to understand how they influence convergence. Based on a qualitative analysis of a localized surface wave propagating over a compact bathymetry feature of a similar size, we suggest a necessary condition for convergence is that the ratio of the surface wave amplitude to the bathymetry amplitude $\hat{\eta}/\hat{\beta}$ be at most $\mathcal{O}(10^{-2})$ when bathymetry amplitudes are $\mathcal{O}(10^{-1})$. Additionally, we observe that a relatively large number of observations N_{obs} is necessary for convergence. $N_{obs} = 45$ was the optimal number of observations in the present study.

Perhaps the most significant observation of this study was that that the surface wave $\eta(\beta^{(b)})$ has relatively low sensitivity to errors in the reconstructed bathymetry. We showed that the free surface error was orders of magnitude smaller than the bathymetry reconstruction error as a function of N_{obs}, $\hat{\beta}$ and $\hat{\eta}$ respectively. Reconstructing the bathymetry with a relative error of about 10% is sufficiently accurate for surface wave modelling in most cases.

If this can be verified in higher dimensions and with inclusion of additional complexities such as turbulence and Coriolis effect, this can enhance tsunami forecast models by quantifying exact tolerance levels for the prediction. Additionally, tolerances for smaller scales in bathymetry reconstruction may be quantified.

As the present analysis was primary qualitative, in chapters 2 and 3 we conduct more rigorous sensitivity analyses to verify these results. Our objective is to derive analytical results that quantify the sensitivity of response functions (such as the error in the surface wave given the reconstructed bathymetry) to the observations, or to the bathymetry and the initial conditions amplitudes. In Chapter 4, our analysis is based on works like Wang et al. [63], using second order adjoint methods (i.e. the Hessian of the cost function $\mathcal{J}$) to derive the sensitivity of response functions to observations. In Chapter 5, we use an alternative technique, Global Sensitivity Analysis (GSA), to decompose the variance of the bathymetry and surface wave error respectively, and generate sensitivity indices quantifying the influence of the initial conditions and bathymetry amplitudes, and the position of the bathymetry relative to observations.

Our focus on how the bathymetry and the initial conditions shapes affect reconstruction of the bathymetry from free surface observations is a valuable contribution to the literature reviewed in section 3.2. Additionally, we hope the sensitivity analyses and results observed here can pave the way for more refined approaches to incorporate bathymetry estimates in tsunami models. Extending this approach to 2-D and verifying the main conclusions observed here would contribute to efforts for accurate bathymetry in tsunami models. Similar analyses are presented in Chapter 2, where we have extended the 1-D variational data assimilation for the initial conditions reconstruction in Kevlahan et al. [27], to 2-D. we observe that some of the main conclusions from 1-D are still relevant in the 2-D case. It is possible that this may be true for the 2-D bathymetry assimilation as well.

Chapter 4

Second order adjoint sensitivity analysis for variational data assimilation of tsunami models

4.1 Introduction

Data assimilation is integral to accurate climate, atmosphere and ocean modelling. Variational data assimilation algorithms such as 3D-VAR, 4D-VAR, and Kalman filtering techniques like the EnKF (Ensemble Kalman filter) are regularly used for numerical weather prediction and forecasts of climate trends. For example, data assimilation is used by the ECMWF for climate reanalysis, where archived observations are “reanalysed”, in order to create a comprehensive global dataset describing the recent history of the earth’s climate, atmosphere and oceans [12]. Additionally, data assimilation is used in tsunami forecast models, where observations of surface waves are used to reconstruct missing information such as initial conditions, and subsequently predict impact at coastlines [35]. The observations have a significant impact on the results of the data assimilation scheme, as demonstrated in Chapter 3. Therefore, it is interesting to quantify the sensitivity of the assimilation algorithm results to perturbations in the observation operator. In this study, we use methods outlined in Shutyaev et al. [53] and Shutyaev et al. [52], to derive the sensitivity of arbitrary response functions to perturbations in observations for (i) the data assimilation scheme for initial conditions outlined in

Kevlahan et al. [27], and (ii) the data assimilation scheme for bathymetry given in chapter 3. We do this by analytically deriving the Hessian of the cost function minimised in the data assimilation scheme, and take advantage of its properties to derive expressions for the sensitivity of the data assimilation to perturbations in the observations. We present a numerical implementation of this algorithm for the bathymetry data assimilation. The focus is on the latter instead of the initial conditions assimilation, as our main objective is to use the present analysis to further investigate the qualitative results observed in chapter 3. Specifically, these are (i) the link between low sensitivity and the accuracy of the data assimilation scheme, and (ii) the effect of perturbations in the model parameters and the observations on the sensitivity. Ultimately we wish to see whether the conclusions from the bathymetry assimilation in chapter 3 can be better understood using the results of the sensitivity analysis in this chapter.

In section 4.2 we summarise the data assimilation schemes. Section 4.4 gives the analytical derivation of the Hessian and subsequent sensitivity analysis for initial conditions assimilation results, and section 4.5 gives the derivation for bathymetry assimilation results. Section 4.6 presents the numerical implementation of the sensitivity analysis for the bathymetry data assimilation. Finally, section 4.7 summarises the main results and suggests further considerations for future analyses.

4.2 Initial conditions reconstruction

We briefly summarise the data assimilation scheme for initial conditions reconstruction. The 1-D SWE are

$$\frac{\partial \eta}{\partial t} + \frac{\partial}{\partial x}\Big((1+\eta-\beta)u\Big) = 0, \tag{4.2.1a}$$

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left(\frac{1}{2}u^2+\eta\right) = 0, \tag{4.2.1b}$$

$$\eta(x,0) = \ \phi(x), \tag{4.2.1c}$$

$$u(x,0) = \ 0. \tag{4.2.1d}$$

We assimilate a set of measurements $y^{(o)}(t)$, which are observations of the true height perturbation $\eta(x,t)$ at positions $\{x_j\}, j = 1, ..., N_{obs}$. We assume that the initial conditions $\phi(x)$ are compactly supported, and that we have periodic boundary conditions on some domain $\Omega = \Big\{x \in \mathbb{R}; -L \leq x \leq L\Big\}$. In the initial conditions reconstruction, the system is integrated in time for $t \in [0\ T]$, where $T = 2$ is chosen such that the free surface wave does not reach the boundary. This is appropriate considering that our objective is to reconstruct the initial conditions of a tsunami wave in real time.

We assume that we do not have complete information about $\phi(x)$, and our objective is to minimise the error between the forecast solution $\eta^{(f)}(x,t)$ given some guess for ϕ, and the observations $y^{(o)}(t)$. We define this error in terms of a cost function,

$$\mathcal{J}(\phi) = \frac{1}{2}\int_0^T \sum_{i=1}^{M} \Big[H\eta^{(f)}(x_j, t; \phi) - y_j^{(o)}(t)\Big]^2 \, dt, \tag{4.2.2}$$

where H is the linear operator from the state to the observation space. This is equivalent to solving

$$\nabla\mathcal{J}^{L^2}(\phi^{(b)}) = 0, \tag{4.2.3}$$

Where $\phi^{(b)}$ is the optimal approximation for $\phi(x)$, and the local minimiser of (4.2.2). We formulate a Lagrangian constrained by (4.2.1) and some arbitrarily chosen adjoint variables (Lagrange multipliers) (η^*, u^*) that are solutions of

$$\begin{cases} \dfrac{\partial \eta^*}{\partial t} + u\dfrac{\partial \eta^*}{\partial x} + \dfrac{\partial u^*}{\partial x} = H\Big(\eta^{(f)}(x,t;\phi) - y^{(o)}(t)\Big), & \text{(4.2.4a)} \\ \dfrac{\partial u^*}{\partial t} + (1+\eta)\dfrac{\partial \eta^*}{\partial x} + u\dfrac{\partial u^*}{\partial x} = 0, & \text{(4.2.4b)} \\ \eta^*(x,T) = 0, & \text{(4.2.4c)} \\ u^*(x,T) = 0. & \text{(4.2.4d)} \end{cases}$$

We use the Riesz representation theorem and the Gateaux derivative representation of $\mathcal{J}'(\phi;\eta')$ given an arbitrary perturbation η' to derive

$$\nabla^{L^2}\mathcal{J}(\phi) = -\eta^*(x,0), \tag{4.2.5}$$

where

$$\begin{aligned}\mathcal{J}'(\phi;\eta') &= \langle \nabla \mathcal{J}(\phi), \eta' \rangle_{L^2(\Omega)} \\ &= \int_{-L}^{L} \nabla^{L_2}\mathcal{J}(\phi)\ \eta'\ dx. \end{aligned} \tag{4.2.6}$$

For a more detailed analysis of this derivation we refer the reader to Kevlahan et al. [27]. Finally, the optimal reconstruction of the initial conditions $\phi^{(b)}$ is where

$$\nabla^{L^2}\mathcal{J}(\phi^{(o)}) = -\eta^*(x,0) = 0. \tag{4.2.7}$$

In the assimilation scheme we use a gradient descent algorithm to iteratively find the optimal reconstruction of the initial conditions $\phi^{(b)}$ given some initial guess, such that (4.2.2) is minimised.

4.3 Bathymetry reconstruction

We recall from Chapter 3 that the data assimilation scheme to recover missing bathymetry information $\beta(x)$ in (4.2.1) is similar to the initial conditions case, except now $\phi(x)$ is known, and the cost function we minimise is

$$\mathcal{J}(\beta) = \frac{1}{2}\int_0^T \sum_{i=1}^{M} \left[\eta^{(f)}(x_j,t;\beta) - y_j^{(o)}(t)\right]^2 dt. \tag{4.3.1}$$

We derived the adjoint system

$$\begin{cases} \dfrac{\partial \eta^*}{\partial t} + u\dfrac{\partial \eta^*}{\partial x} + \dfrac{\partial u^*}{\partial x} = H\Big(\eta^{(f)}(x,t;\beta) - y^{(o)}(t)\Big) \qquad , & (4.3.2a) \\ \dfrac{\partial u^*}{\partial t} + (1+\eta-\beta)\dfrac{\partial \eta^*}{\partial x} + u\dfrac{\partial u^*}{\partial x} = 0, & (4.3.2b) \\ \eta^*(x,T) = 0, & (4.3.2c) \\ u^*(x,T) = 0. & (4.3.2d) \end{cases}$$

We found the following equivalent expression for the Gâteaux derivative,

$$\begin{aligned}\mathcal{J}'(\beta;\beta') &= \langle \nabla \mathcal{J}(\beta), \beta' \rangle_{L^2(\Omega)}\\ &= \int_{-L}^{L} \nabla^{L_2}\mathcal{J}(\beta)\ \beta'\ dx\\ &= \int_{-L}^{L}\int_0^T u\frac{\partial \eta^*}{\partial x}\beta'\ dt\ dx. \end{aligned} \tag{4.3.3}$$

Therefore the Riesz representation of the gradient is

$$\nabla^{L^2}\mathcal{J} = \int_0^T u\frac{\partial \eta^*}{\partial x}\ dt, \tag{4.3.4}$$

and the optimal approximation for $\beta(x)$, defined as $\beta^{(b)}$ is where

$$\nabla^{L^2}\mathcal{J}(\beta^{(b)}) = \int_0^T u\frac{\partial \eta^*}{\partial x}\ dt = 0. \tag{4.3.5}$$

For the bathymetry reconstruction, as we have no constraints on the assimilation time, the final time is chosen to be $T = 2L$.

4.4 Initial conditions reconstruction sensitivity analysis

To derive the sensitivity of some arbitrary response function using the methods outlined in Shutyaev et al. [53] and Shutyaev et al. [52], we need to formulate expressions for the Hessian of the cost functions (4.2.2) and (4.3.1) formulated in section 4.2. [53, 52] give a general method that utilises properties of the Hessian, however they do not provide a derivation, which means we need to extend our variational adjoint analysis used to find the gradient of $\mathcal{J}$, to find the Hessian of $\mathcal{J}$.

While works such as Wang et al. [63] provide a derivation of the Hessian vector product for the initial conditions assimilation, their derivation is for the finite dimensional case, and assumes a vector form for both the state variables and the control variable. In our case the derivation of the first order adjoint is for the infinite dimensional case in the space $L^2(\Omega)$ over some domain Ω, where we used the L^2 inner product and the Riesz respresentation theorem to extract our gradient $\nabla^{L^2}\mathcal{J}(\phi)$. For that reason, a derivation of the Hessian in the same functional space as for our first order adjoint is appropriate, and in doing so we aim to extract the "Gâteaux Hessian" for Hilbert spaces with the following definition:

If f is twice Gâteaux differentiable at x, we can identify $D^2f(x)$ with the operator $\nabla^2 f(x) \in \mathcal{B}(\mathcal{H})$ in the sense that

$$(\forall y \in \mathcal{H})(\forall z \in \mathcal{H}),$$
$$\Big(D^2 f(x)y\Big)z = \Big\langle z, \nabla^2 f(x)y\Big\rangle_{\mathcal{B}(\mathcal{H})}, \tag{4.4.1}$$

where we call $\nabla^2 f(x)$ the (Gâteaux) Hessian of f at x, $\mathcal{B}(\mathcal{H})$ is the space of continuous linear functionals in $\mathcal{H}$, and $Df(x)y$ is the Gâteaux derivative of f in the direction y.

In the remainder of this section we present the derivation of the Hessian of $\mathcal{J}(\phi)$ and subsequently the sensitivity analysis for the initial conditions assimilation, using methods outlined in [53]. Parallel results for bathymetry assimilation are presented in section 4.5 following the methods given in [52].

4.4.1 Hessian of $\mathcal{J}(\phi)$ for the initial conditions reconstruction

To derive a form for the Hessian in a Hilbert space we use the fact that we derived the following form for the Gâteaux derivative of $\mathcal{J}$ with respect to the initial conditions and some perturbation direction η' in Kevlahan et al. [27],

$$\mathcal{J}'(\phi;\eta') = -\int_{-L}^{L} \eta^*(x,0;\phi)\eta'\,dx. \tag{4.4.2}$$

Then, if we consider a second perturbation of $\mathcal{J}'(\phi;\eta')$, η'' where we have $\phi \to \phi+\varepsilon\eta''$, the second order Gâteaux derivative of $\mathcal{J}$ can be expressed as

$$\begin{aligned}\mathcal{J}''(\phi;\eta';\eta'') =& \lim_{\varepsilon\to 0}\frac{\mathcal{J}'(\phi+\varepsilon\eta'';\eta')-\mathcal{J}'(\phi;\eta')}{\varepsilon}\\ =&\frac{d}{d\varepsilon}\mathcal{J}'(\phi+\varepsilon\eta'';\eta')\Big|_{\varepsilon=0}\\ =&\frac{d}{d\varepsilon}\Big\{-\int_{-L}^{L}\eta^*(x,0;\phi+\varepsilon\eta'')\eta'\ dx\Big\}.\end{aligned}$$

We consider a regular perturbation expansion of $\eta^*(x,0;\phi+\varepsilon\eta'')$, approximating it by the series $f_0+f_1\varepsilon+\mathcal{O}(\varepsilon^2)$. We can see this is equivalent to a Taylor expansion about $\varepsilon=0$. Then

$$\begin{aligned}\frac{d}{d\varepsilon}\mathcal{J}'(\phi+\varepsilon\eta'';\eta')\Big|_{\varepsilon=0} =& -\frac{d}{d\varepsilon}\int_{-L}^{L}(f_0+f_1\varepsilon)\eta'\ dx\Big|_{\varepsilon=0}\\ =& -\int_{-L}^{L}f_1\eta'\ dx.\end{aligned}$$

In order to find the term f_1, we recognise that this is equivalent to the coefficient of the linear term in the Taylor approximation, and $f_1=\frac{d}{d\varepsilon}\eta^*(x,0;\phi+\varepsilon\eta'')\Big|_{\varepsilon=0}$.

To find f_1, we need the adjoint variable η^* at $t=0$ given the perturbed initial conditions $\phi+\varepsilon\eta''$. Substituting this back into our forward system (4.2.1)and adjoint system (4.2.4), and gathering terms of $\mathcal{O}(\varepsilon)$ will give us our perturbed SWE and second order adjoint (SOA) system. We assume this perturbation in the initial conditions brings about the following perturbations to our state and adjoint variables;

- $u\to u+\varepsilon\hat{u}$ and $\eta\to\eta+\varepsilon\hat{\eta}$ for the shallow water system (4.2.1).
- $u^*\to u^*+\varepsilon\bar{u}$ and $\eta^*\to\eta^*+\varepsilon\bar{\eta}$ for the adjoint system (4.2.4).

The resulting perturbed model for the SWE is

$$\begin{cases} \dfrac{\partial \hat{\eta}}{\partial t} + \dfrac{\partial}{\partial x}\Big(u\hat{\eta}\Big) + \dfrac{\partial}{\partial x}\Big((\eta+1)\hat{u}\Big) = 0, & \text{(4.4.3a)} \\ \dfrac{\partial \hat{u}}{\partial t} + \dfrac{\partial\Big(\hat{u}u\Big)}{\partial x} + \dfrac{\partial \hat{\eta}}{\partial x} = 0, & \text{(4.4.3b)} \\ \hat{\eta}(x,0) = \eta'', & \text{(4.4.3c)} \\ \hat{u}(x,0) = 0, & \text{(4.4.3d)} \end{cases}$$

and the second order adjoint (SOA) model is

$$\begin{cases} \dfrac{\partial \bar{\eta}}{\partial t} + \hat{u}\dfrac{\partial \eta^*}{\partial x} + u\dfrac{\partial \bar{\eta}}{\partial x} + \dfrac{\partial \bar{u}}{\partial x} = -H\hat{\eta}, & \text{(4.4.4a)} \\ \dfrac{\partial \bar{u}}{\partial t} + (\eta+1)\dfrac{\partial \bar{\eta}}{\partial x} + \hat{\eta}\dfrac{\partial \eta^*}{\partial x} + u\dfrac{\partial \bar{u}}{\partial x} + \hat{u}\dfrac{\partial u^*}{\partial x} = 0, & \text{(4.4.4b)} \\ \bar{\eta}(x,T) = 0, & \text{(4.4.4c)} \\ \bar{u}(x,T) = 0. & \text{(4.4.4d)} \end{cases}$$

Going back to the definition of f_1, we can rewrite this as

$$\begin{aligned} f_1 =& \frac{d}{d\varepsilon}\Big(\eta^*(x,0;\phi) + \varepsilon\bar{\eta}(x,0;\eta'')\Big)\Big|_{\varepsilon=0} \\ =& \ \bar{\eta}(x,0;\eta''), \end{aligned}$$

And thus we have

$$\mathcal{J}''(\eta_0;\eta';\eta'') = -\int_{-L}^{L} \bar{\eta}(x,0;\eta'')\eta' \ dx = \tag{4.4.5}$$

$$\langle -\bar{\eta}(x,0;\eta''), \ \eta'\rangle_{L^2(\Omega)} \tag{4.4.6}$$

and by the definition of the Gâteaux Hessian, we get $\nabla^2\mathcal{J}(\eta_0;\eta'') = -\bar{\eta}(x,0;\eta'')$.

We define successive solutions of (4.4.3) and (4.4.1) by the operator $\mathcal{H}$ acting on η''. Thus the Hessian of the cost function given a perturbation η'' can be defined as

$$\mathcal{H}\eta'' = \ -\bar{\eta}(x,0;\eta''). \tag{4.4.7}$$

This form of the Hessian of the cost function can be verified using a modified form of the kappa test analysis outlined in Kevlahan et al. [27]. To ensure our derivation is correct, we define

$$\kappa(\varepsilon) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \frac{\mathcal{J}'(\phi + \varepsilon\eta''; \eta') - \mathcal{J}'(\phi; \eta')}{-\int_{-L}^{L} \bar{\eta}(x, 0; \eta'')\eta' \, dx}, \tag{4.4.8}$$

where

$$\mathcal{J}'(\eta_0 + \varepsilon\eta''; \eta') \tag{4.4.9}$$

$$= \lim_{\tau \to 0} \frac{\mathcal{J}'(\phi + \varepsilon\eta'' + \tau\eta') - \mathcal{J}'(\phi + \varepsilon\eta'')}{\tau}, \tag{4.4.10}$$

and

$$\mathcal{J}'(\phi; \eta') = \lim_{\tau \to 0} \frac{\mathcal{J}'(\phi + \tau\eta') - \mathcal{J}'(\phi)}{\tau}. \tag{4.4.11}$$

If the derivations are correct, $\kappa(\varepsilon)$ should converge to 1 as $\varepsilon \to 0$.

A key observation is that we have only derived the action of the Hessian on some perturbation η'', whereas in section 4.2 we were able to derive the gradient of $\mathcal{J}(\phi)$ for any arbitrary perturbation η'. It does not seem possible to proceed as before and find the Hessian for any arbitrary η'' using variational methods. This is because the gradient $\nabla\mathcal{J}$ is an element of the Hilbert space $L^2(\Omega)$, whereas the Hessian $\mathcal{H}$ is an operator defined on this space. Indeed, Wang et al. [63] were also limited to derivation of a “Hessian vector product” in the finite dimensional case. Despite this, the current derivation is sufficient for the following sensitivity analysis.

4.4.2 Sensitivity analysis for the initial conditions reconstruction

Define the optimality system as the successive solution of

$$\begin{cases} \dfrac{\partial \eta}{\partial t} + \dfrac{\partial}{\partial x}\Big((1+\eta)u\Big) = 0, & (4.4.12a) \\ \dfrac{\partial u}{\partial t} + \dfrac{\partial}{\partial x}\Big(\dfrac{1}{2}u^2 + \eta\Big) = 0, & (4.4.12b) \\ \eta(x,0) = \ \phi(x), & (4.4.12c) \\ u(x,0) = \ 0. & (4.4.12d) \end{cases}$$

$$\begin{cases} \dfrac{\partial \eta^*}{\partial t} + u\dfrac{\partial \eta^*}{\partial x} + \dfrac{\partial u^*}{\partial x} = H\Big(\eta - m\Big), & (4.4.13a) \\ \dfrac{\partial u^*}{\partial t} + (1+\eta)\dfrac{\partial \eta^*}{\partial x} + u\dfrac{\partial u^*}{\partial x} = \ 0, & (4.4.13b) \\ \eta^*(x,T) = \ 0, & (4.4.13c) \\ u^*(x,T) = \ 0, & (4.4.13d) \end{cases}$$

Where $\phi(x)$ is the optimal reconstruction of the initial conditions $\eta(x,0)$ giving

$$\nabla^{L^2}\mathcal{J}(\phi^{(b)}) = -\eta^*(x,0) = 0, \tag{4.4.14}$$

and $m(t)$ are the observations taken at positions $\{x_j\}$ at for $j = 1, ..., N_{obs}$ at continuous times t. Note that for the initial conditions reconstruction we assume a flat bathymetry $\beta(x) = 0$. Let us consider some arbitrary response function

$$\mathcal{G}(\eta, u, \phi). \tag{4.4.15}$$

This function may represent any quantity of interest given the optimal results of the data assimilation. In the numerical implementation we choose $\mathcal{G}$ to be the least squares error in the free surface wave given the optimal control variable, as shown in (4.6.1).

By the chain rule, the sensitivity of $\mathcal{G}$ to perturbations in the observations m can be defined as

$$\frac{d\mathcal{G}}{dm} = \frac{d\mathcal{G}}{d\eta}\frac{d\eta}{dm} + \frac{d\mathcal{G}}{du}\frac{du}{dm} + \frac{d\mathcal{G}}{d\phi}\frac{d\phi}{dm}. \tag{4.4.16}$$

Let us now consider a perturbation in the observations $m \to m + \hat{m}$ giving

- $u \to u + \hat{u}$ and $\eta \to \eta + \hat{\eta}$ for the shallow water system (4.4.12).
- $u^* \to u^* + \tilde{u}^*$ and $\eta^* \to \eta^* + \tilde{\eta}^*$ for the adjoint system (4.4.13).
- $\phi \to \phi + \hat{\phi}$ for the optimal initial condition.

Our perturbed system becomes

$$\begin{cases} \dfrac{\partial \hat{\eta}}{\partial t} + \dfrac{\partial}{\partial x}\Big(u\hat{\eta}\Big) + \dfrac{\partial}{\partial x}\Big((\eta+1)\hat{u}\Big) = 0, & \text{(4.4.17a)} \\ \dfrac{\partial \hat{u}}{\partial t} + \dfrac{\partial\Big(\hat{u}u\Big)}{\partial x} + \dfrac{\partial \hat{\eta}}{\partial x} = 0, & \text{(4.4.17b)} \\ \hat{\eta}(x,0) = \hat{\phi}(x), & \text{(4.4.17c)} \\ \hat{u}(x,0) = 0, & \text{(4.4.17d)} \end{cases}$$

$$\begin{cases} \dfrac{\partial \tilde{\eta}^*}{\partial t} + \hat{u}\dfrac{\partial \eta^*}{\partial x} + u\dfrac{\partial \tilde{\eta}^*}{\partial x} + \dfrac{\partial \tilde{u}^*}{\partial x} = H(\hat{m} - \hat{\eta}(x_i,t;\hat{\phi}), & \text{(4.4.18a)} \\ \dfrac{\partial \tilde{u}^*}{\partial t} + (\eta+1)\dfrac{\partial \tilde{\eta}^*}{\partial x} + \hat{\eta}\dfrac{\partial \eta^*}{\partial x} + u\dfrac{\partial \tilde{u}^*}{\partial x} + \hat{u}\dfrac{\partial u^*}{\partial x} = 0, & \text{(4.4.18b)} \\ \tilde{\eta}^*(x,T) = 0, & \text{(4.4.18c)} \\ \tilde{u}^*(x,T) = 0. & \text{(4.4.18d)} \end{cases}$$

$$-\hat{\eta}^*(x,0) = 0. \tag{4.4.19}$$

Then we can say

$$\left\langle \frac{d\mathcal{G}}{dm}, \hat{m} \right\rangle_{Y_{obs}} = \left\langle \frac{\partial \mathcal{G}}{\partial \eta}, \hat{\eta} \right\rangle_{Y} + \left\langle \frac{\partial \mathcal{G}}{\partial u}, \hat{u} \right\rangle_{Y} + \left\langle \frac{\partial \mathcal{G}}{\partial \phi}, \hat{\phi} \right\rangle_{Y_p} \tag{4.4.20}$$

where

$$Y_{obs} = L^2\Big([-L, L] \times [0, T]\Big), \tag{4.4.21}$$

$$Y = L^2\Big([-L, L] \times [0, T]\Big), \tag{4.4.22}$$

$$Y_p \in L^2\Big([-L, L]\Big), \tag{4.4.23}$$

are the observation space, state space, and the initial conditions space respectively. Now, let us introduce some adjoint variables P_i, $i = 1, ..., 5$, where $P_i \in Y$, $i = 1, ..., 4$, and $P_5 \in Y_{obs}$. Then if we take the inner product of P_1 and P_2 with the systems (4.4.17), inner product of P_3 and P_4 with (4.4.18), and P_5 with (4.4.19), we get the following duality relation

$$\begin{aligned}
0 = & \\
& \int_0^T \int_{-L}^{L} \Bigg\{ P_1 \left[\frac{\partial \hat{\eta}}{\partial t} + \frac{\partial}{\partial x}\Big(u\hat{\eta}\Big) + \frac{\partial}{\partial x}\Big((\eta + 1)\hat{u}\Big) \right] \\
& + P_2 \left[\frac{\partial \hat{u}}{\partial t} + \frac{\partial \Big(\hat{u}u\Big)}{\partial x} + \frac{\partial \hat{\eta}}{\partial x} \right] \\
& + P_3 \left[\frac{\partial \tilde{\eta}^*}{\partial t} + \hat{u}\frac{\partial \eta^*}{\partial x} + u\frac{\partial \tilde{\eta}^*}{\partial x} + \frac{\partial \tilde{u}^*}{\partial x} - H(\hat{m} - \hat{\eta}) \right] \\
& + P_4 \left[\frac{\partial \tilde{u}^*}{\partial t} + (\eta + 1)\frac{\partial \tilde{\eta}^*}{\partial x} + \hat{\eta}\frac{\partial \eta^*}{\partial x} \right. \\
& \left. + u\frac{\partial \tilde{u}^*}{\partial x} + \hat{u}\frac{\partial u^*}{\partial x} \right] \Bigg\} dt dx \\
& + \int_{-L}^{L} P_5 \Big[-\hat{\eta}^*(x, 0) \Big] dx.
\end{aligned} \tag{4.4.24}$$

integrating the double integral in (4.4.24) by parts in space and time, we are able to transfer the derivatives onto the adjoint variables P_i, $i = 1, ..., 5$ instead of on $\hat{u}, \hat{\eta}, \tilde{u}^*, \tilde{u}^*$. Since the choice of adjoint variables is arbitrary, we pick the following systems for P_i,

$$\begin{cases} \dfrac{\partial P_3}{\partial t} + \dfrac{\partial}{\partial x}\Big(uP_3\Big) + \dfrac{\partial}{\partial x}\Big((\eta+1)P_4\Big) = 0, & \text{(4.4.25a)} \\ \dfrac{\partial P_4}{\partial t} + \dfrac{\partial\Big(uP_4\Big)}{\partial x} + \dfrac{\partial P_3}{\partial x} = 0, & \text{(4.4.25b)} \\ P_3(x,0) = -P_5 & \text{(4.4.25c)} \\ P_4(x,0) = 0 & \text{(4.4.25d)} \end{cases}$$

$$\begin{cases} \dfrac{\partial P_1}{\partial t} + P_4\dfrac{\partial \eta^*}{\partial x} + u\dfrac{\partial P_1}{\partial x} + \dfrac{\partial P_2}{\partial x} - HP_3 = -\dfrac{\partial \mathcal{G}}{\partial \eta}, & \\ \dfrac{\partial P_2}{\partial t} + (\eta+1)\dfrac{\partial P_1}{\partial x} + P_3\dfrac{\partial \eta^*}{\partial x} + u\dfrac{\partial P_2}{\partial x} + P_4\dfrac{\partial u^*}{\partial x} = -\dfrac{\partial \mathcal{G}}{\partial u}, & \text{(4.4.26a)} \\ P_1(x,T) = P_5 - \dfrac{\partial \mathcal{G}}{\partial \phi}, & \text{(4.4.26b)} \\ P_2(x,T) = 0. & \text{(4.4.26c)} \end{cases}$$

We subsequently eliminate P_5 by using the fact that $P_1(x,T) = P_5 - \frac{\partial \mathcal{G}}{\partial \phi}$, and $P_3(x,0) = -P_5$, and define $P_3(x,0) = \nu$, where the auxiliary variable ν is defined as

$$\nu = \frac{\partial \mathcal{G}}{\partial \phi} - P_1(x,0). \tag{4.4.27}$$

Then subsequently, as a result of integration by parts and the choice of systems for P_i, (4.4.24) reduces to

$$\Big\langle HP_3, \hat{m} \Big\rangle_{Y_{obs}} = \Big\langle \frac{\partial \mathcal{G}}{\partial \eta}, \hat{\eta} \Big\rangle_Y + \Big\langle \frac{\partial \mathcal{G}}{\partial u}, \hat{u} \Big\rangle_Y + \Big\langle \frac{\partial \mathcal{G}}{\partial \phi}, \hat{\phi} \Big\rangle_{Y_p}. \tag{4.4.28}$$

By the Riesz representation theorem and equivalence of inner products in (4.4.20) and (4.4.28), we can define the sensitivity of the response function $\mathcal{G}(\eta, u, \phi)$ to perturbations in the observations m as

$$\frac{\partial \mathcal{G}}{\partial m} = HP_3(x,t). \tag{4.4.29}$$

Finding P_3 requires solving the systems (4.4.25) (with $P_3(x,0) = \nu$), and (4.4.26).

