

Data assimilation for the two-dimensional shallow water equations: Optimal initial conditions for tsunami modelling

R.A. Khan^{*}, N.K.-R. Kevlahan

McMaster University, 1280 Main St West, Hamilton, ON, Canada

ARTICLE INFO

Keywords:

Shallow water equations
Observations
Characteristics
Grid
Arc

ABSTRACT

Accurate modelling of tsunami waves requires complete boundary and initial 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. We demonstrate that variational data assimilation for the continuous shallow water equations (SWE) is a feasible approach for recovering initial conditions. We showed that the necessary conditions for reconstructing one-dimensional initial conditions in Kevlahan et al. (2019) can be extended to the maximum Euclidean distance between pairwise observations to two-dimensions. We use Sadoury finite-difference finite volume simulations to verify convergence to the true initial conditions can be achieved for observations arranged in multiple configurations, for both isotropic and anisotropic initial conditions, and with realistic bathymetry data in two dimensions. We compare observations arranged in straight lines, in a grid, and along concentric circles, and assess the optimal number and configuration of observation points such that convergence to the true initial conditions is achieved. These idealised results with simplified two-dimensional geometry are a first step towards more physically realistic settings. Recent advances in altimetry observation data now permit much denser measurements of sea surface height than is possible with a fixed buoy network. This provides the opportunity to use the method developed here for more accurate tsunami forecasts in realistic settings.

1. Introduction

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 (Nakamura et al., 2006). Consequently, information on the initial conditions of tsunamis is likely to be incomplete or inaccurate. Tsunamis are series of waves caused by large-scale disturbances in the ocean, such as seismic activity. In this study we focus on tsunami propagation in the deep ocean and not on the interaction of tsunami waves with the coast (where the dispersive Boussinesq equation may be a better model). 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 4000 m, tsunamis can travel faster than 700 km/h, and are capable of generating atmospheric gravity waves that can travel into the upper atmosphere.

In this study we formulate a variational data assimilation scheme for the two-dimensional shallow water equations (SWE), where observations of sea surface height are used to reconstruct missing initial conditions data. The two-dimensional 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 (2.1).

For the one-dimensional SWE considered in Kevlahan et al. (2019), the wavefront is simply a moving point (on one side). Therefore, in one dimension the only degrees of freedom in observation placement are their spacing and their number. However, in the two-dimensional extension considered here, the wavefront is a one-dimensional curve and the observations can be placed along curves or in various two-dimensional arrays. Additional features that increase the complexity of the this case include nonlinear characteristic curves of the hyperbolic partial differential equation, and the wave focusing effects of two-dimensional bathymetry features. Our goal is to investigate if the minimum spacing requirement in Kevlahan et al. (2019) can be extended to two dimensions. We are particularly interested in exploring the effects of greater choice in observation placement and the richer nonlinear wave-front geometry allowed by the two-dimensional problem.

^{*} Correspondence to: Department of Meteorology (MISU), Stockholm University, Stockholm, Sweden.
E-mail address: ramsha.khan@misu.su.se (R.A. Khan).

In practice, [Zou et al. \(1992\)](#) have demonstrated that the discretised two-dimensional 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 ([Wang et al., 2019](#)). 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) ([Gonzalez et al., 1998](#)). The large investment required for such apparatus limits the feasibility of a dense network, and currently the global network consists of approximately 60 systems, spread over the entire Pacific Ocean and located mostly near coastlines, at depths of 1–6 km. Other observations used for tsunami detection and reporting include ocean bottom seismometer (OBS) pressure gauges, however these are sparsely distributed in the Indian ocean ([Wang et al., 2019](#)).

[Angove et al. \(2019\)](#) state that one of the major limitations to accurate tsunami forecasts is their inability to quickly measure and represent the tsunami source. The placement and number of DART tsunameters and ocean bottom seismometers (OBS) is currently not extensive enough to simultaneously provide (i) early warning of impending tsunamis, and (ii) enough observational data to for data assimilation methods to accurately reconstruct the displacement at the source, and subsequently predict wave height and velocity at coastlines. Therefore, our objective with this work is to outline a theoretical data assimilation framework that could make use of satellite altimetry data, where considerable progress has been made in data availability and precision.

Existing data assimilation schemes are able to utilise multiple sources and techniques for assimilating off-shore observations. [Maeda et al. \(2015\)](#) 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. \(2019\)](#) 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.

Kalman filtering techniques have also been used extensively for both parameter estimation and initial condition reconstruction in ocean models. [Mayo et al. \(2014\)](#) 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. \(2019\)](#) 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. \(2019\)](#) 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 equivalent for both ([Fletcher,](#)

[2017; Lorenc, 1986](#)). 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. \(1992\)](#) and [Maeda et al. \(2015\)](#), and Kalman filtering techniques. Consequently, our data assimilation algorithm is independent of the discretisation used in its numerical implementation.

Our model for tsunami wave propagation is based on the two-dimensional non-dispersive SWE, however 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. \(2013\)](#) 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. This idealised configuration with simplified two-dimensional geometry is intended to be a first step for more complex analyses. For this reason we also assume there are no errors in measurements. We aim to validate the basic approach by demonstrating the feasibility of variational data assimilation for tsunami wave prediction, and investigate fundamental questions for the idealised case first before considering more physically realistic settings.

The purpose of the current study is not to determine optimal placement of observations, or to develop new assimilation techniques based on observability criteria. Instead, we extend our one dimensional initial condition variational assimilation technique ([Kevlahan et al., 2019](#)) from one dimension to two dimensions, and analyse results from a variety of numerical experiments assessing its performance and flexibility.

The problem of finding an optimal observation configuration has been considered in [King et al. \(2015\)](#). They derive a measure of how well information about the system state can be inferred using observation sensors, and provide an approximation for a partial observability index. Their findings suggest that optimised observations reduced the root mean square error (RMSE) by approximately 20% compared to equally spaced observations. Similarly, [Kang and Xu \(2014\)](#) also introduce a gradient projection method for maximising partial observability. However, [King et al. \(2015\)](#) provide results showing the efficacy of their observability index for the one-dimensional shallow water equations and optimal observations are arranged in a straight line, with variable spacing. Comparisons are given with equidistant points, and the extension to higher dimensions is not discussed. [Kang and Xu \(2014\)](#) also only provide results for the one-dimensional Burger's equation. We note that the core contribution of these works is the mathematical approximation of partial observability, framing it as an eigenvalue optimisation problem, whereas our goal is not to develop an algorithm to determine optimal placement of observations, but to extend the one-dimensional algorithm derived in [Kevlahan et al. \(2019\)](#) to two-dimensions, and verify its performance in a situation where there is much more freedom in observation placement.

In [Kevlahan et al. \(2019\)](#), we have proven sufficient conditions for optimal reconstruction of the true initial conditions using sparse observations for the one-dimensional SWE. Our objective was to determine the optimal number and locations of wave height measurements, such that the optimally reconstructed initial conditions obtained via data assimilation converged to the true form. In order to extend the analysis to more realistic dynamics in tsunami models, here we implement an analogous scheme on the two-dimensional 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 two-dimensional case is to find the minimum information, in the form of observations of surface wave height (such as their configuration, number, and spacing between adjacent points), required for convergence to the true initial conditions.

This focus also differentiates the current work from existing literature on data assimilation for the two-dimensional SWE like Zou et al. (1992), 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 two-dimensional SWE system, a novelty of our approach is that we consider the infinite dimensional case, unlike previous works on 4D-VAR. Our purpose is to extend the one-dimensional data assimilation scheme derived in Kevlahan et al. (2019) to two-dimensional.

Subsequently, we investigate whether the conditions for convergence observed for the one-dimensional case, also hold for the more complex two-dimensional system. We recall that for the one-dimensional 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 one-dimensional case we were able to exploit the fact that the linear one-dimensional SWE system can be formulated as the one-dimensional 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 two-dimensional case, as the analytical solution to the adjoint system is not easily found. Existing works, such as Iacono (2005) on analytical solutions of the two-dimensional SWE, involve drastic simplifications of the equilibrium problem. Such simplifications would not help us adequately investigate the minimum observation information required for tsunami models with high dimensionality.

Therefore, our approach is focused on a qualitative investigation of the applicability of results from the one-dimensional analysis for the two-dimensional data assimilation, and on a comparison of different observation spacings, configurations, and numbers. We begin in Section 2 by extending the data assimilation to two-dimensional, and implement this algorithm numerically in Section 3. The results and complementary analyses are divided into the following three topics.

1. Section 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 two-dimensional SWE.
2. Section 4.1: Investigate whether the sufficient conditions for convergence in Kevlahan et al. (2019) for the one-dimensional case extends to the two-dimensional data assimilation.
3. Section 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 6.

2. Derivation

The two-dimensional shallow water equations (SWE) are

$$\frac{\partial \eta}{\partial t} + \frac{\partial}{\partial x}((\eta + H - \beta)u) + \frac{\partial}{\partial y}((\eta + H - \beta)v) = 0, \quad (2.1a)$$

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + \frac{\partial \eta}{\partial x} = 0, \quad (2.1b)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + \frac{\partial \eta}{\partial y} = 0, \quad (2.1c)$$

$$\eta(x, y, 0) = \phi(x, y), \quad (2.1d)$$

$$u(x, y, 0) = 0, \quad (2.1e)$$

$$v(x, y, 0) = 0, \quad (2.1f)$$

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 1.

