Article:

Geoscientific
Model Development

Application of the adjoint approach to optimise the initial conditions of a turbidity current with the AdjointTurbidity 1.0 model

Samuel D. Parkinson[1], **Simon W. Funke**[2], **Jon Hill**[3], **Matthew D. Piggott**[1,4], **and Peter A. Allison**[1]

[1]Department of Earth Science and Engineering, Imperial College London, London, UK
[2]Scientific Computing, Simula Research Laboratory, Fornebu, Norway
[3]Environment Department, University of York, York, UK
[4]Grantham Institute – Climate Change and the Environment, Imperial College London, London, UK

Correspondence to: Simon W. Funke (simon@simula.no)

Abstract. Turbidity currents are one of the main drivers of sediment transport from the continental shelf to the deep ocean. The resulting sediment deposits can reach hundreds of kilometres into the ocean. Computer models that simulate turbidity currents and the resulting sediment deposit can help us to understand their general behaviour. However, in order to recreate real-world scenarios, the challenge is to find the turbidity current parameters that reproduce the observations of sediment deposits.

This paper demonstrates a solution to the inverse sediment transportation problem: for a known sedimentary deposit, the developed model reconstructs details about the turbidity current that produced the deposit. The reconstruction is constrained here by a shallow water sediment-laden density current model, which is discretised by the finite-element method and an adaptive time-stepping scheme. The model is differentiated using the adjoint approach, and an efficient gradient-based optimisation method is applied to identify the turbidity parameters which minimise the misfit between the modelled and the observed field sediment deposits. The capabilities of this approach are demonstrated using measurements taken in the Miocene Marnoso-arenacea Formation (Italy). We find that whilst the model cannot match the deposit exactly due to limitations in the physical processes simulated, it provides valuable insights into the depositional processes and represents a significant advance in our toolset for interpreting turbidity current deposits.

1 Introduction

Turbidity currents are density currents driven by sediment particles that are suspended by turbulence in the containing fluid (Lowe, 1979). They occur frequently throughout the Earth's oceans and are one of the main processes by which sediment is moved from the continental shelf to the deep ocean. The largest turbidity currents can involve several hundred cubic kilometres of sediment (Talling et al., 2007c) and can travel for hundreds of kilometres across the seabed at speeds of tens of metres per second (Heezen and Ewing, 1952).

The vast majority of the available data on turbidity currents are contained in the sedimentary deposits that they leave behind. Significant effort is spent on attempting to diagnose details about the turbidity current that produced these deposits. Talling et al. (2007a, 2012) describe the current theories on how deposits found in the field are formed. The experimental evidence cannot yet validate all of these theories. Computer models, along with laboratory experiments, have been useful tools in improving our understanding of the dynamics of turbidity currents (Talling et al., 2007a; Kneller and Buckee, 2000; Parkinson et al., 2014). However, computer models have not often been directly applied to recreating deposits found in the field, despite their capacity to do so. They are generally applied to idealised cases to understand a specific physical mechanism. It is useful to directly apply models in an attempt to recreate real-world deposits (Fukushima et al., 1985; Huang et al., 2009; Doyle et al., 2010), but this requires good knowledge of the initial
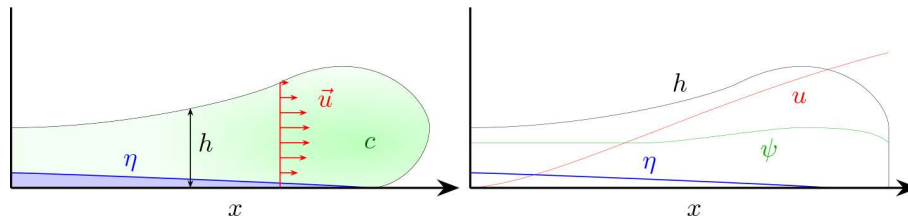
Figure 1. This schematic representation of a dense gravity current (left) and a corresponding depth-averaged shallow water approximation (right) shows current height h, volume fraction c, depth-averaged volume fraction ψ, velocity $\boldsymbol{u}$, the forward component of the depth-averaged velocity u, and deposit depth η.

and boundary conditions and accurate estimates of values for other controlling model parameters, which are often hard to determine (Talling et al., 2007a).

The task of obtaining a set of model input parameters based upon a desired model output represents an inverse problem. It can also be interpreted as an optimisation problem for which model parameters are sought to minimise the misfit between the deposit profile generated by the model and a target deposit profile, which is produced from measurements taken in the field.

In this paper, a shallow water model is used to simulate turbidity currents. The shallow water equations are a set of partial differential equations (PDEs). The optimisation of PDE-based models occurs throughout science and engineering and is already applied, for instance, in ocean science (Menemenlis et al., 2008), renewable energy (Funke et al., 2014), and design problems (Giles and Pierce, 2000). In addition, there is a growing interest in applying inverse modelling techniques to the modelling of turbidity currents (Lesshafft et al., 2011; Naruse, 2013; Rowley, 2013). In particular, Lesshafft et al. (2011) applied a gradient-free optimisation method to reconstruct parameters for a turbidity model.

PDE models of turbidity currents require the definition of initial and boundary conditions. In the simplest case, this could involve the definition of a static lock-release laboratory configuration with uniform sediment depth and a single, uniform sediment grain size. Such a simple configuration would at least require the definition of the initial depth of the current, the concentration of sediment in the fluid, the ratio of initial depth to length, and the parameters controlling the particle settling velocity and flow front speed. More realistic initial conditions would be an inflow condition with time-varying depth, velocity, and concentrations of a wide range of sediment grain sizes along with information defining the topography of the bed, its composition, the parameterisations governing bed erosion rates, flow rheology, and bed-load transport. As the model complexity and the choices of boundary and initial conditions increase, the range of deposit shapes that can be generated by the model also increases such that the model is capable of better recreating a range of deposits found in the field. However, with this added complex-

ity, the parameter space grows and the manual tuning of the parameters becomes a greater challenge.

This paper presents a shallow water sediment-laden density current model, released under the name AdjointTurbidity 1.0, that uses a novel finite-element mixed discontinuous Galerkin function space with adaptive time stepping (Sect. 2). The model implementation is verified through comparison with analytical solutions and convergence analyses (Sect. 2.5). The model is then differentiated using the adjoint method, which is an efficient way of computing the sensitivity of a model output to many input parameters (Sect. 3). This enables the use of fast-converging gradient-based optimisation techniques. Finally, a gradient-based optimisation technique is applied to minimise the data misfit between the modelled sediment deposit and field measurements taken in the Miocene Marnoso-arenacea Formation (Sect. 4). To the best of the authors' knowledge, this paper represents the first published work in which adjoint-based optimisation is applied to turbidity currents and demonstrates the usefulness of these techniques for interpreting sedimentary successions that have been deposited by turbidity currents.

2   Model

Shallow water models solve the Navier–Stokes equations in depth-averaged form (Fig. 1). They are a valid approximation when the horizontal length scale, or the length of the current, is much larger than the vertical length scale, or the height of the current. This is the case for all sediment-laden density currents a short period after an initial release. In this case, the vertical pressure gradients are in near hydrostatic balance. Sediment in the current is assumed to be well mixed by the turbulence in the flow such that there is a vertically uniform sediment distribution.

Shallow water sediment-laden density current models come in a variety of forms. Parker et al. (1986) proposed the "four-equation" model. This is a complex model which accounts for entrainment of sediment from the bed and entrainment of ambient fluid into the flow. It has an extra equation for the internal kinetic energy of the flow, which is translated into potential energy through these mixing processes. A drag force is applied along the length of the current which takes

into account the viscous forces impeding the flow motion at the base and top of the flow. This model has been applied to large-scale turbidity currents (Fukushima et al., 1985; Huang et al., 2009). It is dependent upon the selection of numerous governing parameters and hence is a good case for inverse modelling. A similar but slightly simplified model was used by Doyle et al. (2008, 2010) for modelling the plume of a dense pyroclastic basal flow. This model also included a dense underflow that has been applied in direct comparison with field measurements (Doyle et al., 2010).

Bonnecaze et al. (1993) proposed one- and two-layer sediment-laden shallow water density current models. The two-layer model includes equations for the motion of the ambient fluid through which the density current is propagating. This is important when the ambient fluid depth is similar to the initial current depth. The one- and two-layer models presented by Bonnecaze et al. (1993) use a coordinate system that adapts relative to the length of the flow. The moving coordinate system allows the speed to be prescribed at the front of the current. This speed can be well approximated using the Froude number, the height, and the volume fraction of sediment in the current. This is a good approach, as the speed of the front of a gravity current is governed by dynamics that cannot be resolved by a vertically averaged model. The moving coordinate system also results in a discretisation that scales with the horizontal length scale of the flow. This is beneficial for capturing the important flow features. This model has been used extensively in understanding turbidity current flow characteristics (Harris et al., 2001; Hogg, 2006), including the effects of modelling polydisperse suspensions (Harris et al., 2002; Garcia, 1994) and the effects of external flow (Hallworth et al., 1998). The model used in this paper is based upon the single layer shallow water model of Bonnecaze et al. (1993).

2.1 Governing equations

The equations governing the current column height, h, and the vertically integrated momentum, $q = uh$, with u being the depth-averaged current velocity, are described in non-dimensionalised conservative form as

$$\frac{\partial h}{\partial t} + \frac{\partial q}{\partial x} = 0, \tag{1}$$

$$\frac{\partial q}{\partial t} + \frac{\partial}{\partial x}\left(\frac{q^2}{h} + \frac{\varphi h}{2}\right) = 0 \tag{2}$$

with boundary conditions

$$q = 0 \text{ at } x = 0, \tag{3}$$

$$q = \dot{x}_N h \text{ at } x = x_N(t), \tag{4}$$

given that

$$\dot{x}_N = Fr\varphi^{1/2} \text{ at } x = x_N(t), \tag{5}$$

where x_N is the location of the front of the current, $\dot{x}_N$ is the velocity of the front of the current, Fr is the Froude

number, and $\varphi = \psi h$ is the vertically integrated volume fraction of sediment where ψ is the depth-averaged volume fraction of sediment within the flow. Through experimentation, the Froude number for a density current with a head height < 0.075 of the total water depth has been found to be 1.19 (Huppert and Simpson, 1980). The evolution of φ is described using

$$\frac{\partial \varphi}{\partial t} + \frac{\partial}{\partial x}\left(\frac{q\varphi}{h}\right) = -\beta\frac{\varphi}{h}, \tag{6}$$

where β is a constant particle-settling parameter. Hence, the gravitational forcing term in (Eq. 2) (the last term on the left-hand side) changes with time as φ is advected and settles out of the column.

This single layer model ignores the effect of the motion of the overlying fluid on the current. This approximation is valid for flows in which the maximum column height is significantly less than the depth of the ambient fluid (Bonnecaze et al., 1993; Hogg, 2006). Viscous forces are also ignored. For high Reynolds number flows, the viscous forces will be negligible in relation to the buoyancy forces. Bonnecaze et al. (1993) found that this was valid while the Reynolds number was greater than $\mathcal{O}(1)$.

The amount of deposited sediment, η, is also recorded and is calculated using

$$\frac{\partial \eta}{\partial t} = \beta\frac{\varphi}{h}. \tag{7}$$

The model is non-dimensionalised with the length, time, and velocity scales h_0, $(h_0/g_0')^{1/2}$, and $(h_0 g_0')^{1/2}$ respectively. Here, h_0 is the dimensional depth of the initial sediment release, $g_0' = \psi_0 g (\rho_p - \rho_a)/\rho_a$ is the initial reduced gravity of the current, ρ_p is the sediment particle density, ρ_a is the ambient fluid density, which is assumed to equal the interstitial fluid density, ψ_0 is the initial volume fraction of sediment, and g is the acceleration due to gravity. Finally, the volume fraction is scaled such that $\int_0^{x_N(0)} \varphi \mathrm{d}x = 1$.