This is a coupled system of four variables, with two initial time conditions and two final time conditions, making it challenging to solve. However, this difficulty can be avoided by recognising that (4.4.25) is equivalent to the perturbed system for the Hessian $\nabla^2\mathcal{J}(\phi)$ (4.4.3), and that (4.4.26) is equivalent to the second order adjoint (SOA) system (4.4.1) with forcing term $\left(-\frac{\partial\mathcal{G}}{\partial\eta}, -\frac{\partial\mathcal{G}}{\partial u}\right)^T$. Shutyaev et al. (2017) show that the solutions to the adjoint systems (4.4.25) and (4.4.26) are then equivalent to solving

$$\mathcal{H}\nu = \mathcal{F}, \tag{4.4.30}$$

where $\mathcal{F}$ is defined as

$$\mathcal{F} = \frac{\partial\mathcal{G}}{\partial\phi} + \psi(x,0), \tag{4.4.31}$$

and ψ is the solution of the forced first order adjoint system

$$\begin{cases} \dfrac{\partial\psi}{\partial t} + u\dfrac{\partial\psi}{\partial x} + \dfrac{\partial\varphi}{\partial x} = \dfrac{\partial\mathcal{G}}{\partial\eta}, & \text{(4.4.32a)} \\ \dfrac{\partial\varphi}{\partial t} + (1+\eta)\dfrac{\partial\psi}{\partial x} + u\dfrac{\partial\varphi}{\partial x} = \dfrac{\partial\mathcal{G}}{\partial u}, & \text{(4.4.32b)} \\ \psi(x,T) = \ 0, & \text{(4.4.32c)} \\ \varphi(x,T) = \ 0. & \text{(4.4.32d)} \end{cases}$$

The significance of recognising the systems (4.4.26) and (4.4.25) as the Hessian with external forcing is that if we may use properties of the Hessian to solve for ν uniquely. If $\mathcal{H}$ is the hessian of $\mathcal{J}$, then since $\mathcal{J}$ is convex, $\mathcal{H}$ is positive definite. [53] demonstrate that the operator equation (4.4.30) has a solution everywhere in Y_p, and for every $\mathcal{F}$, we can find a unique ν such that (4.4.30) holds. Then the sensitivity $\partial\mathcal{G}/\partial m$ can be found by the steps in algorithm 4.

Algorithm 4 Calculation of Second Order Adjoint Sensitivity $\frac{\partial \mathcal{G}}{\partial m}$ for Initial Condition Assimilation

```
1: Define $\mathcal{F} = \frac{\partial \mathcal{G}}{\partial \phi} + \psi(x, 0)$, where $\psi$ is the solution of (4.4.32).
2: Solve $\mathcal{H}\nu = \mathcal{F}$ for $\nu$.
3: Solve the system (4.4.25) using $P_3(x, 0) = \nu$ to find $P_3(x, t)$.
4: Define $\frac{\partial \mathcal{G}}{\partial m} = HP_3(x, t)$, where $H$ is the operator mapping the $\eta$ from state space $Y$ to the observation space $Y_{obs}$ .
```

4.5 Bathymetry reconstruction sensitivity analysis

4.5.1 Hessian of $\mathcal{J}(\beta)$ for bathymetry reconstruction

The derivation of the Hessian for $\mathcal{J}(\beta)$ and subsequent sensitivity analysis for the bathymetry assimilation described in section 4.3 a is similar procedure to the initial conditions case. As we now use results of parameter data assimilation instead of initial conditions assimilation, the adjoint equations are slightly different, however the derivation steps remain the same. The sensitivity analysis is based on the procedure for parameter assimilation described in [52].

We know from section 4.3 that the Gâteaux derivative of $\mathcal{J}$ with respect to the bathymetry and some perturbation direction β' is

$$\mathcal{J}'(\beta; \beta') = -\int_{-L}^{L} \left(\int_0^T u \frac{\partial \eta^*}{\partial x} \, dt \right) \beta' \, dx. \tag{4.5.1}$$

Consider a second perturbation of $\mathcal{J}'(\beta; \beta')$, $\hat{\beta}$ where we have $\beta \to \beta + \varepsilon\hat{\beta}$. Then the second order Gâteaux derivative of $\mathcal{J}$ is

$$\begin{aligned}\mathcal{J}''(\beta; \beta'; \hat{\beta}) =& \frac{d}{d\varepsilon} \mathcal{J}'(\beta + \varepsilon\hat{\beta}; \beta') \Big|_{\varepsilon=0} \\ =& \frac{d}{d\varepsilon} \left\{ \int_{-L}^{L} \int_0^T u(x, t; \beta + \varepsilon\hat{\beta}) \frac{\partial \eta^*}{\partial x}(x, t; \beta + \varepsilon\hat{\beta}) \, dtdx \right\} \Big|_{\varepsilon=0}.\end{aligned}$$

Considering a regular perturbation expansion of the integrand as before, we approximate it by the series $f_0 + f_1\varepsilon + \mathcal{O}(\varepsilon^2)$. Then

$$\begin{aligned}
&\frac{d}{d\varepsilon}\mathcal{J}'(\beta+\varepsilon\hat{\beta};\beta')\Big|_{\varepsilon=0} \\
&= -\frac{d}{d\varepsilon}\int_{-L}^{L}\Big(\int_0^T (f_0+f_1\varepsilon)\ dt\Big)\beta'\ dx\Big|_{\varepsilon=0} \\
&= -\int_{-L}^{L}\Big(\int_0^T f_1\ dt\Big)\beta'\ dx.
\end{aligned}$$

We have $f_1 = \frac{d}{d\varepsilon}u(x,t;\beta+\varepsilon\hat{\beta})\frac{\partial}{\partial x}\eta^*(x,t;\beta+\varepsilon\hat{\beta})\Big|_{\varepsilon=0}$, and require u and adjoint variable η^* given the perturbed bathymetry $\beta+\varepsilon\hat{\beta}$. To find the resulting forward and adjoint systems given the perturbation, we assume this perturbation in the initial conditions produces the following perturbations to our state and adjoint variables;

- $u \to u + \hat{u}$ and $\eta \to \eta + \hat{\eta}$ for the shallow water system.
- $u^* \to u^* + \bar{u}$ and $\eta^* \to \eta^* + \bar{\eta}$ for the adjoint system.

The resulting perturbed model for the state variables $\hat{u}, \hat{\eta}$ is

$$\begin{cases}
\dfrac{\partial \hat{\eta}}{\partial t} + \dfrac{\partial}{\partial x}\Big((1+\eta-\beta)\hat{u}\Big) + \dfrac{\partial\Big(u\hat{\eta}\Big)}{\partial x} - \dfrac{\partial\Big(\hat{\beta}u\Big)}{\partial x} = 0, & (4.5.2a)\\
\dfrac{\partial \hat{u}}{\partial t} + \dfrac{\partial\Big(\hat{u}u\Big)}{\partial x} + \dfrac{\partial \hat{\eta}}{\partial x} = 0, & (4.5.2b)\\
\hat{\eta}(x,0) = 0, & (4.5.2c)\\
\hat{u}(x,0) = 0. & (4.5.2d)
\end{cases}$$

and the second order adjoint (SOA) model is

$$\begin{cases}
\dfrac{\partial \bar{\eta}}{\partial t} + \hat{u}\dfrac{\partial \eta^*}{\partial x} + u\dfrac{\partial \bar{\eta}}{\partial x} + \dfrac{\partial \bar{u}}{\partial x} = -\hat{\eta}(x_i,t;\hat{\beta}), & (4.5.3a)\\
\dfrac{\partial \bar{u}}{\partial t} + (1+\eta-\beta)\dfrac{\partial \bar{\eta}}{\partial x} + (\hat{\eta}-\hat{\beta})\dfrac{\partial \eta^*}{\partial x} + u\dfrac{\partial \bar{u}}{\partial x} + \hat{u}\dfrac{\partial u^*}{\partial x} = 0, & (4.5.3b)\\
\bar{\eta}(x,T) = 0, & (4.5.3c)\\
\bar{u}(x,T) = 0. & (4.5.3d)
\end{cases}$$

Giving us

$$\nabla^2 \mathcal{J}(\beta;\hat{\beta}) = \int_0^T \Big(\hat{u}\frac{\partial \eta^*}{\partial x} + u\frac{\partial \bar{\eta}}{\partial x}\Big)dt. \tag{4.5.4}$$

We define the Hessian $\mathcal{H}$ acting on the perturbation $\hat{\beta}$ as the successive solution of the perturbed and SOA models such that

$$\mathcal{H}\hat{\beta} = \int_0^T \Big(\hat{u}\frac{\partial \eta^*}{\partial x} + u\frac{\partial \bar{\eta}}{\partial x}\Big)dt. \tag{4.5.5}$$

This derivation can be verified using the kappa test with

$$\kappa(\varepsilon) = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \frac{\mathcal{J}'(\beta+\varepsilon\hat{\beta};\beta') - \mathcal{J}'(\beta;\eta')}{-\int_{-L}^{L}\Big(\int_0^T \Big(\hat{u}\frac{\partial \eta^*}{\partial x} + u\frac{\partial \bar{\eta}}{\partial x}\Big)dt\Big)\beta'\ dx}, \tag{4.5.6}$$

As before, if the derivations are correct, $\kappa(\varepsilon)$ should converge to 1 as $\varepsilon \to 0$. The results for the kappa test for the Hessian $\mathcal{H}$ are presented in section 4.6 in figure 4.2.

4.5.2 Sensitivity analysis for bathymetry reconstruction

Define the optimality system for bathymetry assimilation as the successive solution of

$$\begin{cases} \dfrac{\partial \eta}{\partial t} + \dfrac{\partial}{\partial x}\Big((1+\eta-\beta)u\Big) = 0, & (4.5.7a)\\ \dfrac{\partial u}{\partial t} + \dfrac{\partial}{\partial x}\Big(\dfrac{1}{2}u^2+\eta\Big) = 0, & (4.5.7b)\\ \eta(x,0) = \ \eta_0(x), & (4.5.7c)\\ u(x,0) = \ 0. & (4.5.7d) \end{cases}$$

$$\begin{cases} \dfrac{\partial \eta^*}{\partial t} + u\dfrac{\partial \eta^*}{\partial x} + \dfrac{\partial u^*}{\partial x} = \ H\Big(\eta-m\Big), & (4.5.8a)\\ \dfrac{\partial u^*}{\partial t} + (1+\eta-\beta)\dfrac{\partial \eta^*}{\partial x} + u\dfrac{\partial u^*}{\partial x} = \ 0, & (4.5.8b)\\ \eta^*(x,T) = \ 0, & (4.5.8c)\\ u^*(x,T) = \ 0, & (4.5.8d) \end{cases}$$

Where $\lambda(x)$ is the optimal reconstruction of the bathymetry $\beta(x)$ giving :

$$\int_0^T u\frac{\partial \eta^*}{\partial x}\,dt = 0. \tag{4.5.9}$$

As before, we have periodic boundary conditions in space, where the domain is $\Omega = \Big\{x \in \mathbb{R}; x \in [-L, L]\Big\}$.

Let us consider some arbitrary response function

$$\mathcal{G}(\eta, u, \lambda). \tag{4.5.10}$$

Then by the chain rule, the sensitivity of $\mathcal{G}$ to perturbations in the observations m can be defined as

$$\frac{d\mathcal{G}}{dm} = \frac{d\mathcal{G}}{d\eta}\frac{d\eta}{dm} + \frac{d\mathcal{G}}{du}\frac{du}{dm} + \frac{d\mathcal{G}}{d\lambda}\frac{d\lambda}{dm}. \tag{4.5.11}$$

Give some perturbation of the observations such that given $m \to m + \hat{m}$ we have

- $u \to u + \hat{u}$ and $\eta \to \eta + \hat{\eta}$ for the shallow water system (4.5.7).
- $u^* \to u^* + \tilde{u}^*$ and $\eta^* \to \eta^* + \tilde{\eta}^*$ for the adjoint system (4.5.8).
- $\lambda \to \phi + \hat{\lambda}$ for the optimal initial condition.

Our perturbed system becomes

$$\begin{cases} \dfrac{\partial \hat{\eta}}{\partial t} + \dfrac{\partial}{\partial x}\Big(u\hat{\eta}\Big) + \dfrac{\partial}{\partial x}\Big((\eta + 1 - \lambda)\hat{u}\Big) - \dfrac{\partial}{\partial x}\Big(\hat{\lambda} u\Big) - \dfrac{\partial}{\partial x}\Big(\lambda\hat{u}\Big) = 0, & (4.5.12a) \\ \dfrac{\partial \hat{u}}{\partial t} + \dfrac{\partial\Big(\hat{u}u\Big)}{\partial x} + \dfrac{\partial \hat{\eta}}{\partial x} = 0, & (4.5.12b) \\ \hat{\eta}(x, 0) = \hat{\phi}(x), & (4.5.12c) \\ \hat{u}(x, 0) = 0, & (4.5.12d) \end{cases}$$

$$\begin{cases} \dfrac{\partial\tilde{\eta}^*}{\partial t} + \hat{u}\dfrac{\partial\eta^*}{\partial x} + u\dfrac{\partial\tilde{\eta}^*}{\partial x} + \dfrac{\partial\tilde{u}^*}{\partial x} = H(\hat{m} - \hat{\eta}(x_i, t; \hat{\lambda}), & \text{(4.5.13a)} \\ \dfrac{\partial\tilde{u}^*}{\partial t} + (\eta + 1 - \lambda)\dfrac{\partial\tilde{\eta}^*}{\partial x} + \hat{\eta}\dfrac{\partial\eta^*}{\partial x} & \text{(4.5.13b)} \\ + u\dfrac{\partial\tilde{u}^*}{\partial x} + \hat{u}\dfrac{\partial u^*}{\partial x} - \hat{\lambda}\dfrac{\partial\eta^*}{\partial x} = 0, & \\ \tilde{\eta}^*(x, T) = 0, \tilde{u}^*(x, T) = 0. & \text{(4.5.13c)} \end{cases}$$

$$\int_0^T \big(\hat{u}\frac{\partial\eta^*}{\partial x} + u\frac{\partial\tilde{\eta}^*}{\partial x}\big)dt. \tag{4.5.14}$$

Then we can say

$$\left\langle \frac{d\mathcal{G}}{dm}, \hat{m} \right\rangle_{Y_{obs}} = \left\langle \frac{\partial\mathcal{G}}{\partial\eta}, \hat{\eta} \right\rangle_{Y} + \left\langle \frac{\partial\mathcal{G}}{\partial u}, \hat{u} \right\rangle_{Y} + \left\langle \frac{\partial\mathcal{G}}{\partial\lambda}, \hat{\lambda} \right\rangle_{Y_p} \tag{4.5.15}$$

where

$$Y_{obs} = L^2\Big([-L, L] \times [0, T]\Big), \tag{4.5.16}$$

$$Y = L^2\Big([-L, L] \times [0, T]\Big), \tag{4.5.17}$$

$$Y_p = L^2\Big([-L, L]\Big) \tag{4.5.18}$$

are as before in section 4.4. Let us introduce some adjoint variables P_i, $i = 1, ..., 5$, where $P_i \in Y$, $i = 1, ..., 4$, and $P_5 \in Y_{obs}$. Then if we take the inner product of P_1 and P_2 with the systems (4.5.12), inner product of P_3 and P_4 with (4.5.13), and

P_5 with (4.5.14), we get the following duality relation

$$
\begin{align}
0 = & \nonumber\\
& \int_0^T \int_{-L}^{L} \Bigg\{ P_1 \Bigg[\frac{\partial \hat{\eta}}{\partial t} + \frac{\partial}{\partial x}\big(u\hat{\eta}\big) + \frac{\partial}{\partial x}\big((\eta + 1 - \lambda)\hat{u}\big) \nonumber\\
& - \frac{\partial}{\partial x}\big(\hat{\lambda} u\big) - \frac{\partial}{\partial x}\big(\lambda \hat{u}\big)\Bigg] \nonumber\\
& + P_2 \Bigg[\frac{\partial \hat{u}}{\partial t} + \frac{\partial\big(\hat{u}u\big)}{\partial x} + \frac{\partial \hat{\eta}}{\partial x}\Bigg] \nonumber\\
& + P_3 \Bigg[\frac{\partial \tilde{\eta}^*}{\partial t} + \hat{u}\frac{\partial \eta^*}{\partial x} + u\frac{\partial \tilde{\eta}^*}{\partial x} + \frac{\partial \tilde{u}^*}{\partial x} \tag{4.5.19}\\
& - H(\hat{m} - \hat{\eta}(x_i, t; \hat{\lambda})\Bigg] \nonumber\\
& + P_4 \Bigg[\frac{\partial \tilde{u}^*}{\partial t} + (\eta + 1 - \lambda)\frac{\partial \tilde{\eta}^*}{\partial x} + \hat{\eta}\frac{\partial \eta^*}{\partial x} \tag{4.5.20}\\
& + u\frac{\partial \tilde{u}^*}{\partial x} + \hat{u}\frac{\partial u^*}{\partial x} - \hat{\lambda}\frac{\partial \eta^*}{\partial x}\Bigg] \nonumber\\
& + P_5 \Bigg[\hat{u}\frac{\partial \eta^*}{\partial x} + u\frac{\partial \tilde{\eta}^*}{\partial x}\Bigg]\Bigg\} dt dx. \tag{4.5.21}
\end{align}
$$

Integrating (4.5.21) by parts in space and time, we are able to transfer the derivatives onto the adjoint variables P_i, $i = 1, ..., 5$ instead of on $\hat{u}, \hat{\eta}, \tilde{u}^*, \tilde{u}^*$. We pick the following systems for P_i,

$$
\begin{cases}
\dfrac{\partial P_3}{\partial t} + \dfrac{\partial}{\partial x}\big((1 + \eta - \lambda)P_4\big) - \dfrac{\partial\big(uP_5\big)}{\partial x} - \dfrac{\partial\big(\lambda P_4\big)}{\partial x} = 0, & \text{(4.5.22a)}\\
\dfrac{\partial P_4}{\partial t} + \dfrac{\partial\big(P_4 u\big)}{\partial x} + \dfrac{\partial P_3}{\partial x} = 0, & \text{(4.5.22b)}\\
P_3(x, 0) = 0, & \text{(4.5.22c)}\\
P_4(x, 0) = 0. & \text{(4.5.22d)}
\end{cases}
$$

$$
\begin{cases}
\dfrac{\partial P_1}{\partial t} + P_4\dfrac{\partial \eta^*}{\partial x} + u\dfrac{\partial P_1}{\partial x} + \dfrac{\partial P_2}{\partial x} - HP_3 = -\dfrac{\partial \mathcal{G}}{\partial \eta}, & \text{(4.5.23a)}\\
\dfrac{\partial P_2}{\partial t} + (1 + \eta - \lambda)\dfrac{\partial P_1}{\partial x} + (P_3 - P_5)\dfrac{\partial \eta^*}{\partial x} + u\dfrac{\partial P_2}{\partial x} + P_4\dfrac{\partial u^*}{\partial x} = -\dfrac{\partial \mathcal{G}}{\partial u}, & \text{(4.5.23b)}\\
\qquad P_1(x, T) = 0, & \text{(4.5.23c)}\\
\qquad P_2(x, T) = 0. & \text{(4.5.23d)}
\end{cases}
$$

$$\int_0^T \Big(P_4 \frac{\partial \eta^*}{\partial x} + u\frac{\partial P_1}{\partial x}\Big) dt = \frac{\partial \mathcal{G}}{\partial \lambda}. \tag{4.5.24}$$

Subsequently as a result of integration by parts and the choice of systems for P_i, (4.5.21) reduces to

$$\Big\langle HP_3, \hat{m} \Big\rangle_{Y_{obs}} = \Big\langle \frac{\partial \mathcal{G}}{\partial \eta}, \hat{\eta} \Big\rangle_{Y} + \Big\langle \frac{\partial \mathcal{G}}{\partial u}, \hat{u} \Big\rangle_{Y} + \Big\langle \frac{\partial \mathcal{G}}{\partial \lambda}, \hat{\lambda} \Big\rangle_{Y_p}. \tag{4.5.25}$$

By the Riesz representation theorem and equivalence of inner products in (4.4.20) and (4.4.28), we define the sensitivity of the response function $\mathcal{G}(\eta, u, \lambda)$ to perturbations in the observations m as

$$\frac{\partial \mathcal{G}}{\partial m} = HP_3(x,t). \tag{4.5.26}$$

As in the initial conditions case, we observe that (4.5.22) is equivalent to the perturbed system for the Hessian $\nabla^2 \mathcal{J}(\phi)$ (4.5.1) with $P_5 = \hat{\beta}(x)$, and (4.5.23) is equivalent to the second order adjoint (SOA) system (4.5.1) with forcing term $\Big(-\frac{\partial \mathcal{G}}{\partial \eta}, -\frac{\partial \mathcal{G}}{\partial u}\Big)^T$. Let us replace P_5 with the auxiliary variable ν. Then Shutyaev et al. [52] show that the solutions to the adjoint systems (4.5.22) and (4.5.23) are equivalent to solving

$$\mathcal{H}\nu = \mathcal{F}, \tag{4.5.27}$$

where $\mathcal{F}$ is defined as

$$\mathcal{F} = \frac{\partial \mathcal{G}}{\partial \lambda} - \int_0^T u \frac{\partial \gamma}{\partial x} dt, \tag{4.5.28}$$

and γ is the solution of the forced first order adjoint system

$$\begin{cases} \dfrac{\partial \gamma}{\partial t} + u\dfrac{\partial \gamma}{\partial x} + \dfrac{\partial \psi}{\partial x} = -\dfrac{\partial \mathcal{G}}{\partial \eta}, & (4.5.29\text{a}) \\ \dfrac{\partial \psi}{\partial t} + (1+\eta-\beta)\dfrac{\partial \gamma}{\partial x} + u\dfrac{\partial \psi}{\partial x} = -\dfrac{\partial \mathcal{G}}{\partial u}, & (4.5.29\text{b}) \\ \gamma(x,T) = 0, & (4.5.29\text{c}) \\ \psi(x,T) = \ 0. & (4.5.29\text{d}) \end{cases}$$

As before, under the assumption $\mathcal{H}$ is positive definite (as our cost function $\mathcal{J}$ is convex), we can find a unique ν for every $\mathcal{F}$ such that $\mathcal{H}\nu = \mathcal{F}$ [52]. We are then able to find $\frac{\partial \mathcal{G}}{\partial m}$ by the steps outlined in algorithm 5.

Algorithm 5 Calculation of Second Order Adjoint Sensitivity $\frac{\partial \mathcal{G}}{\partial m}$ for Bathymetry Assimilation

```
Define $\frac{\partial \mathcal{G}}{\partial \lambda} - \int_0^T u\frac{\partial \gamma}{\partial x}dt$, where $\gamma$ is the solution of (4.5.29).
Solve $\mathcal{H}\nu = \mathcal{F}$ for $\nu$.
Solve the system (4.4.25) using $P_5(x) = \nu$ to find $P_3(x,t)$.
Define $\frac{\partial \mathcal{G}}{\partial m} = HP_3(x,t)$, where $H$ is the operator mapping the $\eta$ from state
    space $Y$ to the observation space $Y_{obs}$ .
```

4.6 Numerical Implementation

The results of this analysis would be a significant extension and complement to the work undertaken in Kevlahan et al. [27], and the results in chapter 3. In this section, we present the numerical implementation of the second order adjoint sensitivity analysis for the bathymetry assimilation. While the initial conditions assimilation certainly merits a similar analysis and can provide key insights for tsunami models, our primary focus is to extend and possibly confirm the results observed for data assimilation for bathymetry in chapter 3. To summarise, we noted that the surface wave resulting from the optimally reconstructed bathymetry was insensitivie to the reconstruction error. We also observed that a greater number of observation points, placed before the support of the bathymetry resulted in better convergence of the L^2 error in bathymetry reconstruction. Additionally, we investigated the effect of varying the amplitudes of the bathymetry and initial condition, and concluded that a necessary condition for data assimilation convergence was that the amplitude of the initial conditions be at least two orders of magnitude smaller than the normalised amplitude of the bathymetry. The sensitivity analysis in the present study is a significant addition to this work, since it demonstrates how results can vary as the observation operator is perturbed.

In tsunami modelling we wish to choose the number and positions of the observation network such that the accuracy of the surface wave forecast is maximised. Therefore selecting an observation operator that does not impact the surface wave accuracy is an important consideration, and so we quantify how sensitive the surface prediction given some reconstructed bathymetry is to the choice of observation operator. We define the response function $\mathcal{G}$ as the least squares error between the surface wave given the true bathymetry, and the wave resulting from the optimally reconstructed bathymetry,

$$G(\eta, u, \beta^{(b)}) = \int_0^T \left[\eta(\beta^{(t)}) - \eta(\beta^{(b)})\right]^2 dt, \tag{4.6.1}$$

where we have implicit dependence on the surface wave velocity $u(x,t)$ through coupling in the shallow water equations. We recall that $m_j(t)$ are the observations taken at positions $\{x_j\}$ for $j = 1, ..., N_{obs}$ at continuous times t. Then,

$d\mathcal{G}/dm$ can be defined as the sensitivity of the error in the surface wave produced by the reconstructed bathymetry (4.6.1), to perturbations in the observations $\{m_j(t)\}$.

These perturbations include changes in the positions $\{x_j\}$, or any perturbations in the parameters of the system (4.2.1) that result in a change in $\{m_j(t)\}$. Subsequently $d\mathcal{G}/dm$ is an $N_{obs} \times N_t$ matrix, where N_t is the temporal resolution, and the $m_{i,j}$ entry of $d\mathcal{G}/dm$ represents the sensitivity of the surface wave error (4.6.1) to the i-th observation at the j-th time step.

Our analysis aims to answer the following three questions:

I Is there a correlation between low sensitivity $\partial\mathcal{G}/\partial m$ and optimal convergence of the reconstructed bathymetry to the exact shape?

II How does changing parameters in the data assimilation scheme (such as the number of observation points or bathymetry shape) affect the sensitivity $\partial\mathcal{G}/\partial m$?

III Can our conclusions from chapter 3 be verified by the results of the sensitivity analysis?

We examine these questions one by one, beginning with question I.