Let us define $m_j^{(o)}(t)$ to be observations of the surface wave at positions $\{(x, y)_j\}$ for $j = 1, \dots, N_{obs}$ at continuous times t . 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 conditions ϕ . We express this as a cost functional $\mathcal{J} : L^2(\Omega) \rightarrow \mathbb{R}$, constrained by the system (2.1),

$$\mathcal{J}(\phi) = \frac{1}{2} \int_0^T \sum_{j=1}^M \left[\eta^{(f)}(x_j, y_j, t; \phi) - m_j^{(o)}(t) \right]^2 dt. \quad (2.2)$$

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

$$\phi^{(b)} = \operatorname{argmin}_{\phi \in L^2(\Omega)} \mathcal{J}(\phi). \quad (2.3)$$

Since the minimum of (2.2) is achieved when

$$\nabla^{L^2} \mathcal{J}(\phi^{(b)}) = 0, \quad (2.4)$$

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), given ϕ .

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

$$\mathcal{J}'(\phi; \phi') = \lim_{\epsilon \rightarrow 0} \frac{\mathcal{J}(\phi + \epsilon \eta') - \mathcal{J}(\phi)}{\epsilon}. \quad (2.5)$$

Expanding the perturbation to $\mathcal{O}(\epsilon)$, we can reformulate (2.5) as

$$\mathcal{J}'(\phi; \phi') = - \int_0^T \left(\eta^{(f)}(x_j, y_j, t; \phi) - m^{(o)}(t) \right) \eta' dt, \quad (2.6)$$

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 $\mathcal{O}(\epsilon)$ system.

As the Gateaux derivative is a directional derivative in the direction of the perturbation ϕ' , we can express (2.6) as the inner product between $\nabla \mathcal{J}$ and ϕ ,

$$\mathcal{J}'(\phi; \phi') = \langle \nabla \mathcal{J}, \phi' \rangle_{L^2(\Omega)} = \int_{\Omega} \nabla^{L^2} \mathcal{J} \eta' d\Omega. \quad (2.7)$$

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

$$\mathcal{J}'(\phi; \phi') = - \int_0^T \left(\eta^{(f)}(x_j, y_j, t; \beta) - m^{(o)}(t) \right) \eta' dt = \int_{\Omega} \nabla^{L^2} \mathcal{J} \phi' d\Omega. \quad (2.8)$$

If we form a Lagrangian of our linearised shallow water system with some arbitrary adjoint variables (η^*, u^*, v^*) ,

$$\begin{aligned} \int_0^T \int_{\Omega} \eta^*(x, y, t) \left[\frac{\partial \eta'}{\partial x} + \frac{\partial}{\partial x} (u \eta' + (H + \eta - \beta) u') \right. \\ \left. + \frac{\partial}{\partial y} (v \eta' + (H + \eta - \beta) v') \right] \\ + u^*(x, t) \left[\frac{\partial u'}{\partial t} + u \frac{\partial u'}{\partial x} + v \frac{\partial u'}{\partial y} + \frac{\partial \eta'}{\partial x} \right] \\ + v^*(x, t) \left[\frac{\partial v'}{\partial t} + u \frac{\partial v'}{\partial x} + v \frac{\partial v'}{\partial y} + \frac{\partial \eta'}{\partial y} \right] d\Omega dt = 0, \end{aligned} \quad (2.9)$$

Table 1

Notation used in the derivation of data assimilation scheme of the two-dimensional SWE to find the optimal initial conditions.

Symbol	Definition
$\eta(x, y, t)$	General solution for the height perturbation
$\phi^{(i)}(x, y)$	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 $\{x_j, y_j\}, j = 1, \dots, N_{obs}$
$\eta^{(f)}(x, y, t)$	Approximate ("forecast") solution generated by approximate bathymetry
$\mathcal{J}^{(n)}$	Cost function at iteration n
$(\cdot)^*$	Adjoint

we observe that by formulating the following system in terms of the adjoint variables (η^*, u^*, v^*) ,

$$\frac{\partial \eta^*}{\partial t} + u \frac{\partial \eta^*}{\partial x} + v \frac{\partial \eta^*}{\partial y} + \frac{\partial u^*}{\partial x} + \frac{\partial v^*}{\partial y} = m(x_i, y_i, t) - H\eta(x, y, t; \beta), \quad (2.10a)$$

$$\frac{\partial u^*}{\partial t} + u \frac{\partial u^*}{\partial x} + v \frac{\partial u^*}{\partial y} + (H + \eta - \beta) \frac{\partial \eta^*}{\partial x} = 0, \quad (2.10b)$$

$$\frac{\partial v^*}{\partial t} + u \frac{\partial v^*}{\partial x} + v \frac{\partial v^*}{\partial y} + (H + \eta - \beta) \frac{\partial \eta^*}{\partial y} = 0, \quad (2.10c)$$

$$\eta^*(x, y, T) = 0, \quad (2.10d)$$

$$u^*(x, y, T) = 0, \quad (2.10e)$$

$$v^*(x, y, T) = 0, \quad (2.10f)$$

and integrating (2.9) by parts in space and time, due to the periodicity of the boundary conditions the Lagrangian (2.9) is reduced to

$$\int_0^T \int_{\Omega} (\eta^{(f)}(x_j, y_j, t; \phi) - m^{(o)}(t)) \eta' d\Omega dt = - \int_{\Omega} \eta^* \eta' \Big|_{t=0} d\Omega, \quad (2.11)$$

Combining this result with the equivalence given by (2.8), we have

$$- \int_{\Omega} \eta^* \eta' \Big|_{t=0} d\Omega = - \int_{\Omega} \eta^* \Big|_{t=0} \phi' d\Omega = \int_{\Omega} \nabla^{L^2} \mathcal{J} \phi' d\Omega, \quad (2.12)$$

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 for the equivalence of inner products to extract $\nabla^{L^2} \mathcal{J}$, giving us

$$\nabla^{L^2} \mathcal{J}(\phi) = -\eta^*(x, y, 0). \quad (2.13)$$

For a detailed derivation we refer the reader to Kevlahan et al. (2019). We utilise an iterative steepest descent algorithm to find our minimiser $\phi^{(b)}$ yielding $\nabla^{L^2} \mathcal{J} = 0$, given some starting guess $\phi^{(g)}$. We also used a Polak-Ribière conjugate gradient descent method, and found results to be equivalent to steepest descent. The steps of the process at each iteration are summarised in algorithm 1.

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. The Sadourny scheme ensures that this discretisation inherits the conservation properties (such as energy or enstrophy conservation) of the original system (Sadourny, 1975).

If the SWE are expressed in the following equivalent form,

$$\frac{\partial \eta}{\partial t} + \frac{\partial}{\partial x} ((\eta + H - \beta)u) + \frac{\partial}{\partial y} ((\eta + H - \beta)v) = 0 \quad (3.3)$$

$$\frac{\partial u}{\partial t} - \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) v + \frac{\partial}{\partial x} \left(g\eta + \frac{1}{2}u^2 + \frac{1}{2}v^2 \right) = 0 \quad (3.4)$$

$$\frac{\partial v}{\partial t} + \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) v + \frac{\partial}{\partial y} \left(g\eta + \frac{1}{2}u^2 + \frac{1}{2}v^2 \right) = 0, \quad (3.5)$$

Algorithm 1 Data Assimilation Algorithm for initial conditions Estimation for the two-dimensional SWE.

- 1: Pick initial estimate for $\phi^{(g)}$.
- 2: Solve the initial value problem for (η, u, v) from $t = 0$ to $t = T$.
- 3: Solve adjoint problem for (η^*, u^*, v^*) backwards in time from $t = T$ to $t = 0$ to find $\eta^*(x, y, 0)$.
- 4: Define $\nabla^{L^2} \mathcal{J} = -\eta^*(x, y, 0)$.
- 5: Compute the optimal time step τ_n at the n -th iteration through a line minimisation algorithm

$$\tau_n = \operatorname{argmin}_{\tau \in \mathbb{R}} \mathcal{J}(\phi^{(n)}(x, y) - \tau \nabla^{L^2} \mathcal{J}(\phi^{(n)}(x, y))). \quad (3.1)$$
- 6: Use a gradient descent algorithm to compute the guess for ϕ at the next time step

$$\phi^{(n+1)}(x, y) = \phi^{(n)}(x, y) - \tau_n \nabla^{L^2} \mathcal{J}(\phi^{(n)}(x, y)). \quad (3.2)$$
- 7: Repeat until $\|\eta^*(x, y, 0)\| \approx 0$.
- 8: Set $\phi^{(b)}(x, y) := \phi^{(n)}(x, y)$.