Following Bonnecaze et al. (1993), a coordinate transformation from (x, t) to (y, τ) is applied where $y = x/x_N(t)$ and $t = \tau$. This is a convenient form for the equations as the front of the current is always at the right-hand boundary of a fixed computational domain, and hence the boundary condition at the front of the flow is applied at the right-hand side of the domain. The transformed derivatives are given by

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial \tau} - \frac{y\dot{x}_N}{x_N}\frac{\partial}{\partial y}, \tag{8}$$

$$\frac{\partial}{\partial x} = \frac{1}{x_N}\frac{\partial}{\partial y}. \tag{9}$$

Applying this coordinate transformation, but keeping t in place of τ for notational simplicity following Bonnecaze et al. (1993), produces the system of equations

$$\frac{\partial h}{\partial t} = \frac{1}{x_N}\left(y\dot{x}_N\frac{\partial h}{\partial y} - \frac{\partial q}{\partial y}\right), \tag{10}$$

$$\frac{\partial q}{\partial t} = \frac{1}{x_N}\left(y\dot{x}_N\frac{\partial q}{\partial y} - \frac{\partial}{\partial y}\left(\frac{q^2}{h} + \frac{\varphi h}{2}\right)\right), \tag{11}$$

$$\frac{\partial \varphi}{\partial t} = \frac{1}{x_N}\left(y\dot{x}_N\frac{\partial \varphi}{\partial y} - \frac{\partial}{\partial y}\left(\frac{q\varphi}{h}\right)\right) - \beta\frac{\varphi}{h}, \tag{12}$$

$$\frac{\partial \eta}{\partial t} = \frac{1}{x_N}\left(y\dot{x}_N\frac{\partial \eta}{\partial y}\right) + \beta\frac{\varphi}{h}, \tag{13}$$

$$\frac{\partial x_N}{\partial t} = \dot{x}_N \tag{14}$$

with boundary conditions

$$q = 0 \text{ at } y = 0, \tag{15}$$

$$q = \dot{x}_N h \text{ at } y = 1, \tag{16}$$

$$\eta = 0 \text{ at } y = 1, \tag{17}$$

given that $\dot{x}_N = Fr\varphi^{1/2}$ at $y = 1$.

Note that Eq. (14) for x_N has been introduced to close the system.

It is now shown that the boundary conditions (see Eqs. 15–17) are sufficient to uniquely solve this system. Equations (10)–(13) are a hyperbolic system of PDEs. For such a system to be well posed, there must be a boundary condition for each inwardly propagating characteristic. This system of equations has four characteristic velocities:

$$\frac{dy}{dt} = c_{\pm} := \frac{1}{x_N}\left(u - y\dot{x}_N \pm \varphi^{1/2}\right), \tag{18}$$

$$\frac{dy}{dt} = c := \frac{1}{x_N}\left(u - y\dot{x}_N\right), \tag{19}$$

$$\frac{dy}{dt} = c_\eta := -\frac{1}{x_N}y\dot{x}_N. \tag{20}$$

These are obtained using the method of characteristics, where $c_{\pm}$ is the characteristic velocity of waves in shallow water, c is the advection velocity of sediment, and c_η is the advection velocity of deposited sediment which is advected away from the current head as the domain length increases.

Due to the boundary conditions on momentum, the following is true: $u = q/h = 0$ at $y = 0$ and $u = q/h = \dot{x}_N$ at $y = 1$. Hence, $c = 0$ at both $y = 0$ and $y = 1$. Therefore, there are three inwardly propagating characteristics: $c_+ = \varphi^{1/2}/x_N$ at $y = 0$, $c_- = -\varphi^{1/2}/x_N$ at $y = 1$, and $c_\eta = -y\dot{x}_N/x_N$ at $y = 1$. Hence, three boundary conditions are required for the problem to be well posed such that the three boundary conditions (Eqs. 15–17) are exactly what is required.

2.2 Discretisation and numerical method

As the cell size grows throughout the simulation, it is possible to use a much larger time step at the end of the simulation than at the start of the simulation. To exploit this property, an adaptive time-stepping scheme is used in this model. A new time-dependent variable is introduced, Δt, which will vary according to a CFL criteria, C, based upon a velocity scale, $\dot{x}_N$, and the mesh element size such that

$$\Delta t = C\frac{\Delta_x}{\dot{x}_N} = C\frac{x_N\Delta_y}{\dot{x}_N}, \tag{21}$$

where Δ_x is the mesh element size in x, and Δ_y is the mesh element size in the transformed coordinate system y. The time-dependent model variables are defined as a vector

$$U = \left[h, q, \varphi, \eta, x_N, \dot{x}_N, \Delta t\right]^T. \tag{22}$$

The system is discretised in time using a second-order explicit Runge–Kutta time discretisation (Cockburn and Shu, 2001). An implicit term is added to the semi-discrete system in order to solve for the diagnostic variables $\dot{x}_N$ and Δt. With U^n as the solution at the beginning of the time step, U^{n+1} as the solution at the end of the time step, and $U^{(0)}$, $U^{(1)}$, and $U^{(2)}$ as intermediate values, the system of equations (discretised in time) can be written as

$$\begin{aligned}
U^{(0)} &= U^n, \\
U^{(1)} &= A(U^{(0)}) + L(U^{(0)}) + K(U^{(1)}), \\
U^{(2)} &= A(U^{(1)}) + L(U^{(1)}) + K(U^{(2)}), \\
U^{n+1} &= \frac{1}{2}U^{(0)} + \frac{1}{2}U^{(2)},
\end{aligned} \tag{23}$$

where

$$A(U) = \left[h, q, \varphi, \eta, x_N, 0, 0\right]^T, \tag{24}$$

$$L(U) = \tag{25}$$

$$\Delta t\left(\frac{1}{x_N}\left(y\dot{x}_N\frac{\partial f_1(U)}{\partial x} - \frac{\partial f_2(U)}{\partial x}\right) + f_3(U)\right),$$

$$f_1 = \left[h, q, \varphi, \eta, 0, 0, 0\right]^T, \tag{26}$$

$$f_2 = \left[q, \frac{q^2}{h} + \frac{\varphi h}{2}, \frac{q\varphi}{h}, 0, 0, 0, 0\right]^T, \tag{27}$$

$$f_3 = \left[0, 0, -\beta\frac{\varphi}{h}, \beta\frac{\varphi}{h}, 0, 0, 0\right]^T. \tag{28}$$

$A(U)$ is non-zero where there is a time derivative term. $L(U)$ is the explicit right-hand side term multiplied by Δt. Note that $K(U)$ contains the implicit right-hand side terms. $K(U)$ can only be easily described well in weak form, so this is defined later.

The spatially weak form of the semi-discrete system (Eq. 23) is obtained as Ψ and integrated over the domain Ω.

This gives Ω. For all Ψ in an appropriately chosen test space,

$$
\int_\Omega \Psi \cdot U^{(0)} \mathrm{d}\Omega = \int_\Omega \Psi \cdot U^n \mathrm{d}\Omega,
$$

$$
\int_\Omega \Psi \cdot U^{(1)} \mathrm{d}\Omega = \int_\Omega \Psi \cdot A(U^{(0)}) \, \mathrm{d}\Omega \tag{29}
$$

$$
+ \int_\Omega \Psi \cdot L(U^{(0)}) \, \mathrm{d}\Omega + \int_\Omega \Psi \cdot K(U^{(1)}) \, \mathrm{d}\Omega,
$$

$$
\int_\Omega \Psi \cdot U^{(2)} \mathrm{d}\Omega = \int_\Omega \Psi \cdot A(U^{(1)}) \mathrm{d}\Omega
$$

$$
+ \int_\Omega \Psi \cdot L(U^{(1)}) \, \mathrm{d}\Omega + \int_\Omega \Psi \cdot K(U^{(2)}) \, \mathrm{d}\Omega,
$$

$$
\int_\Omega \Psi \cdot U^{n+1} \mathrm{d}\Omega = \frac{1}{2} \int_\Omega \Psi \cdot \left(U^n + U^{(2)} \right) \mathrm{d}\Omega.
$$

Piecewise linear discontinuous Galerkin (DG) elements are used to discretise the spatially varying state variables. Thus, the spatial and temporal discretisations both have second-order accuracy. DG element types are known to be particularly suitable for advection dominated problems (Peraire and Persson, 2008). They are good at preserving discontinuities as they produce stable discretisations without the need for diffusive stabilisation strategies, such as streamline upwinding (Peraire and Persson, 2008). These are important features in shallow water particle-laden density current models.

In order to construct a DG formulation, a regular partition $\mathcal{T}_h = \{e\}$ of Ω into non-overlapping subdomains $\Omega_e \in \Omega$ with boundaries $\partial\Omega_e$ is considered. The piecewise linear DG function space is denoted $\mathbf{DG}_1$. For this function space, piecewise linear test functions with no global continuity requirement are considered; i.e. functions that have the potential to be double valued on $\partial\Omega_e$. x_N, $\dot{x}_N$, and Δt are defined on a function space, $\mathbf{R}$, which is constant throughout the spatial domain. Therefore, the vector of model unknowns U is defined on a mixed function space, $\mathcal{V}_h = \mathbf{DG}_1{}^4 \times \mathbf{R}^3$. The function in the mixed function space is denoted with $\Psi_h \in \mathcal{V}_h$ and the discretised approximation of the state variable with $U_h \in \mathcal{V}_h$.

Notice that $L(U)$ contains derivatives of discontinuous functions. Its undiscretised weak form is

$$
\int_\Omega \Psi \cdot L(U) \mathrm{d}\Omega = \tag{30}
$$

$$
\Delta t \int_\Omega \Psi \cdot \left(\frac{1}{x_N} \left(y \dot{x}_N \frac{\partial f_1(U)}{\partial x} - \frac{\partial f_2(U)}{\partial x} \right) + f_3(U) \right) \mathrm{d}\Omega.
$$

The discretised DG formulation of Eq. (30) is then

$$
\sum_{e \in \mathcal{T}_h} \int_{\Omega_e} \Psi_h \cdot L(U_h) \mathrm{d}\Omega = \Delta t \sum_{e \in \mathcal{T}_h} \int_{\Omega_e} \Psi_h \tag{31}
$$

$$
\cdot \left(\frac{1}{x_N} \left(y \dot{x}_N \frac{\partial f_1(U_h)}{\partial x} - \frac{\partial f_2(U_h)}{\partial x} \right) + \Psi_h f_3(U_h) \right) \mathrm{d}\Omega.
$$

Integrating the gradient terms by parts and slightly rearranging yields

$$
\sum_{e \in \mathcal{T}_h} \int_{\Omega_e} \Psi_h \cdot \frac{x_N}{\Delta t} L(U_h) \mathrm{d}\Omega =
$$

$$
- \sum_{e \in \mathcal{T}_h} \int_{\Omega_e} \frac{\partial}{\partial x} (\Psi_h y \dot{x}_N) \cdot f_1(U_h) \, \mathrm{d}\Omega \tag{32}
$$

$$
+ \sum_{e \in \mathcal{T}_h} \int_{\partial\Omega_e} \widehat{\Psi_h} \cdot y \, \dot{x}_N \, \widehat{f_1}(U_h) \widehat{n} \mathrm{d}\sigma + \sum_{e \in \mathcal{T}_h} \int_{\Omega_e} \frac{\partial \Psi_h}{\partial x} \cdot f_2(U_h) \, \mathrm{d}\Omega
$$

$$
- \sum_{e \in \mathcal{T}_h} \int_{\partial\Omega_e} \widehat{\Psi_h} \cdot \widehat{f_2}(U_h) \widehat{n} \, \mathrm{d}\sigma + \sum_{e \in \mathcal{T}_h} \int_{\Omega_e} \Psi_h \cdot x_N f_3(U_h) \, \mathrm{d}\Omega,
$$

where $\widehat{\cdot}$ indicates that the function is double valued and special attention is required. The various summations can now be rewritten as integrals over the entire domain Ω, all element interfaces Σ_h, and the domain boundaries Γ_h. Note that

$$
\sum_{e \in \mathcal{T}_h} \int_{\partial\Omega_e} \Psi_h \cdot \widehat{U_h} \, \mathrm{d}\sigma \equiv \int_{\Sigma_h} \Psi_h \cdot \widehat{U_h} \, \mathrm{d}\sigma + \int_{\Gamma_h} \Psi_h \cdot U_0 \, \mathrm{d}\sigma, \tag{33}
$$

and

$$
\sum_{e \in \mathcal{T}_h} \int_{\Omega_e} \Psi_h \cdot U_h \, \mathrm{d}\Omega \equiv \int_\Omega \Psi_h \cdot U_h \, \mathrm{d}\Omega. \tag{34}
$$