Question I

We wish to quantify the relationship between the convergence in bathymetry assimilation and the sensitivity $\partial\mathcal{G}/\partial m$ derived as a function of the assimilation results. If

If the error in the bathymetry reconstruction is defined as

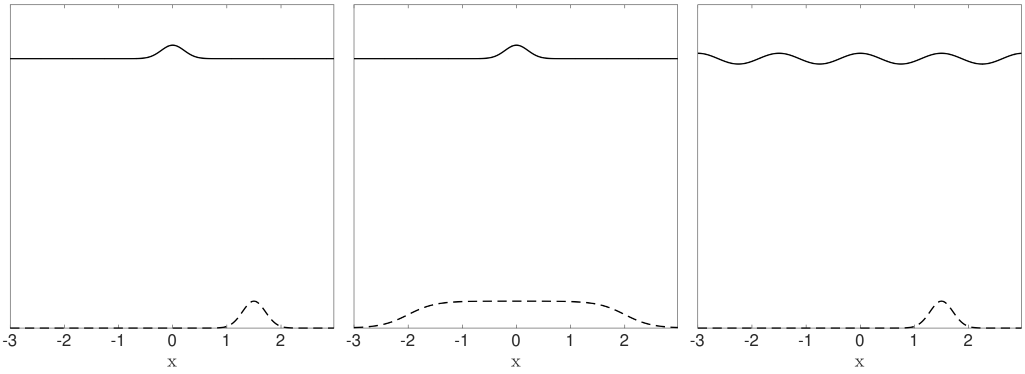
$$\frac{\| \beta^{(t)} - \beta^{(b)} \|_{L^2}}{\| \beta^{(t)} \|_{L^2}}, \tag{4.6.2}$$

(where $\beta^{(t)}$ is the true bathymetry and $\beta^{(b)}$ is the optimal reconstruction), then our objective is to understand whether a relatively higher error for the bathymetry reconstruction (4.6.2) correlated with increased sensitivity $\partial\mathcal{G}/\partial m$ where $\mathcal{G}$ is given by (4.6.1). This question is significant because high sensitivity $\partial\mathcal{G}/\partial m$ is undesirable, as it implies that variations in where observations are measured and their total number, have a large effect on the accuracy of surface wave predictions

To answer this question, we compare the convergence of the relative L^2 error in convergence to the exact bathymetry of the optimal reconstruction as seen in chapter 3, and the resulting sensitivity $\partial\mathcal{G}/\partial m$ for each case considered, These are summarised in table 4.1, and can be seen in figure 4.1.

Case	Bathymetry	Initial Conditions
I	Gaussian	Gaussian
II	Gaussian	Sinusoidal
III	Sandbar	Gaussian

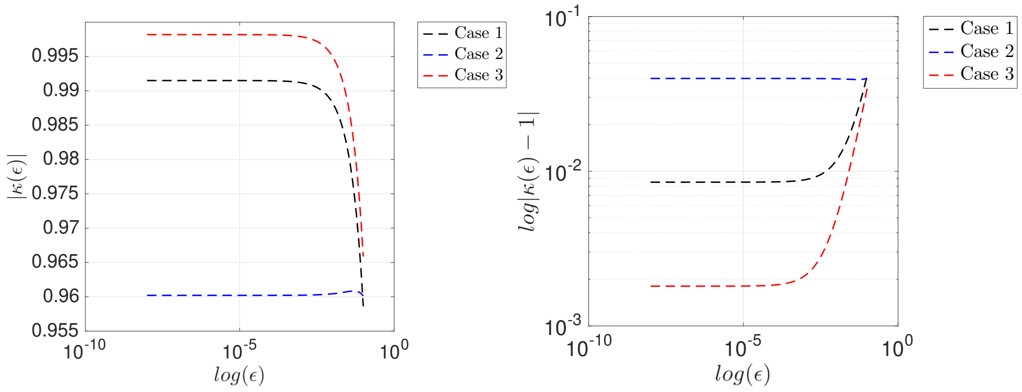
Table 4.1: Cases considered for data assimilation algorithm.



(a) Case I (b) Case II (c) Case III

Figure 4.1: The three test cases for bathymetry $\beta(x)$ (dashed line) and initial conditions $\phi(x)$ (solid line) for the data assimilation. The surface wave initial conditions $\hat{\eta}$, bathymetry $\hat{\beta}$, and average depth H are not to scale in these diagrams, as $\hat{\eta}$ was restricted to 1% of $\hat{\beta}$ across most of the numerical analyses, and $\beta = 0.1$.

We recall that these cases were chosen to analyse convergence in scenarios where the support of $\phi(x)$ and the support of $\beta^{(t)}(x)$ overlap or are disjoint, and to evaluate the effect of a surface wave with compactly supported initial conditions (cases I and II) or periodic initial conditions (case III). We consider Gaussian and sandbar profiles for the bathymetry as a 1-D approximation for peaks and ridges characterising ocean bathymetry. The results of the kappa test for each case are given in figure 4.2.



(a) Kappa convergence (b) Kappa convergence error

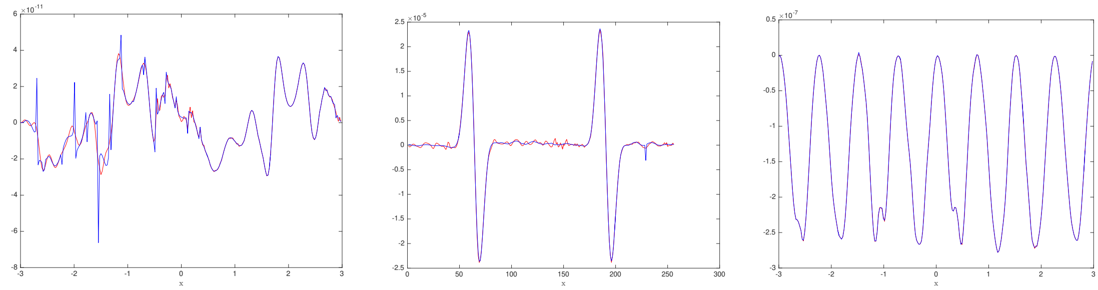
FIGURE 4.2: The results of the kappa test in (4.5.1), to verify the numerical calculation of the Hessian $\mathcal{H}$. (a) presents the log convergence to 1 of (4.5.1) as $\varepsilon \to 0$, and (b) presents the log error in the convergence.

Using these data assimilation results, we implement algorithm 5 to find the sensitivity of the cost function G to the observations m.

Recall the results for the convergence error in the optimal data assimilation results in figure 3.4(c). As we can see, the best convergence is for case I. For Cases II and III, the optimal results do not accurately recover the exact bathymetry and the error in the reconstructed bathymetry is relatively larger, $\mathcal{O}(10^{-2})$ and $\mathcal{O}(10^{-1})$ respectively.

To find the optimal v such that $\mathcal{H}\nu = \mathcal{F}$ as in (4.5.27), we used the Bi-Conjugate Gradient Stabilised Method (BICGSTAB) to solve the system for v. Figure 4.3 illustrates the convergence of $\mathcal{H}\nu$ to the right hand side $\mathcal{F}$ for the optimal ν.

The BICGSTAB method is a Krylov linear solver designed for non-symmetric linear operators, and was found to prove better convergence in this analysis than with comparable solvers such as GMRES (Generalised Minimum Residual Method). We observe that the error in convergence in case I (figure 4.3(a)) is due to noise in the right hand side $\mathcal{F}$, and that $\mathcal{H}\nu$ in figure 4.3(b) shows some overfitting. A reasonable tolerance for convergence of BICGSTAB is a relative residual error less than $1e-4$, however due to the small-scale noise shown in figure 4.3, this



(a) Case I (b) Case II (c) Case III

Figure 4.3: $\mathcal{H}\nu$ (red) and $\mathcal{F}$ (blue) for each case where ν is the solution of $\mathcal{H}\nu = \mathcal{F}$ found using the matlab linear solver `bicgstabl`.

tolerance was not met. Convergence is improved for BICGSTAB when the the system is preconditioned [4]. Consequently, errors may be due to the fact that preconditioners for the operator $\mathcal{H}$ were not easily computable in the algorithm, and hence were not used. As the error in figure 4.3 seems to be due to a small amount of noise only, for the purpose of this study we consider the convergence to be satisfactory.

Thus, assuming that we have optimally found ν (representing the adjoint variable P_5) given $\mathcal{F}$ and the operator $\mathcal{H}$, we can derive the sensitivity $\partial \mathcal{G}/\partial m = HP_3(x,t)$ as the solution of the system (4.4.25) using ν as the initial condition, where H is the observation operator.

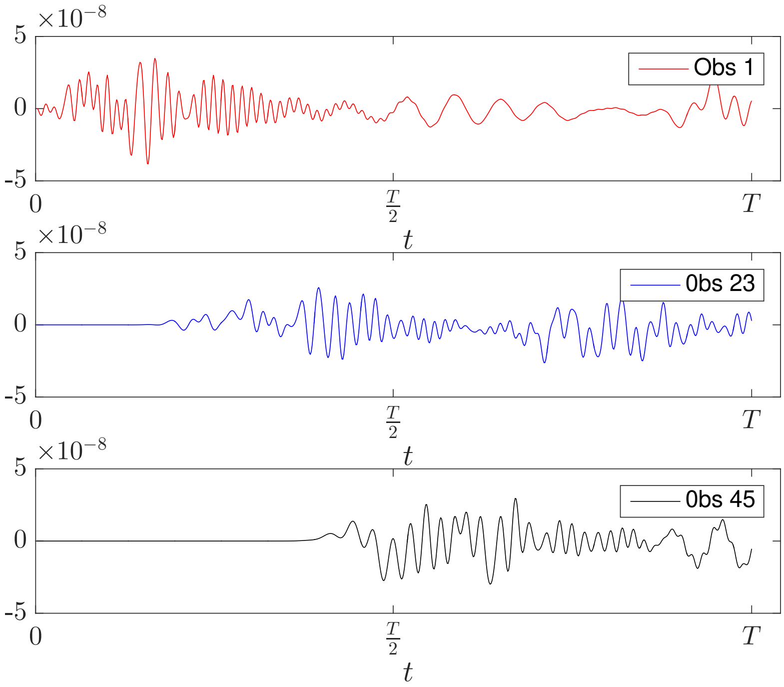


FIGURE 4.4: The sensitivity $d\mathcal{G}/dm$ as a function of time (with final time $t = T$), for assimilation results for Case I. There are $N_{obs} = 45$ observations, equidistantly spaced with $\Delta x = 0.06$ and with the first point at $0.1L$. Results show $d\mathcal{G}/dm$ at three distinct observation points m_j, where $j = 1$ (first observation), $\lceil \frac{N_{obs}}{2} \rceil$ (the median observation), and N_{obs} (the last observation).

The sensitivity for case I is presented in figure 4.4 as a function of time, highlighted at three observation points. We derived sensitivity for assimilation results using $N_{obs} = 45$, as this value produced optimal convergence in chapter 3. These are equidistantly spaced with $\Delta x = 0.06$ and with the first point at $0.1L$. We recall that the sensitivity $d\mathcal{G}/dm$ is an $N_{obs} \times N_t$ matrix. To be concise, in figure 4.4 we only plot results for $d\mathcal{G}/dm$ for three distinct observation points m_j, where $j = 1$ (first observation), $\lceil \frac{N_{obs}}{2} \rceil$ (the median observation), and N_{obs} (the last observation). Figure 4.4 shows that at each m_j the sensitivity is $\mathcal{O}(10^{-8})$. Comparing the trend across observation points, we see increased oscillation at the time steps where the surface wave is presumably observed by the measurement points. However, the amplitude of the sensitivity remains low. Given the relative error for the case I bathymetry reconstruction was $\mathcal{O}(10^{-4})$, our initial hypothesis is that the lower error in the reconstructed bathymetry is correlated with low sensitivity of the surface wave error given the reconstructed bathymetry, to the observations.

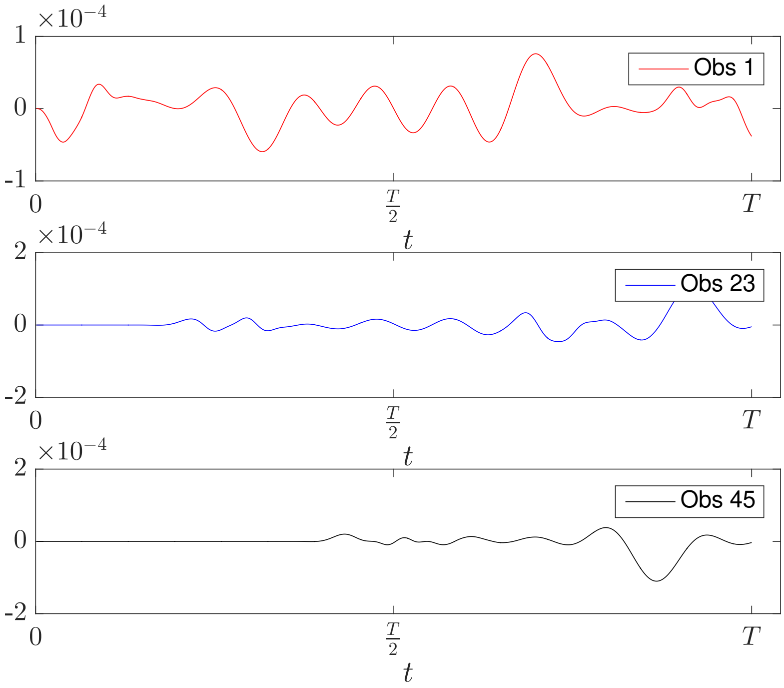


FIGURE 4.5: The sensitivity $d\mathcal{G}/dm$ as a function of time (with final time $t = T$), for assimilation results for Case II. There are $N_{obs} = 45$ observations, equidistantly spaced with $\Delta x = 0.06$ and with the first point at $0.1L$. Results show $d\mathcal{G}/dm$ at three distinct observation points m_j, where $j = 1$ (first observation), $\lceil \frac{N_{obs}}{2} \rceil$ (the median observation), and N_{obs} (the last observation).

To substantiate this further, we observe the results for cases II and III, as presented in figures 4.5 and 4.6. The number, position, and spacing of observations is the same as for figure 4.4. For case II we see that the sensitivity is much higher, $\mathcal{O}(10^{-4})$ at m_1, m_{23} and m_{45}. The oscillations indicating the passing wave are present for case I as well, and have a relatively lower frequency compared to figure 4.4, possibly due to flatter curvature of the sandbar bathymetry compared to a Gaussian bathymetry, resulting in a more gradual effect on the surface wave. We note that the bathymetry reconstruction error (4.6.2) for case II (as shown in figure **??**) is $\mathcal{O}(10^{-2})$, and subsequently an order of magnitude higher than for case I. Subsequently the increased sensitivity in figure 4.5 for case II compared to case I is correlated to a higher reconstruction error in bathymetry with a sandbar profile that with a Gaussian profile.

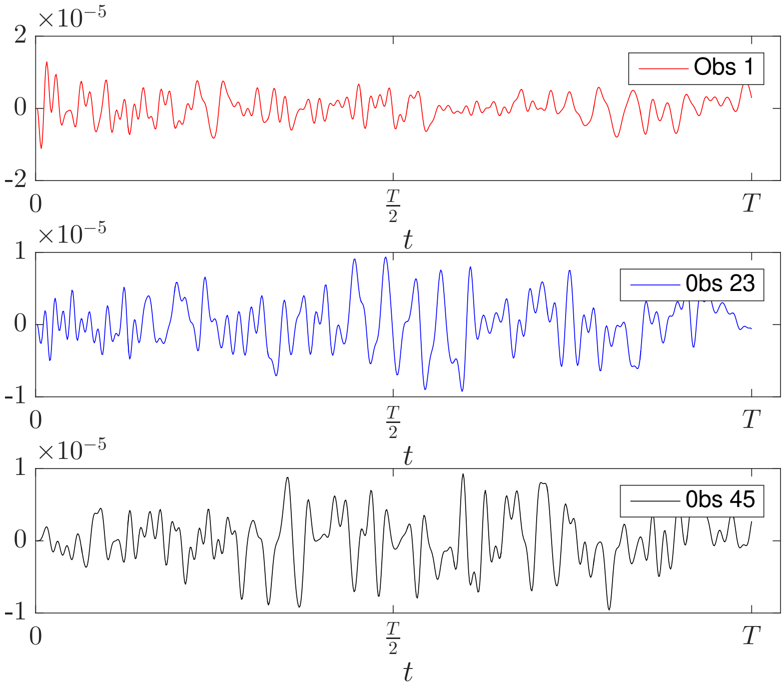
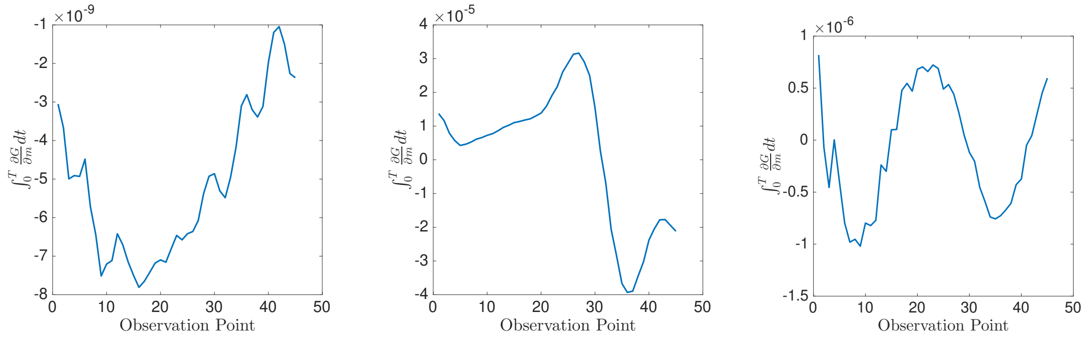


FIGURE 4.6: The sensitivity $d\mathcal{G}/dm$ as a function of time (with final time $t = T$), for assimilation results for Case III. There are $N_{obs} = 45$ observations, equidistantly spaced with $\Delta x = 0.06$ and with the first point at $0.1L$. Results show $d\mathcal{G}/dm$ at three distinct observation points m_j, where $j = 1$ (first observation), $\lceil \frac{N_{obs}}{2} \rceil$ (the median observation), and N_{obs} (the last observation

For case III, we see that the sensitivity is intermediate, $\mathcal{O}(10^{-5})$ at all three observation points, and the oscillatory behaviour is present at all time steps, due to the periodic nature of the surface wave. For case III, the relative error in the bathymetry reconstruction is also much larger than case I, $\mathcal{O}(10^{-1})$. therefore we observe that for cases II and rom3, the fact that the sensitivity in these cases is at least three orders of magnitude higher than with case I further confirms our initial hypothesis that a higher error in bathymetry reconstruction is related to higher sensitivity of the surface wave error to perturbations in the observations.

To gauge the sensitivity across all observation points, in figure 4.7 we consider the time integrated sensitivity across all 45 observation points for each case. As we can see, the hypothesis is confirmed across all cases, where the sensitivity of the surface wave error to perturbations in the observation points is much higher for the cases with a higher L^2 error in the reconstructed bathymetry (cases II and III).



(a) Case I (b) Case II (c) Case III

FIGURE 4.7: The time integrated sensitivity $\int_0^T \frac{\partial \mathcal{G}}{\partial m} dt$ at each observation point.

We summarise the results comparing sensitivity of the surface wave error to perturbations in the observations and relative error in bathymetry reconstruction in table 4.2. We note that the sensitivity is lower for case III compared to case II despite the fact that III has a relatively higher error in the bathymetry reconstruction. however, this value is still high relative to case I, and so we stipulate that our hypothesis conclusion has merit, even though the correlation may not be completely linear.

Case	Error	Sensitivity
I	$\mathcal{O}(10^{-3})$	$\mathcal{O}(10^{-9})$
II	$\mathcal{O}(10^{-2})$	$\mathcal{O}(10^{-5})$
III	$\mathcal{O}(10^{-1})$	$\mathcal{O}(10^{-6})$

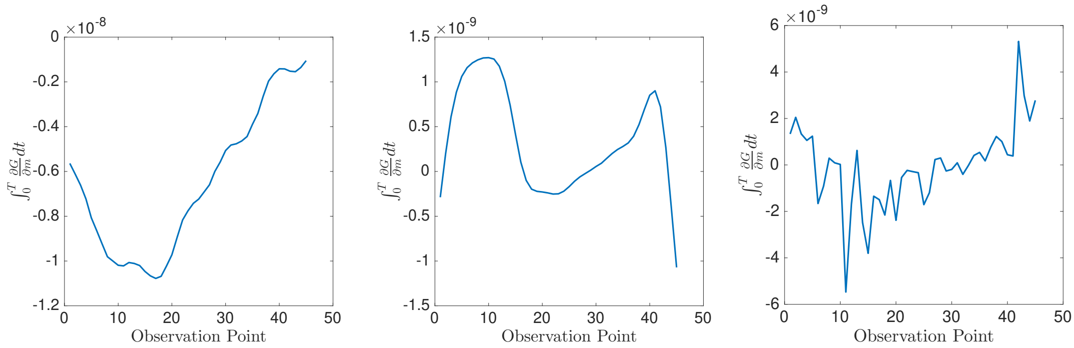
TABLE 4.2: Comparison of the relative L^2 reconstruction error (4.6.2) in the bathymetry as shown in figure 3.5(c), and the time integrated sensitivity $\int_0^T d\mathcal{G}/dm\ dt$ of the surface wave error to the observations.

We now extend the sensitivity analysis to question II, and observe the trend when we perturb influential parameters in the algorithm, such as the placement and number of observation points.

Question II

We now analyse the dependence of the sensitivity $d\mathcal{G}/dm$ on the position and number of observation points, as well as on the shape of the bathymetry. Our objective is to determine if there is an optimal spacing, or number of observations such that the sensitivity of the surface wave error (4.6.1) to observations $\{m_j(t)\}$ remains low. For each of cases I, II, and III, we vary the (i) spacing of observation points, (ii) number of observation points and (iii) shape of the bathymetry. The latter is motivated by the significant difference in sensitivity we observed between case I, where we have a gaussian bathymetry with relatively low standard deviation, and case II, where we have a sandbar bathymetry with more “spread”.

First, we consider the sensitivity of the surface wave error to the spacing of the observation points. Our objective is to find the spacing of a fixed number of evenly spaced observation points that minimizes the sensitivity of the surface wave error to perturbations in the observations $\{m_j(t)\}$ for $j = 1, ..., N_{obs}$, where $N_{obs} = 45$. We iteratively run the data assimilation algorithm, where in each iteration the position of the last observation point x_{45} is fixed, and as Δx is varied, the the first observation point is shifted back (as in figures 4.8 (d), (e) and (f)). We wish to observe how the resulting sensitivity of the surface wave error to observations changes, and thus answer the question: does increasing the spacing of the observations (and subsequently increasing coverage of the domain by observation points) lead to a decrease in the sensitivity $d\mathcal{G}/dm$? If so, the implication for tsunami models is that, given a fixed number of evenly spaced observation points, it is best to choose a relatively wide spacing that covers as much of the bathymetry feature as possible.



(a) $x_1 = 0.3$, $\Delta x = 0.06$ (b) $x_1 = -1.19$, $\Delta x = 0.09$ (c) $x_1 = -2.83$, $\Delta x = 0.13$

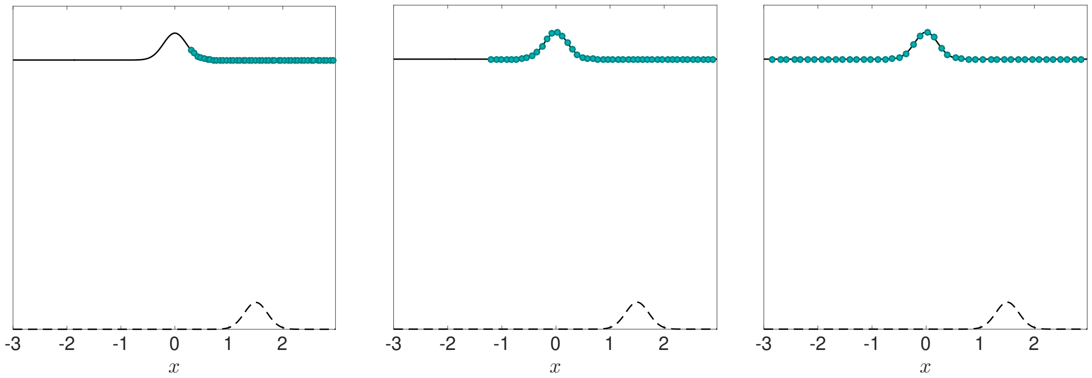
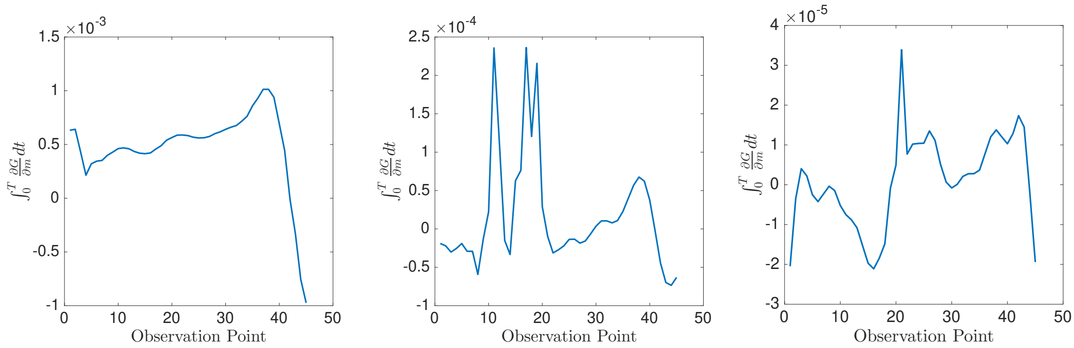


Figure 4.8: Case I: The time integrated sensitivity of the surface wave error $\int_0^T \frac{\partial G}{\partial m} dt$ as the location of the first observation point is varied such that the observation points cover a greater proportion of the domain and the initial conditions support.

Figure 4.8, showing the time integrated sensitivity $\int_0^T \frac{\partial G}{\partial m} dt$ of the surface wave error for case I, corresponding to different spacings Δx for 45 observation points. The values of Δx and the resulting observation configurations are given in the bottom panels of figure 4.8. We observe that for case I, the amplitude of the sensitivity $d\mathcal{G}/dm$ is small for each choice of Δx, at most $\mathcal{O}(10^{-8})$. This suggests that with a Gaussian bathymetry and initial conditions (as seen in figure 4.8(d)-(f)), larger spacing of observations does not significantly increase the sensitivity of the error in the surface wave produced by the reconstructed bathymetry, to the observations $m_j(t)$.



(a) $x_1 = 0.3$, $\Delta x = 0.06$ (b) $x_1 = -1.19$, $\Delta x = 0.09$ (c) $x_1 = -2.83$, $\Delta x = 0.13$

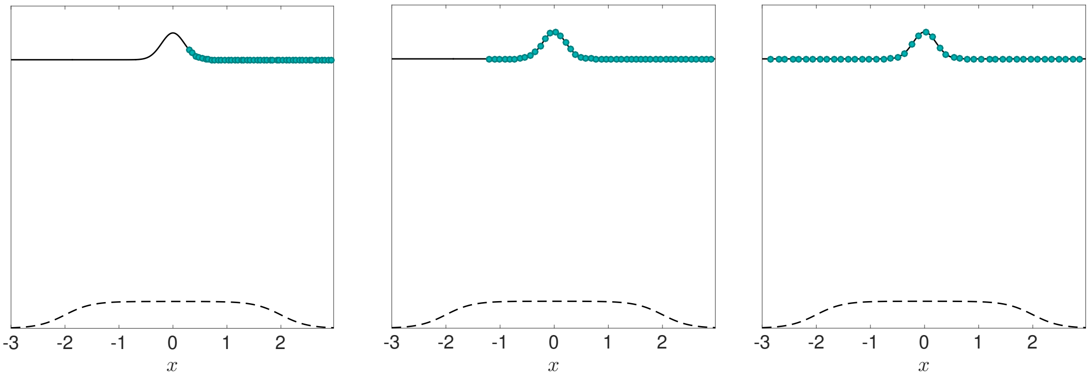
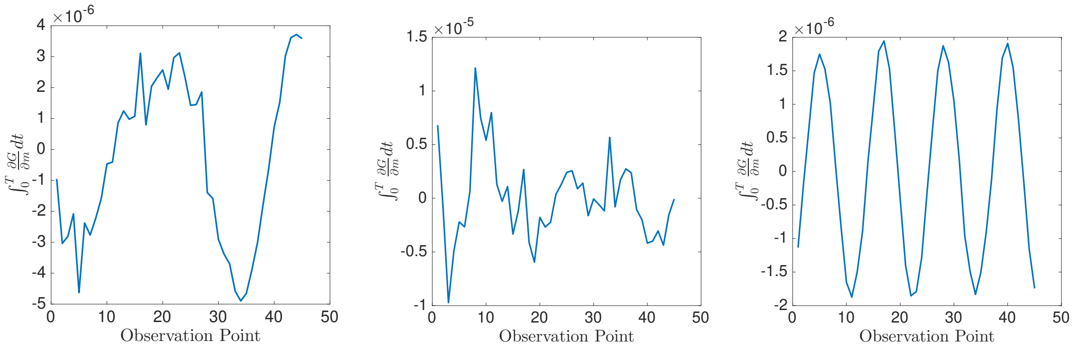


FIGURE 4.9: Case II: The time integrated sensitivity of the surface wave error $\int_0^T \frac{\partial G}{\partial m} dt$ as the location of the first observation point is varied such that the observation points cover a greater proportion of the domain and the initial conditions support.

