Synthetic benchmarking of concentrated pyroclastic current models

V. Gueugneau, S. Charbonnier, T. Esposti Ongaro, M De' Michieli Vitturi, Marc Peruzzetto, A. Mangeney, François Bouchut, A. Patra, Karim Kelfoun

Synthetic benchmarking of concentrated pyroclastic current models

[1,*]Gueugneau V., [1]Charbonnier S., [2]Esposti Ongaro T., [2,7] de'Michieli Vitturi M., [3,8] Peruzzetto M., [3]Mangeney A., [4]Bouchut F., [5]Patra A., [6]Kelfoun K.

[1] School of Geosciences, University of South Florida, Tampa, FL, USA
[2] Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Pisa, Pisa, Italy
[3] Université de Paris, Institut de Physique du Globe, CNRS, Paris, France
[4] Université Gustave Eiffel, Laboratoire d'Analyse et de mathématiques appliquées, CNRS, Paris, FR
[5] Computational Data Science and Engineering, University of Buffalo, Buffalo, NY, USA
[6] Université Clermont Auvergne, Laboratoire Magma et Volcans, CNRS-IRD-OPGC, Clermont-Ferrand, France
[7] University at Buffalo, Buffalo, NY, USA
[8] Bureau de Recherches Géologiques et Minières, Orléans, France

*Corresponding author: vgueugneau@usf.edu

Abstract

Validation and benchmarking of pyroclastic current (PC) models is required to evaluate their performance and their reliability for hazard assessment. Here we present results of a benchmarking initiative built to evaluate four models commonly used to assess concentrated PC hazard: SHALTOP, TITAN2D, VolcFlow and IMEX_SfloW2D. The benchmark focuses on the simulation of channelized flows with similar source conditions over five different synthetic channel geometries: 1) a flat incline plane, 2) a channel with a sharp 45° bend, 3) a straight channel with a break-in-slope, 4) a straight channel with an obstacle, and 5) a straight channel with a constriction. Several outputs from 60 simulations using three different initial volume fluxes were investigated to evaluate the performance of the four models when simulating valley-confined PC kinematics, including overflows induced by topographic changes. Quantification of the differences obtained between model outputs at t = 100 s allowed us to identify: 1) issues with the Voellmy-Salm implementation of TITAN2D and 2) small discrepancies between the three other codes that are either due to various curvature and velocity formulations and/or numerical frameworks. Benchmark results were also in agreement with field observations of natural PCs: a sudden change in channel geometries combined with a high-volume flux are keys to generate overflows. The synthetic benchmarks proved to be useful for evaluating model performance, needed for PCs hazard assessment. The overarching goal is to provide an interpretation framework for volcanic mass flow hazard assessment studies to the geoscience community.

Keywords

numerical modeling – synthetic benchmarking – depth-averaged models – concentrated pyroclastic currents – hazard assessment

Declarations

Acknowledgements

I.   Introduction

1.1. Motivation

Our current ability to simulate the behavior of pyroclastic currents (PCs) is limited by our incomplete knowledge of their internal dynamics. These fast-moving flows composed of hot volcanic particles and gas represent a threat for infrastructure and populations surrounding volcanoes (Neri et al. 2015; Brown et al. 2017). Advances in our knowledge have been hindered because of: (i) the intrinsic dangers and costs of performing field studies on natural deposits just after their emplacement; (ii) the difficulties in investigating their internal structure and performing in-situ measurements; and (iii) the complications of linking deposits with their unsteady and non-uniform flow behavior (Dufek et al. 2015). Both numerical modeling (Dartevelle 2004a,b; Dufek and Bergantz 2007; Esposti Ongaro et al. 2008, 2012; Dufek et al. 2009; Benage et al. 2016; Kelfoun et al. 2017; Sweeney and Valentine, 2017; Valentine and Sweeney 2018) and experimental modeling (Dellino et al. 2010; Roche et al. 2008, 2010; Roche, 2012, 2015; Andrews, 2014; Andrews and Manga, 2011, 2012; Breard et al. 2016; Sulpizio et al 2016; Breard and Lube 2017; Smith et al. 2018; Dellino et al. 2019; Brosch and Lube, 2020) approaches have progressively emerged as one key alternative to study these hazardous flows and further enhancing the sedimentological and physical models of PCs, as summarized in recent review papers (Sulpizio et al. 2014; Dufek et al. 2015; Dufek, 2016; Lube et al. 2020).