Additionally, note that within domain boundary integrals, the $\widehat{\cdot}$ notation is dropped as the function is single valued at this location. U_h is also replaced with U_0, which is either the boundary value if a Dirichlet boundary condition is present or the function value at the boundary if it is not. Note that in the case of the boundary condition for q at $y = 1$, U_0 is still a function of U_h. Applying Eqs. (33) and (34) to (32) yields

$$
\int_\Omega \Psi_h \cdot \frac{x_N}{\Delta t} L(U_h) \mathrm{d}\Omega = - \int_\Omega \frac{\partial}{\partial x} (\Psi_h y \dot{x}_N) \cdot f_1(U_h) \, \mathrm{d}\Omega
$$

$$
+ \int_{\Sigma_h} \widehat{\Psi_h} \cdot y \dot{x}_N \, \widehat{f_1}(U_h) \widehat{n} \mathrm{d}\sigma + \int_{\Gamma_h} \Psi_h \cdot y \, \dot{x}_N \, f_1(U_0) n \, \mathrm{d}\sigma \tag{35}
$$

$$
+ \int_\Omega \frac{\partial \Psi_h}{\partial x} \cdot f_2(U_h) \, \mathrm{d}\Omega - \int_{\Sigma_h} \widehat{\Psi_h} \cdot \widehat{f_2}(U_h) \widehat{n} \, \mathrm{d}\sigma
$$

$$
- \int_{\Gamma_h} \Psi_h \cdot f_2(U_0) \, n \, \mathrm{d}\sigma + \int_\Omega \Psi_h \cdot x_N f_3(U_h) \, \mathrm{d}\Omega. \tag{36}
$$

A choice of flux term must be made to handle the double-valued terms. This will involve some coupling between the

elements on either side of the interface. An upwind flux is used for the advection term, $\widehat{f_1}$, and based upon experience an average flux works well for $\widehat{f_2}$. This gives

$$
\int_\Omega \Psi_h \cdot \frac{x_N}{\Delta t} L(U_h) \mathrm{d}\Omega = -\int_\Omega \frac{\partial}{\partial x}\left(\Psi_h y \dot{x}_N\right) \cdot U_h \mathrm{d}\Omega \tag{37}
$$

$$
+ \int_{\Gamma_h} \Psi_h \cdot y\,(\dot{x}_N n)_{\text{down}} f_1(U_0) \mathrm{d}\sigma + \int_{\Sigma_h}\left(\Psi_h^+ - \Psi_h^-\right) \cdot
$$

$$
y\left(f_1^+(U_h)(\dot{x}_N\,n^+)_{\text{up}} + f_1^-(U_h)(\dot{x}_N n^-)_{\text{up}}\right) \mathrm{d}\sigma
$$

$$
+ \int_\Omega \frac{\partial \Psi_h}{\partial x} \cdot f_2(U_h) \mathrm{d}\Omega - \int_{\Gamma_h} \Psi_h \cdot f_2(U_0)\, n\, \mathrm{d}\sigma
$$

$$
- \int_{\Sigma_h}\left(\Psi_h^+ - \Psi_h^-\right) \cdot \frac{1}{2}\left(f_2(U_h)^+ + f_2(U_h)^-\right) n^+ \mathrm{d}\sigma
$$

$$
+ \int_\Omega \Psi_h \cdot x_N f_3(U_h) \mathrm{d}\Omega,
$$

where $(\cdot)^+$ and $(\cdot)^-$ indicate the function values on either side of an interior element boundary. $(\cdot)_{\text{up}}$ is equal to $(\cdot)$ where $\dot{x}_N n^\pm > 0$ and 0 otherwise. Conversely, $(\cdot)_{\text{down}}$ is equal to $(\cdot)$ where $\dot{x}_N n^\pm < 0$ and 0 otherwise. $K(U)$ can be described in weak discretised form as

$$
\int_\Omega \Psi_h \cdot K(U_h)\mathrm{d}\Omega = \tag{38}
$$

$$
\int_\Omega \Psi_h \cdot K_\Omega(U_h)\mathrm{d}\Omega + \int_{\partial\Omega_{\mathrm{R}}} \Psi_h \cdot K_\sigma(U_h)\mathrm{d}\sigma,
$$

$$
K_\Omega(U_h) = \left[0,0,0,0,0,0, C\frac{x_N \Delta_y}{\dot{x}_N}\right], \tag{39}
$$

$$
K_\sigma(U_h) = \left[0,0,0,0,0, Fr(\varphi)^{1/2}, 0\right], \tag{40}
$$

where $\partial\Omega_{\mathrm{R}}$ is the right-hand boundary at $y = 1$ such that a solution for $\dot{x}_N$ is obtained by solving only at the front of the current.

Using Eqs. (29), (37), and (38) and applying Eq. (34), the full weak, discontinuous form of Eq. (23) can be obtained.

Find $U_h^{(0)}, U_h^{(1)}, U_h^{(2)}, U_h^{(n+1)} \in \mathcal{V}_h$ such that $\forall \Psi_h \in \mathcal{V}_h$

$$
\int_\Omega \Psi_h \cdot U_h^{(0)}\mathrm{d}\Omega = \int_\Omega \Psi_h \cdot U_h^n \mathrm{d}\Omega,
$$

$$
\int_\Omega \Psi_h \cdot U_h^{(1)}\mathrm{d}\Omega = \int_\Omega \Psi_h \cdot A(U_h^{(0)})\mathrm{d}\Omega + \int_\Omega \Psi_h \cdot L(U_h^{(0)})\mathrm{d}\Omega
$$

$$
+ \int_\Omega \Psi_h \cdot K(U_h^{(1)})\mathrm{d}\Omega, \tag{41}
$$

$$
\int_\Omega \Psi_h \cdot U_h^{(2)}\mathrm{d}\Omega = \int_\Omega \Psi_h \cdot A(U_h^{(1)})\mathrm{d}\Omega + \int_\Omega \Psi_h \cdot L(U_h^{(1)})\mathrm{d}\Omega
$$

$$
+ \int_\Omega \Psi_h \cdot K(U_h^{(2)})\mathrm{d}\Omega,
$$

$$
\int_\Omega \Psi_h \cdot U_h^{n+1}\mathrm{d}\Omega = \frac{1}{2}\int_\Omega \Psi_h \cdot \left(U_h^{(0)} + U_h^{(2)}\right)\mathrm{d}\Omega.
$$

This set of equations is solved for each time step of the simulation as a nonlinear variational problem using Newton's method with an LU decomposition solver for the linear problems.

2.3 Slope limiting

Discontinuous Galerkin discretisations for convection dominated problems can suffer from over- and undershoots at discontinuities that can cause instability problems (Kuzmin, 2010; Cockburn and Shu, 2001). Slope limiting can be applied to solve this problem, but this typically involves discontinuous operations, which are problematic in a gradient-based optimisation framework. Therefore, we do not use slope limiting here and limit ourselves to the assumption of smooth initial conditions where slope limiting is not necessary. It would be possible to formulate a continuous slope-limiting function to overcome this limitation if it was required.

2.4 Implementation

The shallow water sediment-laden density current model described above was built using the FEniCS framework (Logg et al., 2012), an open-source software project that provides features for the automated, efficient solution of differential equations. Using a high-level interface, the model partial differential equations are described in variational form using UFL (Unified Form Language) (Alnæs et al., 2012). This can be achieved in Python or C++ code in a way that is remarkably similar to how one would describe the equations on paper. At runtime, this model description is compiled into efficient C++ kernels that handle the assembly of the required matrices to generate the systems of equations that are then solved using PETSc (Balay et al., 2014).
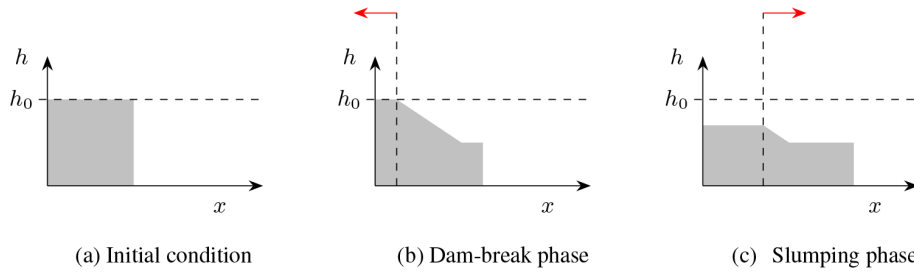
(a) Initial condition   (b) Dam-break phase   (c) Slumping phase

Figure 2. Schematic diagram of the lock-release static initial condition **(a)** and the following dam-break **(b)** and slumping **(c)** phases with the shock wave propagation direction indicated by ($\longrightarrow$).

2.5   Forward model verification

Many laboratory experiments and computer models are based around the classical lock-release static initial condition (Fig. 2a). Following the release of the lock-gate, the current accelerates forwards. This is known as the dam-break stage (Ungarish, 2010) (Fig. 2b). As the lock-gate is released, a shock forms which travels in the opposite direction to the front of the current. This shock carries information that sets the fluid in motion. Once this shock reaches the rear wall, all of the fluid behind the lock-gate is in motion. This marks the point of transfer from the dam-break to the slumping phase (Ungarish, 2010) (Fig. 2c). For a non-depositional current (i.e. $\beta = 0$) with initial h and $x_N = 1$, the slumping phase begins at $t = 1$. The current front height and velocity remain approximately constant during this phase of motion. The rear propagating shock is reflected off the no-flow boundary and travels faster than the front of the current. A short while later, it reaches the front of the current, marking the end of the slumping phase. The current is now able to "forget" the initial condition and begins adjusting to self-similar propagation (Ungarish, 2010). For a non-depositional current (i.e. $\beta = 0$), the reflected shock reaches the front of the current at $t = 3$ (Ungarish, 2010).

Hoult (1972) showed that a similarity solution, which is a solution that looks the same at all times or at all length scales, could be obtained for a single layer shallow water density current model during the self-similar phase of propagation. This is described as

$$x_N = \kappa t^{2/3}, \ u = \frac{2}{3}\kappa t^{-1/3}u_s, \ h = \kappa^{-1}t^{-2/3}h_s, \ \psi = 1, \ \ (42)$$

where

$$y = x/\kappa t^{2/3}, \ \kappa = \left(\frac{27Fr^2}{12 - 2Fr^2}\right)^{1/3}, \tag{43}$$

$$u_s = y, \ h_s = \frac{4}{9}\kappa^3\left(\frac{y^2}{4} - \frac{1}{4} + \frac{1}{Fr^2}\right). \tag{44}$$

The domain is unit length, as in all cases for this model. This solution is valid for the model described in this paper so long as the settling velocity of particles, β, is equal to
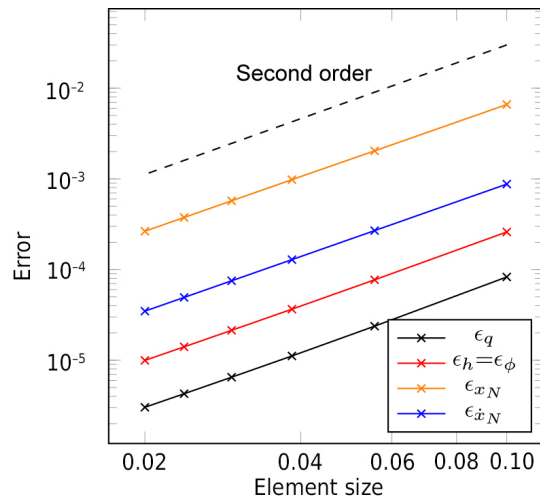

Figure 3. Similarity convergence analysis. All variables are shown to converge on the correct solution at the correct order. $\epsilon_{(*)}$ indicates the L_2 norm of the error in the solution obtained for variable ($*$).