Then the discretisation

$$\frac{\partial u}{\partial t} = \frac{1}{\Delta x} \delta_i (g\eta + \frac{1}{2}u^2 + \frac{1}{2}v^2) + \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) \overline{\overline{\overline{(1 + \eta - \beta)^j v}}}, \quad (3.6)$$

$$\frac{\partial v}{\partial t} = \frac{1}{\Delta y} \delta_i \left(g\eta + \frac{1}{2}u^2 + \frac{1}{2}v^2 \right) + \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) \overline{\overline{\overline{(1 + \eta - \beta)^i u}}}, \quad (3.7)$$

where

$$\delta_i f = f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}}, \quad (3.8)$$

$$\overline{\overline{\overline{f}}}^i = \frac{1}{2} \left(f_{i+\frac{1}{2}} + f_{i-\frac{1}{2}} \right), \quad (3.9)$$

ensures that the total energy is conserved, such that

$$\begin{aligned} 0 = & \sum_{ij} \Delta x \Delta y \left[g\eta \frac{\partial \eta}{\partial t} + \left(\frac{1}{2}u^2 + \frac{1}{2}v^2 \right) \frac{\partial}{\partial t} (1 + \eta - \beta) \right] \\ & + \sum_{ij} \Delta x \Delta y \left[(1 + \eta - \beta) u \frac{\partial u}{\partial t} \right] \\ & + \sum_{ij} \Delta y \Delta x \left[(1 + \eta - \beta) v \frac{\partial v}{\partial t} \right] \end{aligned} \quad (3.10)$$

where the staggered centre difference operator (3.8) and centre linear interpolation operator (3.9) 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. 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 (Spiteri and Ruuth, 2002). 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

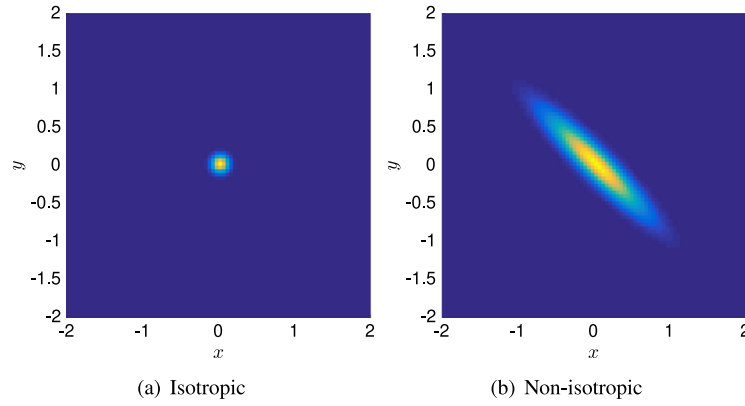


Fig. 1. Planar view of the true initial conditions $\phi^{(i)}$.

time steps. The line minimisation for 3.1s carried out using the Matlab function `fminunc`.

We consider the following isotropic Gaussian initial conditions, where the Gaussian has the same variance and physical properties in every direction,

$$\phi^{(i)}(x, y) = \frac{1}{20} \exp\left(-\frac{x^2 + y^2}{0.1^2}\right). \quad (3.11)$$

For cases where anisotropic initial conditions are used, the variance of the Gaussian in the x and y directions is different, and the Gaussian has been rotated by $\pi/4$ radians.

$$\phi^{(i)}(x, y) = \frac{1}{20} \exp\left(-\frac{x_r^2 + (y_r/5)^2}{0.15^2}\right), \quad (3.12)$$

$$x_r = x \cos\left(-\frac{\pi}{4}\right) - y \sin\left(-\frac{\pi}{4}\right), \quad (3.13)$$

$$y_r = x \cos\left(-\frac{\pi}{4}\right) + y \sin\left(-\frac{\pi}{4}\right). \quad (3.14)$$

Fig. 1 shows the isotropic and anisotropic initial conditions. In the results, we indicate an approximation of this support (spatial domain where the initial condition is non-zero) to highlight the position of 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.

The results in Section 4 are for assimilation with flat bathymetry $\beta(x, y) = 0$ as in Kevlahan et al. (2019). However, in Section 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.

3.1. Verification test

To verify our numerical implementation, we conduct a 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 $\varepsilon \rightarrow 0$,

$$\kappa(\varepsilon) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \frac{\mathcal{J}(\phi + \varepsilon \phi') - \mathcal{J}(\phi)}{\int_{\Omega} -\eta^*(x, y, 0) \phi' d\Omega}. \quad (3.15)$$

The results of the verification test are shown in Fig. 2, with a non-zero bathymetry.

4. Numerical results

The analysis for the two-dimensional data assimilation is more complex than the one-dimensional case due to the more complex geometry of nonlinear wavefronts propagating in two dimensions and the increased 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. We classify a result as convergent if the relative L^2 error is lower than 10% based on convergence observed in Kevlahan et al. (2019), noting that all results where the minimum distance criteria (4.1) was not met had a relative convergence error higher than 10% in the 1-D case.

In Kevlahan et al. (2019), 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 in most cases our initial conditions are 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 that to capture the isotropic nature of the initial conditions it may be beneficial to place observations along an arc 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 qualitatively investigate this claim by analysing results for three observation configurations. (i) In a square grid centred at $(0,0)$, (ii) along the characteristics $x = y$ and $x = -y$, and (iii) on circular arcs along $x^2 + y^2 = r_i^2$ for $i = 1, 2$. In Section 4.1 we extend the analysis for observation spacing from the one-dimensional case as in Kevlahan et al. (2019) for each configuration. For all cases, we assume the initial guess $\phi^{(g)} = 0$. In Section 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 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.

4.1. Observation spacing

In order to investigate the effect of observation spacing for the two-dimensional initial conditions assimilation, we use the results of Kevlahan et al. (2019) as a benchmark. We recall that for the one-dimensional assimilation, optimal convergence to the true initial conditions $\phi^{(i)}$ 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

$$|m_1 - m_2| \leq \frac{\pi}{k_{\max}}, \quad (4.1)$$

where k_{\max} is the largest effective wavenumber of $\phi^{(i)}$. In the absence of a rigorous equivalent result for the two-dimensional assimilation, we qualitatively investigate whether convergence improves when (4.1) is satisfied for the two-dimensional wavenumbers k and l . To find

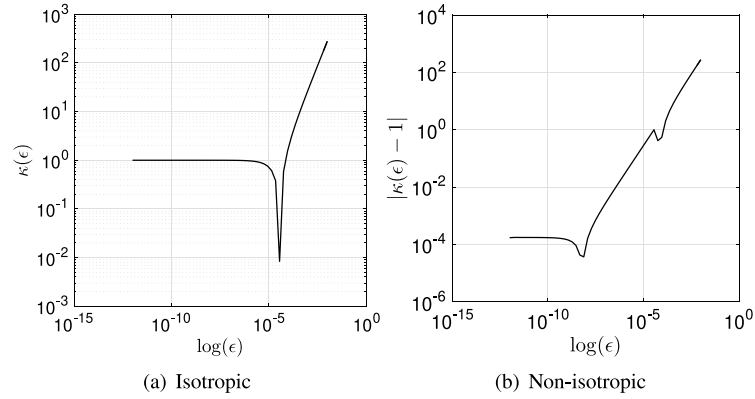


Fig. 2. Verification test for numerical implementation. (a) is the loglog graph of κ as $\epsilon \rightarrow 0$, and (b) is the difference $|\kappa(\epsilon) - 1|$ as $\epsilon \rightarrow 0$.

k_{max} and l_{max} , we examined the energy spectrum of the true isotropic two-dimensional initial conditions. We observed 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 one-dimensional case, at least one pair of points was required to satisfy (4.1). However, it may be that a greater number of points satisfying this condition is correlated with improved convergence in the two-dimensional 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 $\Delta x = \Delta y$ (Δr for the arc configurations) in a grid, along characteristics $x = \pm y$, and in radial arcs centred on the initial conditions. We also investigate convergence with different numbers of pairwise observations with a Euclidean distance fulfilling (4.1). We begin with observations along characteristic lines $x = \pm y$.

4.1.1. 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 Kevlahan et al. (2019) where $N_{obs} \leq 5$ was considered. However, given the higher dimensionality of the two-dimensional 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 (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 (4.1) for $\Delta x = 0.2$.

The characteristic configurations can be seen in Fig. 3(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 $\phi^{(i)} \neq 0$ may be slightly larger, and corresponds to Fig. 1. The configurations in 3(a), 3(b), and 3(c) satisfy (4.1), and those in 3(d) do not. We wish to determine whether convergence is worse where spacing is larger than 0.0785.

The convergence results for both cases are given in 3(e), the relative L^2 error in the reconstruction after 1000 iterations. We observe that the assimilation converges 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 (4.1) (and therefore are consistent with the minimum spacing theorem in Kevlahan et al. (2019)), 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.

4.1.2. 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 Fig. 4(a), (b), (c), and (d) with the same values of Δx as for the characteristic case. The trend in the relative error decrease over 1000 iterations are presented in Fig. 4(e).

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 4(d) has a relative error at the final iteration of $\mathcal{O}(10^{-2})$. This suggests that given the convergence criteria of a relative L^2 error lower than 10%, each spacing considered in Fig. 4 converges, 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 (4.1), whereas the ones with $\Delta x = 0.1$ and $\Delta x = 0.2$ had none, confirming the hypothesis that (4.1) is a sufficient condition for convergence with a grid configuration, for the two-dimensional 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 one-dimensional theorem in Kevlahan et al. (2019).

It should be noted however that in Fig. 4, 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 spacing Δx and Δy between adjacent observations. Consequently in Section 4.2 we investigate the effect of increasing the number of observation points, and subsequently in Section 4.3 the effect of removing all points that lie in the initial conditions support.