Results for case II are shown in figure 4.9. We note that increasing the spacing Δx improves the time integrated sensitivity $\int_0^T d\mathcal{G}/dm\, dt$. We note that this may be due to the increased "coverage" of the bathymetry support by the observation points. We observe that configurations of observations for case II in figures 4.9(b) and (c) show larger coverage of the bathymetry feature by observations than in

4.9(a), where only half of the bathymetry is observed by the right propagating surface wave.



(a) $x_1 = 0.3$, $\Delta x = 0.06$ (b) $x_1 = -1.19$, $\Delta x = 0.09$ (c) $x_1 = -2.83$, $\Delta x = 0.13$

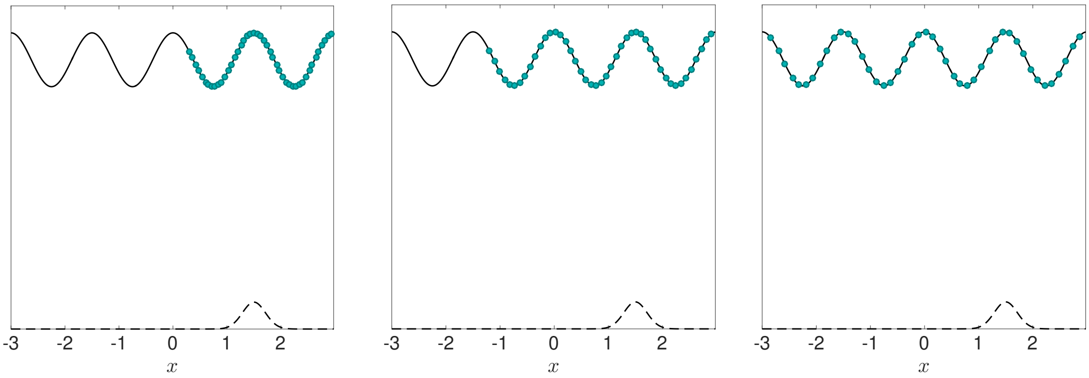
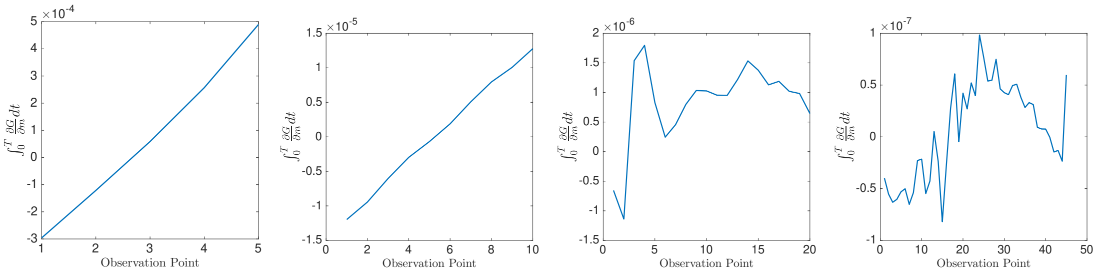
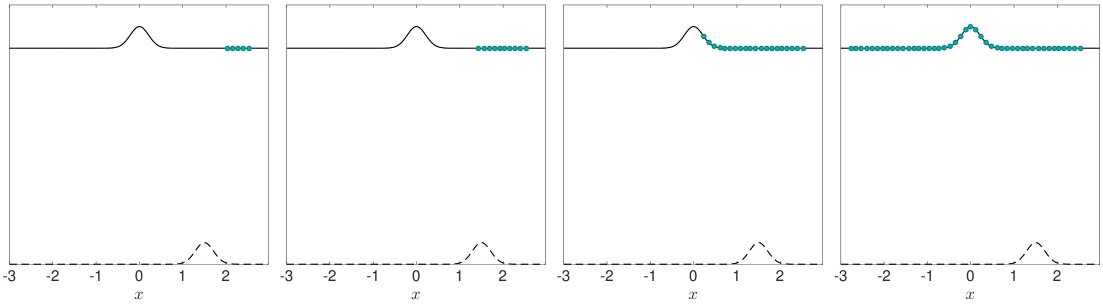


Figure 4.10: Case III: The time integrated sensitivity $\int_0^T \frac{\partial G}{\partial m} dt$ as the location of the first observation point is varied such that the observation points cover a greater proportion of the domain and the initial conditions support.

The results for case III are summarised in figure 4.10. There does not appear to be a significant change in the time integrated sensitivity shown in figure 4.10(a)-(c). Our hypothesis that the position of the observation points relative to the support of the bathymetry significantly affects the sensitivity $d\mathcal{G}/dm$ like for case II, is strengthened by the results in 4.10. In each configuration 4.10(d)-(f), the first observation point is positioned to the left of the bathymetry perturbation, and the bathymetry is always covered by the "observed zone" for each Δx. This was also true for case I, but not for case II, and as a greater proportion of the bathymetry support was covered by the observation points, the sensitivity decreased. Based on this result we suggest that while varying the spacing Δx does not have a significant impact on $\int_0^T d\mathcal{G}/dm\ dt$, increasing the coverage of the observation network such that it covers the support of the bathymetry, results in lower sensitivity of the response function G to the observations $m_j(t)$.



(a) $N_{obs} = 5$ (b) $N_{obs} = 10$ (c) $N_{obs} = 20$ (d) $N_{obs} = 45$

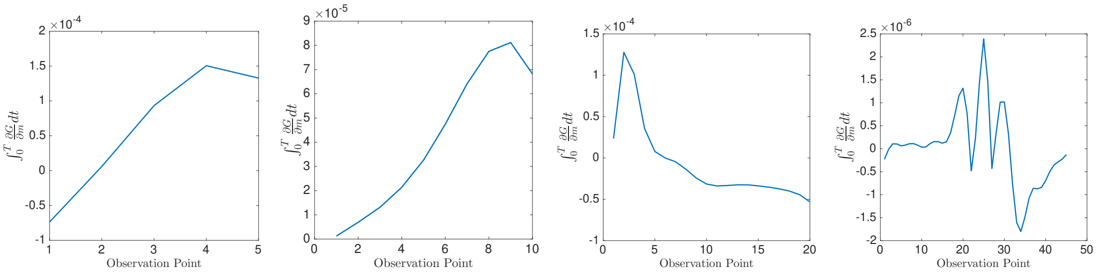


(e) (f) (g) (h)

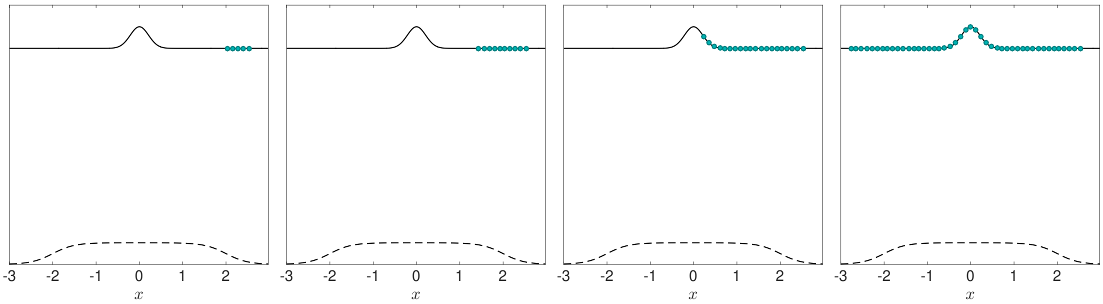
FIGURE 4.11: Case I: The time integrated sensitivity $\int_0^T \frac{\partial G}{\partial m} dt$ as the number of the observation points is varied.

Having analysed the *placement* of the observation points, we now consider results when varying the *number* of observation points. We conduct trials of the data assimilation algorithm, where each trial has an increasing number of observation points, and the spacing Δx is fixed at 0.12. Like the analyses for varied spacing, the position of the final observation point is $x_{N_{obs}} = 2.52$, and as N_{obs} increases, the "coverage zone" of the observations does as well. We conduct trials for $N_{obs} = 5, 10, 20,$ and 45. The results for each case are summarised in figures 4.11, 4.12 and 4.13 respectively.

For case I there is a clear decrease in sensitivity as more observation points are considered. We see the sensitivity go from $\mathcal{O}(10^{-4})$ to $\mathcal{O}(10^{-7})$ as N_{obs} is increased from 5 to 45. For case II, we see in figure 4.12 that while the decrease in sensitivity is not as large as for case I, we do see a decrease by two orders of magnitude as we increase N_{obs} from 5 to 45.



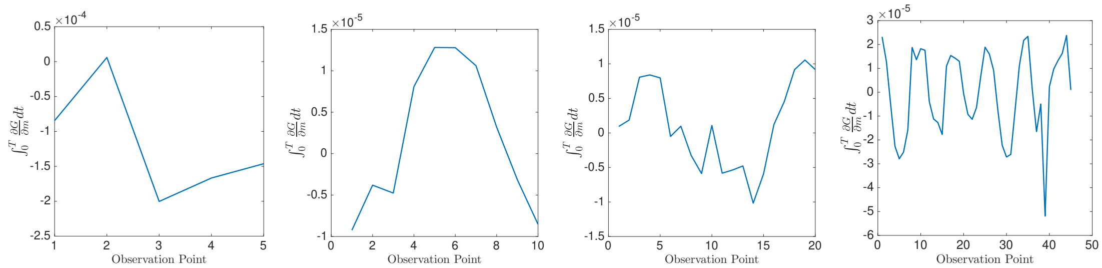
(a) $N_{obs} = 5$ (b) $N_{obs} = 10$ (c) $N_{obs} = 20$ (d) $N_{obs} = 45$



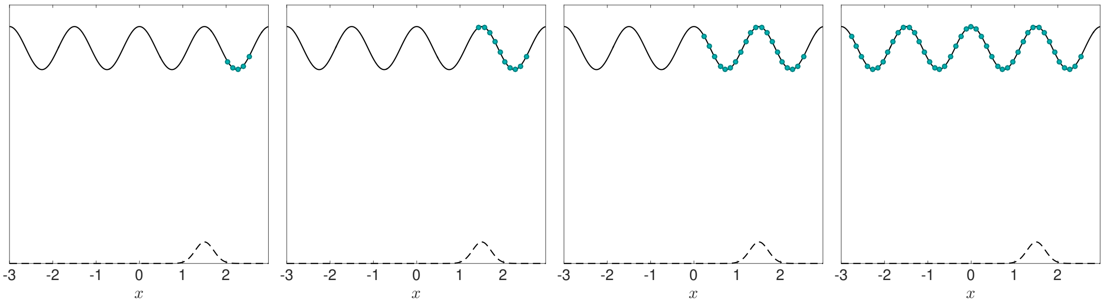
(e) (f) (g) (h)

FIGURE 4.12: Case II: The time integrated sensitivity $\int_0^T \frac{\partial G}{\partial m} dt$ as the number of the observation points is varied.

The results for case III are presented in figure 4.12. We note that like case II, there is a decrease in sensitivity by an order of magnitude when increasing N_{obs} from 5 to 10, however there does not seem to a further increase when increasing N_{obs} from 10 to 20 or 45. This suggests that for case III, there is little change in the surface wave error given the reconstructed bathymetry, to perturbations in observations with varying number of observation points. We note that this is not dissimilar to the qualitative analyses of the bathymetry data assimilation results for case III, where changing the number of observation points did not have a significant improvement in the bathymetry reconstruction error.



(a) $N_{obs} = 5$ (b) $N_{obs} = 10$ (c) $N_{obs} = 20$ (d) $N_{obs} = 45$

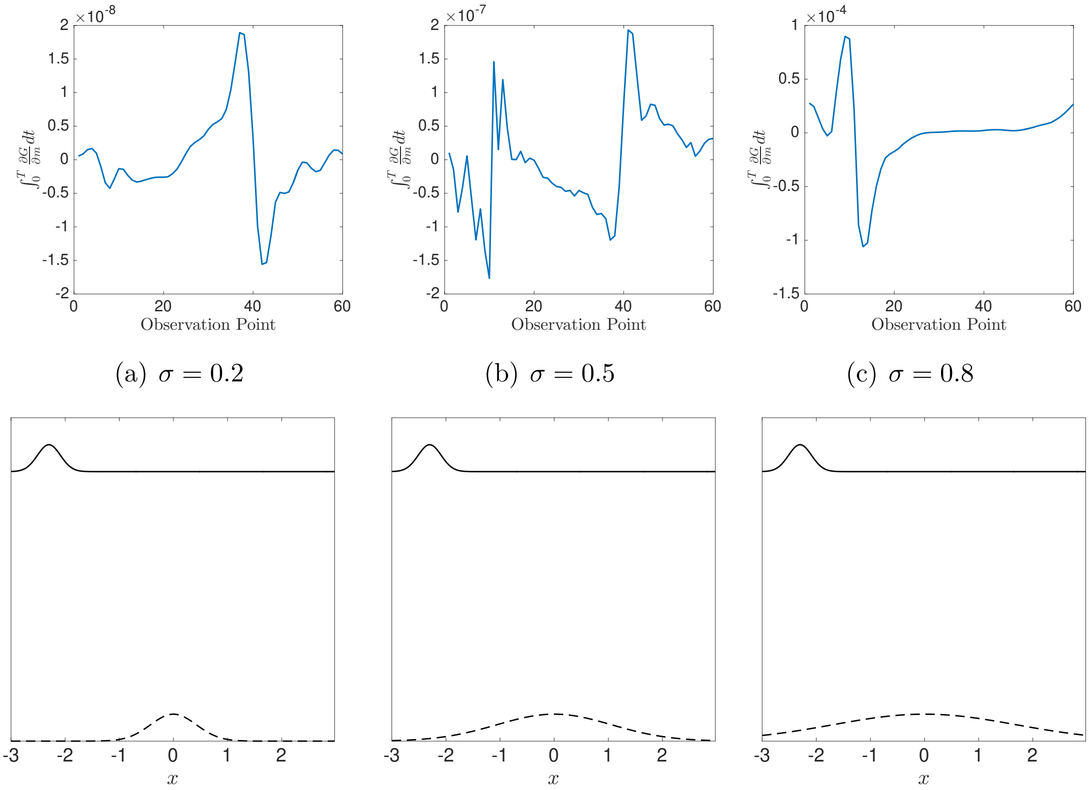


(e) (f) (g) (h)

FIGURE 4.13: Case III: The time integrated sensitivity $\int_0^T \frac{\partial G}{\partial m} dt$ as the number of the observation points is varied.

Based on these results, we can see a contrast emerging between case I and case II, where we have the same initial conditions (Gaussian) but differing bathymetry. Subsequently in our next analysis, we run trials of the data assimilation scheme where we begin with a Gaussian as in case I, and iteratively increase the standard deviation of the bathymetry such as it becomes closer in shape qualitatively to a sandbar type bathymetry, with flatter curvature. We consider two different positions for the initial condition, given in figures 4.14 and 4.15 respectively. This is because for case I the bathymetry is to the right of the support of the initial condition, and so we replicate this positioning in figure 4.14. However, in case II the support for both the bathymetry and the initial conditions is centred at $x = 0$, and so we replicate this for the set of results in figure 4.15.

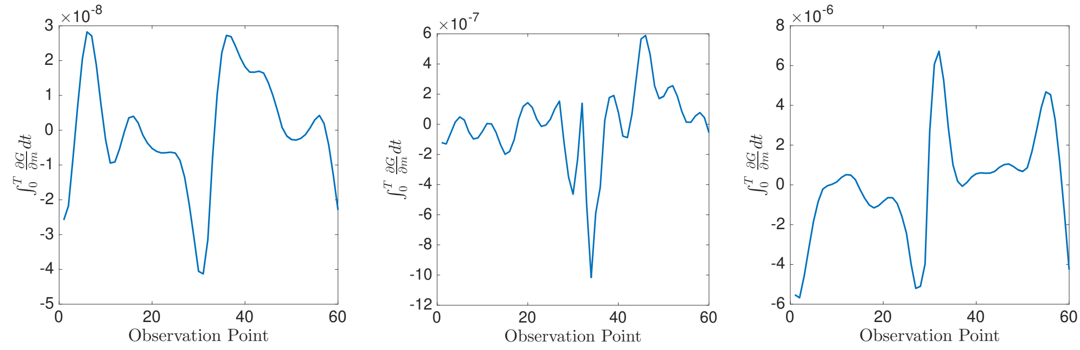
In both 4.14 and 4.15 it is clear that the flatter the bathymetry shape, the higher the sensitivity of the response function to observations. The similarity between the 4.14 and 4.15 suggests that the position of the bathymetry relative to the initial conditions support does not influence sensitivity, but the curvature of the bathymetry does. Each trial was conducted with $N_{obs} = 45$ and $\Delta x = 0.1$, with the position of the first observation at $x_1 = -L$ (complete coverage of the domain by the observations).



(a) $\sigma = 0.2$ (b) $\sigma = 0.5$ (c) $\sigma = 0.8$

FIGURE 4.14: Comparison with Case I: $\int_0^T \frac{\partial G}{\partial m} dt$ as the standard deviation of the bathymetry Gaussian is increased.

Having gained some insight into the influence of spacing and number of points on sensitivity, our ultimate objective is to gauge whether these conclusions regarding the sensitivity of the surface wave error produced by the bathymetry reconstruction to observations $\partial \mathcal{G}/\partial m$, can be linked back to the main conclusions of chapter 3. This analysis is given in Question III.



(a) $\sigma = 0.2$ (b) $\sigma = 0.5$ (c) $\sigma = 0.8$

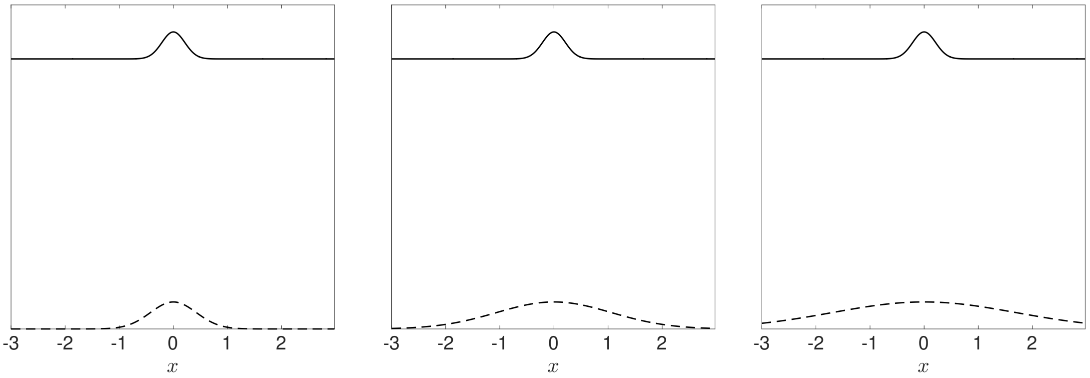


FIGURE 4.15: Comparison with Case II: $\int_0^T \frac{\partial G}{\partial m} dt$ as the standard deviation of the bathymetry Gaussian is increased.

Question III

In chapter 3, we presented results for optimally reconstructing bathymetry using surface wave observations in a variational data assimilation algorithm. We concluded that the data assimilation results are improved by (i) increasing the number of observation points, and (ii) maintaining an optimal ratio between the amplitude of the initial conditions and bathymetry relative to the average depth. We also observed the error in the surface wave produced by the reconstructed bathymetry was orders of magnitude smaller than the error in the bathymetry reconstruction, suggesting low sensitivity of the surface wave to bathymetry reconstruction error.

In terms of the current sensitivity analysis, the question we wish to address is:

Does implementation of the aforementioned conditions (i) and (ii) also result in low sensitivity of the surface wave error (4.6.1) *to perturbations in the observations* $\{m_j(t)\}$*, and is this sensitivity orders of magnitude lower than the reconstruction error in the bathymetry?*

We provide an answer by verifying the following conclusions using the sensitivity analysis results.

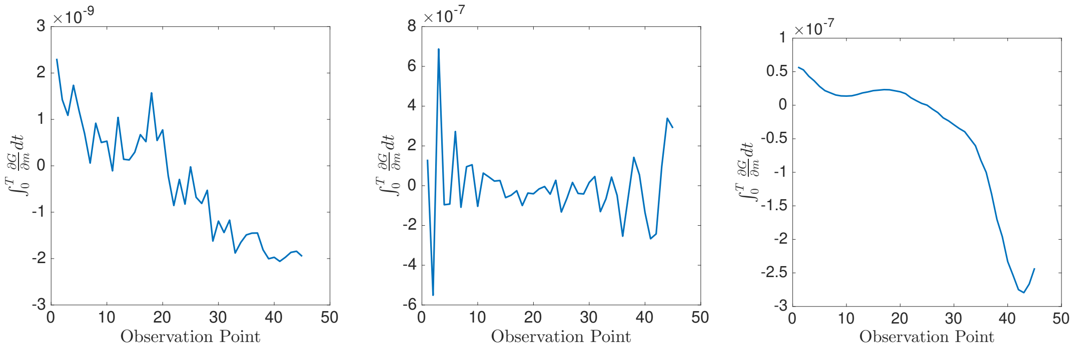
C. i. Does increasing the number of observations (with fixed spacing Δx) result in lower sensitivity of the surface wave error to the observations?

C. ii. Does a sub-optimal ratio (that lead to non-convergent results in chapter 3) between the amplitudes of the initial conditions and bathymetry result in greater sensitivity of the surface wave error to observations?

C. iii. If the reconstruction error in the bathymetry is high (i.e. 10%), does that imply that the sensitivity of the resulting error in the surface wave to observations is also proportionately high?

We already observed the results of (**C. i**) in the previous section where our analyses in figures 4.11, 4.12, and 4.13 indicated lower sensitivity across all three cases when the number of observation points was increased (with fixed spacing $\Delta x = 0.12$). Thus our hypothesis that increasing the number of observation points improves our data assimilation results is supported by the second order adjoint sensitivity analysis.

(**C. ii**) is based on results in chapter 3, where we analysed the affect of varying $\hat{\eta}/\hat{\beta}$ and $\hat{\beta}/H$ on the resulting reconstruction error in the bathymetry. Here $\hat{\eta}$ represents the amplitude of the initial condition, $\hat{\beta}$ is the amplitude of the bathymetry, and H is the average sea depth (normalised to $H = 1$). We observed that convergence was sub-optimal when $\hat{\eta}/\hat{\beta}$ was greater than $\mathcal{O}(10^{-2})$ and bathymetry amplitudes were large (over 10% of the average sea depth H).



(a) $\hat{\beta}/H = 1\%$ (b) $\hat{\beta}/H = 15\%$ (c) $\hat{\beta}/H = 30\%$

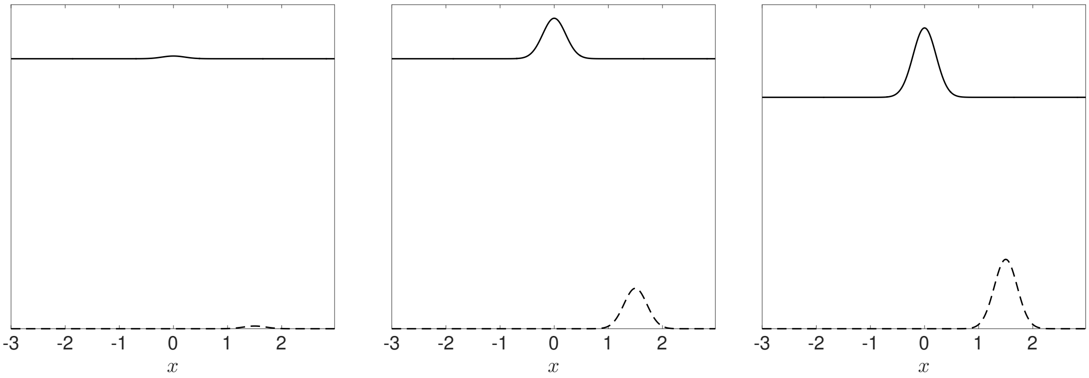
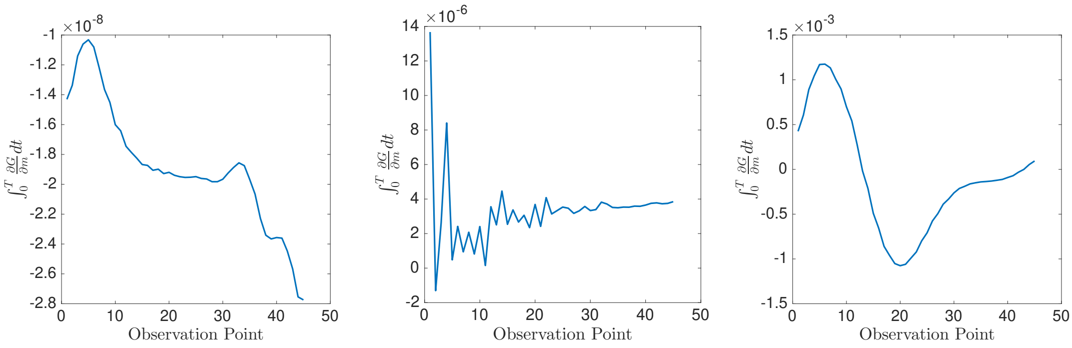


FIGURE 4.16: Case I: Time integrated sensitivity $\int_0^T \frac{\partial G}{\partial m} dt$ as the relative amplitude of the bathymetry is increased.

As having $\hat{\eta}/\hat{\beta} \leq \mathcal{O}(10^{-2})$ is necessary for convergence of the data assimilation, we do not explicitly violate this condition in our analysis. Instead, we analyse the effect on the sensitivity of the surface wave error to observations, as $\hat{\beta}$ is varied (and $\hat{\eta}$ is fixed at 1% of $\hat{\beta}$). The relationship between the ratio $\hat{\beta}/H$ and the sensitivity $\partial \mathcal{G}/\partial m$ can be observed in figures 4.16, 4.17, and 4.18. The figures correspond to cases I, II and III respectively. For each case, we present the time-integrated sensitivity as $\hat{\beta}$ is varied from 0.01 to c, where c is the approximate cut-off value beyond which the surface wave error became larger than 0.1% (as shown in chapter 3). In figure 4.16(a) - (c), results are shown for case I with $\hat{\beta}/H = 1\%, 15\%$, and 30% respectively. We observe that when the bathymetry is 1% of the depth, the sensitivity is $\mathcal{O}(10^{-9})$, and as $\hat{\beta}$ increases to 30% of the depth, the sensitivity increases to $\mathcal{O}(10^{-7})$.



(a) $\hat{\beta}/H = 2\%$ (b) $\hat{\beta}/H = 8\%$ (c) $\hat{\beta}/H = 20\%$

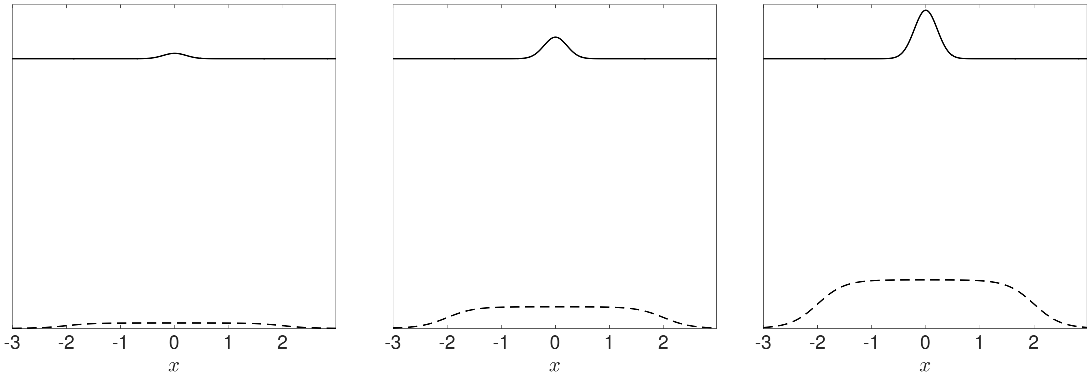
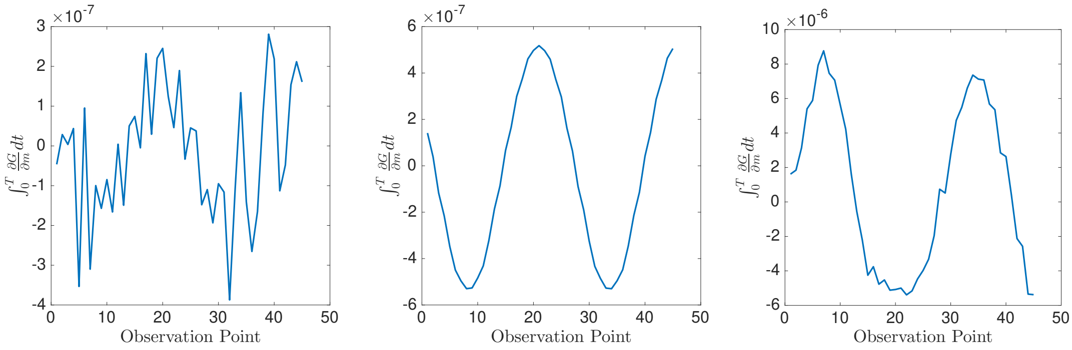


Figure 4.17: Case II: Time integrated sensitivity $\int_0^T \frac{\partial G}{\partial m} dt$ as the relative amplitude of the bathymetry is increased.