0 (i.e. no particle settling). This analytical solution is useful in verifying the implementation of the governing equations and boundary conditions for this model. The solution to the model PDEs should converge on this analytical solution as the mesh resolution is refined at the correct rate for the temporal and spatial discretisation. The use of piecewise discontinuous linear elements and a second-order time-stepping regime means that the convergence order should be quadratic in both space and time.

For the convergence test, the analytical solution is projected onto the model function space forming the initial condition at $t = 3$. At $t = 10$, the L_2 norm of the difference between the model variables and the analytical solution is obtained and used to measure convergence. The analysis shows that all variables converge on the analytical solution at the correct order (Fig. 3). Note that the time step is adaptive and will therefore decrease along with the element size such that this test checks both spatial and temporal convergence. This verifies that the model equations are implemented correctly (Farrell et al., 2011). A qualitative comparison shows that the solution matches the analytical solution very well (Fig. 4).
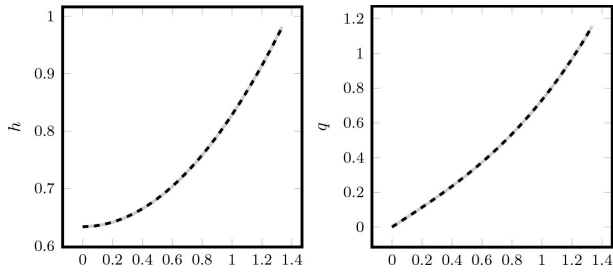
Figure 4. Similarity results for the finest resolution mesh (solid lines) compared against the analytical results (dashed lines) at $t = 10.0$.

3 The adjoint model

Here, we describe the adjoint model and its derivation generally rather than specifically applying it to this model.

Consider a problem with N input parameters forming a vector, m. Let $F(U, m) = 0$ denote the set of PDEs that describe a model where U represents the model variables throughout time. Note that U can be seen as an implicit function of m, $U = U(m)$, by finding a solution to $F(U, m) = 0$. Suppose now that the aim is to minimise an objective functional, $J(U, m)$, by optimising m. Here, $J(U, m)$ will be defined as a function measuring the difference between a deposit profile generated by the model and a target deposit profile. Where optimisation is required to a tolerance of δ_{mi}, with i being the index of each parameter and where each parameter has bounds spanning a range Δ_{mi}, optimisation through a brute force approach will require $\prod_i^N \Delta_{mi}/\delta_{mi}$ evaluations of the model to find the solution. N may be very large for a sediment-laden density current model, which could potentially have time-varying boundary conditions for sediment concentration, velocity and height, uncertainty in the elevation profile, the friction coefficient of the surface over which the current is flowing, and uncertainty in the parameters that govern the physics of the flow, such as entrainment of ambient fluid, front speed, and sediment erosion. Many of these parameters vary over space and time such that the parameter space grows as the resolution in time or space increases. Such a large potential parameter space motivates the use of a more advanced and efficient optimisation strategy.

Numerous algorithms have been developed to improve this brute force approach. These optimisation algorithms begin with an initial guess of the input parameters and iterate, then generate improved estimates until they terminate, hopefully at the optimised solution. The authors refer the reader to Jorge and Stephen (1999) for an extensive description of the range of numerical optimisation methods.

Most of these optimisation algorithms require the gradient of the objective functional with respect to the input parameters, dJ/dm. Approximation techniques, such as finite differencing, could be used to evaluate the gradient, but this will require an excessive number of PDE evaluations and may suffer from noise (Jorge and Stephen, 1999). Here, the adjoint model is used to efficiently calculate the gradient. This approach is favoured as it calculates dJ/dm for any number of input parameters with a single evaluation of the adjoint model.

Obtaining the adjoint model begins by applying the chain rule to dJ/dm:

$$\frac{dJ(U(m), m)}{dm} = \left\langle \frac{\partial J}{\partial U}, \frac{dU}{dm} \right\rangle + \frac{\partial J}{\partial m}. \quad (45)$$

$\partial J/\partial U$ and $\partial J/\partial m$ are both vectors, and they are typically straightforward to compute as J is typically a given analytical function of U and m. dU/dm, on the other hand, is a matrix that is typically dense and expensive to compute. A relationship for dU/dm can be obtained by taking the total derivative of $F(U, m) = 0$ with respect to m:

$$0 = \frac{dF(U(m), m)}{dm} = \frac{\partial F}{\partial U}\frac{dU}{dm} + \frac{\partial F}{\partial m}, \quad (46)$$

$$\Rightarrow \frac{\partial F}{\partial U}\frac{dU}{dm} = -\frac{\partial F}{\partial m}. \quad (47)$$

Equation 47 is termed the *tangent linear equation*. $\partial F/\partial U$ and $\partial F/\partial m$ are both matrices. The solution of this equation is obtained by solving N systems of equations. When there are many functionals, J, and a small set of parameters, m, then this equation can be useful for obtaining dJ/dm via Eq. (45). With a large set of parameters and only one functional, as is the case here, this is not an efficient approach.

However, suppose that $\partial F/\partial U$ in Eq. (47) is invertible so that one can obtain

$$\frac{dU}{dm} = -\left(\frac{\partial F}{\partial U}\right)^{-1}\frac{\partial F}{\partial m}. \quad (48)$$

This expression can be substituted for dU/dm directly into Eq. (45) to obtain

$$\frac{dJ(m)}{dm} = -\left\langle \frac{\partial J}{\partial U}, \left(\frac{\partial F}{\partial U}\right)^{-1}\frac{\partial F}{\partial m} \right\rangle + \frac{\partial J}{\partial m}. \quad (49)$$

A simple property of inner products, $\langle y, Ax \rangle = \langle A^*y, x \rangle$, where A^* is the conjugate transpose or adjoint of A, can be used to shift $(\partial F/\partial U)^{-1}$ to the left-hand side of the inner product:

$$\frac{dJ(m)}{dm} = -\left\langle \left(\frac{\partial F}{\partial U}\right)^{-*}\frac{\partial J}{\partial U}, \frac{\partial F}{\partial m} \right\rangle + \frac{\partial J}{\partial m}. \quad (50)$$

Gathering the left-hand side of the inner product into a new variable,

$$\lambda := \left(\frac{\partial F}{\partial U}\right)^{-*}\frac{\partial J}{\partial U}, \quad (51)$$

yields the linear system of equations that can be solved for the adjoint variable, λ:

$$\left(\frac{\partial F}{\partial U}\right)^* \lambda = \frac{\partial J}{\partial U}. \tag{52}$$

Equation (52) is termed the *adjoint equation*. The right-hand side is a vector and only one evaluation is required to obtain λ for a specific functional, J. Once Eq. (52) is solved, dJ/dm can easily be computed with respect to any parameter m by substituting the value of λ into Eq. (49).

As mentioned above, $\partial J/\partial U$ and $\partial J/\partial m$ are typically straightforward to compute. However, $(\partial F/\partial U)^*$ and $\partial F/\partial m$ still need to be derived and implemented, which is not a simple task for a large set of complex PDEs. The challenge of obtaining these matrices is the main obstacle to using the adjoint model. However, the high-level abstraction of the coding provided by FEniCS to create this model makes calculating $(\partial F/\partial u)^*$ and $\partial F/\partial m$ an automatable task using an additional tool, dolfin-adjoint (Farrell et al., 2013). This powerful tool automatically derives the discrete adjoint and tangent linear models from a forward model written in FEniCS. This makes differentiating the forward model and solving the adjoint equation to obtain the derivative of the objective functional a much simpler task. Additionally, dolfin-adjoint contains tools for carrying out the optimisation of the model parameters by interfacing with IPOPT (Wächter and Biegler, 2006) optimisation algorithms (Funke and Farrell, 2013).

4 Estimation of the parameters for the turbidity current that generated Bed 1.1 in the Marnoso-arenacea Formation

The Marnoso-arenacea Formation spans 17 to 7 Ma (Late Burdigalian to Tortonian) and is over 3500 m thick (Talling et al., 2007b). Deposition occurred from two sources: the northwestern Alpine source and the southwestern Apennine source (Lucchi and Valmori, 1980; Gandolfi et al., 1983). The depositional environment was an elongated foreland basin adjacent to the Apennine thrust belt with turbidites deposited in a relatively wide (> 60 km) basin in a non-channelised manner (Talling et al., 2007b; Lucchi and Valmori, 1980; Gandolfi et al., 1983). The formation provides the most extensive and detailed correlation of flow deposits (beds) in any ancient turbidite system and is therefore a natural laboratory for studying turbidite depositional processes (Amy and Talling, 2006). It has been extensively mapped with more than 100 sections accurately recorded over a correlated distance of more than 120 km (Amy and Talling, 2006). Bed volumes range from $\mathcal{O}(10^{-3})$ km^3 to several km^3 (Talling et al., 2007a). It contains extensive data for evaluating the performance of the adjointed turbidity current model described here.
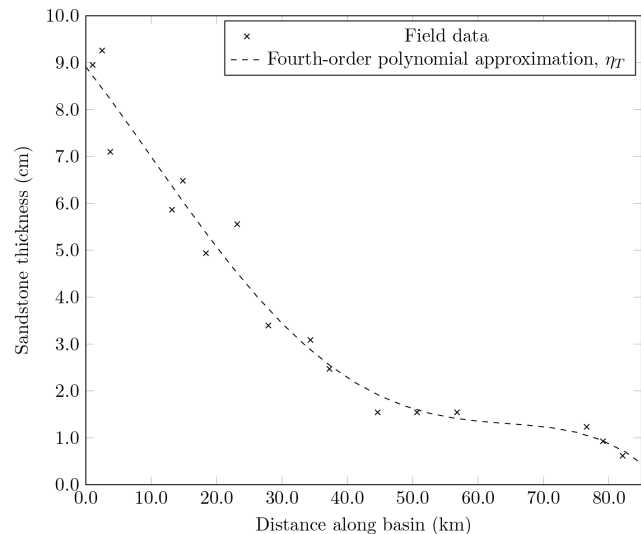

Figure 5. The sandstone depths measured for Bed 1.1 along the Pietralunga and Ridracoli structural elements orientated approximately parallel to the palaeoflow. This has been reconstructed from Fig. 5 in Talling et al. (2007a). A fourth-order polynomial approximation of the deposit profile, η_T, is also shown. This is used as a target for the optimisation algorithm. The base of the bed is shown as a horizontal datum in order to illustrate lateral changes in deposit thickness. Note that a different datum is used in the source figure, which uses the top of Bed 1.2.2 rather than the top of Bed 1. The palaeoelevation of the base of the bed would have varied spatially, reflecting the basin floor relief.

In this section, an optimisation algorithm is used to select model parameters that produce an output deposit that best matches part of Bed 1.1 in the Marnoso-arenacea Formation, as recorded by Amy and Talling (2006). This is defined as a small volume of flow deposit with a total sediment volume of ≈ 0.215 km^3 (Talling et al., 2007a). Talling et al. (2007a) produced an approximate one-dimensional deposit parallel to the palaeoflow (Fig. 5). The shape of the deposit strongly resembles that of very low concentration currents in laboratory tests, and it also resembles the shape of bed profiles generated by the Bonnecaze et al. (1993) model. This implies that the flow that created this deposit was a very low concentration current. The model used in this chapter is very simple. It does not model any stratification or particle–particle interactions in the flow. As such, its application is limited to very low concentration flows, and hence Bed 1.1 is a good candidate as a case study for this model.