In this study, we adopt the term 'pyroclastic current' in its most general sense as proposed by Palladino (2017). Pyroclastic currents display a strong vertical stratification of the volumetric particle concentration ranging from a concentrated regime (between 10 to 60 vol.%) dominated by particle-particle interactions, to a dilute regime (less than a few vol.%; Weit et al. 2018) dominated by gas-particle interactions (Lube et al. 2020). When these two regimes coexist in a single PC, the flow is named *concentrated pyroclastic current* (CPC), which displays a concentrated basal zone and a dilute upper zone referred here as the 'ash-cloud surge'. The interface between the two zones has intermediate and complex dynamics, dominated by exchanges of mass and momentum and by particle clustering (Bréard and Lube, 2016; Lube et al. 2020). Although it is recognized that both the concentrated and dilute systems coexist in most PCs, in some cases no concentrated basal zone is observed (Valentine, 2020), and the flow is named *dilute pyroclastic current* (DPC). While the study of the two endmembers, i.e. CPC and DPC, is essential to build a comprehensive PC model and to help in the interpretation of natural PC deposits, this study focuses on CPCs only.

1.2. A key process: PC overspilling

Small-volume CPCs, which display a volume inferior to 10^8 m^3 usually, are remarkably sensitive to the topography and stay mostly channelized into deep valleys (Cole et al 2002; Tierz et al. 2016).

Under specific circumstances, they can overspill from these valleys and inundate the surrounding slopes, often reaching inhabited areas away from the channelized flow paths. These inhabited areas may be unprepared for these hazards and flow overspill events can cause damages and death. Here, two processes can be distinguished: (i) the 'CPC overspill' for which the CPC, often accompanied by its upper ash-cloud surge but not always, escapes the volcanic valley, like at Merapi during the 2006 and 2010 eruptions (Charbonnier and Gertisser, 2008; Lube et al. 2011; Gertisser et al. 2012; Charbonnier et al. 2013), at Volcán de Colima (Mexico) in 2015 (Macorps et al. 2018) or recently at Fuego volcano in 2018 (Charbonnier et al, 2019; Albino et al, 2020); (ii) the 'ash-cloud surge detachment' for which only the dilute upper zone of the PC detaches and escape the valley, like at Montserrat (Loughlin et al. 2002; Ogburn et al. 2014), Unzen (Nakada and Fujii 1993), or Merapi (Komorowski et al. 2013).

As we restricted our study to CPCs only, we focus here on the CPC overspill process. Several field studies have shown that CPC overspill is likely to be controlled by two main parameters (Charbonnier and Gertisser, 2008; Lube et al. 2011; Gertisser et al. 2012; Ogburn et al., 2014; Macorps et al. 2018):

(i) The morphology of the valley. Valleys in volcanic landscape display a wide range of morphologies and CPC overspill events usually occur when the flow encounters a sudden topographical change (Gertisser et al. 2012; Ogburn et al. 2014). A modification of the channel geometry (both from natural causes and/or the result of human intervention) can potentially reduce the channel capacity (i.e. the maximum volume flux supported by a valley at a specific location), causing CPC to overspill. At least four main topographic features have been identified to have a significant impact on the CPC dynamics: a sharp valley bend (Ogburn et al. 2014; Macorps et al. 2018), a well-defined break in slope along the valley (Bourdier and Abdurackmann 2001; Charbonnier and Gertisser 2012), a sudden constriction of the valley width (Charbonnier and Gertisser 2008, 2011; Jenkins et al. 2013), and an obstacle obstructing the valley (i.e. sabo dam, lava ridges, bridges; Charbonnier and Gertisser, 2008; Lube et al. 2011).

(ii) the CPC local volume flux into that valley at the overspill site. The capacity of a CPC to overspill channel confines is also controlled by how fast and how long it takes for the entire CPC mass to be transported down the channel slope. A large CPC volume flux, exacerbated by the pulsating behavior of CPCs in some eruptions, can locally exceed the channel capacity of a valley and allow the flow to overspill on the surrounding slopes. Previous studies at Soufriere Hills Volcano (Ogburn et al. 2014), Merapi (Charbonnier and Gertisser 2008; Cronin et al. 2013; Jenkins, 2013) or Volcán de Colima (Macorps et al. 2018) highlighted the direct link between the increase of the local CPC volume flux (calculated along the cross-sectional area of the channel) and the occurrence of CPC overspilling and/or ash-cloud surge detachment and decoupling phenomenon. A small-volume CPC generated by a short explosion or a small dome collapse may not generate any of these processes, but a voluminous and

fast CPC generated by a large collapse of a fast-growing lava dome may generate flow overspilling and/or ash-cloud surge decoupling phenomenon, as its volume flux would be higher and exceeds the channel capacity in some areas.