Similarly in figure 4.17, we consider the sensitivity for case II with $\hat{\beta}/H = 2\%, 8\%$, and 20%. We see that as $\hat{\beta}$ increases, the sensitivity increases from $\mathcal{O}(10^{-8})$ by at at least two orders of magnitude for each value of $\hat{\beta}$, indicating a clear correlation between the normalised height of the bathymetry $\hat{\beta}$ and the sensitivity of the surface wave error to observations.



(a) $\hat{\beta}/H = 1\%$ (b) $\hat{\beta}/H = 5\%$ (c) $\hat{\beta}/H = 10\%$

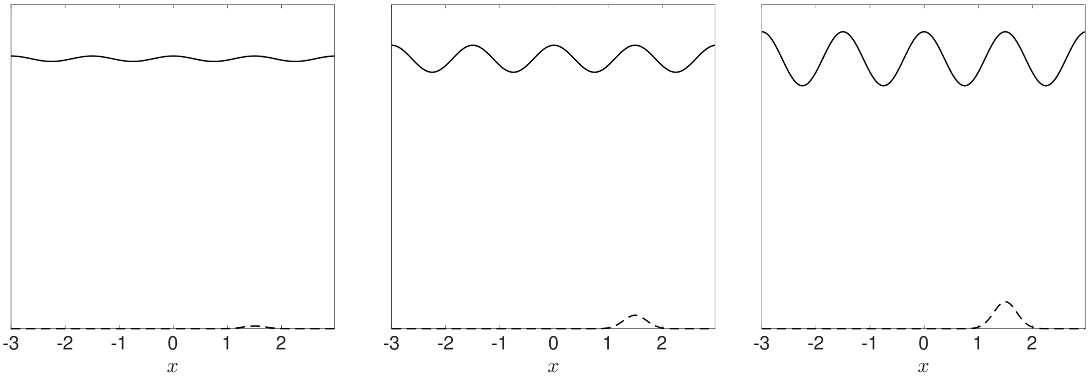


Figure 4.18: Case III: Time integrated sensitivity $\int_0^T \frac{\partial G}{\partial m} dt$ as the relative amplitude of the bathymetry is increased.

Figure 4.18 shows the results for case III, for $\hat{\beta}/H = 1\%, 5\%$, and 10. The error does not decrease from $\mathcal{O}(10^{-7})$ for the first two $\hat{\beta}$ values, but we observe a decrease to $\mathcal{O}(10^{-6})$ when $\hat{\beta} = 0.1$.

To summarise, in each of the cases I, II, and III sensitivity increases with $\hat{\beta}/H$. We note that for cases I and III especially, this increase occurs when $\hat{\beta}/H$ increases from 1% to 10%. For case I sensitivity did not vary significantly when $\hat{\beta}/H$ was 10% or larger (see figure 4.16(b) and figure 4.16(c)). Similarly for case III, figure 4.18 shows constant sensitivity at $\mathcal{O}(10^{-7})$ when $\hat{\beta}/H$ is 1%. Based on these observations, we conclude that a lower relative bathymetry height $\hat{\beta}/H$ decreases the sensitivity of $\mathcal{G}$ to changes in the observations $\{m_j(t)\}$.

Perhaps the most significant question we aim to address with this analysis is

(**C. iii**). In chapter 3 we found that the error in the surface wave was orders of magnitude lower than the error in the bathymetry, suggesting low sensitivity of the surface wave to bathymetry reconstruction error. This implies relatively large tolerance levels for bathymetry reconstruction error in tsunami models may be acceptable, at least in the one-dimensional case. We now wish to make this conclusion more rigorous, and verify whether the second order adjoint sensitivity of the surface wave error to the observations $\partial\mathcal{G}/\partial m$ also exhibits the same behaviour. If $\partial\mathcal{G}/\partial m$ were the same order of magnitude as bathymetry reconstruction error, this would imply that if the bathymetry reconstruction is sub-optimal, the accuracy of surface wave prediction is sensitive to changes in observations. This is undesirable from a forecasting perspective as it indicates predictions may vary greatly based on small changes in the observations. Due to the ill-posedness of inverse problems of this kind, errors in measurements can be amplified and subsequently have a large affect on the surface wave error.

However, based on our analyses we observe clearly that in all cases considered the sensitivity $\partial\mathcal{G}/\partial m$ is orders of magnitude lower than the optimal values for the bathymetry reconstruction error presented in table 4.2. This relationship does not change even for the worst sensitivity results (observed for case II in figure 4.17(c), where the sensitivity is $\mathcal{O}(10^{-3})$. This is still an order of magnitude smaller than the lowest error observed for case II, which was $\mathcal{O}(10^{-2})$. Additionally, we observed in our analysis that even when the bathymetry reconstruction error is sub-optimal the sensitivity can be relatively far smaller, as in figures 4.16(a) and 4.17(a) where the sensitivity was $\mathcal{O}(10^{-9})$ and $\mathcal{O}(10^{-8})$ respectively, while the bathymetry reconstruction error was $\mathcal{O}(10^{-1})$.

4.7 Conclusion

In this chapter, we analytically derived the sensitivity of a response function $\mathcal{G}$ (a function of optimal data assimilation results) to the observations m, using second order adjoint techniques and the Hessian of the original cost function $\mathcal{J}$. We did this for the initial conditions assimilation and for the bathymetry assimilation respectively. In each case we first derive the Hessian product $\mathcal{H}\nu$ given some arbitrary perturbation of the control variable ν, and then demonstrate that deriving the sensitivity $\partial\mathcal{G}/\partial m$ involves solving the forced equation $\mathcal{H}\nu = \mathcal{F}$ with the right hand side $\mathcal{F}$ dependent on the optimal assimilation results. For the present study, we chose the response $\mathcal{G}$ to be the relative L^2 error in the surface wave, produced by the reconstructed bathymetry.

We numerically implemented the sensitivity algorithm for the bathymetry assimilation case, in order to further investigate and confirm the conclusions of the bathymetry assimilation from chapter 3. We formulated three questions. Question I asks whether there is a link between the sensitivity $\partial\mathcal{G}/\partial m$ and the convergence of the reconstructed bathymetry to the exact form. The results for the three bathymetry and initial conditions configurations outlined in table 4.1 showed that a higher error in bathymetry reconstruction is associated with higher sensitivity of the surface wave error to perturbations in the observations.

Question II asks how changing parameters in the data assimilation scheme (such as the spacing and number of observation points, or shape of the bathymetry) affects the sensitivity. We analysed the sensitivity $\partial\mathcal{G}/\partial m$ as we varied the observation point placement and number of points. We concluded that for cases I and III , the spacing Δx of a fixed number of observation points was non-influential and that for all values of Δx considered, the resulting observation configuration covered the entire support of the bathymetry. This was not true for case II where only the largest spacing $\Delta x = 0.13$ corresponded to coverage of the entire bathymetry support by observations, and we observed that the sensitivity decreased as the observation points spanned a larger interval of the bathymetry support.

To gain further insight into the differences between cases I and II and their differing bathymetry shapes (same localized surface wave in each case but compact and “spread out” bathymetry respectively), we conducted trials where we incrementally increased the standard deviation of the Gaussian bathymetry in case I, until its curvature qualitatively resembled the sandbar bathymetry in case II. We observed that the higher standard deviation led to an increase in the time-integrated sensitivity, regardless of the position of the bathymetry relative to the initial condition.

Finally for question III, we investigated whether our previous conclusions could be confirmed by the sensitivity analysis. One of the main conclusions from chapter 3 was that convergence improved with more observation points. Therefore, we investigated whether increasing the number of observations also resulted in lower sensitivity of the surface wave error to the observations. The results given in figures 4.11, 4.12, and 4.13 indicated lower sensitivity across all three cases when the number of observation points was increased, confirming our earlier conclusion.

Our second conclusion from chapter 3 addressed the relationship between the bathymetry and initial conditions amplitudes $\hat{\beta}$ and $\hat{\eta}$, and the average depth H. We observed that a necessary condition for convergence of the reconstructed bathymetry was that the relative height of the initial conditions compared to the bathymetry height $\hat{\eta}/\hat{\beta}$ be less than or equal to $\mathcal{O}(10^{-2})$ when $\hat{\beta}/H$ was larger than 10% of the depth. In the sensitivity analysis we varied the relative bathymetry amplitude $\hat{\beta}/H$ (keeping $\hat{\eta}$ at 1% of $\hat{\beta}$, to investigate the effect of the relative bathymetry amplitude on the sensitivity of the surface wave response $\mathcal{G}$ to the observations. In each of the cases I, II, and III we saw a general increase in sensitivity as the relative bathymetry height $\hat{\beta}/H$ increased.

Finally, we verified the observation from chapter (3) that the error in the surface wave is orders of magnitude lower than the error in the reconstructed bathymetry, suggesting low sensitivity of the surface wave to reconstruction error. We therefore investigated whether the sensitivity $\partial\mathcal{G}/\partial m$ exhibited the same behaviour, i.e. whether the sensitivity of the surface wave error to observations was orders of magnitude lower than the bathymetry reconstruction error. We clearly observed that in all cases considered, the sensitivity $\partial\mathcal{G}/\partial m$ was orders of magnitude lower than the optimal values of the reconstruction error presented in table 4.2.

In conclusion, the analyses in this chapter confirm the results observed for the data assimilation in chapter 3. We have shown that the necessary conditions for convergence of the bathymetry reconstruction error, also correspond to low sensitivity of the surface wave error to the observations. By showing that $\partial\mathcal{G}/\partial m$ is orders of magnitude lower than the bathymetry reconstruction error, we stipulate that even if the reconstruction of the bathymetry is sub-optimal, the forecast of surface wave exhibits low sensitivity to changes in observations. High sensitivity of the surface wave error to observations implies that predictions may vary greatly based on small changes in the observations. As errors in measurements can be amplified and subsequently have a large affect on the surface wave error, the low sensitivity observed in these analyses is encouraging from a forecasting perspective, especially in situations where the relative bathymetry reconstruction error is sub-optimal (larger than 10%).

Improvements to the current analysis could be made by ensuring better convergence of $\mathcal{H}\nu = \mathcal{F}$, such that the residual error is decreased further. Equivalent results for the initial conditions assimilation may also shed more light on the sensitivity of surface wave propagation to observations.

We note that these results are for an idealised 1-D case. The next step would be verify the conclusions of this analysis for the 2-D data assimilation, allowing us to move away from the idealised case and use available ocean bathymetry data and observation data in our analysis. In chapter 3 we provided a concise overview on the importance of bathymetry for accurate tsunami modelling. Consequently, the results from the current analysis help us quantify the accuracy required for the reconstructed bathymetry, such that the sensitivity of the the error in the surface wave to the observations remains low. We also gain a better understanding of how large bathymetry features impact the surface wave accuracy, and the effects of the number and placement of observations. Therefore, the results observed here provide an encouraging first step towards a more realistic implementation for tsunami models, and may serve as a benchmark for future analyses.

The analyses in this chapter quantified the sensitivity of the surface wave error given the reconstructed bathymetry to observations. To gain further insight on the sensitivity of the surface wave error to parameters in the models (like the bathymetry and initial conditions amplitudes $\hat{\eta}$ and $\hat{\beta}$), in chapter 5 we use Global Sensitivity Analysis (GSA) techniques to derive sensitivity indices quantifying the variation in the surface wave error resulting from the respective parameters.

Chapter 5

Global Sensitivity Analysis (GSA)

5.1 Introduction

The second order adjoint sensitivity analysis in Chapter 4 quantified the sensitivity of response functions to perturbations in the observations. However, results from Chapter 3 for data assimilation for bathymetry reconstruction demonstrated the influence of other parameters such as initial conditions and bathymetry amplitude. Additionally, while we discussed the impact of number of observations and spacing in detail in Chapter 4, we wish to gain more insight on how the position of the bathymetry relative to observation points effects the reconstruction error and resulting error in the surface wave. In the present study, we use Global Sensitivity Analysis (GSA) to quantify the influence of these additional parameters on the error in bathymetry reconstruction. We also consider the sensitivity of the surface wave error to model parameters given the reconstructed bathymetry. We focus our analysis on a localized surface wave propagating over a compact bathymetry. This work is motivated by the observation in Chapter 3 that there is low sensitivity of surface waves to noise in bathymetry reconstruction. Using GSA we aim to quantify, and subsequently rank the influence of these parameters on the bathymetry and surface wave errors respectively.

We begin with a review of GSA and applications to Earth System Modelling (ESM) in section 5.2.1. In section 5.2.2 we implement a variance-based sensitivity analysis (VBSA) scheme to define sensitivity indices, using the Sobol' method. We decompose the variance of the model output into the respective contributions of each of the model parameters. We present a brief overview of the Sobol' method, and subsequently discuss considerations such as choice of sampling method for the input parameters, base sample size and accuracy of numerical approximations for model variance. In section 5.3 we present the results for the variance-based Sobol' analysis, attempting to quantify the influence of three significant model parameters. We show that the probability distribution of the surface wave error is highly skewed, suggesting that variance-based analysis may not accurately represent the influence of the parameters on model outputs. Therefore, in section 5.4 we perform density based sensitivity analysis (DBSA) as an alternative, and implement the algorithm outlined in Saltelli et al. [48]. This method uses the conditional and unconditional cumulative density functions of the model to derive sensitivity indices, instead of using moments such as variance. Section 5.5 presents the DBSA results, and we demonstrate that density-based sensitivity indices pass tests for robustness and convergence. We show that the most influential parameter for the surface wave error is bathymetry amplitude, whereas the position of the bathymetry ψ (where ψ is the x-coordinate of the Gaussian bathymetry peak) relative to the fixed position of the observation points was most influential on the bathymetry reconstruction error. Finally, in 5.6 we summarise the results and discuss their connection to conclusions of the bathymetry data assimilation presented in Chapter 3.

5.2 Sensitivity analysis methods, derivation, and sampling considerations

5.2.1 Review of GSA

Saltelli et al. [48] define GSA as the set of mathematical techniques used to assess the propagation of uncertainty in a numerical model. In practice, a set of synthetic indices are derived, that quantify the relative contribution to output variance from different input parameters. These are known as sensitivity indices. Liu and Homma [29] propose that good sensitivity indices should exhibit the following properties:

1. They are global, and consider the influence of model inputs on the entire output range.
2. They are quantifiable, and thus can be computed and reproduced numerically.
3. They are not conditional on any assumed input values.
4. The value of the sensitivity index for some input gives an easy interpretation of the sensitivity of the model to the parameter.
5. They are consistent across multiple samples and simulations.
6. They are ideally moment-independent, and not reliant on a specific quantitative measure of the output distribution.

GSA methods can be categorised as one-at-a-time methods (OAT), where output variations are induced by varying one input at a time, or all-at-a-time methods (AAT), where all input factors are varied simultaneously. A well known methods is the Elementary Effect Test (EET), which computes output perturbations from multiple points in the input space, and quantifies global sensitivity by aggregating individual sensitivities. Another is the Fourier Amplitude Sensitivity Test (FAST), a variance-based method that uses the Fourier expansion of the output for the sensitivity index derivation. For more extensive descriptions we refer the reader to Pianosi et al. [40].

GSA methods are especially prevalent in Earth Systems Modelling (ESM). An example is the work conducted by the Modelling, Observations, Identification for Environmental Sciences (MOISE) project [1] on variance-based sensitivity analysis on a marine eco-system model of the Ligurian sea, using the Sobol' method (which we implement in section 5.2.2).

A concise review of the contributions of GSA on advancement of ESM is provided in Wagener and Pianosi [61]. They highlight the surge in computational capacity for Earth and climate models, and address the resulting problem of increased interaction between model components and parameters, even when representing a relatively low number of physical processes. This issue is particularly problematic in ESM where incomplete knowledge and the lack of relevant exact solutions makes model validation very difficult. In addition, such models are forced with noisy and inaccurate observations. Typical uses of GSA include:

- Ranking the influence of the parameters on model output from highest to lowest.
- Creating a threshold criteria for sensitivity allowing us to determine parameters of negligible influence (screening).
- Finding thresholds in the input parameter values that map into specific output regions (factor mapping).

Thus, the primary objective of this study is to use GSA to analyse the data assimilation model for bathymetry detection presented in Chapter 3, in order to rank, screen and factor map the influence of its parameters. A review of the data assimilation input-output model used in the analysis, and the VBSA algorithm used to derive the sensitivity indices is provided in the following section.

5.2.2 Variance Based Sensitivity Analysis (VBSA)

We begin with a brief review of the model, and provide a derivation of the variance-based sensitivity indices using the Sobol' method [56]. In Chapter 3, we measured the relative L^2 reconstruction error in the bathymetry

$$Y_\beta = \frac{\| \beta^{(t)} - \beta^{(b)} \|_{L^2(\Omega_X)}}{\| \beta^{(t)} \|_{L^2(\Omega_X)}}, \tag{5.2.1}$$

where $\beta^{(t)}$ is the "true" bathymetry and $\beta^{(b)}$ is the "best" reconstruction obtained via the data assimilation scheme. We also found that the relative L^2 error in the propagating surface wave $\eta(x,t)$ given the reconstructed bathymetry $\beta^{(b)}$

$$Y_\eta = \frac{\| \eta(\beta^{(t)}) - \eta(\beta^{(b)}) \|_{L^2(\Omega_{X\times T})}}{\| \eta(\beta^{(t)}) \|_{L^2(\Omega_{X\times T})}}, \tag{5.2.2}$$

was orders of magnitude lower than the bathymetry error (5.2.1), suggesting low sensitivity to the bathymetry error, as a function of the bathymetry amplitude

$\hat{\beta}$. Based on this initial observation, we wish to derive a more rigorous way to measure sensitivity of results to parameters $\hat{\eta}$ (amplitude of the surface wave), $\hat{\beta}$ and ψ (amplitude and position of the bathymetry respectively) such that

$$\eta^{(t)}(x,0) = \hat{\eta}e^{-(10x)^2} \tag{5.2.3}$$

$$\beta^{(t)}(x) = \hat{\beta}e^{-\left(10(x-\psi)\right)^2}. \tag{5.2.4}$$

We solved the 1-D SWE system (3.3.1) using a second order finite difference approximation in space, and a four stage third order Runge-Kutta scheme in time. The mean depth H is normalised to 1. The resolution of our spatial grid is $N = 256$, and our spatial domain is $\Omega_X = \{x \in \mathbb{R}; -L \leq x \leq L\}$. The system is integrated over $\Omega_T = \{t \in \mathbb{R}; 0 \leq t \leq T\}$, with final time $T = 2L$. Periodic boundary conditions are imposed at $x = L$ and $x = -L$ where $L = 3$. We assume we have no background information for bathymetry a priori, and set our initial guess for the bathymetry $\beta^{(g)}(x)$ to 0. We focus on case I from Chapters 3 and 4, where we have compactly supported Gaussian initial conditions centred at zero, and a localised Gaussian bathymetry centred at $x = \psi$. We use 45 equidistant observation points with spacing Δx, and the first observation is located at $x = 0.3$. An example configuration with $N_{obs} = 45$ and $\psi = 1.5$ is shown in figure 5.2.2.

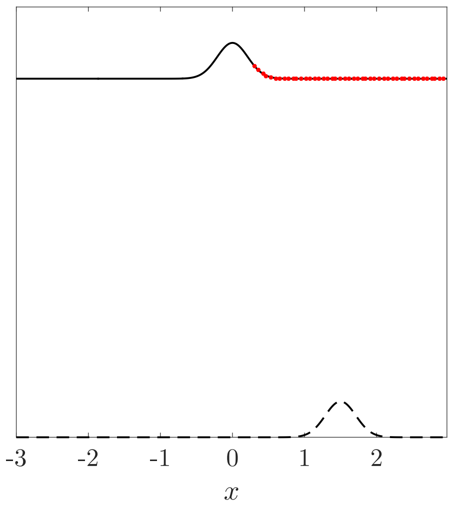


FIGURE 5.1: Gaussian initial conditions and localised Gaussian bathymetry centred at $\psi = 1.5$ (bathymetry and initial conditions amplitudes not to scale), with 45 observation points. Note that the value of ψ is the position of the peak of the Gaussian bathymetry.

The key property of ψ is that it determines the location of the bathymetry relative to the observation points. We see from figure 5.2.2 that when ψ is small, a smaller proportion of observation points observes the surface wave before it interacts with the bathymetry, and for ψ approximately less than 1, the observations do not span the entire bathymetry support. In Chapter 4 section 4.6, we observed that having a larger number of observation points placed to the right of the bathymetry support (and also spanning it) corresponded to low sensitivity of the surface wave error to observations, suggesting that this relative position is a significant parameter for surface wave accuracy. We wish to extend this analysis and discover whether the influence of shifting the bathymetry position ψ relative to the fixed observation points confirms this result.

Using the variance based sensitivity analysis (VBSA) outlined by Sobol [56], we formulate our model as $f(\bar{\eta}, \bar{\beta}, \psi) = Y$, where output Y is either the error in the bathymetry (5.2.1) or error in the surface wave (5.2.2). We represent (5.2.1) by Y_β and (5.2.2) by Y_η.

If we denote the input variables $(\bar{\eta}, \bar{\beta}, \psi)$ as $x = (x_1, x_2, x_3)$ then we can decompose the variance of $f(x_1, x_2, x_3)$ into fractions attributed to each parameter individually and to coupled interactions,

$$f = f_0 + f_1 + f_2 + f_3 + f_{1,2} + f_{1,3} + f_{2,3} + f_{1,2,3}. \tag{5.2.5}$$

If f is a square integrable function, it follows that the fractions (5.2.5) are as well, and thus we denote the constants

$$D = \int f^2(x)\ dx - f_0^2 \tag{5.2.6}$$

$$Di_1..i_s = \int f^2_{i_1..i_s}(x_{i_1}...x_{i_s})\ dx_{i_1}..dx_{i_s} \tag{5.2.7}$$

as variances of the function f and its fractional decomposition respectively, where $i_1 < .. < i_s \leq 3$. We observe that

$$D = \sum_{i=1}^{3} \sum_{i_1<..<i_s}^{3} D_{i_1..i_s}. \tag{5.2.8}$$

As stipulated in Sobol [56], if x is a uniformly distributed random point in the 3-dimensional parameter space, then $f(x)$ and $f_{i_1..i_s}(x_{i_1}...x_{i_s})$ are random variables with variance D and $Di_1..i_s$ respectively. Subsequently, we define the global sensitivity indices as the ratios

$$Si_1..i_s = \frac{Di_1..i_s}{D}. \tag{5.2.9}$$

Sensitivity indices $S_{i_1..i_s}$ are the proportion of total variance represented by each component. All correctly derived sensitivity indices should be non-negative, and should sum to 1, such that

$$S_1 + S_2 + S_3 + S_{1,2} + S_{1,3} + S_{2,3} + S_{1,2,3} = 1 \tag{5.2.10}$$

We use the sensitivity indices $Si_1..i_s$ as a useful proxy to investigate the relative variance contributed to our model $f(x)$ by each parameter. We can also compute sensitivity indices for subsets of our parameter space, for example calculating the sensitivity index $S_1^{(Tot)}$ which quantifies the total variance contributed by the parameter x_1. If we consider a subset $y = (x_{k_1}, ..., x_{k_m})$, of our parameter space such that $1 \leq m < 3$ and $1 \leq k_1 \leq ... \leq k_m \leq 3$, then we can partition our parameters such that we have $x = (y, z)$ where z is the set of $3 - m$ complementary variables. Then the variance corresponding to the subset y is

$$D_y = \sum_{i=1}^{3} \sum_{(i_1 < .. < i_s) \in K} f_{i_1..i_s}(x_{i_1}...x_{i_s}), \quad K = (k_1, ..., k_m), \tag{5.2.11}$$

and by defining D_z in a similar manner, the *total* variance contributed by y is

$$D_y^{(Tot)} = D - D_z, \tag{5.2.12}$$

We demonstrate the difference between the two with an example in the current context.

Assume $y = (x_1, x_3)$ and $z = (x_2)$. Then,

$$D_y = D_1 + D_3 + D_{(1,3)}, \tag{5.2.13}$$

$$D_y^{(Tot)} = D - D_z = D - D_2 = D_1 + D_{(1,2)} + D_{(1,2,3)} + D_{(2,3)} + D_3 + D_{(1,3)}. \tag{5.2.14}$$

And so we define two global sensitivity indices for the set of parameters in y,

$$S_y^{(Tot)} = \frac{D_y^{(Tot)}}{D} \tag{5.2.15}$$

$$S_y = \frac{D_y}{D}. \tag{5.2.16}$$

The values of $S_y^{(Tot)}$ and S_y respectively cast light on the influence of the subset y; for example if $S_y^{(Tot)} = S_y = 1$, this implies the total variations of the model $f(x)$ can be explained by y alone, and hence $f(x)$ has no dependence on z. By this procedure, we can **screen** which variables can be neglected from the model. We call the indices S_1, S_2, S_3 the *first order indices* and $S_{(1,2)}, S_{(1,3)}, S_{(2,3)}$ the *second order indices* respectively. The former are relatively simpler to compute, and can be used to **rank** the input parameters by the sensitivity index values by ordering them from most to least influential on the model variance. However the latter can also be useful, for example, in determining whether the interaction of two variables has a greater effect on the model output than the individual variables alone.

The efficacy of such analyses hinges on the ease of numerical approximation of model variance, and subsequently the sensitivity indices. Sobol [56] outlines a Monte Carlo algorithm that obtains estimates for the variances D, D_y, D_z. Independent points are chosen randomly from the uniformly distributed parameter space and, by conducting enough trials, Monte Carlo estimates are formed for the variances which converge stochastically to the integrals in (5.2.6) and (5.2.7). Further details of the numerical approximation of the sensitivity indices and sampling strategy is given in the following section.

5.2.3 Numerical approximation and sampling strategy

To approximate the variance of the model, and subsequently derive the sensitivity indices, we must sample input parameters such that the model evaluated at these points adequately represent the spread of the probability distribution of Y. One of the most commonly used sampling strategy in GSA is Latin Hypercube Sampling [40].

Latin Hypercube sampling (LHS) is a stratified sampling technique, where the population the model parameters are sampled from is partitioned into homogeneous subpopulations (strata) before sampling [38]. Thus, every sample point can be assigned to a single stratum. In LHS, the sample is generated such that each of the d (uniformly distributed) parameters is stratified into N equal strata. For example when $d = 2$, the sample space is partitioned into a "latin square" where each row and each column contains only a single sample. The latin hypercube is the generalisation of this concept to higher dimensions. When sampling, the range of each of the d parameters is broken up into N equally probable intervals, from which N sample points are chosen in accord with the latin hypercube requirements. The value of the sample point can be randomly generated from a uniform distribution, or it can be chosen at a fixed location within the interval, such as the midpoint. When $N >> d$, the LHS technique can be very effective in capturing the influence of each parameter [49].

For the current analysis, we use a tailored approach considered appropriate for VSBA where a sample of size $2N$ is generated using LHS, then a further dN samples are built using recombinations of the base sample vectors. These are subsequently used to derive Monte Carlo estimates of the sensitivity indices. The latter strategy and computational scheme was outlined by Saltelli et al. [48] and we present a general overview as follows.

Recall that we defined $S_y^{(Tot)} = \frac{D_y^{(Tot)}}{D}$ and $S_y = \frac{D_y}{D}$, where $x = (y, z)$ is a partition of our input parameters, and $D_y^{(Tot)} = D - D_z$. Then to compute the indices $S_y^{(Tot)}$ and S_y we need to approximate the following integrals,

$$\int f(x)\ d(x), \quad \int f^2(x)\ dx, \quad \int f(x)f(y,z')\ dydz', \quad \int f(x)f(y',z)\ dy'dz\ . \tag{5.2.17}$$

The Monte Carlo algorithm outlined in Sobol [56] stipulates that, if we consider two independent random points $\xi = (\gamma, \zeta)$ and $\xi' = (\gamma', \zeta')$ uniformly distributed in the 3-dimensional parameter space, then with N trials the following Monte Carlo estimates can be obtained,

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^{N} f(\xi_i) \xrightarrow{P} f_0, \\
&\frac{1}{N}\sum_{i=1}^{N} f^2(\xi_i) \xrightarrow{P} D + f_0^2 \\
&\frac{1}{N}\sum_{i=1}^{N} f(\xi_i)f(\gamma_j, {\zeta_j}') \xrightarrow{P} D_y + f_0^2 \\
&\frac{1}{N}\sum_{i=1}^{N} f(\xi_i)f({\gamma'}_j, \zeta_j) \xrightarrow{P} D_z + f_0^2,
\end{aligned} \tag{5.2.18}$$

where each trial requires three computations of the model, $f(\gamma, \zeta)$, $f(\gamma', \zeta)$, and $f(\gamma, \zeta')$, and $\xrightarrow{P}$ implies stochastic convergence.