The deposit consists of sandstone and mudstone components. The focus here will be on attempting to recreate only the sandstone portion of the deposit. It is likely that ponding effects have influenced the shape of the mud deposit in this bed (Talling et al., 2007b), which this model cannot replicate. The outcrop quality also deteriorates beyond the extent of the sandstone deposit. Therefore, no attempt is made to model this portion of the bed.

4.1 Choice of initial conditions and parameters

The initial conditions are based upon the analytical solution for a non-depositional flow at a non-dimensional time, $t = t_s = 3$, after a column collapse as described by Eq. (42). Some assumptions are therefore made as to the initial shape, sediment concentration profile, and velocity profile of the flow.

The non-dimensional particle-settling velocity, β, is calculated using the standard Stokes settling law for a particle in suspension (Lamb, 1993) non-dimensionalised by $(h_0 g_0')^{1/2}$ to give

$$\beta = \frac{g' D^2}{18\nu (h_0 g_0')^{1/2}} = \frac{g'^{1/2} D^2}{18\nu (h_0 \psi_0)^{1/2}}, \tag{53}$$

where D is the average sediment diameter. The sediment-reduced gravity, $g' = g_0'/\psi_0 = (\rho_p - \rho_a)g/\rho_a = 16$, is based upon the reduced gravity of silica in water. Using these initial conditions, there are three unknowns: h_0, the dimensional length scale of the current; ψ_0, the initial sediment concentration throughout the current; and D, the mean sediment diameter. These become the set of input parameters that will be optimised with $m = (h_0, \psi_0, D)^T$.

The beginning of the basin is defined as being at the front of the current at $t = t_s$ such that the current is not in the basin prior to the start of the simulation. The current enters the domain as soon as the simulation starts. The end of the simulation, t_f, is defined as the time at which the total suspended sediment is less than 1 % of the starting quantity.

4.2 Choice of optimisation functional

The aim here is to reduce the difference between the deposit profile generated by this model and the target deposit profile from field measurements. To do this, we need to map the non-dimensional, transformed results from the model back to the observation space. We also only measure the variation over the length of the measured deposit. Therefore, the functional that we will aim to minimise, J, has the form

$$J(U(m), m) = \int_0^{\hat{x}_{max}} (\widetilde{\eta} - \eta_T)^2 d\hat{x}, \tag{54}$$

where η_T is the dimensional target deposit profile, $\widetilde{\eta} = \psi_0 h_0 \eta$ is the dimensionalised modelled deposit, $\hat{x}_{max} = 82\,000$ is the extent of the measured data, and $\hat{x} = \widetilde{x} - \widetilde{x}_N(t_s)$ is a coordinate transformation such that $\hat{x} = 0$ at the front of the current at $t = t_s$, $\widetilde{x} = y\,\widetilde{x}_N(t_f)$ is the dimensionalised reverse of the coordinate transformation outlined in Sect. 2.1; $\widetilde{x}_N(t) = h_0 x_N(t)$ is the dimensional length of the current.

To calculate this functional, η_T must be a function of $\hat{x}$. The deposit is approximated using a fourth-order polynomial

$$\eta_T = \sum_{i=0}^{4} c_i \hat{x}^i, \tag{55}$$

where c_i is the ith coefficient. The coefficients are obtained using the least squares method. The fourth-order approximation fits the measured data points well (Fig. 5).

It is important to note that at the end of the simulation, $t = t_f$, the length of the current does not necessarily match the length of the deposit, or $\hat{x}_N \neq \hat{x}_{max}$, where $\hat{x}_N(t) = \widetilde{x}_N(t) - \widetilde{x}_N(t_s)$ is the dimensional length of the modelled deposit within the basin. This complicates the calculation of the above integral.

The calculation of J is split into two components, an integral over the lesser of the length of the modelled current (or the length of the measured data, J_0) and an integral over any remaining length of measured data, J_1, such that

$$J = J_0 + J_1. \tag{56}$$

The first integral takes the form

$$J_0 = \int_0^{\min(\hat{x}_N, \hat{x}_{max})} \left(\eta_T(\hat{x}) - \widetilde{\eta}(\hat{x})\right)^2 d\hat{x}. \tag{57}$$

This can be approximately transformed into the model coordinate system as

$$J_0 = \int_\Omega \left(\gamma_0 \eta_T(\hat{x}(y)) - \gamma_0 \widetilde{\eta}(\hat{x}(y))\right)^2 dy, \tag{58}$$

where γ_0 is a scaled filter. This filter is 0 in the regions $\hat{x} < 0$ and $\hat{x} > \hat{x}_{max}$. Elsewhere, the filter value is a constant such that the integral of the filter over the domain is equal to the length of the dimensional integral, $\min(\hat{x}_N, \hat{x}_{max})$. The filter defines the region of the domain over which the integral is evaluated and scales the resultant value appropriately. The filter is defined as

$$\gamma_0(\hat{x}) = \min(\hat{x}_N, \hat{x}_{max}) \frac{\exp\left(\min(\hat{x} - \hat{x}_{max}, \hat{x}, 0)\right)}{s_{\gamma_0}}, \tag{59}$$

$$s_{\gamma_0} = \int_\Omega \exp\left(\min(\hat{x} - \hat{x}_{max}, \hat{x}, 0)\right) d\hat{x}. \tag{60}$$

It is important that the functional is differentiable. Therefore, the min and max functions are replaced by smooth approximations f_{min} and f_{max} defined as

$$f_{min}(a, b) = \ln(\exp(10a) + \exp(10b))/10, \tag{61}$$

$$f_{max}(a, b) = f_{min}(-a, -b). \tag{62}$$

The second integral, J_1, takes the form

$$J_1 = \int_{\min(\hat{x}_N, \hat{x}_{max})}^{\hat{x}_{max}} \eta_T(\hat{x})^2 d\hat{x}, \tag{63}$$

Table 1. Taylor remainders $R_0 = |\widehat{J}(x_N(0) + \delta x_N)|$ and $R_1 = |\widehat{J}(x_N(0) + \delta x_N) - \widehat{J}(x_N(0)) - \mathrm{d}\widehat{J}(x_N(0))/\mathrm{d}u\delta x_N|$ for the with functional given by $J(\eta) = \int_\Omega \eta(t = t_\mathrm{f})^2 \mathrm{d}\Omega$.

δx_N	$R_0(\delta x_N)$	order	$R_1(\delta x_N)$	order
1.0	1.16×10^{-8}		2.67×10^{-9}	
0.5	5.11×10^{-9}	1.18	6.54×10^{-10}	2.03
0.25	2.39×10^{-9}	1.01	1.56×10^{-10}	2.07
0.0125	1.15×10^{-9}	1.05	3.54×10^{-11}	2.14
0.00625	5.64×10^{-10}	1.03	6.97×10^{-12}	2.34

such that J_1 integrates the target deposit volume beyond the extent of the modelled deposit. If the modelled deposit length exceeds the length of the measured data, this integral will be 0. Again, this can be approximately transformed into the model coordinate system as

$$J_1 = \int_\Omega (\gamma_1 \eta_\mathrm{T}(\widetilde{x}(y)))^2 \mathrm{d}y, \tag{64}$$

$$\widetilde{x}(y) = \min(\hat{x}_N + y(\hat{x}_\mathrm{max} - \hat{x}_N), \hat{x}_N), \tag{65}$$

where γ_1 is a scaled filter similar to γ_0 defined as

$$\gamma_1(\hat{x}) = (\hat{x}_\mathrm{max} - \min(\hat{x}_N, \hat{x}_\mathrm{max})) \tag{66}$$
$$\cdot \frac{\exp(\min(\hat{x} - \hat{x}_\mathrm{max}, 0))}{s_{\gamma_0}},$$

$$s_{\gamma_0} = \int_\Omega \exp(\min(\hat{x} - \hat{x}_\mathrm{max}, 0)) \, \mathrm{d}\hat{x}. \tag{67}$$

Again, min and max are replaced by smooth differentiable alternatives.

4.3 Verification of the gradient calculation

The gradient computation was verified using the Taylor remainder convergence test. Let $\widehat{J}(m) \equiv J(U(m), m)$, a pure function of m. The zero-order Taylor expansion states that

$$\left|\widehat{J}(m + \delta m) - \widehat{J}(m)\right| = \mathcal{O}(\|\delta m\|), \tag{68}$$

while the first-order Taylor expansion states that

$$\left|\widehat{J}(m + \delta m) - \widehat{J}(m) - \frac{\mathrm{d}\widehat{J}(m)}{\mathrm{d}m}\delta m\right| = \mathcal{O}\left(\|\delta m\|^2\right). \tag{69}$$

Even small errors in the derivative destroy the second-order convergence in Eq. (69). Therefore, testing the convergence of these expansions with the gradient calculated from the adjoint yields a strong indicator of whether the adjoint gradient computation is correct.

The above convergence was succesfully carried out for the implemented adjoint model with a number of different controls and functionals. As an example, Table 1 shows the results with m as the initial deposit length. The convergence of

the remainder term is second order with respect to varying magnitudes of δm, providing strong evidence that the adjoint model and gradient computation are implemented correctly.

4.4 Optimisation of a model with one sediment class

With confidence that the forward and backward models are working, optimisation of the input parameters, $m = [h_0, \psi_0, D]$, to minimise the objective functional J can now be performed as

$$\min_m = J(U(m), m), \tag{70}$$

with the the following bounding constraints on the input parameters:

$$10\,\mathrm{m} \leq h_0 \leq 10\,\mathrm{km}, \tag{71}$$
$$0.001\,\% \leq \psi_0 \leq 50\,\%, \tag{72}$$
$$1\,\mu\mathrm{m} \leq D \leq 1\,\mathrm{mm}. \tag{73}$$

These bounding constraints are chosen based upon very loose limits of expected values that each parameter may possibly take. The principal purpose of these bounds is to avoid invalid negative values being generated for any of the parameters.