1.3. <u>Benchmarking of numerical models for concentrated pyroclastic currents</u>

Because of the complex physics of PCs, various numerical codes have been developed throughout the years (more than 30 since Valentine and Wohletz (1989), see Table 1), while no rigorous PC model inter-comparison has been conducted yet. The urgent need for a community-wide PC model benchmark clearly arises today not only to better assess the applicability and performance of the various models available, but also to support and improve PC hazards assessment worldwide. A first attempt of a CPC model inter-comparison was conducted by Charbonnier and Gertisser (2012) using two of the most widely used mass flow models, i.e. VolcFlow and TITAN2D, based on the reproduction of the 2010 Merapi eruption. A second attempt was conducted recently by Ogburn and Calder (2017) who compared a larger variety of models (i.e. TITAN2D, VolcFlow, LAHARZ, and PFz) in their ability to reproduce a series of well-recorded block-and-ash flows from Soufrière Hills Volcano (Montserrat, West Indies, UK). Their approach was based on a best-fit procedure of field observations, using different sources, rheologies, and boundary conditions. Here, we present a benchmark study aiming at assessing model-related uncertainties by comparing flow simulations performed under similar source and boundary conditions, following the validation framework proposed by Esposti Ongaro et al. (2020) based upon a hierarchical procedure commonly adopted for complex engineering systems (Oberkampf and Trucano, 2002).

Following Esposti Ongaro et al. (2020), we here distinguish between *verification* (i.e. the assessment of the mathematical correctness of a numerical model) and *validation* (i.e. the assessment of model reliability/performance with respect to the natural phenomenon). The validation procedure can be subdivided in four *validation tiers*, at an increasing level of complexity. At each level, the successful comparison of model results with reference datasets is a *confirmation* of model reliability (see Esposti Ongaro et al. 2020, section 'Confirmation'). In this work, we have tested models that have already been verified, and confirmed at the lowest validation Tier 3, against some reference *Unit problems* (i.e. simple experiments to test some fundamental physical behavior; see for example Mangeney et al. 2007; Gueugneau et al. 2017). We focus on Tier 2, *Benchmarks,* i.e. standardized problems having some degree of complexity, mainly concerning geometrical and scaling complications, for which full-scale experiments can be designed. In cases where experimental datasets are not (yet) available, synthetic benchmarks (also called inter-comparison studies) can be conceived to define the differences/similarities of the numerical models. In this framework, a benchmark is a preliminary step before the validation of models against a natural case. Numerous benchmark studies of this type have

Approach	Code acronym	References
Statistical correlations	ECM	Malin and Sheridan (1982)
	LAHARZ	Schilling et al. (1998)
	PFz	Widiwajawanti et al. (2009)
Cellular automata	*MOLASSES*	Richardson et al. (2016)
Kinematic	PYROFLOW	Wadge et al. (1998)
Depth-averaged (1D, steady-state, mixture, incompressible)	B&W96	Bursik and Woods (1996)
	ISH07	Ishimine (2005)
	SHI19	Shimizu et al. (2019)
	DOY08	Doyle et al. (2008)
Depth-averaged (2D, transient, mixture, incompressible)	TITAN2D	Patra et al. (2005)
	VOLCFLOW	Kefloun and Druitt (2005)
	IMEX-SFlow2D	De'Micheli Vitturi et al. (2019)
	SHALTOP	Mangeney et al. (2005)
	DAN3D	McDougall and Hungr (2004)
Multiphase (3D, transient, compressible)	MFIX	Valentine and Wohletz (1989) ; Dufek and Bergantz (2007) ; Sweeney and Valentine (2017) ; Breard et al. (2019)
	PDAC	Neri et al. (2003); Esposti Ongaro et al. (2007)
	DOR10	Doronzo et al. (2010)

Table 1: Compilation of the main modeling approaches and the related numerical models used to simulate concentrated pyroclastic currents. Shading highlights models tested in this benchmark.

already been conducted in geosciences for volcanic plume models (Suzuki et al. 2016 in 3D and Costa et al. 2016, in 1D), lava flow models (Cordonnier et al. 2015; Diettrich et al. 2017), landslide/debris flow models (Landslide benchmarking initiative by the JTC1, Hungr et al. 2007 and Pastor et al. 2018), tsunamis (Horrillo et al. 2014), ocean processes (Martinec et al. 2018), or geothermal modelling (Wang et al. 2020).