4.1.3. Observations along arcs

Current observations used for detection and forecasting of tsunamis such as DART buoys are usually arrayed in an arc-like formation along coastlines (Gonzalez et al., 1998). Therefore, we consider a circular

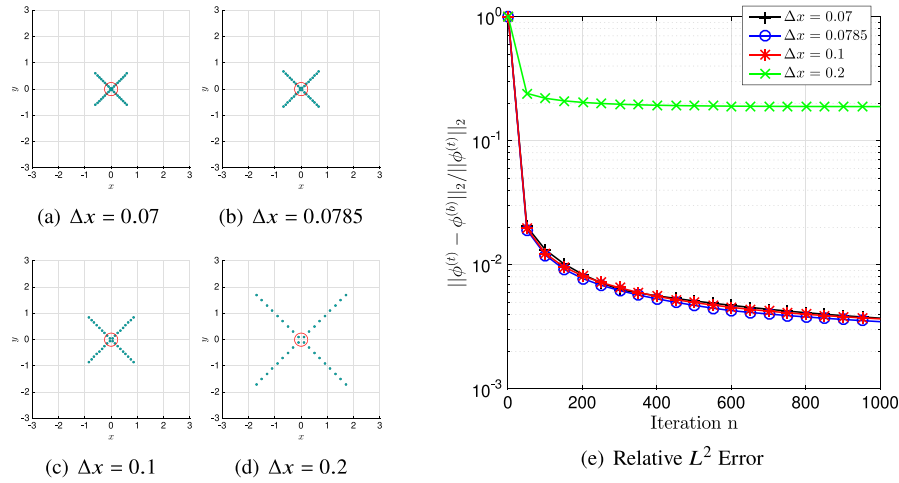


Fig. 3. Observations along characteristics $x = \pm y$ with varying spacing such that $\Delta x = \Delta y$, and $N_{obs} = 36$. The green circles represent the observation points, and the area inside the red circle approximates the support of the initial conditions. The assimilation time is $t \in [0, 2]$. (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.

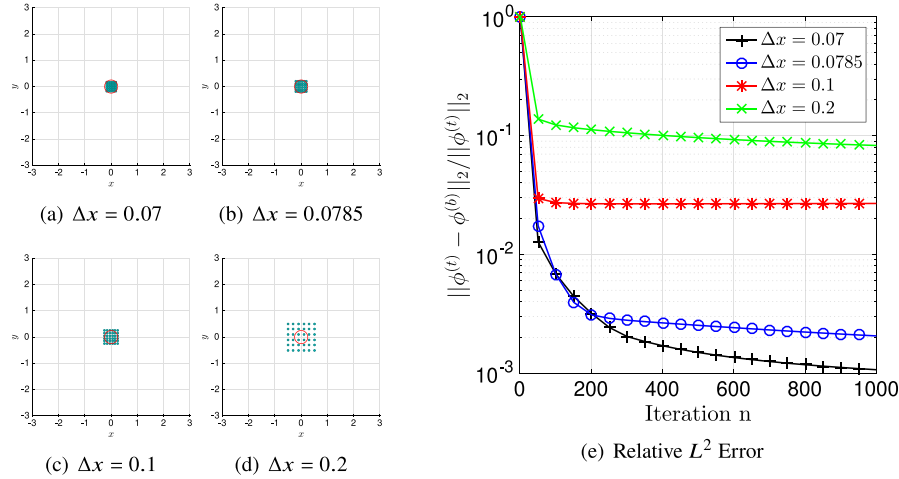


Fig. 4. Observation configuration in a grid format with varying spacing such that $\Delta x = \Delta y$, and $N_{obs} = 36$. (d) shows the convergence of the cost function after 1000 iterations and (e) represents the relative L^2 error in the initial conditions reconstruction. All configurations converge (L^2 error less than 10%), however it is marginal for (d).

arc configuration for observations, as shown in Fig. 5(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 Figs. 5(a) and 5(b) satisfy (4.1) and Figs. 5(c) and 5(d) do not, similar to the characteristic and grid configurations.

We observe in Fig. 5(e) 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 (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 Fig. 5(f)–(i). 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 satisfying (4.1) is necessary. In the following section we investigate the minimum number of observations necessary for convergence, and assess whether convergence improves as we increase the number of observations in each of the three configurations.

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 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.

The results for the error convergence given a characteristic configuration of the observations are presented in Fig. 6(a). We observe that the relative reconstruction error 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

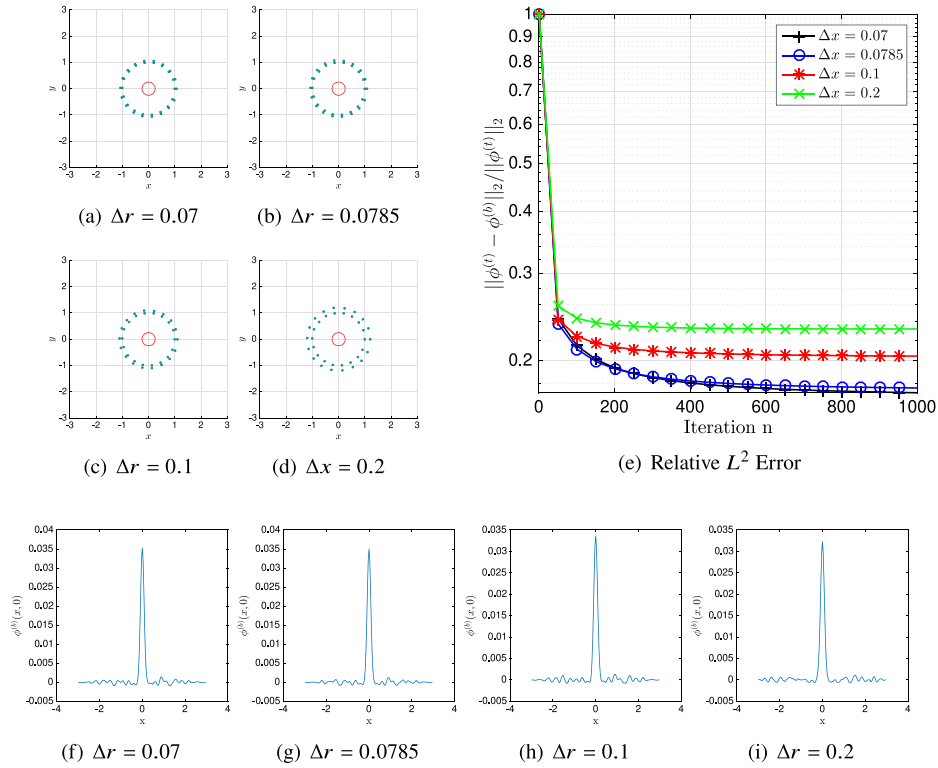


Fig. 5. Observations along the two concentric circles $x^2 + y^2 = r_i^2$ for $i = 1, 2$ and with spacing between circles $\Delta r = |r_1 - r_2|$ and $N_{obs} = 36$. (d) shows the convergence of the cost function after 1000 iterations and (e) represents the relative L^2 error in the initial conditions reconstruction. f-h: Cross-sections of the reconstructed initial conditions at $y = 0$ for cases in Fig. 5 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.

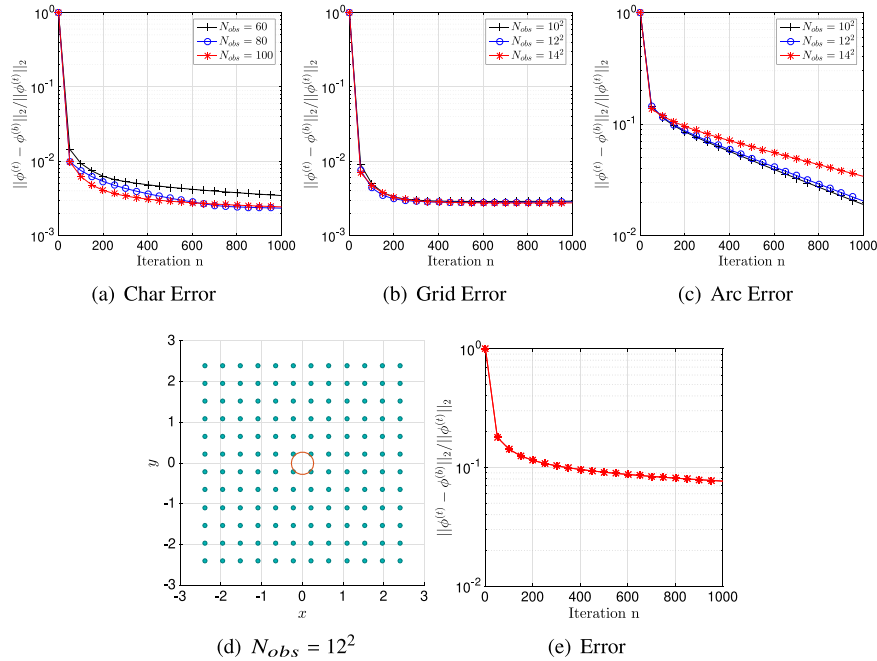


Fig. 6. (a)–(c) present the convergence of the relative L^2 error with increased number of points and $\Delta x = 0.0785$ for each configuration of observations. (d) shows a grid formation with increased number of points and $\Delta x = 0.44$, and (e) highlights the nonconvergence of the relative L^2 error.

in Fig. 3(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.

The error convergence results for observations placed in a grid format with varying number of observation points are given in Fig. 6(b).