The nonlinear optimisation library, IPOPT (Wächter and Biegler, 2006), is used to solve this problem. This library implements a primal-dual interior point algorithm which has good global and local convergence properties (Wächter and Biegler, 2005). The interface to this library is supplied by dolfin-adjoint (Funke and Farrell, 2013). The initial input parameters are set to

$$m = \begin{pmatrix} h_0 & = & 2.3\,\mathrm{km} \\ \psi_0 & = & 0.07\,\% \\ D & = & 200.0\,\mu\mathrm{m} \end{pmatrix}. \tag{74}$$

The aim is to recreate the sand deposit by modelling only the sand in the flow using a single average grain size, D. The value of ψ_0 is based upon a combined initial volumetric concentration for the sand and mud mixture of 0.5 %, with 86 % of the mixture being mud. The starting value for h_0 is based upon the area of the two-dimensional deposit profile and the value of ψ_0. The average sediment grain size is a reasonable estimate of the average grain size based upon the information provided by Talling et al. (2007a). The input parameters provided to the optimisation algorithm, $\overline{m}$, are normalised such that they are all equal to 1. Thus,

$$m_i = \overline{m}_i \, m_0, \tag{75}$$

where m_0 indicates the initial parameter values in Eq. (74), and m_i indicates the value of m after optimisation iteration i. This scaling helps the optimisation algorithm work effectively (Jorge and Stephen, 1999).
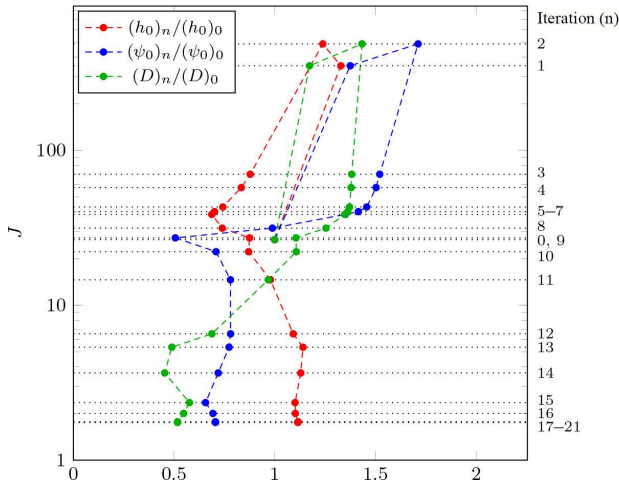
Figure 6. Values of the parameters over the optimisation iterations against the value of the objective functional, J, that we are aiming to minimise. The values shown are normalised by their starting values. $(*)_n$ is the value of parameter $*$ at the start of iteration n.
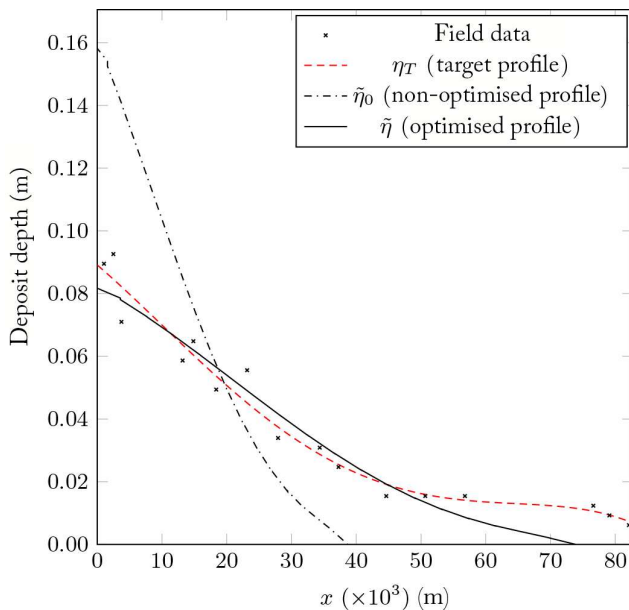

Figure 7. The dimensional deposit output $\widetilde{\eta}_0$ from the initial parameter guess (Eq. 74) and the optimised dimensional deposit output $\widetilde{\eta}$ from the optimised parameter (Eq. 77) shown against the field measurement from Bed 1.1 (Talling et al., 2007a) and the fourth-order polynomial target deposit profile, η_T.

The criteria for finishing the parameter optimisation is based upon the relative change in J between iterations such that

$$\frac{|J_i - J_{i-1}|}{J_i} < 1.0 \times 10^{-5}, \tag{76}$$

where J_i is the value of J after the ith iteration. The optimisation is completed in 21 iterations with a final functional

value of $J = 1.75$ (Fig. 6). The optimised deposit profile, η, compares relatively well with η_T (Fig. 7). Most notably, there is a significant variation in the thickness towards the end of the deposit. This will be addressed later. The final optimised values are

$$m = \begin{pmatrix} h_0 &=& 2.56\,\text{km} \\ \psi_0 &=& 0.0494\,\% \\ D &=& 103\,\mu\text{m} \end{pmatrix}. \tag{77}$$

These optimised values are not completely acceptable. The value for h_0 represents the initial height of the current starting from a static lock-release initial condition. This translates to an initial current height of 993.3 m at the start of this simulation and as the current enters the basin plain. This value appears to be quite large for a relatively small turbidity current. Additionally, the average sediment diameter of 103 μm is lower than expected. Talling et al. (2007a) defines the sandstone interval as dominated by sediment grains estimated to be larger than $\approx 125\,\mu$m.

With the exception of the sediment diameter, the optimised values are fairly similar to those chosen as input values. This confirms that the input parameters chosen were sensible predictions of the starting conditions for the gravity current. To test this hypothesis, we ran the same situation starting from a number of alternative initial conditions. We found that there are indeed a number of local minima. An optimisation with initial conditions

$$m = \begin{pmatrix} h_0 &=& 3.0\,\text{km} \\ \psi_0 &=& 0.02\,\% \\ D &=& 100.0\,\mu\text{m} \end{pmatrix} \tag{78}$$

was optimised to

$$m = \begin{pmatrix} h_0 &=& 3.95\,\text{km} \\ \psi_0 &=& 0.02\,\% \\ D &=& 154.0\,\mu\text{m} \end{pmatrix}. \tag{79}$$

Figure 8 shows a comparison of the two generated deposit profiles. The two profiles are very comparable, even though the alternative profile is created by a much larger, but much less dense, initial current.

The existence of alternative minima must always be considered when running optimisations of this type. It is important to have a good understanding of the problem to choose sensible initial starting conditions and also to assess the resultant optimised values. A regularisation approach would avoid this problem but assumes prior knowledge about the target profile.

A clear omission from the model is the presence of mud in the suspension. The presence of mud will significantly alter the energy budget of the flow. A mud sediment class can easily be included so that the model produces more realistic optimised values. This is detailed below.
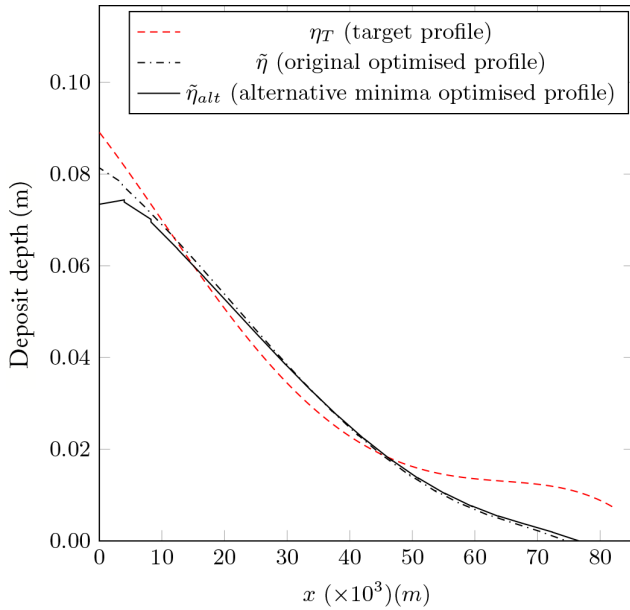
Figure 8. The dimensional deposit output $\widetilde{\eta}$ from the optimised deposit output from the initial parameter guess (Eq. 74) and the dimensional deposit output $\widetilde{\eta}_{alt}$ from an alternative minima achieved by using the initial parameter guess (Eq. 78) shown against the fourth-order polynomial target deposit profile, η_T.

4.5 Extending the model to include an additional sediment class for the mud in suspension

Investigating the effect of including mud in the sediment mixture can be achieved relatively simply by including an additional transport equation with a form identical to Eq. (12) (Dorrell et al., 2013),

$$\frac{\partial \varphi_m}{\partial t} = \frac{1}{x_N}\left(y\dot{x}_N\frac{\partial \varphi_m}{\partial y} - \frac{\partial}{\partial y}\left(\frac{q\varphi_m}{h}\right)\right) - \beta_m\frac{\varphi_m}{h}, \quad (80)$$

where $\varphi_m = \psi_m h$ is the vertically integrated volume fraction of mud in the suspension. ψ_m is the depth-averaged volume fraction of mud within the flow, and β_m is the settling velocity of the mud particles. Using a single tracer equation, we approximate the distribution of mud particle sizes using a single mud diameter, the same way the distribution of sand is modelled in the flow. We neglect the flocculation of mud particles. Assuming that the density of both sediment classes is the same, (Eq. 11) is modified to include this new sediment class in the gravity term:

$$\frac{\partial q}{\partial t} = \frac{1}{x_N}\left(y\dot{x}_N\frac{\partial q}{\partial y} - \frac{\partial}{\partial y}\left(\frac{q^2}{h} + \frac{(\varphi + \varphi_m)h}{2}\right)\right). \quad (81)$$

Finally, φ and φ_m are scaled such that at the start of the simulation $\varphi + \varphi_m = 1$, where previously $\varphi = 1$. The aim is still to recreate the deposit of sand, and hence the equation for η stays the same. We term the sand deposit generated by

this modified model η_2. The discretisation for Eq. (80) is consistent with the rest of the model, as presented in Sect. 2.2.

The initial condition needs to be altered to include the new sediment class. The initial vertically averaged volume fraction of sand is changed to $\psi = f_s$, and a new initial condition for the vertically averaged volume fraction of mud is introduced as $\psi_m = 1 - f_s$. The sand fraction, f_s, is estimated by Talling et al. (2007a) to be 0.14 and is kept fixed.

β_m must also be calculated. This is done in the same way as for β, except that a different sediment diameter parameter is used and optimised: D_m, the mean diameter of mud particles in the flow. The equation for β_m is therefore

$$\beta_m = \frac{g'^{1/2}D_m^2}{18\nu(h_0\psi_0)^{1/2}}. \quad (82)$$

Note that ψ_0 is now the combined initial volume fraction of sand and mud in the flow.

4.6 Optimisation for a model with two sediment classes

The set of optimised input parameters is redefined as $m = [h_0, \psi_0, D, D_m]^T$. An additional bounding constraint is added for D_m such that the new bounding constraints for m are

$$10\,\text{m} \le h_0 \le 10.0\,\text{km}, \quad (83)$$
$$0.001\,\% \le \psi_0 \le 50\,\%, \quad (84)$$
$$1.0\,\mu\text{m} \le D \le 1.0\,\text{mm}, \quad (85)$$
$$1.0\,\mu\text{m} \le D_m \le 100.0\,\mu\text{m}. \quad (86)$$

The initial input parameters are set to

$$m = \begin{pmatrix} h_0 &=& 2.1\,\text{km} \\ \psi_0 &=& 5\,\% \\ D &=& 200.0\,\mu\text{m} \\ D_m &=& 20.0\,\mu\text{m} \end{pmatrix}. \quad (87)$$

The input parameters are normalised as detailed in Sect. 4.4 before being passed to the optimisation algorithm. The criteria for finishing the optimisation are consistent with the previous optimisation (see Eq. 76).

The optimisation of the model with two sediment classes is completed in 17 iterations with a final functional value of $J = 2.13$ (see Fig. 9). Therefore, the fit is quantitatively slightly worse when mud is included in the model. This is a surprising result, as the model now more closely matches reality. Qualitatively, it is very hard to determine which model fits the data better. The resultant deposit is very similar in shape to that obtained when only modelling sand in the flow (see Fig. 10). The fit appears to be worse at the start of the deposit. The runout length is slightly longer when mud is included such that the fit towards the end of the deposit is slightly improved.