Saltelli et al. [48] implement this approximation using two sample matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ of size $N \times d$, having elements a_{ij} and b_{ij} respectively where $i = 1, ..., N$ and $j = 1, ..., d$. The matrix $\boldsymbol{A}_{\boldsymbol{B}}^{(\boldsymbol{i})}$ is introduced as the matrix with columns from $\boldsymbol{A}$ except the i-th column which is from $\boldsymbol{B}$. Then the following is an improved approximation for D_y and $D_y^{(Tot)}$

$$D_y = \frac{1}{N}\sum_{j=1}^{N} f(\boldsymbol{B})_j(f(\boldsymbol{A}_{\boldsymbol{B}}^{(i)})_j - f(\boldsymbol{A})_j) \tag{5.2.19}$$

$$D_y^{(Tot)} = \frac{1}{2N}\sum_{j=1}^{N}(f(\boldsymbol{A})_j - f(\boldsymbol{A}_{\boldsymbol{B}}^{(i)})_j)^2 \ . \tag{5.2.20}$$

$2N$ model evaluations are required to compute Y given the input sample matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, and a further dN evaluations are needed to compute Y for the sample matrix $\boldsymbol{A}_{\boldsymbol{B}}^{(i)}$. The total number of evaluations is then $N_{total} = N(d+2)$.

The choice of an appropriate base sample rate N has been widely debated in current GSA literature. Given that N determines the number of model evaluations we require, it determines the computational cost of the analysis (computation of the indices themselves is a relatively quicker post-processing step). As previously mentioned, with a LHS strategy we require $N >> d$ to accurately capture the influence of the parameters on Y. For VBSA especially, current results in literature suggest that the appropriate base sample size varies from one application to another, and often very large base sample rates are needed to achieve reliable results, especially for input factors that exhibit high sensitivity [40]. Generally, it is accepted that the appropriate base rate is dependent on the number of parameters d. For $d = 3$ as in our case, convergence has been reported across multiple VBSA experiments in literature for N_{total} between 10^3 and 10^4 [50]. Once the appropriate base rate is selected, then generally all pairs of first-order and total-effect indices can be computed, and any single interaction terms can be computed at the additional cost of N model evaluation. Saltelli et al. [49] estimate that for a model with 10 input parameters and $N_{total} = N(d+2)$, at half a minute of CPU per run, a good characterisation of the first-order and total-effects indices can be obtained at a cost of about 50h of CPU.

For the present analysis, a base rate of $N = 3000$ was selected, and using LHS, sample matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ of size 3000×3 were generated, corresponding to our three (uniformly distributed) input variables $\hat{\eta}, \hat{\beta}, \psi$. Based on the conclusions of the bathymetry data assimilation in Chapter 3, we picked the parameter intervals to be $\hat{\beta} \in [0.01, 0.35]$, $\hat{\eta} \in [0.001, 0.0035]$, and $\psi \in [0.5, 2]$.

Using the resampling strategy outlined above, the resampled matrix $\boldsymbol{A_B}$ is created of size 9000×3. Thus the number of evaluations N_{total} needed of our model Y is 15000.

For the numerical calculation of the sensitivity indices, we used the SAFE toolbox for MATLAB developed by Francesca Pianosi, Fanny Sarrazin and Thorsten Wagener at the Department of Civil Engineering at the University of Bristol [41].

It should be noted that the actual base sample rate N used to calculate sensitivity indices was slightly less than 3000. This is because post model evaluation, results for Y where the data assimilation failed to converge, were removed. Such instances were expected, as one of the observations from Chapter 3 was that a necessary condition for convergence of the data assimilation is that $\hat{\eta}/\hat{\beta} < 0.1$. This condition is not satisfied by combinations of $\hat{\eta}$ and $\hat{\beta}$ from the extreme ends of the parameter space intervals, such as for $\hat{\beta} = 0.01$ and $\hat{\eta} = 0.0035$. However, the removal did not cause a significant decrease in the base sample rate and the adjusted base sample rate was $N = 2986$.

5.3 Results of VBSA

The objectives for the analysis are to determine first order sensitivity indices for the main and total effects for the three parameters $\hat{\beta}$, $\hat{\eta}$, and ψ. In doing so, we may rank the parameters in order of influence on the bathymetry reconstruction error Y_β, and also the resulting error in the surface wave Y_η. Additionally, we wish identify non-influential parameters based on some sensitivity tolerance levels. Our primary goal is to determine conditions for optimal convergence in the data assimilation scheme that lead to improved knowledge for tsunami modelling, e.g. understanding the relationship between the relative position of the observation points and bathymetry (through input parameter ψ) and the effect it has on the accuracy of tsunami wave model prediction.

We present results for the two different model outputs, $f(\hat{\eta}, \hat{\beta}, \psi) = Y_\beta$ and $f(\hat{\eta}, \hat{\beta}, \psi) = Y_\eta$, where Y_η and Y_β are the error norms given by (5.2.1) and (5.2.2) respectively. Figure 5.2 illustrates the initial results for the first order sensitivity indices S_i and $S_i^{(Tot)}$ for $f(\hat{\eta}, \hat{\beta}, \psi) = Y_\beta$.

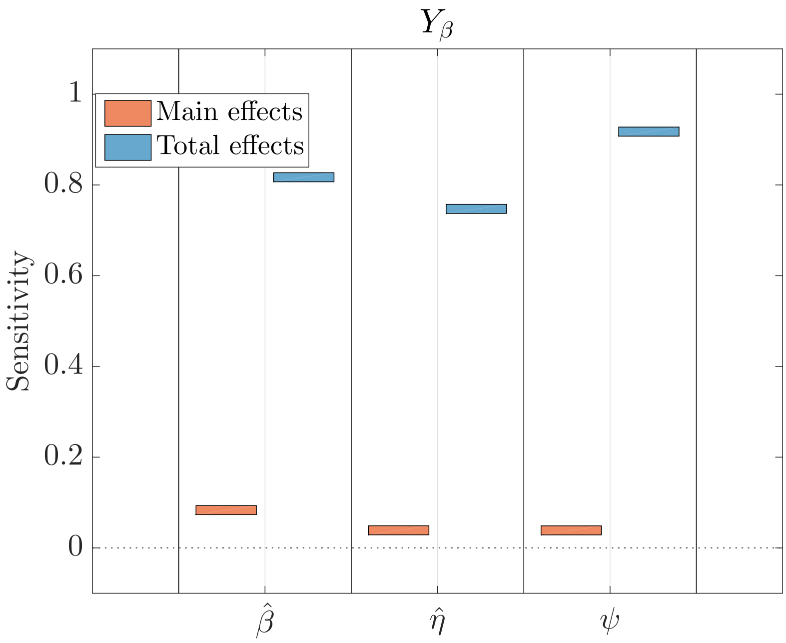


Figure 5.2: Sensitivity indices S_i (Main effects) and $S_i^{(Tot)}$ (total effects) for the three input parameters, and model output Y_β. We note that based on these results, the combined effects of the parameters is much more significant than their effects individually.

The sensitivity indices for main effects S_i are very low for each input parameter, at $\mathcal{O}(10^{-2})$. However the total effect $S_i^{(Tot)}$ in each case is relatively very high. The exact values are presented in table 5.1. For VBSA, input parameters with sensitivity indices below 1% are generally considered non-influential [40]. Thus, all the main and total effects indices in figure 5.2 may be considered significant. We observe that the main effects for $\hat{\beta}$ are more than double the sensitivity of $\hat{\eta}$ and ψ respectively, and that ψ exhibits the highest total effect. These initial results suggest that while none of these input parameters present significant contributions to the model variability on their own, their interactions with other sources of variability in the bathymetry reconstruction error is high.

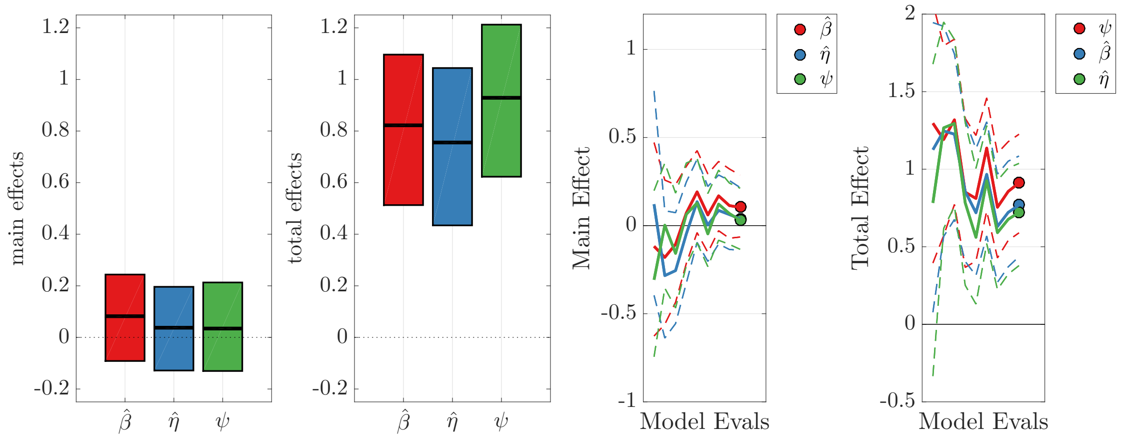
However, before we accept these conclusions we must assess the credibility of these results. Pianosi et al. [40] provide a useful checklist for assessing the convergence and robustness of our results. The derived sensitivity indices are an approximation, and thus susceptible to approximation error due to factors such as sample size. Subsequently we use a bootstrapping method to resample the data, allowing us to (i) derive confidence intervals for the index estimates, and (ii) assess the convergence of indices as the size of the sample is varied.

Input	S_i	$S_i^{(Tot)}$	$\bar{S}_i$	$S_i^{(\bar{Tot})}$
$\hat{\beta}$	0.0833	0.8169	0.0819	0.8166
$\hat{\eta}$	0.0388	0.7472	0.0386	0.7450
ψ	0.0387	0.9175	0.0412	0.9136

TABLE 5.1: Main (S_i) and Total ($S_i^{(Tot)}$) sensitivity indices for each input parameter. $\bar{S}_i$ and $S_i^{(\bar{T}ot)}$ represent main and total sensitivity indices averaged over 700 bootstrap resamples. Asinput parameters with sensitivity indices above 0.1 are generally considered influential, each parameter is significant in terms of both total and main effects. However, we note that the total effect indices are much larger, indicating that interaction effects are high.

The bootstrapping technique involves generating multiple resamples by drawing randomly (with replacement) from the original sample of the model. Because we are using the existing sample for which input/output has already been generated, there are no further model evaluations required. We can analyse behaviour of the index estimates as the sample size is varied (convergence), and assess whether estimates are independent of a specific input-output sample (robustness).

To assess convergence and robustness we create confidence intervals for the sensitivity index distribution. For the present analysis we compute 95% confidence intervals (5% significance level). Results for the main and total effects generated over 700 bootstrap resamples with confidence intervals are given in figure 5.3(a), and the convergence over the different resamples ranging from $N = 300$ to $N = 2986$ is shown in figure 5.3(b). Index results for S_i and $S_i^{(Tot)}$ averaged over the 700 bootstrap resamples are also given in table 5.1.



(a) Bootstrapping (b) Convergence

FIGURE 5.3: Extended sample with bootstrapping and convergence for Y_β. Main and total effects generated over 700 bootstrap resamples are shown in figure (a), and the convergence over different resamples ranging from $N = 300$ to $N = 2986$ is shown in (b).

As a summary statistic to test convergence, we compute the maximum width of the confidence intervals for all the model parameters

$$\mathrm{Stat}_{cf} = \max_{i=1,..,d} |S_i^{ub} - S_i^{lb}| \tag{5.3.1}$$

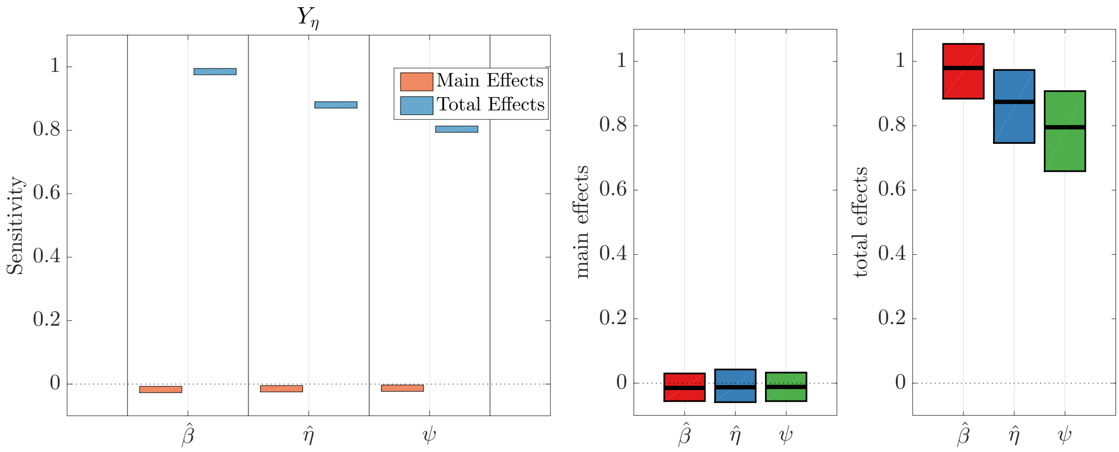
Where S_i^{lb} and S_i^{ub} represent the lower-bound and upper-bound of the i-th input parameter respectively. Confidence intervals with a width approaching zero indicate convergence of the index to a fixed value Sarrazin et al. [50]. We present an overview of the distance $|S_i^{ub} - S_i^{lb}|$ for each parameter in table 5.2.

Input	$\|S_i^{ub} - S_i^{lb}\|$	$\|(S_i^{(Tot)})^{ub} - (S_i^{(Tot)})^{lb}\|$
$\hat{\beta}$	0.3264	0.6001
$\hat{\eta}$	0.3454	0.6051
ψ	0.3341	0.5817

TABLE 5.2: Summary statistics Stat_{cf} for each parameter, corresponding to the confidence intervals in figure 5.3(a). These values indicate a margin of error far greater than desired.

As we see from table 5.2, Stat$_{cf}$ statistics for S_i and $S_i^{(Tot)}$ are 0.3454 and 0.6051 respectively. If assume the confidence intervals are around the mean of the predicted sensitivity index, we observe that the maximal margin of error for S_i is 0.1727 and for $S_i^{(Tot)}$ it is 0.3025. As the sensitivity indices are normalised between 0 and 1, this indicates a margin of error of 17% and 30% respectively, which are far greater than desired. The convergence in figure 5.3(b) further reflects this as we see the confidence interval widths (represented by the area between the dotted lines for each input) are very high, and it is not clear that the indices are approaching a fixed value as the resample size increases.

We conduct similar analyses for the model $f(\hat{\eta}, \hat{\beta}, \psi) = Y_\eta$. We approximate the sensitivity of the surface wave error to the input parameters using the same sample discussed above. Results for the initial approximations for the indices, and results with bootstrapping and confidence intervals are given in figure 5.4(a) and 5.4(b).



(a) Initial Results (b) With Bootstrapping

FIGURE 5.4: (Sensitivity indices for Y_η, with an extended sample. Main (S_i) and Total ($S_i^{(Tot)}$) effects over 700 bootstrap resamples are shown in figure (a), and the convergence over different resamples ranging from $N = 300$ to $N = 2986$ is shown in (b).

A key detail in figure 5.4(a) is that the sensitivity indices for each parameter are negative. In principle, the exact sensitivity indices should always take values in $[0,1]$. Negative values are a result of approximation error, and can occur if the value of the index is lower than the estimation error. Consider figure 5.4(b) where we have used bootstrapping to derive confidence intervals. Our summary statistic $Stat_{cf}$ for Y_η can be calculated using the values of $|S_i^{ub} - S_i^{lb}|$ and $|(S_i^{(Tot)})^{ub} - (S_i^{(Tot)})^{lb}|$ given in table 5.3.

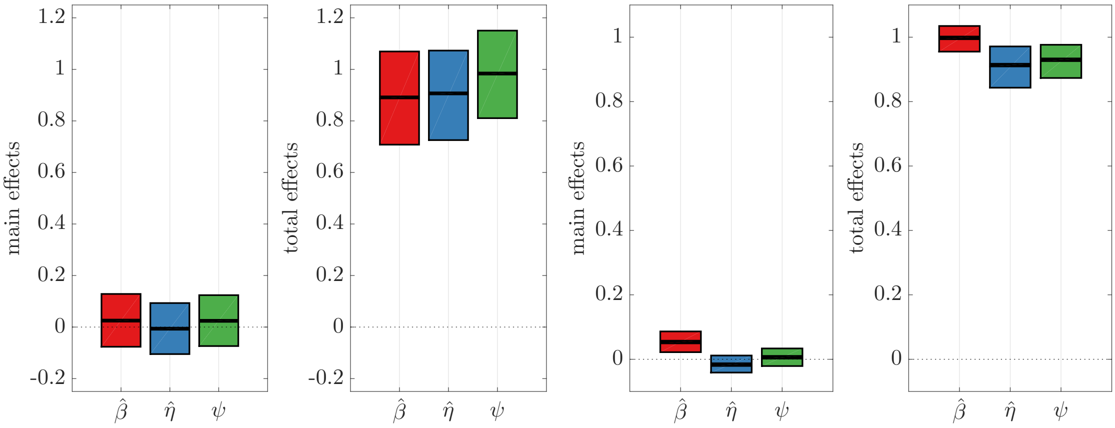
Input	$\|S_i^{ub} - S_i^{lb}\|$	$\|(S_i^{(Tot)})^{ub} - (S_i^{(Tot)})^{lb}\|$
$\hat{\beta}$	0.0868	0.1701
$\hat{\eta}$	0.0985	0.2201
ψ	0.0919	0.2548

Table 5.3: Summary statistics for confidence intervals as in figure 5.4(b) for model output Y_η. If the true values for S_i and $S_i^{(Tot)}$ are less than the error margins indicated in the table, the approximated indices have negative values.

$Stat_{cf}$ statistics for S_i and $S_i^{(Tot)}$ are 0.0985 and 0.2548 respectively. Thus the maximal margins of error for each are 0.0493 and 0.1274. Thus can see that if the true values for S_i and $S_i^{(Tot)}$ were less than these margins, our approximated indices could have negative values.

In light of this analysis, we clearly must address the lack of robustness and convergence present in our results, and find the underlying cause. In many cases these issues are a reflection of inadequate sample size [40, 50]. High confidence interval summary statistics, lack of convergence to fixed values over multiple resamples, and negative approximations for the indices can in theory be mitigated by increasing the sample size.

To try to remedy these shortcomings, we generated 6000 additional samples using LHS and the recombination method outlined in Saltelli et al. [48]. When adjusted for removal of sample points where the data assimilation didn't converge, the extended base rate is $N = 8956$, and the extended resampled matrix $\boldsymbol{A_B}$ is size $26,868 \times 3$. Sensitivity indices and confidence intervals are approximated using bootstrapping over 700 resamples for both model outputs Y_β and Y_η. The results are summarised in figure 5.5, and the equivalent convergence analysis as the resample size is increased from 900 to 8956 is presented in figure 5.6.



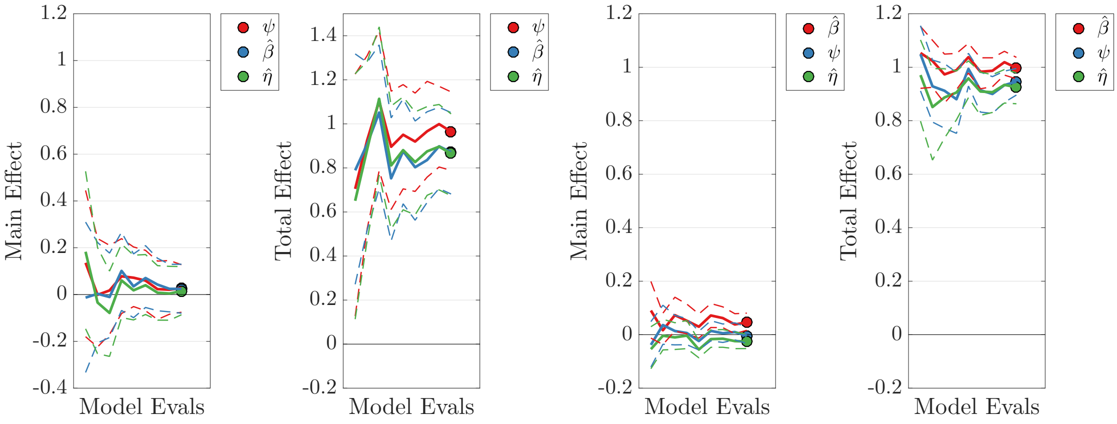
(a) Y_β

(b) Y_η

Figure 5.5: Sensitivity indices for main and total effects for (a) Y_β, and (b) Y_η. Extended sample size is $N = 8956$, and confidence intervals are shown for 700 bootstrap resamples.

Looking at figure 5.5(a) and comparing with the results for the model output Y_β in figure 5.3(a), we see that there is some decrease in confidence interval width, and $\text{Stat}_{cf} = 0.2057$ for S_i, and $\text{Stat}_{cf} = 0.3528$ for $S_i^{(Tot)}$. However, this is not a significant decrease in the margin of error, and we note that the index for $\hat{\eta}$ is negative, indicating estimation error larger than the value of the index. Results for model output Y_η in figure 5.5(b) indicate a similar trend; while the summary statistic has been reduced, it is still far above an acceptable tolerance, and all three input parameters have negative sensitivity indices.

To verify whether this is because the sample size is still inadequate, we consider the convergence analysis for both outputs in figure 5.6(a) and 5.6(b). Observe that convergence has actually improved relative to figure 5.3 for Y_β, and it is clear in figure 5.6(b) that for Y_η, resamples of increasing size do not have a big effect on the indices as the sample size increases. Despite this, the confidence interval widths are still too high to consider the index values to be reliable.



(a) Y_β (b) Y_η

Figure 5.6: Convergence of the main and total effect sensitivity indices for (a) Y_β, and (b) Y_η. The dotted lines represent confidence for resamples ranging from $N = 300$ to $N = 8956$.

Given these results, we can assume that the large confidence intervals and approximation errors are due to factors other than the size of the sample. Pianosi et al. [40] outline the following principles that VBSA relies on:

1. The variance of the output distribution is a good proxy of output uncertainty.
2. The contribution to the output variance from a given input parameter is a measure of uncertainty.

It is therefore possible that our data assimilation model does not meet these requirements, and so we investigate potential scenarios where variance is *not* a good measure of sensitivity, and check whether our model exhibits any of these characteristics.

Sobol [55] observe that the VBSA sensitivity indices are related to the terms of the variance decomposition of the model output, and as such reflect the structure of the model itself. Thus, they are subject to two necessary conditions for variance analysis; (i) The input factors are independent and uncorrelated, and (ii) The output distribution is not highly skewed or multi-modal, otherwise the assumption that variance captures model uncertainty is broken, and subsequently variance is no longer a meaningful indicator of sensitivity [40, 42, 28].

Subsequently, we check our model input factors for correlation, and analyse the output probability density distributions for Y_β and Y_η for skewness. To check correlation between $\hat{\beta}$, $\hat{\eta}$ and ψ, we compute the Spearman correlation coefficient on our input sample, defined in terms of the covariance between two random variables A and B,

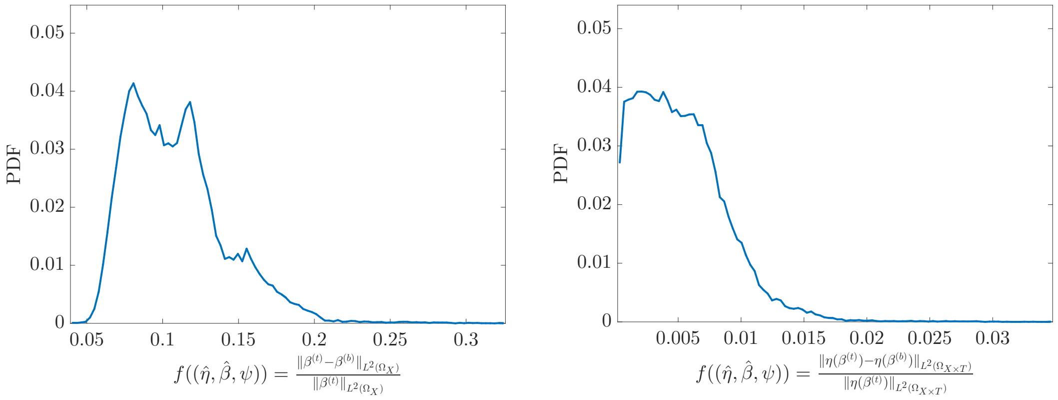
$$\rho(A,B) = \frac{\mathrm{cov}(A,B)}{\sigma_A \sigma_B}, \tag{5.3.2}$$

where $\mathrm{cov}(A,B)$ is the covariance and σ is the standard deviation. $\rho(A,B)$ takes values between 0 and ± 1, where $\rho = \pm 1$ indicates complete positive or negative correlation and $\rho = 0$ indicates no correlation. We define the 3×3 correlation coefficient matrix $\mathbf{R}$ containing all pairwise correlation coefficients for the input parameters.

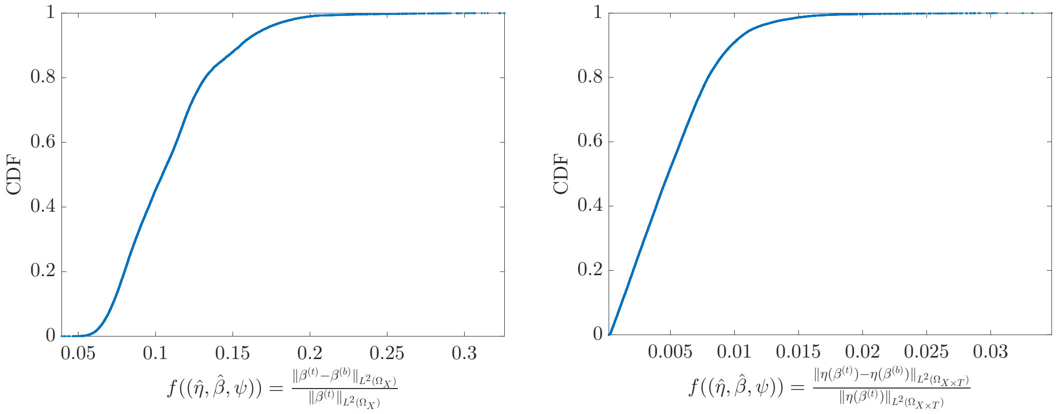
$$\boldsymbol{R} = \begin{bmatrix} 1.0000 & -0.0066 & -0.0060 \\ -0.0066 & 1.0000 & 0.0040 \\ -0.0060 & 0.0040 & 1.0000 \end{bmatrix}. \tag{5.3.3}$$

The order of the input parameters in each row/column is $\hat{\beta}$, $\hat{\eta}$ and ψ, e.g. the $r_{1,2}$ entry corresponds to $\rho(\hat{\beta}, \hat{\eta})$. As each variable is perfectly correlated with itself the diagonal entries are always 1. We can see the correlation coefficients for each pair are very low, below 1%. This is not surprising, as in reality the amplitude of the initial conditions and the amplitude or position of bathymetry observed outside the support of the initial conditions are independent. However, within the data assimilation algorithm we observed that optimal results occurred when there was a dependency between $\hat{\beta}$ and $\hat{\eta}$, and since we discarded some sample points where the necessary condition relating these was broken, it was possible the values could be correlated within the sample. However $\mathbf{R}$ clearly indicates that this is not the case, and so we may conclude our input parameters are independent and uncorrelated.

To check for skewness, we present the probability density functions (PDFs) and cumulative density functions (CDFs) for Y_β and Y_η in figure 5.7.



(a) PDF Y_β (b) PDF Y_η



(c) CDF Y_β (d) CDF Y_η

FIGURE 5.7: Probability distributions of Y.

We see immediately that both PDFs in figure 5.7(a) and 5.7(b) are skewed to the left. The PDF for Y_η especially is highly skewed, and as such does not fulfill the requirements of VBSA. It is evident that the skewed distribution is the culprit for the high error margins in our sensitivity index approximations. Confidence intervals use the variability of the model output to assess the accuracy of the approximation. If the distribution is skewed, this explains the high values for Stat_{cf}.