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 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 Fig. 6(e) that the assimilation does not converge, and therefore large

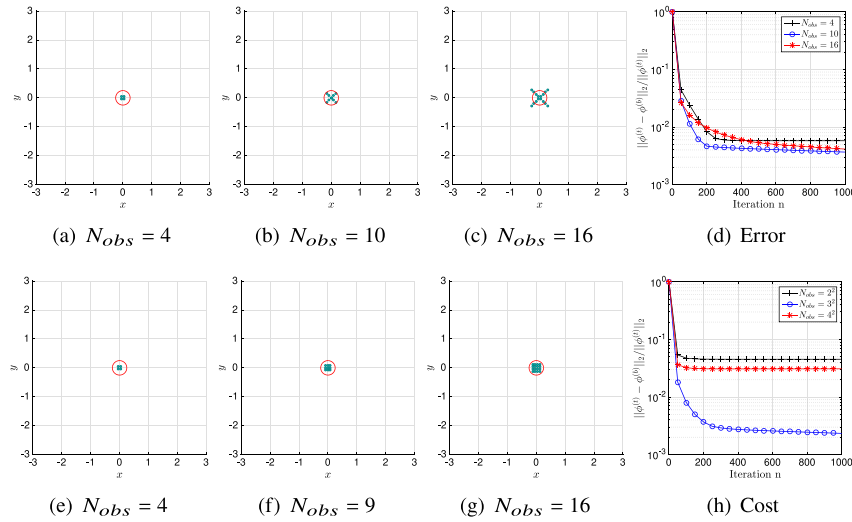


Fig. 7. (a)–(d) Characteristic formation, and (e)–(h) Grids formation, with a small number of points and $\Delta x = 0.0785$.

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 4.2.1.

We observed in Fig. 5 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 Fig. 6(c). The relative L^2 error 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 no points satisfy (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.

4.2.1. 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 Fig. 7(a)–(d). 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 7(e)–(h). 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.

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 results observed thus far, we note that for all cases that converge (such as in Fig. 7), 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 4.3.

4.3. Results with no observations within the support of the initial conditions

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 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 Fig. 8.

We observe that even with $N_{obs} = 28$, the configuration in Fig. 8(c) (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 Fig. 7(e). 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 shown in Fig. 8(a). Consequently, we investigate whether convergence can be achieved when N_{obs} is relatively large.

Fig. 9(a)–(d) 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 9(d). We note that the error is almost two orders of magnitude higher than it was in Fig. 6(a). 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

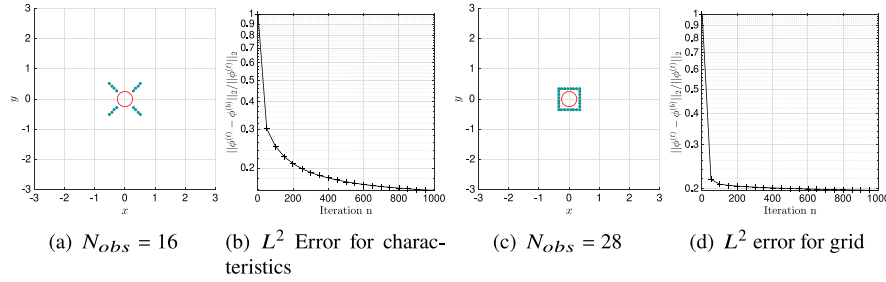


Fig. 8. Convergence of the relative L^2 error for observations in a (a) characteristic configuration, and (b) grid configuration, with small N_{obs} and $\Delta x = 0.07$, and points within the support of the initial conditions removed.

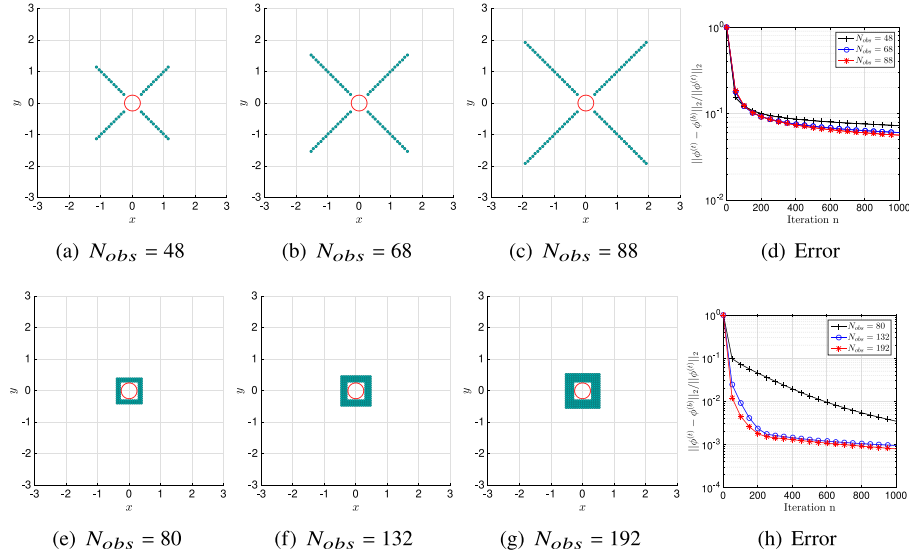


Fig. 9. (a)–(d): Characteristic configuration, and (e)–(h) Grid configuration with points within the support of the initial conditions removed.

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.

The configurations and result for observations in a grid format are presented in Fig. 9(e)–(h). We note some significant differences between the results here and those in Fig. 6(b). First, it is clear from Fig. 9(h) that there is improved 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 Fig. 6(b) did not indicate difference in convergence for the different values of N_{obs} . Secondly, the best convergence in Fig. 9(h) is a relative L^2 error of approximately 0.1%, indicating better convergence than observed in Fig. 6(b), despite the overall number of observations being comparable. It is interesting to note that in Fig. 6(b) the error stagnated after approximately 400 iterations. This suggests that perhaps 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 Fig. 9(h) we observe that the error is still maintaining a negative slope even at the final iteration, suggesting improved convergence. To verify this hypothesis, Fig. 10 shows the energy spectra of the relative absolute error of the reconstructed initial conditions for configurations 6(b) and 9(h) with $N_{obs} = 144$ and $N_{obs} = 132$ respectively.

We can see that the energy of larger wavenumbers in the reconstruction error $\|\phi^{(t)} - \phi^{(b)}\| / \|\phi^{(t)}\|_{L^2}$ is higher for configuration 6(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 9(h) 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%.

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 some pairwise observations satisfies the condition (4.1). The convergent configurations for the grid configuration in Fig. 4 had 64 pairwise observations satisfying the spacing condition (4.1), while the non-convergent cases with spacing $\Delta x \geq 0.1$ had none. Similarly, the non-convergent arc configurations in Fig. 5 had at most 14 pairwise observations with a Euclidean distance less than 0.0785. This may suggest that there is a minimum number of pairwise observations required that satisfy the spacing condition (4.1). However, as the minimum spacing observed in this analysis is a qualitative estimate and not analytically derived, we cannot conclusively quantify the minimum number of observations based on this value alone.

Additionally, 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 Fig. 10.

To extend these results for more realistic conditions, we conducted equivalent analyses with the anisotropic initial conditions given in (3.12).

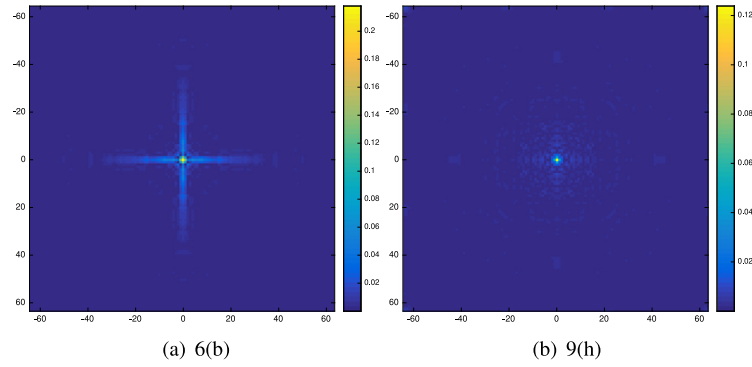


Fig. 10. Energy spectrum of the error $|\phi^{(i)} - \phi^{(b)}| / \|\phi^{(i)}\|_{L^2}$ for the grid configurations in 6(b) (points inside the support of $\phi^{(b)}$) and 9(h) (no points inside the support of $\phi^{(b)}$) with $N_{obs} = 144$ and $N_{obs} = 132$ respectively.

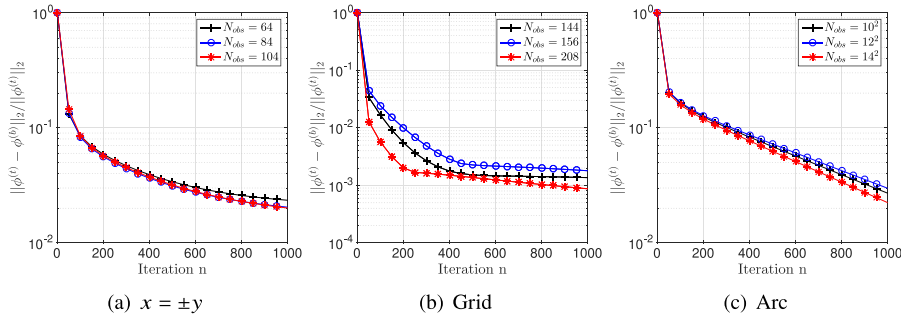


Fig. 11. Convergence of the L^2 error with anisotropic initial conditions, with points within the support of the initial conditions removed, for each observation configuration and $\Delta r = 0.07$.