The fit with the measured data is still poor towards the end of the deposit. Talling et al. (2007a) noted that the distal section of small deposits in the Marnoso-arenacea Formation show evidence of transport in a tractional boundary
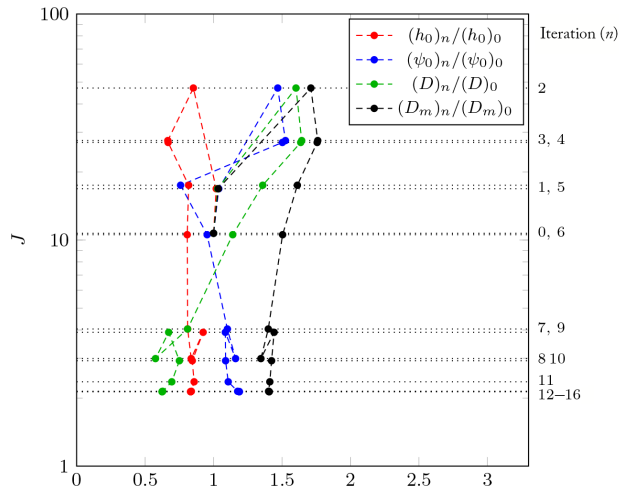
Figure 9. Values of the parameters from the model with both mud and sand sediment classes over the optimisation iterations against the value of the objective functional, J, that we are aiming to minimise. The values shown are normalised by their starting values. $(*)_n$ is the value of parameter $*$ at the start of iteration n.
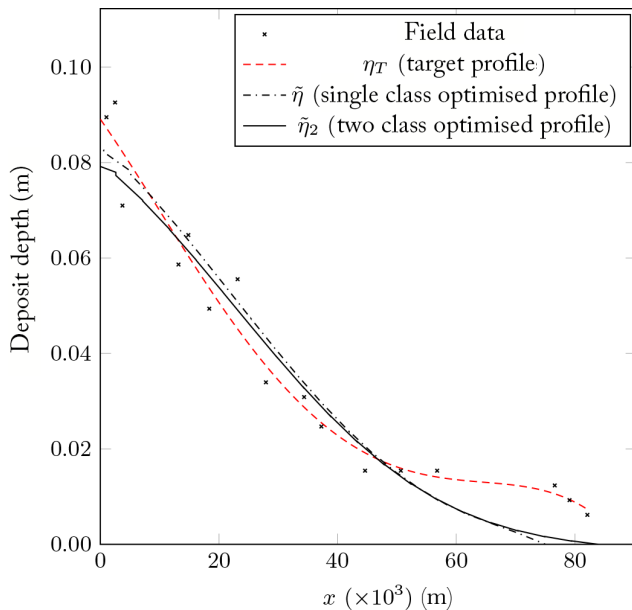

Figure 10. Optimised dimensional deposit output from the model with both mud and sand sediment classes, $\tilde{\eta}_2$, shown against the optimised results from the single sediment class model, $\tilde{\eta}$, the field measurement from Bed 1.1 (Talling et al., 2007a), and the fourth-order polynomial target deposit profile, η_T.

layer. This simulation does not model bedload transport or erosion, which is the likely reason for the difference in the results. The velocity of the head of the turbidity current in this simulation varies between 10 and $2.4\,\mathrm{m\,s^{-1}}$ over the period during which sand is deposited (Fig. 12). At these head velocities, erosion is very likely to occur. Models for erosion

and bedload transport exist (Garcia, 1994; Sequeiros et al., 2009). These could be added in future work.

The final optimised values are

$$
m = \begin{pmatrix} h_0 & = & 1.92\,\mu\mathrm{km} \\ \psi_0 & = & 5.94 \times 10^{-3} \\ D & = & 125\,\mu\mathrm{m} \\ D_\mathrm{m} & = & 28.1\,\mu\mathrm{m} \end{pmatrix}.
\tag{88}
$$

A comparison of these results to those obtained without a mud sediment class shows that the value of h_0 has reduced by 25 % and translates to an initial current height of 745.0 m as the current enters the basin plain. The average sediment diameter has also increased by 21 % to 125 µm, bringing the average diameter in line with the estimates from the field measurements by Talling et al. (2007a). Arguably, the sand and mud classes should be subdivided further. Dorrell et al. (2013) described how polydisperse density currents will have longer runout distances than equivalent currents with uniform sediment at the mean value of the poydisperse current.

It is also interesting to assess the sensitivity of the model to variations in the input parameters by analysing the final gradient of the objective functional,

$$
\frac{\mathrm{d}J}{\mathrm{d}\overline{m}} = \begin{pmatrix} \mathrm{d}J/\mathrm{d}\overline{h}_0 & = & 5.1 \times 10^{-3} \\ \mathrm{d}J/\mathrm{d}\overline{\psi}_0 & = & -1.8 \times 10^{-3} \\ \mathrm{d}J/\mathrm{d}\overline{D} & = & -2.4 \times 10^{-3} \\ \mathrm{d}J/\mathrm{d}\overline{D}_\mathrm{m} & = & 1.5 \times 10^{-6} \end{pmatrix},
\tag{89}
$$

where $\overline{}$ indicates a parameter value normalised by its value on the initial optimisation iteration. The sensitivity of the functional to changes in the mud diameter is several orders of magnitude smaller than the sensitivity to changes in the other variables.

It is indeed found that changing this value has very little effect on the obtained deposit. The same simulation is run with the mud diameter decreased by 2 orders of magnitude such that the input parameter values are

$$
m = \begin{pmatrix} h_0 & = & 1.92\,\mu\mathrm{km} \\ \psi_0 & = & 5.94 \times 10^{-3} \\ D & = & 125\,\mu\mathrm{m} \\ D_\mathrm{m} & = & 0.281\,\mu\mathrm{m} \end{pmatrix}.
\tag{90}
$$

The resulting functional value is $J = 2.13$, which is identical to that obtained for the optimised simulation. There is no discernible difference in the resulting deposit, η_3 (Fig. 11). The head height and velocity only vary a small amount over the period during which sand is deposited (Fig. 12a and b). The current properties vary significantly after the sand has been deposited and mud is still in suspension, but this does not have any effect on the sand deposit.

Although the sandstone deposits generated by the single and two sediment class models are very similar, the properties of the turbidity currents that produced them are very different (Fig. 12). The turbidity current with mud in suspension
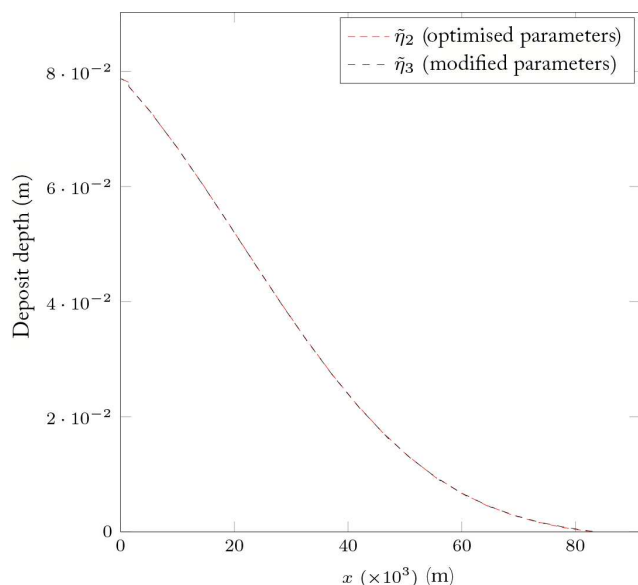
Figure 11. Optimised dimensional deposit output from the model with both mud and sand sediment classes, $\widetilde{\eta}_2$, shown against the results from the same model with the same parameters but a mud-settling velocity reduced by 2 orders of magnitude, $\widetilde{\eta}_3$.

travels approximately twice as quickly due to the increased gravitational forces produced by the sand and mud mixture (Fig. 12b). Sand also drops out of the suspension much more rapidly (Fig. 12d). All of the sand is deposited within approximately 6 h. The model without mud deposits sand over a period of more than 20 h. This is due to the reduced height of the current at the start of the simulation and the faster decrease in the height of the current as a result of the higher head velocity (Fig. 12a and d). Clearly, the presence of mud in the suspension has a significant impact on the resultant flow and must be included in the model.

The simulated turbidity currents that produced η_2 and η_3 deposited sand over a similar time period. After all of the sand had fallen out of suspension, less than 25 % of the mud settled from the flow for both of these currents; the current head is > 50 m tall, and the head is moving at > 1.0 m s^{-1} (Fig. 12). Hence, there is still a significant amount of energy in the flow. The remaining mud suspension will reach the end of the basin ($\widetilde{x}_N \approx 130$ km) and will still have a significant amount of energy left when it does so. It is very hard to predict what will happen after this point. The current may be partly reflected, and ponding of the suspended mud is likely to occur. This result is in agreement with the explanations of Talling et al. (2007b).

The height of the current in the optimised simulation with both sand and mud sediment classes is ≈ 750 m as it enters the basin, although this decreases very quickly as the current propagates. It is possible that including processes such as fluid entrainment, erosion, and bedload transport may reduce the necessity for such a large initial current height in produc-

ing this deposit. More complex initial and boundary conditions may also have a significant impact on this value. It is unclear what effect an inflow boundary condition with time-varying height, sediment concentration, and velocity would have on the results. This would be an interesting addition to the models capabilities.

The model also neglects variations in the bed profile. The gradient of the sea floor in the basin where the Marnoso-arenacea Formation was created was substantially less than 1 degree (Amy and Talling, 2006). Variations in gradient of this magnitude will have a negligible impact on the head velocity (Middleton, 1966). However, small variations will have an impact on the velocity of the body of the current. Future work will address this.

5 Conclusions

This paper has presented a novel implementation of the shallow water equations for modelling density currents using a mixed finite-element formulation. The model has been differentiated to allow for parameter optimisation using gradient-based optimisation techniques and the use of gradient information in sensitivity analyses.

The proposed model is based upon simplifying shallow water sediment-laden density current assumptions and has been used here to recreate a low-volume deposit from the Marnoso-arenacea Formation in Italy with some success. However, the lack of many key flow processes within the current model, including bedload transport and re-entrainment, has arguably led to optimised parameters values which would be improved upon with a more complete underlying model.

This paper has demonstrated the power of gradient-based optimisation methods for determining the set of input parameters that best fits a particular turbidity current deposit. Since the input parameters are rarely known with any accuracy for these flows, optimisation represents a sensible way to better estimate these values.

Future development of the model could enable more complex boundary conditions and add parameterisations for ambient fluid entrainment, bed erosion, and bedload transport. This will increase the capacity for the model to recreate a range of deposits found in the field, while the parameter space will grow significantly. The optimisation techniques presented in this paper will allow for the efficient selection of optimised values for a large parameter space.

6 Code availability

The model implementation and the test setups described in this paper are freely available as a separate git repository on bitbucket: https://github.com/funsim/adjoint-turbidity. This repository contains a README file which guides the user through the installation and how to reproduce the results of the paper. The experiment configuration files can viewed
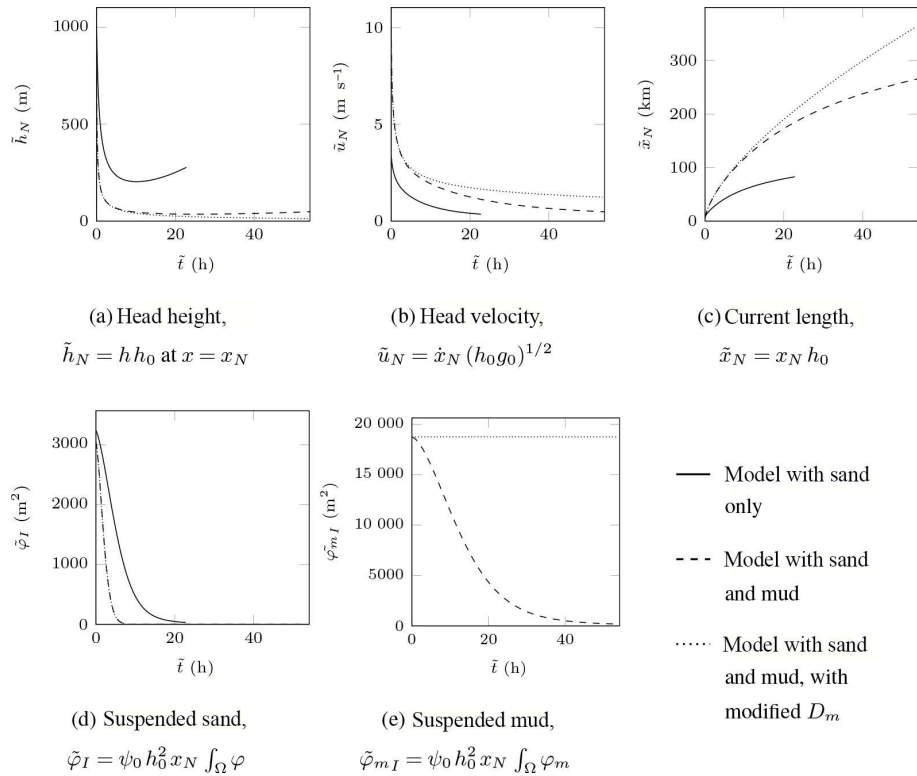
Figure 12. The time evolution of the dimensionalised variables for three simulations: a simulation with a single sand sediment class and optimised input parameters to match the Bed 1.1 sand deposit, a simulation with sand and mud classes and optimised input parameters to match the Bed 1.1 sand deposit, and a simulation with sand and mud classes and the same optimised input parameters but a mud diameter 2 orders of magnitude smaller. The results are shown against the dimensional time, $\widetilde{t} = t \, (h_0/g_0)^{1/2}$.

and changed with the graphical configuration tool spud from the Fluidity project (http://fluidityproject.github.io/). The dynamical core of the model is implemented with the finite-element software FEniCS and its extension dolfin-adjoint. The documentation for FEniCS is available at http://fenicsproject.org/documentation, and the documentation for dolfin-adjoint can be found at http://www.dolfin-adjoint.org.