Pianosi and Wagener [42] observe that in this situation, using variance as a proxy for uncertainty would cause the uncertainty about the model output to increase when uncertainty about one of the inputs is removed, which is a contradiction.

In attempting to pinpoint the cause of the skewed distribution, we note that this may be a result of the parameter space used for sampling. Only those values for $\hat{\eta}$ and $\hat{\beta}$ were considered such that the data assimilation in Chapter 3 converged. As we observed, there was low sensitivity of the surface wave error (Y_η) for these values, and so we see the skew to the left corresponds to a higher frequency of lower values for output Y_η in the probability distribution. Wagener and Pianosi [61] observe that the choice of the input parameter space can have a significant affect on GSA results, and thus using multiple methods to cross-validate results is recommended. Indeed, Pianosi and Wagener [42] note that skewness in output distribution is found in non-linear and environmental models quite often.

Another point to consider is whether the skewed distribution could this be a result of the sampling strategy. While the sampling strategy is an important consideration, the sampling method outlined in Saltelli et al. [48] is considered to be relatively robust [41]. Generally, the sample size is considered more significant that the strategy [40]. However, we observed that even with $\mathcal{O}(10^{-4})$ (within the recommended sample size tolerance for VBSA [49]), the skewed distribution was not altered.

In conclusion, our results do not satisfy the conditions of credibility due to skewness. This is because they violate the principles that VBSA relies on [40]. Because of this, we cannot credibly use these results for screening or ranking of the input parameters. We require a different measure of sensitivity that is not variance based. In the next section, we introduce Density-based methods as an alternative to VBSA.

5.4 Density Based Sensitivity Analysis (DBSA)

Due to the shortcomings of the variance based methods discussed in the previous section, sensitivity analyses have been developed that do not use moments of the model output such as the expected value, variance or skewness, as a measure of uncertainty. Such methods are collectively known as “Density Based Sensitivity Analysis”(DBSA). The idea is to measure sensitivity through variations of the probability density function that occur when the influence of a certain input factor is removed. This is done by computing the difference between unconditional PDFs generated by varying all parameters, and conditional PDFs obtained when fixing individual input parameters at a particular nominal value [40]. The sensitivity index is then computed as a statistic based on this divergence.

For the present analysis, we follow the method outlined in Pianosi and Wagener [42], who improve on this idea by characterising output distributions by their cumulative density functions (CDFs) instead. Known as the PAWN method, density-based sensitivity indices are calculated using conditional and unconditional CDFs, using the fact that these are more efficiently derived than PDFs. We give a brief overview of the PAWN method, the divergence statistic used to compute the indices, and the sampling strategy used in the numerical implementation.

Let us define the unconditional cumulative density distribution of our output Y (where Y can be either Y_β or Y_η) as $F_Y(Y)$, and the conditional cumulative density distribution when the input parameter x_i is fixed as $F_{Y|x_i}(Y)$. As $F_{Y|x_i}(Y)$ represents the case where there is no variability resulting from x_i, the distance between the two functions $F_{Y|x_i}(Y)$ and $F_Y(Y)$ represents the variability in the output induced by x_i. This distance is proportional to the influence of x_i, i.e. if $F_{Y|x_i}(Y) = F_Y(Y)$ then x_i has zero influence on the variability of output Y. In the PAWN method, this distance is computed using the Kolmogorov-Smirnov statistic

$$\mathrm{KS}(x_i) = \max_{Y} \left| F_{Y|x_i}(Y) - F_Y(Y) \right|. \tag{5.4.1}$$

As KS depends on the value of x_i, the PAWN sensitivity index T_i is calculated using a statistic (such as the median or maximum) over all values of x_i,

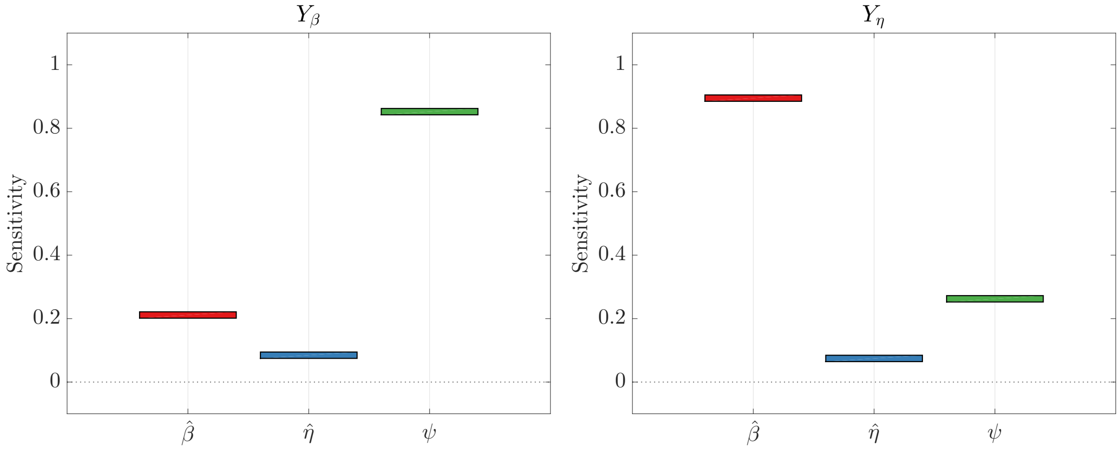
$$T_i = \underset{x_i}{\text{Stat}}\Big[\text{KS}(x_i)\Big]. \tag{5.4.2}$$

Pianosi and Wagener [42] observe that the sensitivity index T_i satisfies all of the properties of a "good" sensitivity index, as it is global, quantitative and model-independent. It has the advantage over VBSA indices, that it is also moment-independent. By definition it is normalised to take values between 0 and 1. We observe that it is comparable to the total effects sensitivity index for VBSA, $S_i^{(Tot)}$, as the divergence between the CDF distributions when x_i is fixed and when it is allowed to vary may potentially include interaction affects.

In the numerical implementation, empirical CDFs are used to compute $F_{Y|x_i}(Y)$ and $F_Y(Y)$, using a sample of the parameter space. PAWN has a big advantage over the sample size requirements in VBSA, as a smaller sample size can be used to affectively approximate the CDFs due to their regularity properties. $F_Y(Y)$ is approximated using N_u model evaluations obtained by sampling over the entire parameter space, while $F_{Y|x_i}(Y)$ is approximated using N_c model evaluations derived by sampling over the non-fixed parameters only, while keeping x_i fixed. Additionally, the conditioning value x_i used to compute T_i in (5.4.2) is replaced by $x_i = \bar{x_i}^{(1)}, \bar{x_i}^{(2)}, .., \bar{x_i}^{(n)}$, n randomly-sampled values for the fixed input x_i. And thus the total number of model evaluations is $N_u + n \times d \times N_c$. Generally these values are chosen by trial and error. Pianosi and Wagener [42] suggest a reasonable choice for n is between 10 and 50. For our analysis, we set $N_u = 200$, $N_c = 150$, and $n = 15$. Thus a total of 6950 model evaluations were obtained, a significant reduction relative to the 45000 total evaluations conducted for VBSA, and subsequently a fraction of the computational cost. As with the variance-based indices, the computation of the density-based indices is a quick post-processing step. To verify whether the sample size is adequate, bootstrapping can be used to analyse robustness of sensitivity index estimates, as in the previous section. We present the results for the DBSA sensitivity indices approximated using the PAWN method in the following section.

5.5 Results of DBSA

Like the VBSA analysis, the number of evaluations used in the analysis was adjusted to exclude sample points where the data assimilation did not converge. N_c was adjusted to 145, and 6725 evaluations were used to approximate T_i. For the initial analysis we chose the *maximum* as the statistic used in 5.4.2.

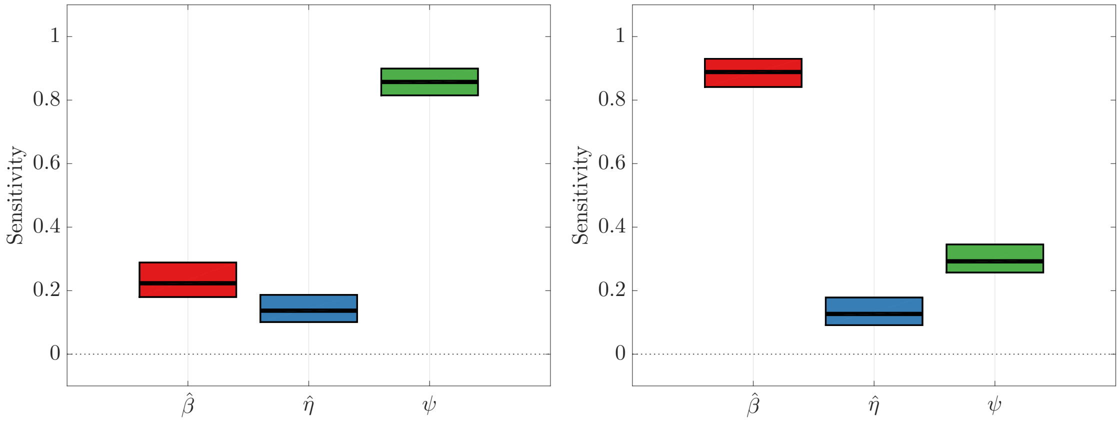


(a) Y_β (b) Y_η

FIGURE 5.8: Approximation of the density-based sensitivity indices for $\hat{\beta}(red)$, $\hat{\eta}$ (blue) and ψ (green). (a) shows the influence of each parameter on Y_β and (b) shows the influence of each parameter on Y_η.

Figure 5.8 presents the initial results for the sensitivity indices T_i for each input parameter. Figure 5.8(a) gives the result for model output Y_β and figure 5.8(b) for Y_η. In both cases, we see a big contrast to the results we observed with VBSA, highlighting how much the skewness of the PDFs can distort the approximated sensitivity. Y_β has the greatest sensitivity to position of the bathymetry ψ (86%). Interestingly, we can see in figure 5.8(b) that the amplitude of the initial conditions $\hat{\eta}$ and amplitude of the bathymetry $\hat{\beta}$ have relatively low influence on the error in the surface wave Y_η compared to ψ, at 22% and 13% respectively. On the other hand, we observe that the amplitude of the bathymetry $\hat{\beta}$ is highly influential for the bathymetry error Y_β, with $T_{\hat{\beta}} = 0.895$.

As before, we check these results for robustness and convergence before we accept them as conclusive. Bootstrapping is done over 700 resamples, and confidence intervals are derived. These can be seen in figure 5.9. We present the distances between upper and lower bounds for each parameter in table a, and subsequently compute Stat_{cf} for each.



(a) Y_β (b) Y_η

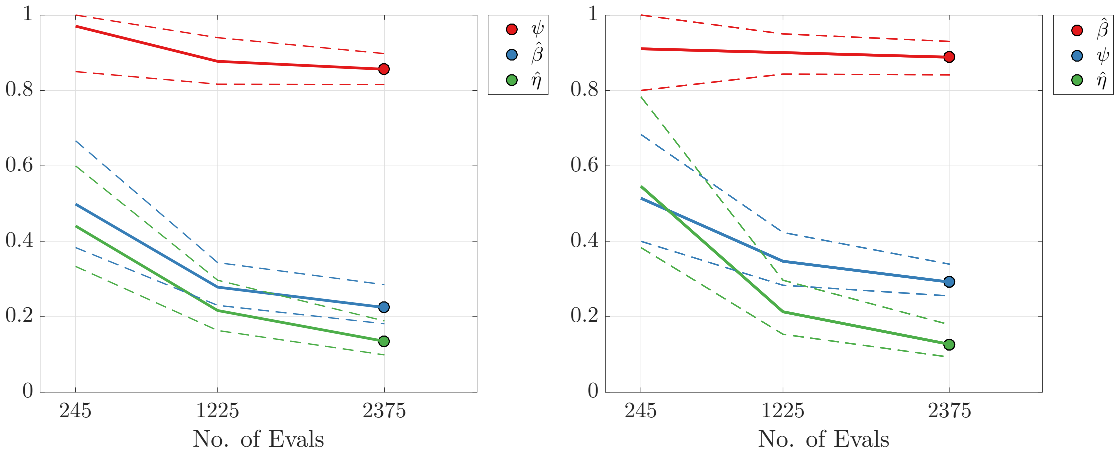
FIGURE 5.9: Approximation of the density-based sensitivity indices for $\hat{\beta}(red)$, $\hat{\eta}$ (blue) and ψ (green), with confidence intervals derived using 700 resamples. (a) shows the influence of each parameter on Y_β and (b) shows the influence of each parameter on Y_η.

Input	$\lvert T_i^{ub} - T_i^{lb}\rvert$ (Y_β)	T_i^m (Y_β)	$\lvert T_i^{ub} - T_i^{lb}\rvert$ (Y_η)	T_i^m (Y_η)
$\hat{\beta}$	0.1081	0.2241	0.0850	0.8878
$\hat{\eta}$	0.0831	0.1335	0.0824	0.1265
ψ	0.0800	0.8583	0.0848	0.2899

TABLE 5.4: Width of the confidence interval $\lvert T_i^{ub} - T_i^{lb}\rvert$, and mean index T_i^m averaged over 700 bootstrap resamples. Results given for model output Y_β and Y_η respectively.

We observe a great improvement in the Stat_{cf} statistics relative to VBSA results, where $\mathrm{Stat}_{cf}(Y_\beta) = 0.1081$ and $\mathrm{Stat}_{cf}(Y_\eta) = 0.0850$. The largest margin of error from the mean T_i^m across all input factors and both outputs is 6.3%. If we set a tolerance level of 0.05 for the error margin, then there is only a single instance ($|T^{ub}_{\hat{\beta}} - T^{m}_{\hat{\beta}}| = 0.0630$) where the error margin is higher. We hypothesize that this could be further improved by taking a slightly larger sample size.

The convergence of the sensitivity indices is given in figure 5.10. We can see that across both errors Y_β and Y_η, the most influential parameters (position ψ and amplitude $\hat{\beta}$ of the bathymetry) appear to converge. For the other less influential parameters in each case, convergence is still relatively smooth. However, we see that the index values are still decreasing slightly even at the largest sample, suggesting that a larger number of samples might be needed for convergence to a fixed point. Nevertheless, given the results of the robustness analysis and convergence observed thus far, we claim that we may still accurately gauge the influence of the parameters on our model outputs with the current sample size, especially in order to determine the most influential factors. Consequently, we proceed to consider ranking, screening, and factor mapping of the input parameters.



(a) Y_β

(b) Y_η

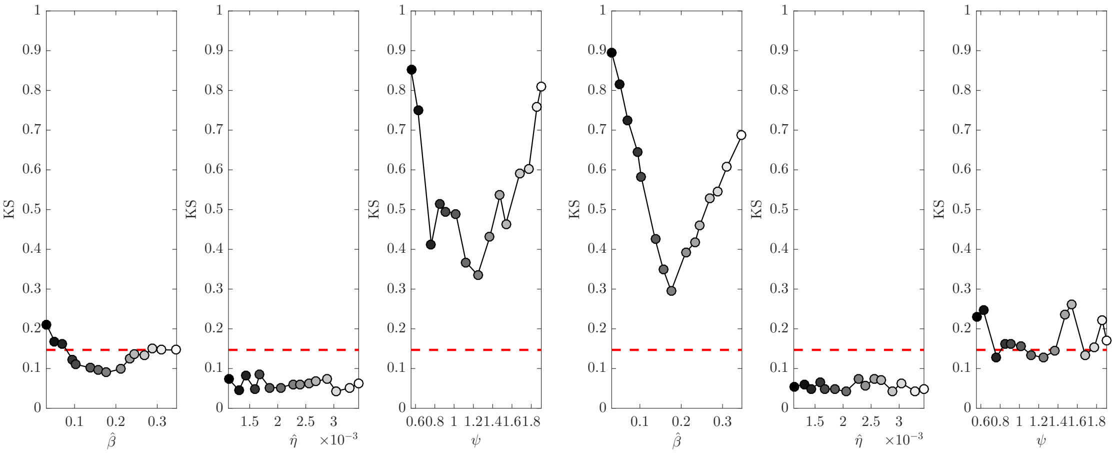
FIGURE 5.10: Convergence analysis for DBSA indices for (a) Y_β, and (b) Y_η, for resamples of size $N = 215$ to $N = 2375$.

The PAWN method can be effectively used to screen for non-influential input factors. Figure 5.11 presents an overview of the different values of KS over the 15 randomly chosen sample points x_i for each input parameter. To determine which input parameters are most influential, the two-sample Kolmogorov-Smirnov test [42] has been implemented, which allows us to reject the hypothesis that $F_Y(Y)$

and $F_{Y|x_i}(Y)$ are the same if

$$KS > c(\alpha)\sqrt{\frac{N_c + N_u}{N_c N_u}}, \tag{5.5.1}$$

where α is the confidence level, and $c(\alpha)$ is a critical value determined in the literature [42]. Figure 5.11(a) presents results for Y_β and figure 5.11(b) for Y_η. The red dotted line represents the threshold value at the confidence level $\alpha = 0.05$. KS statistics for input x_i that fall below this critical threshold indicate that x_i is non-influential.



(a) Y_β

(b) Y_η

Figure 5.11: K-statistics with significance level 0.05. Values below the dotted red line are non-influential.

We observe that for both outputs Y_β and Y_η, $\hat{\eta}$ is non-influential across all 15 choices for the fixed value. KS statistics for $\hat{\beta}$ fall below the threshold at fixed values approximately between 0.1 and 0.25 for the model Y_β. This indication corroborates results from Chapter 3, where we observed that the bathymetry reconstruction error Y_β was higher when the amplitude of the bathymetry $\hat{\beta}$ was either too big or too small relative to the aveerage depth H. In the present analysis the maximum KS statistic is used to compute T_i for each input parameter, and so $\hat{\beta}$ is still considered influential on Y_β. It is interesting to contrast this result with the KS statistics for $\hat{\beta}$ and output Y_η, as shown in figure 5.11(b). While we see a similar trend, where the KS statistic is higher if the bathymetry amplitude $\hat{\beta}$ is at the lower or upper end of the sample interval, the resulting variation is much higher, and all KS statistics are above the influence threshold value.

In a similar manner, while the position of the bathymetry ψ has relatively low KS statistics across all fixed values of the surface wave error Y_η (approximately half of the values are above the cutoff with no clear trend), it has relatively much higher KS values, and subsequently higher influence on the bathymetry reconstruction error Y_β, as observed in the third panel in figure 5.11(a). Once again, KS statistics are much higher at values closer to the endpoints of the parameter space for ψ, [0.5 2]. We suggest this indicates that the placement of the bathymetry *relative to the observation points* has a significant influence on the error in the bathymetry reconstruction. This confirms conclusions of the second order adjoint analysis in Chapter 4 that the coverage of the bathymetry by observations was influential on the surface wave error. This is because as the bathymetry Gaussian placement varies over the interval, so does the time and position where the bathymetry is first observed by the measurement points.

Thus, based on the threshold level for a confidence interval $\alpha = 0.05$, and using the maximum KS statistic to compute T_i, we observe that for both bathymetry reconstruction error Y_β and surface wave error Y_η, the amplitude of the initial conditions $\hat{\eta}$ can be considered non-influential. We note that the choice of a different statistic in (5.4.2) such as the mean or median may change the value of T_i. We provide a comparison of the different indices in tables 5.5. Given the threshold value of 0.147 ($c(\alpha = 0.05)$), we can see that by choosing a mean or median statistic, $\hat{\beta}$ is categorised as a non-influential parameter for the bathymetry reconstruction error Y_β.

Input	Max	Mean	Median
$\hat{\beta}$	0.2114	0.1334	0.1336
$\hat{\eta}$	0.0847	0.0613	0.0610
ψ	0.8524	0.5608	0.5155

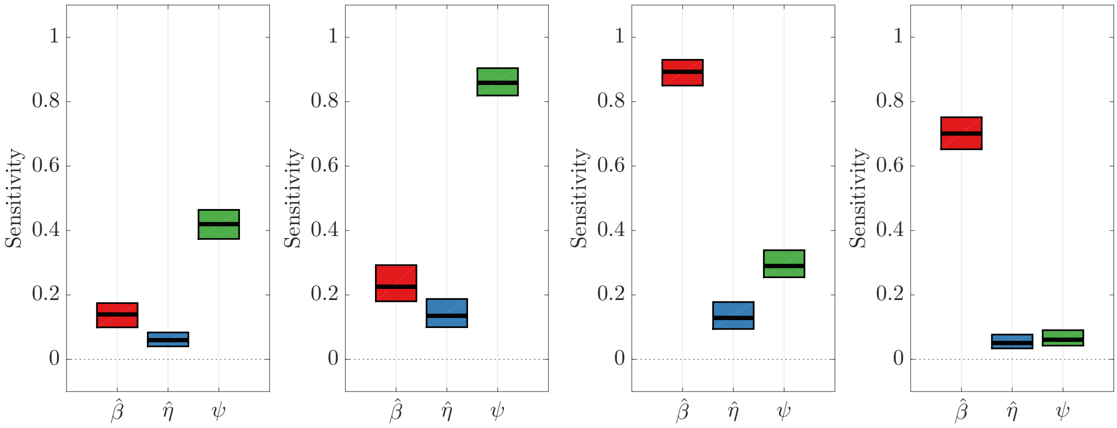
Table 5.5: Model Output Y_β: Indices T_i using different statistics in the definition of the sensitivity index (5.4.2). Values highlighted in red are below the threshold value of 0.147 ($c(\alpha = 0.05) = 1.36$) and are non-influential.

Input	Max	Mean	Median
$\hat{\beta}$	0.8950	0.5582	0.5467
$\hat{\eta}$	0.0745	0.0560	0.0538
ψ	0.2624	0.1780	0.1626

Table 5.6: Model Output Y_η: Indices T_i using different statistics in the definition of the sensitivity index (5.4.2). Values highlighted in red are below the threshold value of 0.147 ($c(\alpha = 0.05) = 1.36$) are non-influential.

As we can see from table 5.5 , the overall ranking of the parameters does not change based on the statistic used for Y_β. The only exception is the parameter $\hat{\beta}$, which is influential with a max statistic. The position of the bathymetry ψ is the most influential parameter for Y_β, and the amplitude of the initial conditions $\hat{\eta}$ is the least. Similarly for model output Y_η, results in table 5.6 indicate that across all choice of statistics, $\hat{\beta}$ is the most influential parameter and $\hat{\eta}$ is non-influential.

So far we have used the results of the PAWN sensitivity analysis to rank and screen our input factors. A third analysis that can prove insightful is factor mapping; We can compute sensitivity indices based on a sub-interval of our output Y given the current sample. For example, over all model evaluations, the maximum surface wave error Y_η is 2.63×10^{-2}, and the minimum is 2.23×10^{-4}. Similarly the maximum bathymetry reconstruction error Y_β is 0.265 and the minimum is 3.75×10^{-2}. To determine the influence of input parameters when the values of Y_η and Y_η are relatively high, we define a parameter M such that we can compute sensitivity indexes for all input-output samples for the cases $Y > M$, and $Y \leq M$ separately. We choose $M = 0.015$ for output Y_β and $M = 0.001$ for Y_η. The results are given in figures 5.12(a) and 5.12(b).



(a) Output Y_β, M = 0.15

(b) Output Y_η, $M = 0.001$

Figure 5.12: Indices T_i for influence of $\hat{\beta}$ (red), $\hat{\eta}$ (blue) and ψ (green) on model outputs $Y > M$ (left panel), and $Y \leq M$ (right panel) for (a) Y_β, and (b) Y_η. Confidence intervals were calculated using 700 bootstrap samples.

While there is not much difference in terms of ranking between the left ($Y > M$) and right ($Y \leq M$) panes for either Y_β or Y_η, it is interesting to note that for the surface wave error in figure 5.12(b) the influence of the bathymetry amplitude $\hat{\beta}$ decreases when $Y_\eta \leq 0.001$ (right panel). This is insightful, as it suggests that the bathymetry is more influential when the errors in the surface wave are large. However, in both intervals of Y_η the influence of $\hat{\beta}$ is high (above 70%). Additionally, the effect of the bathymetry position ψ when Y_η greater than 0.001 is 29%, and subsequently is greater than the influence threshold value. On the other hand, the influence of the initial conditions amplitude $\hat{\eta}$ and bathymetry position ψ drops below the threshold value when $Y_\eta \leq 0.001$.

For the bathymetry reconstruction error Y_β, we see in figure 5.12(a) that the influence of the bathymetry position ψ is approximately doubled when the bathymetry reconstruction error is less than 0.015. The lower sensitivity of each of the input parameters when the bathymetry reconstruction error is higher could be suggestive of other significant influences on the model error, that cannot be attributed to these parameters alone.

5.6 Conclusion and further considerations

In summary, we conclude that variance-based sensitivity analysis is not appropriate for our data assimilation model as the skewness of our model output violates the necessary conditions that VBSA relies on. This is the cause of the large confidence intervals and the negative sensitivity indices observed in section 5.3. As an alternative, we used a density-based method to compute indices that rely on the variation between unconditional and conditional cumulative density functions when particular inputs are fixed, as a measure of sensitivity. Implementing the PAWN algorithm outlined in Pianosi and Wagener [42], we derived sensitivity indices for three input parameters $\hat{\beta}$, $\hat{\eta}$ and ψ, and the model outputs Y_β and Y_η. Our objective was to rank the the influence of the inputs from highest to lowest, screen for non-influential parameters, and finding thresholds in the input parameter values that map into specific output regions (factor mapping). We have summarised the conclusions in table 5.7.

Our results showed that for the model Y_η (the error in the surface wave given the reconstructed bathymetry), the most influential parameter was $\hat{\beta}$, the amplitude of the bathymetry, whereas the amplitude of the surface wave initial conditions $\hat{\eta}$ was categorised as non-influential. This confirms conclusions from Chapter 3 (section 3.6.3), where we observed that the surface wave error increased as the amplitude of the relative bathymetry $\hat{\beta}/H$ became larger. We consider specific output regions for both the bathymetry reconstruction error and the surface wave error. For the latter we consider relative L^2 errors $Y_\eta > 0.001$ and $Y_\eta \leq 0.001$, and found that the influence of all three parameters (the bathymetry and initial conditions amplitudes $\hat{\beta}$ and $\hat{\eta}$, and the bathymetry position ψ), decreases when $Y_\eta \leq 0.001$.

The bathymetry position ψ was also the most influential parameter on the bathymetry reconstruction error Y_β, while the initial conditions amplitude $\hat{\eta}$ was determined to be non-influential. The influence of ψ for small errors $Y_\beta \leq 0.015$ was approximately 85%, however this dropped by almost half for large errors $Y_\beta > 0.015$. Overall the bathymetry reconstruction error was less sensitive to each of the input parameters when the bathymetry reconstruction error was large. This suggests that other parameters that are not considered in this analysis may have a significant influence for larger values of the bathymetry reconstruction error.

Input	$Y_\beta > 0.15$	$Y_\beta \leq 0.15$	$Y_\eta > 0.001$	$Y_\eta \leq 0.001$
$\hat{\beta}$	Non-influential	Influential	Influential	Influential
$\hat{\eta}$	Non-influential	Non-influential	Non-Influential	Non-influential
ψ	Influential	Influential	Influential	Non-influential

TABLE 5.7: Classification of each input parameter as influential or non-influential for the sub-regions of Y_β and Y_η in figure 5.12. Values below the threshold value of 0.147 ($c(\alpha = 0.05) = 1.36$) are non-influential.The entries highlighted in blue are the most influential parameter for Y_β and Y_η respectively.

While the locations of the observation points was kept fixed in the analysis, the choice of the bathymetry position ψ alters the values of the observations, as the point where the bathymetry is observed by the measurement points varies, as does the proportion of observation points that observe the surface wave before it interacts with the bathymetry. We observed that ψ was influential on the surface wave error for all statistics, except when the error in the surface wave Y_η was less than 0.001. As our objective is to determine which parameters are responsible for large errors in the surface wave, we can conclude that the placement of the observation points relative to the reconstructed bathymetry is influential on the surface wave accuracy. This observation confirms conclusions from Chapter 4, where we determined that the position of the observation points (and subsequently the coverage of the bathymetry support) had a significant effect on the sensitivity of the surface wave error to observations.

A key observation is that for both bathymetry reconstruction error and the surface wave error, the amplitude of the initial conditions $\hat{\eta}$ was non-influential for each sub-region. This is in contrast to conclusions from Chapter 3, where we determined that a necessary condition for convergence of the bathymetry assimilation was that the amplitude ratio $\hat{\eta}/\hat{\beta}$ be less than 0.1. However, the lack of sensitivity to $\hat{\eta}$ can be explained by the fact that all values where the amplitude ratio condition was violated (and subsequently the assimilation was non-convergent) were removed from the sample used to derive sensitivity indices. Therefore, we cannot accept the conclusion that $\hat{\eta}$ is non-influential in the bathymetry reconstruction and the resulting free surface wave, on the basis of these results alone.