4.4. Results for anisotropic 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 contrast with results for isotropic Gaussian initial conditions. Our aim 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.

Briefly, we found that there was no significant impact on convergence with anisotropic initial conditions (as given in (3.12)), compared to results in Section 4.3. The convergence of the relative L^2 error for each of the three configurations is presented in Fig. 11. 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} .

4.4.1. Summary of main results for observation configurations (flat bathymetry)

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 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 (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 for convergence, or there is a minimum number of pairwise observations with a Euclidean distance satisfying (4.1), which the arc configuration did not satisfy (having only 14 such points, whereas the convergent results for the grid configurations in Fig. 4 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 Fig. 10.

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).

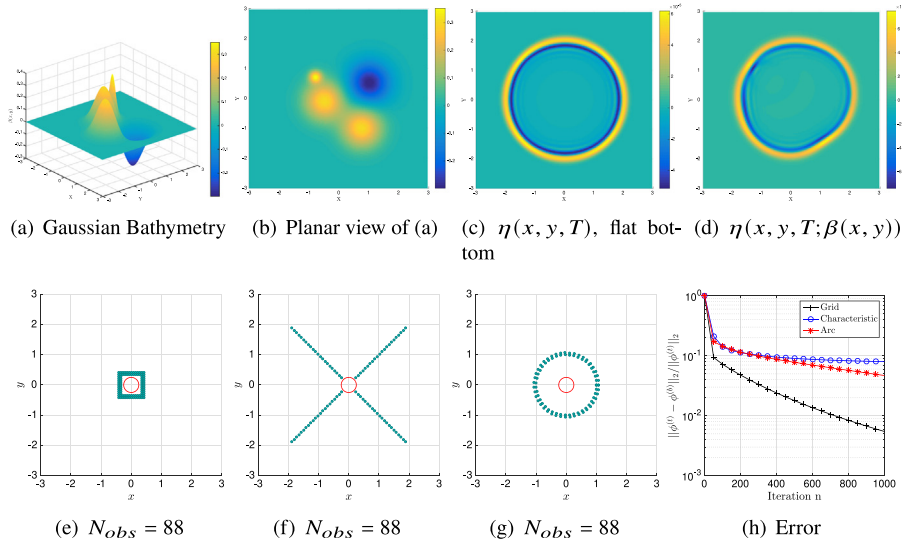


Fig. 12. (a)–(b): Gaussian Bathymetry with 3 peaks and a basin as described in (5.1). The y -axis represents the amplitude of the bathymetry relative to the mean depth H , which has been normalised to $H = 1$. Planar view of the free surface wave (c) without bathymetry, and (d) with bathymetry given in (5.1) at the final time T , showing the effect of the bathymetry in distorting the surface wave. (e)–(h): Results with Gaussian Bathymetry. Note that the configuration in (f) can no longer be referred to as observations placed along characteristics, as these are no longer characteristics of the distorted wave.

The purpose of this section is to take the best results observed thus far, and integrate them into a more realistic model, adding some features of an operational tsunami forecast model. To this end, we limit our results to $N_{obs} = 88$. Even though we observed higher convergence with a greater number of N_{obs} (for example in Fig. 9(h)), there are practical difficulties in employing such dense observation networks. Therefore, we limit the maximum number of observation locations, on the basis 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.

5.1. Gaussian bathymetry

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

$$\begin{aligned} \beta(x, y) = & \frac{3}{10} \exp \left[-\frac{(x+0.5)^2 + (y+0.1)^2}{0.65^2} \right] \\ & - \frac{3}{10} \exp \left[-\frac{(x-1)^2 + (y-0.5)^2}{0.7^2} \right] \\ & + \frac{3}{10} \exp \left[-\frac{(x-0.75)^2 + (y+1)^2}{0.65^2} \right] \\ & + \frac{3}{10} \exp \left[-\frac{(x+0.8)^2 + (y-0.7)^2}{0.2^2} \right] \end{aligned} \quad (5.1)$$

where a representation is given in Fig. 12. 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 Fig. 12, where Fig. 12(c) is the free surface wave $\eta(x, y, t)$ at $t = T$ with a flat bottom bathymetry, and Fig. 12(d) is the free surface $\eta(x, y, t; \beta(x, y))$ at $t = T$ with the bathymetry given in (5.1).

We observe in Fig. 12(d) 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 Fig. 12(a). Similarly the amplitude of the wave is slightly damped in the first quadrant, indicating the effect of the basin.

The results with isotropic initial conditions (and no points inside the initial conditions support) are shown in Fig. 12(h). 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 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 do 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 with realistic bathymetry.

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 (NOAA

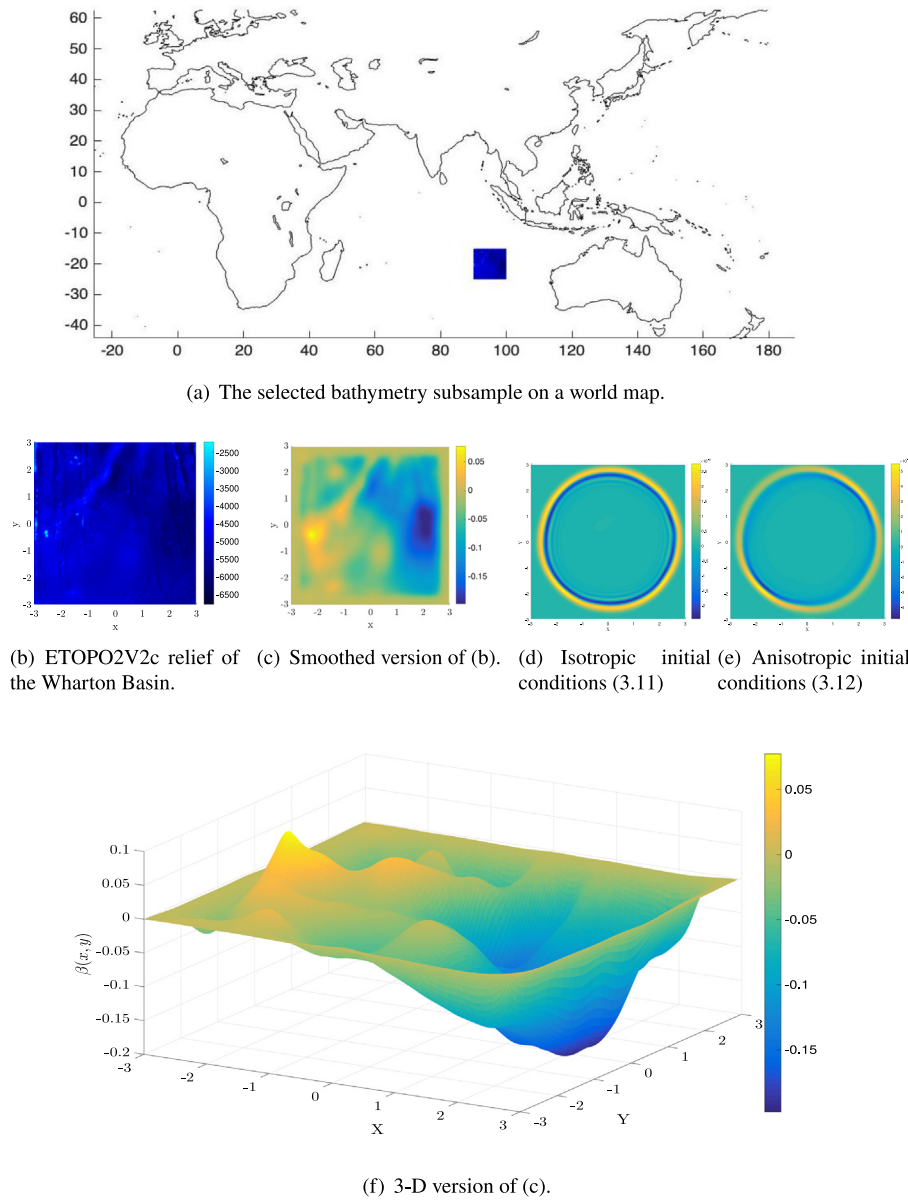


Fig. 13. Smoothed Etopo2v2c bathymetry relief within $[15^{\circ}\text{S}, 25^{\circ}\text{S}]$ and $[90^{\circ}\text{E}, 100^{\circ}\text{E}]$. d-e: Planar view of the free surface wave with ETOPO2 bathymetry and (d) Isotropic initial conditions (3.11), and (e) Anisotropic initial conditions (5.1), at the final time T .

National Geophysical Data Center, 2006). The grid values represent elevation at the cell edges, averaged over the cell area. The horizontal grid spacing is 2-min of latitude and longitude, where 1 min of latitude represents 1.853 km at the Equator, and the vertical precision is 1 m, 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 $[15^{\circ}\text{S}, 25^{\circ}\text{S}]$ and $[90^{\circ}\text{E}, 100^{\circ}\text{E}]$ respectively. This is equivalent to a rectangular grid of approximately 1100 km \times 1100 km (accounting for slight differences between degrees of longitude due to the equatorial bulge, where the maximum variation is approximately 60 km). 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 (Heidarzadeh

et al., 2017). A visual representation of the region and a relief map as taken directly from the ETOPO2v2c database are given in Fig. 13(a) and 13(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 a low pass filter 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 5 m). The smoothed and scaled relief map, and 3-D plot of the bathymetry can be seen in Fig. 13(c) and 13(f).