Dolfin-adjoint and all FEniCS core components are licensed under the GNU LGPL as published by the Free Software Foundation, either version 3 of the licence or (optionally) any later version.

Edited by: A. Sandu
Reviewed by: two anonymous referees

References

Alnæs, M. S., Logg, A., Ølgaard, K. B., Rognes, M. E., and Wells, G. N.: Unified Form Language: A domain-specific language for weak formulations of partial differential equations, ACM T. Math. Software, 40, 9:1–9:37, doi:10.1145/2566630, 2012.

Amy, L. A. and Talling, P. J.: Anatomy of turbidites and linked debrites based on long distance $(120 \times 30 \, \text{km})$ bed correlation, Marnoso Arenacea Formation, Northern Apennines, Italy, Sedimentology, 53, 161–212, doi:10.1111/j.1365-3091.2005.00756.x, 2006.

Balay, S., Adams, M. F., Brown, J., Brune, P., Buschelman, K., Eijkhout, V., Gropp, W. D., Kaushik, D., Knepley, M. G., McInnes, L. C., Rupp, K., Smith, B. F., and Zhang, H.: PETSc Web page (Portable, Extensible Toolkit for Scientific Computation), available at: http://www.mcs.anl.gov/petsc, 2014.

Bonnecaze, R. T., Huppert, H. E., and Lister, J. R.: Particle-driven gravity currents, J. Fluid Mech., 250, 339–339, doi:10.1017/S002211209300148X, 1993.

Cockburn, B. and Shu, C.-W.: Runge–Kutta discontinuous Galerkin methods for convection-dominated problems, J. Sci. Comput., 16, 173–261, doi:10.1023/A:1012873910884, 2001.

Dorrell, R. M., Hogg, A. J., and Pritchard, D.: Polydisperse suspensions: Erosion, deposition, and flow capacity, J. Geophys. Res.-Earth, 118, 1939–1955, doi:10.1002/jgrf.20129, 2013.

Doyle, E. E., Hogg, A. J., Mader, H. M., and Sparks, R. S. J.: Modeling dense pyroclastic basal flows from collapsing columns, Geophys. Res. Lett., 35, L04305, doi:10.1029/2007GL032585, 2008.

Doyle, E. E., Hogg, A. J., Mader, H. M., and Sparks, R. S. J.: A two-layer model for the evolution and propagation of dense and dilute regions of pyroclastic currents, J. Volcanol. Geoth. Res., 190, 365–378, doi:10.1016/j.jvolgeores.2009.12.004, 2010.

Farrell, P. E., Piggott, M. D., Gorman, G. J., Ham, D. A., Wilson, C. R., and Bond, T. M.: Automated continuous verification for numerical simulation, Geosci. Model Dev., 4, 435–449, doi:10.5194/gmd-4-435-2011, 2011.

Farrell, P. E., Ham, D. A., Funke, S. W., and Rognes, M. E.: Automated derivation of the adjoint of high-level transient finite element programs, J. Sci. Comput., 35, C369–C393, doi:10.1137/120873558, 2013.

Fukushima, Y., Parker, G., and Pantin, H. M.: Prediction of ignitive turbidity currents in Scripps Submarine Canyon, Mar. Geol., 67, 55–81, doi:10.1016/0025-3227(85)90148-3, 1985.

Funke, S. W. and Farrell, P. E.: A framework for automated PDE-constrained optimisation, arXiv:1302.3894, 2013.

Funke, S. W., Farrell, P. E., and Piggott, M. D.: Tidal turbine array optimisation using the adjoint approach, Renew. Energ., 63, 658–673, doi:10.1016/j.renene.2013.09.031, 2014.

Gandolfi, G., Paganelli, L., and Zuffa, G. G.: Petrology and dispersal pattern in the Marnoso-arenacea Formation (Miocene, Northern Apennines), Int. J. Sediment. Res., 53, 493–507, doi:10.1306/212F8215-2B24-11D7-8648000102C1865D, 1983.

Garcia, M. H.: Depositional turbidity currents laden with poorly sorted sediment, J. Hydraul. Eng.-ASCE, 120, 1240–1263, doi:10.1061/(ASCE)0733-9429(1994)120:11(1240), 1994.

Giles, M. B. and Pierce, N. A.: An introduction to the adjoint approach to design, Flow Turbul. Combust., 65, 393–415, 2000.

Hallworth, M. A., Hogg, A. J., and Huppert, H. E.: Effects of external flow on compositional and particle gravity currents, J. Fluid Mech., 359, 109–142, doi:10.1017/S0022112097008409, 1998.

Harris, T. C., Hogg, A. J., and Huppert, H. E.: A mathematical framework for the analysis of particle–driven gravity currents, P. Roy. Soc. A-Math. Phy., 457, 1241–1272, doi:10.1098/rspa.2000.0728, 2001.

Harris, T. C., Hogg, A. J., and Huppert, H. E.: Polydisperse particle-driven gravity currents, J. Fluid Mech., 472, 333–371, doi:10.1017/S0022112002002379, 2002.

Heezen, B. C. and Ewing, W. M.: Turbidity currents and submarine slumps, and the 1929 Grand Banks earthquake, Am. J. Sci., 250, 849–873, 1952.

Hogg, A. J.: Lock-release gravity currents and dam-break flows, J. Fluid Mech., 569, 61–87, doi:10.1017/S0022112006002588, 2006.

Hoult, D. P.: Oil spreading on the sea, Annu. Rev. Fluid Mech., 4, 341–368, doi:10.1146/annurev.fl.04.010172.002013, 1972.

Huang, H., Imran, J., and Pirmez, C.: Nondimensional parameters of depth-averaged gravity flow models, J. Hydraul. Res., 47, 455–465, doi:10.1080/00221686.2009.9522021, 2009.

Huppert, H. E. and Simpson, J. E.: The slumping of gravity currents, J. Fluid Mech., 99, 785–799, doi:10.1017/S0022112080000894, 1980.

Jorge, N. and Stephen, J. W.: Numerical optimization, Springer, 1999.

Kneller, B. and Buckee, C.: The structure and fluid mechanics of turbidity currents: a review of some recent studies and their geological implications, Sedimentology, 47, 62–94, doi:10.1046/j.1365-3091.2000.047s1062.x, 2000.

Kuzmin, D.: A vertex-based hierarchical slope limiter for p-adaptive discontinuous Galerkin methods, J. Comput. Appl. Math., 233, 3077–3085, doi:10.1016/j.cam.2009.05.028, 2010.

Lamb, H.: Hydrodynamics, Cambridge University Press, 1993.

Lesshafft, L., Meiburg, E., Kneller, B., and Marsden, A.: Towards inverse modeling of turbidity currents: The inverse lock-exchange problem, Comput. Geosci., 37, 521–529, doi:10.1016/j.cageo.2010.09.015, 2011.

Logg, A., Mardal, K.-A., and Wells, G.: Automated solution of differential equations by the finite element method: The FEniCS book, vol. 84, Springer Science & Business Media, 2012.

Lowe, D. R.: Sediment Gravity Flows, Geology of Continental Slope, 27, 75–82, doi:10.2110/pec.79.27.0075, 1979.

Lucchi, F. R. and Valmori, E.: Basin-wide turbidites in a Miocene, over-supplied deep-sea plain: a geometrical analysis, Sedimentology, 27, 241–270, doi:10.1111/j.1365-3091.1980.tb01177.x, 1980.

Menemenlis, D., Campin, J.-M., Heimbach, P., Hill, C., Lee, T., Nguyen, A., Schodlok, M., and Zhang, H.: ECCO2: High resolution global ocean and sea ice data synthesis, Mercator Ocean Quarterly Newsletter, 31, 13–21, 2008.

Middleton, G. V.: Experiments on density and turbidity currents, I. Motion of the head, Can. J. Earth Sci., 3, 523–546, doi:10.1139/e66-038, 1966.

Naruse, H.: Tsunami-genic turbidity currents: observation and estimation of flow condition by numerical simulations, in: Apennine Workshop: future research directions for submarine flows, AGU Publications, 2013.

Parker, G., Fukushima, Y., and Pantin, H. M.: Self-accelerating turbidity currents, J. Fluid Mech., 171, 145–81, doi:10.1017/S0022112086001404, 1986.

Parkinson, S. D., Hill, J., Piggott, M. D., and Allison, P. A.: Direct numerical simulations of particle-laden density currents with adaptive, discontinuous finite elements, Geosci. Model Dev., 7, 1945–1960, doi:10.5194/gmd-7-1945-2014, 2014.

Peraire, J. and Persson, P. O.: The compact discontinuous Galerkin (CDG) method for elliptic problems, J. Sci. Comput., 30, 1806–1824, doi:10.1137/070685518, 2008.

Rowley, P.: Inverse modelling of a pyroclastic turbidite, in: Apennine Workshop: future research directions for submarine flows, 2013.

Sequeiros, O. E., Cantelli, A., Viparelli, E., White, J. D. L., Garcia, M. H., and Parker, G.: Modeling turbidity currents with nonuniform sediment and reverse buoyancy, Water Resour. Res., 45, 1–28, doi:10.1029/2008WR007422, 2009.

Talling, P. J., Amy, L. A., and Wynn, R. B.: New insight into the evolution of large-volume turbidity currents: comparison of turbidite shape and previous modelling results, Sedimentology, 54, 737–769, doi:10.1111/j.1365-3091.2007.00858.x, 2007a.

Talling, P. J., Amy, L. A., Wynn, R. B., Blackbourn, G., and Gibson, O.: Evolution of turbidity currents deduced from extensive thin turbidites: Marnoso Arenacea Formation (Miocene), Italian Apennines, J. Sediment. Res., 77, 172–196, doi:10.2110/jsr.2007.018, 2007b.

Talling, P. J., Wynn, R. B., Masson, D. G., Frenz, M., Cronin, B. T., Schiebel, R., Akhmetzhanov, A. M., Dallmeier-Tiessen, S., Benetti, S., Weaver, P. P. E., Georgiopoulou, A., Zühlsdorff, C., and Amy, L. A.: Onset of submarine debris flow deposition far from original giant landslide, Nature, 450, 541–544, doi:10.1038/nature06313, 2007c.

Talling, P. J., Masson, D. G., Sumner, E. J., and Malgesini, G.: Subaqueous sediment density flows: Depositional processes and deposit types, Sedimentology, 59, 1937–2003, doi:10.1111/j.1365-3091.2012.01353.x, 2012.

Ungarish, M.: An introduction to gravity currents and intrusions, CRC Press, 2010.

Wächter, A. and Biegler, L. T.: Line search filter methods for nonlinear programming: Motivation and global convergence, J. Optimization, 16, 1–31, doi:10.1137/S1052623403426556, 2005.

Wächter, A. and Biegler, L. T.: On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming, Math. Program., 106, 25–57, doi:10.1007/s10107-004-0559-y, 2006.