Further questions that should be addressed in future work are the inclusion of more parameters in the sensitivity analysis, like the final control time $t = T$, or the resolution of the numerical approximation (both of which would impact the accuracy of the reconstructed bathymetry). Additionally, quantifying the sensitivity to interaction between inputs would also prove a valuable result, especially considering the correlation between $\hat{\eta}$ and $\hat{\beta}$ observed in Chapter 3. As VBSA is not appropriate for this model, using density based methods for interaction affects would need to be explored. Sensitivity analyses on the accuracy of the observations themselves would also provide insight. In atmosphere and ocean models, observation measurements possess a significant degree of uncertainty, and subsequently a measure of their sensitivity to the observation operator would be illuminating for future work [61].

The present analysis proved insightful for an idealised model of data assimilation for the 1-D shallow water equations with Gaussian initial conditions and bathymetry. In order to utilise these results for more realistic forecasts, the inevitable next step for future considerations is the extension to a full 2-D model, and inclusion of multiple forms of bathymetry, such as sandbar or ridge formations.

Chapter 6

Conclusions

In this PhD thesis, we developed and evaluated variational data assimilation techniques for the nonlinear shallow water equations to reconstruct inaccurate or missing initial conditions and bathymetry data, in order to improve tsunami modelling. We use the "optimise then discretise" approach, by deriving the variational assimilation algorithms for the infinite-dimensional case, and then developing a Sadourny finite-difference finite-volume numerical scheme to solve the adjoint equations [47]. In this chapter we discuss the contributions of each chapter in turn, and conclude with the main contributions of our work and perspectives for future research.

In Chapter 2 we derived and implemented a 2-D variational data assimilation algorithm for the optimal reconstruction of initial conditions of surface waves, with the primary aim of extending the results of the 1-D variational assimilation outlined in [27]. We confirmed the feasibility of variationally data assimilation for tsunami waves in idealised 2D configurations, and demonstrated the necessary and/or sufficient conditions for convergence of the reconstructed initial conditions to the true shape. An essential difference between the 1-D and 2-D case is that the observations need to be able to measure the azimuthal symmetry of the surface wave, as well as its minimum length scale. This has implications for the number and configuration of the observation points. We showed that convergence can be achieved for sparse observations arranged in straight lines, grids, and along concentric circular arcs, for both isotropic and anisotropic initial conditions, and with realistic bathymetry data. Importantly, we have extended the 1-D results regarding sufficient conditions for the maximum distance between pairwise observations, and showed that for the 2-D case this is a necessary but not sufficient condition. We found that when there are no observations within the support of the initial conditions, having approximately 80 observations for a grid configuration and 65 for observations along $x = \pm y$, is a necessary condition for convergence. Note that this is still far fewer observation points than the number of grid points in the simulation, $128^2 = 16384$.

In Chapter 3 we derived a data assimilation algorithm for the 1-D nonlinear SWE to reconstruct the bathymetry from a set of observations of the free surface height, by minimising a functional $\mathcal{J}(\beta)$ representing the least squares error between observations and forecast solutions. Initial results showed that $\nabla\mathcal{J} \in L^2(\Omega)$ resulted in small-scale noise in the bathymetry reconstruction and impacted convergence. Consequently, we showed that smoother H^2 gradients are necessary to avoid small-scale errors in the reconstructed bathymetry. With a Gaussian initial condition, we accurately reconstructed the bathymetry for (i) a Gaussian bathymetry, and (ii) a sandbar profile bathymetry. Based on qualitative results for a localized surface wave propagating over a compact bathymetry feature, we observed that a necessary condition for convergence is that the ratio of the initial conditions amplitude to the bathymetry amplitude $\hat{\eta}/\hat{\beta}$ be at most $\mathcal{O}(10^{-2})$ when bathymetry amplitudes are larger than 10% of the average depth. Additionally, we observed that convergence improves with increasing numbers of observation points, and that at least 10 observation points are necessary for convergence of the assimilation algorithm to the true bathymetry with Gaussian initial conditions. We observed that the surface wave $\eta(x, t; \beta^{(b)})$ has relatively low sensitivity to errors in the reconstructed bathymetry, and showed that the free surface error was orders of magnitude smaller than the bathymetry reconstruction error as a function of N_{obs}, $\hat{\beta}$ and $\hat{\eta}$ respectively. We found that reconstructing the bathymetry with a relative error of about 10% is sufficiently accurate for surface wave modelling in most cases.

In order to more precisely quantify the sensitivity of the surface wave produced by the reconstructed bathymetry to perturbations in the observations, in Chapter 4 we analytically derived the sensitivity of the surface wave error ($\mathcal{G}$) to the observations (m), using second order adjoint techniques and the Hessian of the cost function $\mathcal{J}$. We recall that $m_j(t)$ are the observations taken at positions $\{x_j\}$ at for $j = 1, ..., N_{obs}$ at continuous times t, and $d\mathcal{G}/dm$ is an $N_{obs} \times N_t$ matrix, where N_t is the temporal resolution, and the $m_{i,j}$ entry of $d\mathcal{G}/dm$ represents the sensitivity of the surface wave error (4.6.1) to the i-th observation at the j-th time step. We derived the Hessian vector product $\mathcal{H}\nu$ given some arbitrary perturbation of the bathymetry ν, and demonstrated that deriving the sensitivity $\partial\mathcal{G}/\partial m$ involves solving the forced equation $\mathcal{H}\nu = \mathcal{F}$ with the right hand side $\mathcal{F}$ dependent on the optimal assimilation results. Using this sensitivity analysis, we investigated the conclusions of the bathymetry assimilation. We formulated three questions.

Question I asked whether there is a link between the sensitivity $\partial\mathcal{G}/\partial m$ and the convergence of the reconstructed bathymetry to the exact form. Results indicated that a higher error in bathymetry reconstruction is associated with higher sensitivity of the surface wave error to perturbations in the observations.

Question II asked how changing parameters in the data assimilation scheme (such as the spacing and number of observation points, or the shape of the bathymetry) affects the sensitivity. We concluded that for a compact Gaussian bathymetry, the spacing Δx of a fixed number of observations was non-influential and noted that for all values of Δx, the observation points covered the entire support of the bathymetry. This was not true with a sandbar bathymetry, where only the largest spacing resulted in coverage of the entire bathymetry support by observations, and we observed that the sensitivity decreased as the observation points spanned a larger area of the bathymetry support. We conducted trials where the width of the Gaussian bathymetry was increased incrementally, until it qualitatively resembled the sandbar bathymetry. We observed that the larger width led to an increase in the time-integrated sensitivity, regardless of the position of the bathymetry relative to the initial condition. We concluded that more extensive bathymetry features produce surface waves that are more sensitive to perturbations in the observation points.

Finally for question III, we investigated whether the qualitative observations from the bathymetry algorithm presented in Chapter 3 could be verified by the second order adjoint sensitivity analysis. One of the main conclusions from Chapter 3 was that convergence improved with more observation points. Therefore, we investigated whether increasing the number of observations also resulted in lower sensitivity of the surface wave error to the observations. The results indicated lower sensitivity across all cases when the number of observation points was increased, confirming our earlier conclusion.

We investigated the relationship between the bathymetry and initial conditions amplitudes $\hat{\beta}$ and $\hat{\eta}$, and the average depth H, using the second order adjoint sensitivity. We varied the relative bathymetry amplitude $\hat{\beta}/H$, to investigate the effect of the relative bathymetry amplitude on the sensitivity of the surface wave response $\mathcal{G}$ to the observations. We saw a general increase in sensitivity as the relative bathymetry height $\hat{\beta}/H$ increased, again confirming our observations in Chapter 3.

Finally, we verified the observation from Chapter 3 that the error in the surface wave is orders of magnitude lower than the error in the reconstructed bathymetry, suggesting low sensitivity of the surface wave to reconstruction error. We investigated whether the sensitivity $\partial\mathcal{G}/\partial m$ exhibited the same behaviour, i.e. whether the sensitivity of the surface wave error to observations was orders of magnitude lower than the bathymetry reconstruction error. In all cases considered, the sensitivity $\partial\mathcal{G}/\partial m$ was orders of magnitude lower than the minimum reconstruction error, confirming our observations from Chapter 3.

Complementary to the second order adjoint sensitivity analysis, we used Global Sensitivity Analysis (GSA) to derive sensitivity indices quantifying the influence of (i) the location of a Gaussian bathymetry relative to the location of the observation points, and (ii) the amplitudes of the initial conditions and bathymetry, on the bathymetry reconstruction error and the surface wave response respectively. We demonstrated that variance-based sensitivity analysis (VBSA) is not appropriate for our data assimilation model as the skewness of the model output violates necessary conditions that variance-based methods rely on, explaining the lack of convergence and robustness observed in initial results. As an alternative, we used a density-based sensitivity analysis (DBSA) to compute indices that rely on the variation between unconditional and conditional cumulative density functions (CDFs) when particular inputs are fixed, as a measure of sensitivity. We derived sensitivity indices in order to (i) rank the the influence of the inputs from highest to lowest, (ii) screen for non-influential parameters, and (iii) find thresholds in the input parameter values that mapped into specific output regions (factor mapping).

Our results showed that for the error in the surface wave, the most influential parameter was the amplitude of the bathymetry $\hat{\beta}$, whereas the amplitude of the initial conditions $\hat{\eta}$ was categorised as non-influential. This confirmed conclusions from Chapter 3, where we observed that the surface wave error increased as the amplitude of the relative bathymetry $\hat{\beta}/H$ became larger. We considered specific output regions for both the bathymetry reconstruction error and the surface wave error. For the latter we consider relative L^2 errors greater or less than 0.1%, and found that the influence of all three parameters decreases when the error is less than 0.1%.

We varied the position of the bathymetry (where ψ is the position of the peak of the Gaussian bathymetry as seen in figure 5.2.2), such that as ψ was decreased, a decreasing proportion of observation points observed the surface wave before it interacted with the bathymetry. For ψ approximately less than 1, the observations did not span the entire bathymetry support . We observed ψ was influential on the surface wave error for all statistics, when the error in the surface wave was larger than 0.1%. This confirmed conclusions from Chapter 4, where we determined that the position of the observation points (i.e. whether observations spanned the entire bathymetry support or not) had a significant effect on the sensitivity of the surface wave error to observations. These results suggest that in order to maintain low sensitivity of the surface wave response to the observations $\{m_j(t)\}$, the observations should span entire bathymetry features of Gaussian shapes 10% of the average depth, especially in cases where the error in the surface wave produced by the reconstructed bathymetry is predicted to be larger than 0.1%.

The bathymetry position ψ was the most influential parameter on the relative bathymetry reconstruction error, while the initial conditions amplitude $\hat{\eta}$ was non-influential. Overall the bathymetry reconstruction error was less sensitive to each of the input parameters when the bathymetry reconstruction error was large, suggesting that other parameters not considered in this analysis may have significant influence on larger values of the bathymetry reconstruction error.

For both bathymetry reconstruction error and the surface wave error, the amplitude of the initial conditions $\hat{\eta}$ was non-influential for each sub-region. On the other hand, in Chapter 3 we concluded that a necessary condition for convergence of the bathymetry assimilation was that the amplitude ratio $\hat{\eta}/\hat{\beta}$ be less than 0.1. However, the lack of sensitivity to $\hat{\eta}$ can be explained by the fact that all values where the amplitude ratio condition was violated (and subsequently the assimilation was non-convergent) were removed from the sample used to derive sensitivity indices. Therefore, we cannot assume that $\hat{\eta}$ is non-influential in the bathymetry reconstruction and the resulting free surface wave.

These results help us understand both the feasibility and limitations of variational data assimilation to recover initial conditions and bathymetry, with potential to inform future tsunami models in order to (i) maximise accuracy of the surface wave prediction through optimal reconstruction of the necessary data, (ii) attain a priori knowledge of how different bathymetry and initial conditions shapes can effect the surface wave error, and (iii) how these can be mitigated through optimal configuration of the observations. By demonstrating low sensitivity of the surface wave to the reconstruction error, we conclude we only need the reconstructed bathymetry to be "good enough" to produce sufficiently accurate surface waves. We observed that more than 10 observation points are required to ensure the relative error in the bathymetry reconstruction is less than 10%, especially when the bathymetry is more extensive, and having the observations span the entire bathymetry feature (as opposed to partial converage) reduces the sensitivity of the surface wave response to perturbations in the observations of sea surface height $m_j(t)$. The influence of "coverage" of the bathymetry feature by the observation points increases when the error produced in the surface wave by the reconstructed bathymetry is larger than 0.1%. Larger surface wave error is correlated with larger bathymetry amplitudes, however bathymetry features as tall as 30% of the average depth can be accurately reconstructed as long as the amplitude of the initial conditions is less than 1% of the bathymetry height.

In conclusion, we note that these results for bathymetry assimilation are for an idealised case, and the logical next steps are to extend the bathymetry assimilation algorithm to two dimensions, as we did for the initial conditions assimilation.

Secondarily, we could also include other physical effects such as Coriolis force, bottom friction and noisy observations, while retaining computational efficiency and numerical stability. Inclusion of such effects is also a logical next step for the 2-D assimilation for the initial conditions reconstruction. Although we are primarily interested in tsunami wave propagation over variable bathymetry over open ocean, we could also consider interaction with coastlines, or idealized closed geometries, such as channels. Finally, in order to incorporate observations that were effectively measured on a globe, we require a domain as close to realistic propagation as possible. Thus, the 2-D assimilation could be extended to the sphere.

We noted in Chapter 1 that tsunami waves can also be modelled using the Boussinesq water wave approximation. While both Boussinesq and shallow water approximations are widely used for analysing solitary wave propagation, for breaking waves the Boussinesq equations provide a better model of the wave evolution up till breaking point. Incorporation of coastlines motivates further analysis using the Boussineq approximation, for both comparison purposes as well as to verify the results derived for the shallow water model.

A significant consideration that was not addressed in this study is uncertainty propagation. Realistic applications of initial conditions reconstruction or parameter estimation using observations often involve a high degree of uncertainty in both the assimilated data and estimated values. Inclusion of a covariance function that accurately describes the deviation between the true and inferred parameter or observed value is necessary to ascertain the accuracy of the predicted state, but such covariances are not perfectly known. While Kalman filtering is a stochastic process that uses the error covariance of the estimated state and the measurements to generate an optimised estimate at the next time step, variational data assimilation usually has static covariances that do not evolve with future forecasts. However, there is scope for hybrid data assimilation methods, that effectively take flow-dependent covariances from an ensemble Kalman filtering model, and introduce them into the variational algorithm through additional control variables in the cost function [3]. Additionally, it is possible to describe a spatially varying parameter (like bathymetry) by modelling it as a stochastic processes, thereby determination of the covariance function can be achieved by reduced-order representation of the stochastic process (a series of random variables with probability density functions and statistical moments). This approach was successfully implemented by Mayo et al. [34] to estimate Manning's n coefficient (used to represent bottom stress terms), where a spectral representation of the stochastic process was used, with eigenfunctions of the associated covariance used as basis functions. Similar analyses for future bathymetry estimation would help quantify the uncertainty in the estimated parameter field.

Another simplification in of the mathematical model used in our analyses is that it describes the motion of a single fluid layer with a free surface. Additional dynamics may be captured by considering two or more distinct layers of different densities. This is especially suitable to an ocean model with an upper mixed layer and lower ocean, or deep ocean currents flowing under relatively quiescent fluid [58].

Ultimately, we have developed and extensively verified the use of variational data assimilation for the initial conditions and bathymetry reconstruction in a variety of idealized 1D and 2D configurations. We have found valuable insights on the effects of the observation operator and parameters, on convergence of reconstructed data to exact initial conditions or bathymetry shapes. Our verification of 1-D conditions for the 2-D initial condition assimilation suggest that results for the 2-D bathymetry assimilation similar to those we have highlighted in this thesis, may be observed. Additionally, the sensitivity analysis we have derived of the surface wave to bathymetry, should prove widely applicable to other applications in oceanography. Consequently, we believe that our work provides a valuable first step to understand the observability of bathymetry and initial conditions by surface waves, and the subsequent effect of reconstructing innaccurate or missing data, on surface wave accuracy. These results have the potential to faciliate more accurate forecasts of tsunami waves, and mitigate their high socioeconomic impact.

.

References

[1] *Activity Report Project-Team MOISE: Modelling, Observations, Identification for Environmental Sciences.* Université Joseph Fourier (Grenoble), 2014.

[2] R. Banner. ELEMENTARY 4D-VAR DARC Technical Report. *Data Assimilation Research Centre* (2) (2007), 101–109.

[3] D. Barker, A. Lorenc, and A. Clayton. Hybrid variational/ensemble data assimilation (2011).

[4] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. van der Vorst. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods.* Society for Industrial and Applied Mathematics, 1994.

[5] D. E. Barrick and L. S. Fedor. Measurement of ocean wave heights with a satellite radar altimeter. *Eos, Transactions American Geophysical Union* 59(9) (1978), 843–847.

[6] N. G. D. Center. *2-minute Gridded Global Relief Data (ETOPO2) v2. National Geophysical Data Center, NOAA.* Accessed: July 2020. 2006.

[7] W. Craig and C. Sulem. Asymptotics of Surface Waves over Random Bathymetry. *Quarterly of Applied Mathematics* 68(1) (2010), 91–112. ISSN: 0033569X, 15524485.

[8] M. Damen, P. van Dijk, J. Duim, H. van der Werff, B. Krol, B. Masselink, and F. van Ruitenbeek. *Characteristics of Tsunamis.* Accessed: 2020-01-24.

[9] C. Dawson and C. Mirabato. *Shallow Water Equations.* Computational Engineering and Sciences, University of Texas at Austin., 2008.

[10] L. Dongfang et al. Comparison between Boussinesq and Shallow-Water Models in Predicting Solitary Wave Runup on Plane Beaches. *Coastal Engineering Journal* 55(4) (2013).

[11] J. P. Dugan. Bathymetry measurements from long range airborne imaging systems. *Proc. 4th Int. Conf. Remote Sensing for Marine and Coastal Environments. Orlando, FL: ERIM* 1 (Mar. 1997), 451–457.

[12] *ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate.* 2017.

[13] *ETOPO2: Topography and Bathymetry.* `https://sos.noaa.gov/datasets/etopo2-topography-and-bathymetry-natural-colors/`. Accessed: 2020-07-22.

[14] S. Fletcher. *Data Assimilation for the Geosciences.* 1st. Elsevier, 2017. Chap. 20.

[15] H. Ge, H. Liu, and L. Zhang. Accurate Depth Inversion Method for Coastal Bathymetry: Introduction of Water Wave High-Order Dispersion Relation. *Journal of Marine Science and Engineering* 8(3) (Feb. 2020), 153. ISSN: 2077-1312.

[16] A. Gessese and M. Sellier. A Direct Solution Approach to the Inverse Shallow-Water Problem. *Mathematical Problems in Engineering* 2012 (Nov. 2012), 417950. ISSN: 1024-123X.

[17] H. Ghorbanidehno, J. Lee, M. Farthing, T. Hesser, P. K. Kitanidis, and E. F. Darve. Novel Data Assimilation Algorithm for Nearshore Bathymetry. *Journal of Atmospheric and Oceanic Technology* 36(4) (Apr. 2019), 699–715. ISSN: 0739-0572.

[18] F. I. Gonzalez, H. B. Milburn, E. Bernard, and J. Newman. Deep-ocean Assessment and Reporting of Tsunamis (DART): Brief Overview and Status Report. In: 1998.

[19] A. K. Griffith and N. K. Nichols. Data Assimilation Using Optimal Control Theory. *Numerical Analysis Report* 10(2) (1994), 101–109.

[20] S. Grilli and R. Subramanya. Numerical modeling of wave breaking induced by fixed or moving boundaries. *Computational Mechanics* 17 (1996), 374–391.

[21] S. T. Grilli. Depth inversion in shallow water based on nonlinear properties of shoaling periodic waves. *Coastal Engineering* 35(3) (1998), 185–209. ISSN: 0378-3839.

[22] S. T. Grilli. Depth inversion in shallow water based on nonlinear properties of shoaling periodic waves. *Coastal Engineering* 35 (3 Mar. 1998), 185–209.

[23] M. Heidarzadeh, T. Harada, K. Satake, T. Ishibe, and T. Takagawa. Tsunamis from strike-slip earthquakes in the Wharton Basin, northeast Indian Ocean: March 2016 Mw7.8 event and its relationship with the April 2012 Mw 8.6 event. *Geophysical Journal International* 211(3) (Sept. 2017), 1601–1612. ISSN: 0956-540X.

[24] M. P. Hickey. Atmospheric Gravity Waves and Effects in the Upper Atmosphere Associated with Tsunamis. In: *The Tsunami Threat*. Ed. by N.-A. M?rner. Rijeka: IntechOpen, 2011. Chap. 32.

[25] R. Iacono. Analytic solutions to the shallow water equations. *Phys. Rev. E* 72 (1 July 2005), 017302.

[26] T. Jang, S. Han, and T. Kinoshita. An inverse measurement of the sudden underwater movement of the sea-floor by using the time-history record of the water-wave elevation. *Wave Motion* 47(3) (2010), 146–155. ISSN: 0165-2125.

[27] N. Kevlahan, R. Khan, and B. Protas. On the convergence of data assimilation for the one-dimensional shallow water equations with sparse observations. *Adv Comput Math* 45 (2019), 3195–3216.

[28] S. Kucherenko, S. Tarantola, and P. Annoni. Estimation of global sensitivity indices for models with dependent variables. *Computer Physics Communications* 183(4) (2012), 937–946. ISSN: 0010-4655.

[29] Q. Liu and T. Homma. A new computational method of a moment-independent uncertainty importance measure. *Reliability Engineering System Safety* 94(7) (2009). Special Issue on Sensitivity Analysis, 1205–1211. ISSN: 0951-8320.

[30] A. C. Lorenc. Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society* 112(474) (1986), 1177–1194.

[31] M. Losch and C. Wunsch. Bottom Topography as a Control Variable in an Ocean Model. *Journal of Atmospheric and Oceanic Technology* 20(11) (2003), 1685–1696.

[32] S. C. Lubard, J. E. Krimmel, L. Thebaud, D. D. Evans, and O. H. Shemdin. Optical image and laser slope meter intercomparison of high-frequency waves. *J. Geophys. Res* 85 (1980), 4996–5002.

[33] T. Maeda, K. Obara, M. Shinohara, T. Kanazawa, and K. Uehira. Successive estimation of a tsunami wavefield without earthquake source data: A data

assimilation approach toward real-time tsunami forecasting. *Geophysical Research Letters* 42(19) (2015), 7923–7932.

[34] T. Mayo, T. Butler, C. Dawson, and I. Hoteit. Data assimilation within the Advanced Circulation (ADCIRC) modeling framework for the estimation of Manning's friction coefficient. *Ocean Modelling* 76 (2014), 43–58. ISSN: 1463-5003.

[35] K. Nakamura et al. Sequential Data Assimilation: Information Fusion of a Numerical Simulation and Large Scale Observation Data. *Journal of Universal Computing* 12(6) (2006), 608–626.

[36] D. Nicholls and M. Taber. Detection of ocean bathymetry from surface wave measurements. *European Journal of Mechanics B/Fluids* 28 (2 Mar. 2009), 224–233.

[37] D. P. Nicholls and M. Taber. Detection of ocean bathymetry from surface wave measurements. *European Journal of Mechanics - B/Fluids* 28(2) (2009), 224–233. ISSN: 0997-7546.

[38] A. B. Owen. *Monte Carlo theory, methods and examples.* 2013.

[39] M. Ozisik and H. R. B. Orlande. *Inverse Heat Transfer: Fundamentals And Applications.* Taylor & Francis, 2000.

[40] F. Pianosi, K. Beven, J. Freer, J. W. Hall, J. Rougier, D. B. Stephenson, and T. Wagener. Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling Software* 79 (2016), 214–232. ISSN: 1364-8152.

[41] F. Pianosi, F. Sarrazin, and T. Wagener. A Matlab toolbox for Global Sensitivity Analysis. *Environmental Modelling Software* 70 (2015), 80–85. ISSN: 1364-8152.

[42] F. Pianosi and T. Wagener. A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. *Environmental Modelling Software* 67 (2015), 1–11. ISSN: 1364-8152.

[43] C. C. Piotrowski and J. P. Dugan. Accuracy of bathymetry and current retrievals from airborne optical time-series imaging of shoaling waves. *IEEE Transactions on Geoscience and Remote Sensing* 40 (12 Dec. 2002), 2606–2618.

[44] C. C. Piotrowski and J. P. Dugan. Accuracy of bathymetry and current retrievals from airborne optical time-series imaging of shoaling waves. *IEEE Trans. Geoscience and Remote Sensing* 40 (2002), 2606–2618.

[45] B. Protas. Adjoint-based optimization of PDE systems with alternative gradients. *Journal of Computational Physics* 227(13) (2008), 6490–6510. ISSN: 0021-9991.

[46] A. Ribal and I. R. Young. 33 years of globally calibrated wave height and wind speed data based on altimeter observations. *Scientific Data* 6(1) (May 2019), 77. ISSN: 2052-4463.

[47] R. Sadourny. The Dynamics of Finite-Difference Models of the Shallow-Water Equations. *Journal of the Atmospheric Sciences* 32(4) (Apr. 1975), 680–689. ISSN: 0022-4928.

[48] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications* 181(2) (2010), 259–270. ISSN: 0010-4655.

[49] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. *Global Sensitivity Analysis: The Primer*. 2008. ISBN: 978-0-470-05997-5.

[50] F. Sarrazin, F. Pianosi, and T. Wagener. Global Sensitivity Analysis of environmental models: Convergence and validation. *Environmental Modelling Software* 79 (2016), 135–152. ISSN: 1364-8152.

[51] W. P. Schellart, D. R. Stegman, R. J. Farrington, and L. Moresi. Influence of lateral slab edge distance on plate velocity, trench velocity, and subduction partitioning. *Journal of Geophysical Research: Solid Earth* 116(B10) (2011).

[52] V. Shutyaev, F.-X. Le Dimet, and E. Parmuzin. Sensitivity analysis with respect to observations in variational data assimilation for parameter estimation. *Nonlinear Processes in Geophysics* 25(2) (2018), 429–439.

[53] V. Shutyaev, F. Le Dimet, and E. Shubina. Sensitivity with respect to observations in variational data assimilation. *Russian Journal of Numerical Analysis and Mathematical Modelling* 32 (1 2017), 61–71.

[54] W. H. Smith, R. Scharroo, V. V. Titov, D. Arcas, J. SAO, and B. K. Arbic. Satellite Altimeters Measure Tsunami—Early Model Estimates Confirmed. *Oceanography* issue$_v$*olume* (June 2005).

[55] I. M. Sobol. Sensitivity analysis for non-linear mathematical models. *Mathematical Modeling & Computational Experiment (Engl. Transl.)* 1 (1993), 407–414.

[56] I. Sobol. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation* 55(1) (2001). The Second IMACS Seminar on Monte Carlo Methods, 271–280. ISSN: 0378-4754.

[57] R. J. Spiteri and S. J. Ruuth. A New Class of Optimal High-Order Strong-Stability-Preserving Time Discretization Methods. *SIAM Journal on Numerical Analysis* 40(2) (2002), 469–491.

[58] A. Stewart and P. J. Dellar. Multilayer shallow water equations with complete Coriolis force. Part 1. Derivation on a non-traditional beta-plane. *Journal of Fluid Mechanics* 651 (2010), 387–413.

[59] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation.* Society for Industrial and Applied Mathematics, 2005.

[60] W. Tsai and D. K. P. Yue. Computation of Nonlinear Free-Surface Flows. *Annual Review of Fluid Mechanics* 28(1) (1996), 249–278.

[61] T. Wagener and F. Pianosi. What has Global Sensitivity Analysis ever done for us? A systematic review to support scientific advancement and to inform policy-making in earth system modelling. *Earth-Science Reviews* 194 (2019), 1–18. ISSN: 0012-8252.

[62] Y. Wang, T. Maeda, K. Satake, M. Heidarzadeh, H. Su, A. F. Sheehan, and A. R. Gusman. Tsunami Data Assimilation Without a Dense Observation Network. *Geophysical Research Letters* 46(4) (2019), 2045–2053.

[63] Z. Wang, I. Navon, F.-X. Le Dimet, and X. Zou. The second order adjoint analysis: Theory and applications. *Meteorology and Atmospheric Physics* 50 (Jan. 1992), 3–20.

[64] C. Wunsch. *The Ocean Circulation Inverse Problem.* Cambridge University Press, 1996.

[65] Y. Yang, E. M. Dunham, G. Barnier, and M. Almquist. Tsunami Wavefield Reconstruction and Forecasting Using the Ensemble Kalman Filter. *Geophysical Research Letters* 46(2) (2019), 853–860.

[66] X. Zou, I. M. Navon, and F. X. L. Dimet. Incomplete observations and control of gravity waves in variational data assimilation. *Tellus A: Dynamic Meteorology and Oceanography* 44(4) (1992), 273–296.