We implement the data assimilation for observations along the lines $x = \pm y$, a grid, and an arc, with isotropic initial conditions and

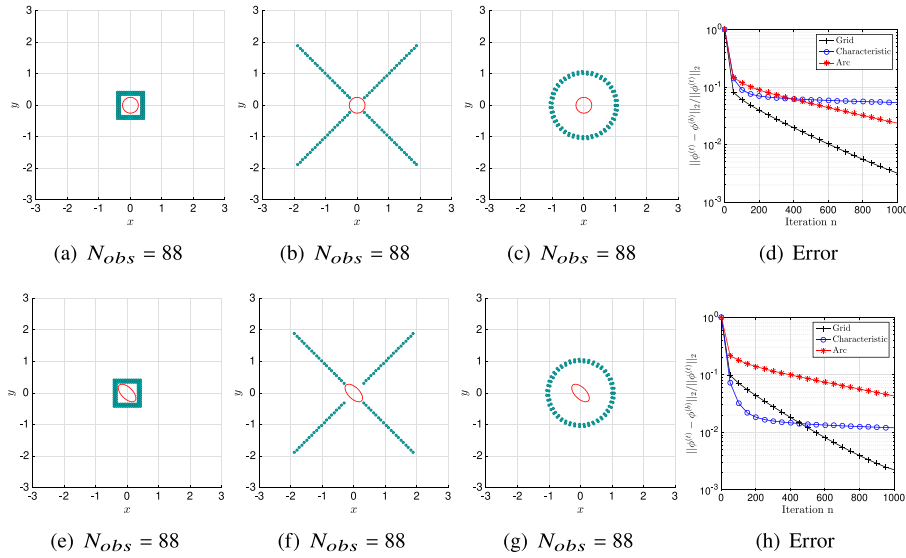


Fig. 14. Results with ETOPO2 and (a)–(d) isotropic initial conditions. We observe better convergence with an arc configuration of observations than for a straight-line configuration, despite the latter being in closer proximity to the initial condition region. (e)–(h): Results with ETOPO2 and anisotropic initial conditions.

anisotropic initial conditions respectively. Figs. 13(d) and 13(e) show the free surface wave at the final time T for each initial conditions case. The effect of the bathymetry on the free surface wave is relatively small compared to the flat bottom case in Fig. 12(c), than for the Gaussian bathymetry as in Fig. 12(d). However, we note that the highest amplitude in the ETOPO2 bathymetry (as shown in Fig. 13(c)) is only approximately 0.06, which is significantly less than the bathymetry in (5.1). Nevertheless, we observe slightly deeper troughs in the free surface wave indicated by the dark blue regions in 13(d), which are indicative of bathymetry effects. We conclude that changes in the wave speed due to 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 Fig. 14. For the isotropic initial conditions, we see in Fig. 14(d) 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 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.

We consider analogous results with a anisotropic initial conditions in Fig. 14(e)–(h). The main trends in the relative L^2 error shown are similar to the results with isotropic initial conditions as shown in Fig. 14(d), 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 4.4 that the arc and straight line configurations show relatively worse convergence with a radially asymmetric surface wave and bathymetry features. The main difference across the two cases is the relatively better performance of the straight line formation. Nevertheless, overall convergence for both isotropic and anisotropic initial conditions is relatively equivalent.

6. Conclusion and further considerations

In this study, we implemented a data assimilation scheme for the two-dimensional SWE to reconstruct the initial conditions, using observations of free surface wave height. Our objective was to investigate

configurations of observations that minimise the reconstruction error in the presence of increased complexity (compared to one dimension). These new features include curved non-linear wave-fronts and bathymetry features. We did this by comparing cases where observations satisfied the minimum distance criterion (4.1) between pairwise observations (Kevlahan et al., 2019), and when they did not. 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 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.

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 Eq. (4.1). While this confirms results from the one-dimensional case in part, the main difference is that for the one-dimensional assimilation, the spacing condition (4.1) was sufficient for convergence. We have demonstrated that this is not the case (as for results in Fig. 5) in the two-dimensional assimilation, and we may require a minimum number of points satisfying (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 Eq. (4.1) was not analytically derived for the two-dimensional 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 initial conditions, 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.

Table 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 $\phi^{(i)}$, with large N_{obs} and $\Delta x = 0.07$ (satisfying the necessary condition for convergence given in (4.1)).

Analysis	Lines $x = \pm y$	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 $\phi^{(i)}$, small N_{obs} , $\Delta x = 0.07$	18%	20%	None
No observations within support of $\phi^{(i)}$, large N_{obs} , $\Delta x = 0.07$	7%	0.1%	2%

Convergence was slightly worse for the characteristic configuration with anisotropic 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.

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

- (1) Is it realistic to assume that we have a sufficient number of observations for convergence, in the case of a grid of observations centred on the initial conditions?
- (2) Is it realistic to have a sufficient number of observations of surface wave heights for potential locations of initial conditions (i.e. associated with known seismically active regions)?

Addressing these, we note that the most destructive tsunamis 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. (2011) 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 (Heidarzadeh et al., 2017), 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 logistical feasibility of such configurations. Angove et al. (2019) provide a comprehensive review of the major limitations to accurate tsunami forecasts today, and detail key areas where uncertainties associated with observing tsunamis can be quantified and reduced. They highlight the need for updated methods to accurately capture initial displacement, and detail possible solutions, such as using data from the ever-expanding Global Navigation Satellite System (GNSS), bottom pressure recorders (BPRs), and limited-area cabled observation systems. They conclude that instrumenting commercial fibre-optic cables with BPRs and accelerometers is the most promising approach, but it comes with complexities and costs that cannot be covered by scientific revenue streams alone.

This is especially problematic when we consider that 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. (2019). 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. Angove et al. (2019) note that the process to constrain a tsunami source from altimetric data is similar to that currently employed, and has the advantage of providing both spatial and temporal variation not present in the static time series reported by tsunameters. Despite this, they stipulate that due to limitations in sensitivity, limited satellite coverage, and extended data processing times, operational use of satellite altimetry in real-time tsunami forecasting has not been a viable solution to date.

However, improvements in remote sensing of ocean waves using altimetry have come a long way since the Geodynamics Experimental Ocean Satellite 3 (GEOS-3) satellite was first launched in 1975. Even then, remote sensing of ocean wave height from space was predicted to provide a quantum increase in forecasting capabilities (Barrick and Fedor, 1978). 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 (2019). With a global spatial resolution of $7 \text{ km} \times 7 \text{ km}$, 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. (2005) 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.

Ultimately, the efficacy of such modelling frameworks in a realistic setting also depends on related factors not considered in the present work, such as positioning of observations such that detection time of tsunami waves is optimised. Ferrolino et al. (2020) demonstrate a promising methodology using population-based algorithms that optimises placement of deep ocean sensors around subduction zones, and finds the minimal time it takes for the disturbance on the source to arrive at sensor locations. While this work does not focus on reconstruction of missing initial condition information, or wave accuracy as it propagates away from the source, it is nonetheless a necessary consideration for realistic tsunami observational needs. Earlier detection can potentially mitigate the delay in receiving and processing altimetry data.

As this work is restricted to the extension of the results in Kevlahan et al. (2019) to two dimensions, future development would benefit from considering the observability framework outlined by Kang and Xu (2014) and King et al. (2015). This may provide a more rigorous mathematical justification for optimal two-dimensional configurations of observations. We recall that the mathematical proof of theorem (4.1)

is applicable only for the one-dimensional case, and consequently the results shown here are qualitative. A similar concept to the minimum distance criterion considered here was provided in Hinson (2014), where a lower bound on the radius of attraction to satisfy observability was found. Future work aimed at determining optimal configurations of observations in two dimensions would benefit from exploring the effects of placing sensors within and outside this observability radius.

The analysis provided here can also be made more rigorous by considering the well-posedness of the inverse problem, and the possible issues of non-uniqueness and discontinuities in the mapping used to reconstruct missing parameters, such as initial conditions, from the model and observations. These considerations are addressed in detail by Navon (1998), where regularisation methods such as Tikhonov regularisation, inclusion of background information, and sensitivity analyses are suggested to address such issues and to identify influential parameters within the system. Subsequently a next step for this work would be to conduct sensitivity analyses as in Kevlahan and Khan (2020) and Khan and Kevlahan (2021), and to include regularisation and background information terms in the cost function optimisation.

Global optimisation may be a more efficient and reliable method to minimise the cost function than the gradient minimisation method used here. In particular, global optimisation may avoid the algorithm getting stuck in a local minimum. Ferreiro-Ferreiro et al. (2020) consider a similar problem for tsunami models, where the shallow water model is discretised using a positivity preserving second-order path-conservative finite volume scheme, and the data assimilation problem is posed in a global optimisation framework. Our work is based on the “optimise then discretise” approach, and our goal is to extend the methodology derived in Kevlahan et al. (2019) to two dimensions. Therefore, data assimilation with a global optimisation algorithm is outside the scope of the current work. However, realistic extensions of our results may benefit greatly from the automatic data assimilation methods provided by Ferreiro-Ferreiro et al. (2020).

In conclusion, we have developed a two-dimensional variational data assimilation algorithm for reconstruction of initial conditions of surface waves, with the primary aim of extending the results of the one-dimensional variational assimilation outlined in Kevlahan et al. (2019). 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 one-dimensional results regarding necessary conditions for the maximum distance between pairwise observations as given by Eq. (4.1) (i.e. resolving the smallest scales of initial surface displacement), and have additionally shown that for the two-dimensional 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.

CRediT authorship contribution statement

R.A. Khan: Conceptualisation, Methodology, Formal analysis, Software, Data curation, Writing – original draft, Visualisation, Investigation. **N.K.-R. Kevlahan:** Supervision, Conceptualisation, Methodology, Validation, Writing – review & editing, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Dr. Bartosz Protas for his ongoing contribution and feedback on this research. Additionally, we would like to thank Dr. Laurent Debeau and Dr. Arthur Vidard at the Laboratoire Jean Kuntzmann for their suggestions.

Funding

This project was funded by the NSERC, Canada Discovery Grant of Dr. Nicholas Kevlahan, and research exchange at the Laboratoire Jean Kuntzmann was facilitated by the McMaster-CNRS, Canada Fund travel grant.

References

- Angove, M., Arcas, D., Bailey, R., Carrasco, P., Coetzee, D., Fry, B., Gledhill, K., Harada, S., von Hillebrandt-Andrade, C., Kong, L., McCreery, C., McCurrach, S.J., Miao, Y., Sakya, A.E., Schindel, F., 2019. Ocean observations required to minimize uncertainty in global tsunami forecasts, warnings, and emergency response. *Front. Mar. Sci.* 6, <https://doi.org/10.3389/fmars.2019.00350>, <https://www.frontiersin.org/article/10.3389/fmars.2019.00350>.
- Barrick, D.E., Fedor, L.S., 1978. Measurement of ocean wave heights with a satellite radar altimeter. *EOS Trans. Am. Geophys. Union* 59 (9), 843–847. <https://doi.org/10.1029/EO059i009p00843>, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/EO059i009p00843>.
- Dongfang, L., et al., 2013. Comparison between Boussinesq and shallow-water models in predicting solitary wave runup on plane beaches. *Coast. Eng. J.* 55 (4).
- Ferreiro-Ferreiro, A., García-Rodríguez, J., López-Salas, J., Escalante, C., Castro, M., 2020. Global optimization for data assimilation in landslide tsunami models. *J. Comput. Phys.* 403, 109069. <https://doi.org/10.1016/j.jcp.2019.109069>, <https://www.sciencedirect.com/science/article/pii/S0021999119307>.
- Ferrolino, A.R., Lope, J.E.C., Mendoza, R.G., 2020. Optimal location of sensors for early detection of tsunami waves. In: Krzhizhanovskaya, V.V., Závodszy, G., Lees, M.H., Dongarra, J.J., Sloot, P.M.A., Brissos, S., Teixeira, J. (Eds.), *Computational Science ICCS 2020*. Springer International Publishing, Cham, pp. 562–575.
- Fletcher, S., 2017. *Data Assimilation for the Geosciences*, first ed. Elsevier.
- Ghorbanidehno, H., Lee, J., Farthing, M., Hesser, T., Kitanidis, P.K., Darve, E.F., 2019. Novel data assimilation algorithm for nearshore bathymetry. *J. Atmos. Ocean. Technol.* 36 (4), 699–715. <https://doi.org/10.1175/JTECH-D-18-0067.1>.
- Gonzalez, F.I., Milburn, H.B., Bernard, E., Newman, J., 1998. Deep-ocean assessment and reporting of tsunamis (dart): brief overview and status report.
- Heidarzadeh, M., Harada, T., Satake, K., Ishibe, T., Takagawa, T., 2017. 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. *Geophys. J. Int.* 211 (3), 1601–1612. <https://doi.org/10.1093/gji/ggx395>.
- Hinson, B., 2014. *Observability-Based Guidance and Sensor Placement* (Ph.D. thesis).
- Iacono, R., 2005. Analytic solutions to the shallow water equations. *Phys. Rev. E* 72, 017302. <https://doi.org/10.1103/PhysRevE.72.017302>, URL <https://link.aps.org/doi/10.1103/PhysRevE.72.017302>.
- Kang, W., Xu, L., 2014. Partial observability for some distributed parameter systems. *Int. J. Dyn. Control* 2, 587–596. <https://doi.org/10.1007/s40435-014-0087-4>.
- Kevlahan, N.K.R., Khan, R.A., 2020. Variational assimilation of surface wave data for bathymetry reconstruction. Part I: algorithm and test cases.
- Kevlahan, N., Khan, R., Protas, B., 2019. On the convergence of data assimilation for the one-dimensional shallow water equations with sparse observations. *Adv. Comput. Math.* 45, 3195–3216. <https://doi.org/10.1007/s10444-019-09733-6>.
- Khan, R.A., Kevlahan, N.K.R., 2021. Variational assimilation of surface wave data for bathymetry reconstruction. part I: algorithm and test cases. *Tellus A* 73, 125. <https://doi.org/10.1080/16000870.2021.1976907>.
- King, S., Kang, W., Xu, L., 2015. Observability for optimal sensor locations in data assimilation. *Int. J. Dyn. Control* 3, 416–424. <https://doi.org/10.1007/s40435-014-0120-7>.
- Lorenc, A.C., 1986. Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.* 112 (474), 1177–1194.
- Maeda, T., Obara, K., Shinohara, M., Kanazawa, T., Uehira, K., 2015. Successive estimation of a tsunami wavefield without earthquake source data: A data assimilation approach toward real-time tsunami forecasting. *Geophys. Res. Lett.* 42 (19), 7923–7932. <https://doi.org/10.1002/2015GL065588>, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015GL065588>.

- Mayo, T., Butler, T., Dawson, C., Hoteit, I., 2014. Data assimilation within the advanced circulation (ADCIRC) modeling framework for the estimation of manning's friction coefficient. *Ocean Model.* 76, 43–58. <http://dx.doi.org/10.1016/j.ocemod.2014.01.001>, URL <http://www.sciencedirect.com/science/article/pii/S146350031400002X>.
- Nakamura, K., et al., 2006. Sequential data assimilation: Information fusion of a numerical simulation and large scale observation data. *J. Univ. Comput.* 12 (6), 608–626.
- National Geophysical Data Center, 2006. 2-minute gridded global relief data (ETOPO2) v2. National Geophysical Data Center, NOAA. <http://dx.doi.org/10.7289/V5J1012Q>, URL <https://www.ngdc.noaa.gov/mgg/global/etopo2.html>. (Accessed July 2020).
- Navon, I., 1998. Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography. *Dyn. Atmos. Oceans* 27, 55–79. [http://dx.doi.org/10.1016/S0377-0265\(97\)00032-8](http://dx.doi.org/10.1016/S0377-0265(97)00032-8), <https://www.sciencedirect.com/science/article/pii/S0377026597000328>.
- NOAA National Geophysical Data Center, 2006. 2-Minute gridded global relief data 955 (etopo2) v2. <http://dx.doi.org/10.7289/V5J1012Q>, Accessed 22 July 2020.
- Ribal, A., Young, I.R., 2019. 33 Years of globally calibrated wave height and wind speed data based on altimeter observations. *Sci. Data* 6 (1), 77. <http://dx.doi.org/10.1038/s41597-019-0083-9>.
- Sadourny, R., 1975. The dynamics of finite-difference models of the shallow-water equations. *J. Atmos. Sci.* 32 (4), 680–689. [http://dx.doi.org/10.1175/1520-0469\(1975\)032<0680:TDOFDM>2.0.CO;2](http://dx.doi.org/10.1175/1520-0469(1975)032<0680:TDOFDM>2.0.CO;2).
- Schellart, W.P., Stegman, D.R., Farrington, R.J., Moresi, L., 2011. Influence of lateral slab edge distance on plate velocity, trench velocity, and subduction partitioning. *J. Geophys. Res. Solid Earth* 116 (B10), <http://dx.doi.org/10.1029/2011JB008535>, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JB008535>.
- Smith, W.H., Scharroo, R., Titov, V.V., Arcas, D., Sao, J., Arbic, B.K., 2005. Satellite altimeters measure tsunami—early model estimates confirmed. *Oceanography* <http://dx.doi.org/10.5670/oceanog.2005.62>.
- Spiteri, R.J., Ruuth, S.J., 2002. A new class of optimal high-order strong-stability-preserving time discretization methods. *SIAM J. Numer. Anal.* 40 (2), 469–491. <http://dx.doi.org/10.1137/S0036142901389025>.
- Wang, Y., Maeda, T., Satake, K., Heidarzadeh, M., Su, H., Sheehan, A.F., Gusman, A.R., 2019. Tsunami data assimilation without a dense observation network. *Geophys. Res. Lett.* 46 (4), 2045–2053. <http://dx.doi.org/10.1029/2018GL080930>, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL080930>.
- Yang, Y., Dunham, E.M., Barnier, G., Almquist, M., 2019. Tsunami wavefield reconstruction and forecasting using the ensemble Kalman filter. *Geophys. Res. Lett.* 46 (2), 853–860. <http://dx.doi.org/10.1029/2018GL080644>, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL080644>.
- Zou, X., Navon, I.M., Dimet, F.X.L., 1992. Incomplete observations and control of gravity waves in variational data assimilation. *Tellus A* 44 (4), 273–296. <http://dx.doi.org/10.3402/tellusa.v44i4.14961>.