We present here the results of the first synthetic benchmark of CPC models. Our inter-comparison is based on four commonly used geophysical mass flow models: VolcFlow (Kelfoun and Druitt 2005), TITAN2D (Patra et al. 2005), SHALTOP (Bouchut and Westdickenberg, 2004; Mangeney-Castelnau et al. 2005; Mangeney et al. 2007) and IMEX_Sflow2D (de' Michieli Vitturi et al. 2019). An overview of PC modeling approaches and constitutive equations of each model is followed by the description of the source and boundary conditions used in our benchmarks. Flow simulation results are presented based upon several outputs, selected to accurately investigate numerical flow dynamics, and models are evaluated in their relative capacity to simulate natural-like equivalent: (i) CPC kinematics, and (ii) CPC overspill processes induced by topographic changes. Models give consistent results for the four topographical cases but display noticeable differences in their degree of interaction with the topographic features. These differences are discussed and put into perspective for potential design of ad-hoc validation experiments and current hazard assessment procedures for PCs.

II.   Modeling concentrated pyroclastic currents

2.1. Overview of PC modeling approaches

The fluid dynamics of PCs is extremely complex, including a broad spectrum of phenomena, with a multi-scale interplay between inertial and dissipative processes occurring at a microscopic (e.g. particle-particle interactions), mesoscopic (e.g., turbulence, particle clustering), and macroscopic scale (e.g., bulk internal and basal friction, interaction with the topography, particle deposition; Freundt and Bursik, 1998; Dartevelle 2004a; Dufek et al. 2015; Dufek, 2016; Lube et al. 2020; Esposti Ongaro et al. 2020). However, when limiting investigation to CPCs, some of the complexity can be reduced by neglecting the role of the multiphase processes (by adopting a mixture theory or neglecting the role of the interstitial fluids and the contribution of the kinetic stress to the bulk stress), heat exchange (assuming an isothermal approximation), turbulence (neglecting energy cascade in laminar flows) and the flow compressibility. Although such approximations can be questioned, it is common practice among volcanologists to consider CPCs to be controlled by inertial processes (pressure and gravity forces), topographic interaction and frictional dissipation (Dufek et al. 2015).

To reproduce partially or entirely these fundamental elements of CPC dynamics, many codes have been developed, using different modeling techniques and approaches (Table 1). They can be divided into two main categories: 1) kinematic/empirical models based on statistical correlations or simple physical principles, preferentially used for uncertainty quantification and hazard inundation forecasting purposes (Iverson et al. 1998; Tierz et al. 2016; 2018; Aravena et al. 2020); and 2) models based on the fundamental laws of fluid dynamics, at different level of approximations, used for both hazard assessment and to study fundamental CPC processes and physical behavior. For the latter, two

main sub-categories can be distinguished: multiphase flow models that consider the three-dimensional Navier Stokes equations for each constituent of the volcanic mixture, and depth-averaged models both for steady-state or transient dynamics, usually considering the eruptive mixture as a single (averaged) phase. For a comprehensive review of the different modeling approaches used to simulate PCs, we refer to Roche et al. (2013); Dufek (2016), and Esposti Ongaro et al. (2020).

2.2. Depth-averaged approach

In this study we focus on two-dimensional, depth-averaged, transient mixture models because they constitute a good compromise between model reliability and computational requirements. Since the first mathematical formulation of the approach by Savage and Hutter (1989), depth-averaged models have been extensively used to model gravity-driven flows, especially for their ability to simulate some of their fundamental processes: flow sedimentation/deposition, sensitivity to the topography, and frictional behavior. Such gravity-driven flows include geophysical flows like landslides, debris flows, or rock avalanches (Denlinger and Iverson 2004; Iverson and Denlinger 2004; Mangeney et al. 2003, 2005, 2007; Kelfoun and Druitt 2005; McDougall and Hungr 2004; Christen et al. 2010; George and Iverson 2014; Lucas et al. 2014; Brunet et al. 2017; Peruzzetto et al. 2019), as well as volcanic flows like CPCs (Sheridan et al. 2004; Kelfoun et al. 2009, 2017; Gueugneau et al. 2019, 2020; Salvatici et al. 2016) or lava flows (Bernabeu et al. 2013; Kelfoun and Vargas 2015).

Introduced by Saint-Venant (1871), the depth-averaged approach considers that for thin flows (i.e. flow length far exceeding flow thickness, which is the case for CPCs), reduction of model dimensionality can be obtained by formally integrating the incompressible fluid dynamics equations along the vertical dimension and by neglecting the vertical component of the acceleration. With this approximation, the pressure reduces to hydrostatic and the equation for the vertical component of momentum can be disregarded. In addition, the energy equation is often neglected, the flow being considered as isothermal.

On a flat surface, in a Cartesian coordinate system with x and y horizontal, z vertical, and where $h(x, y, t)$ is the flow depth parallel to z, the Saint-Venant's equations (also called *shallow-water* equations) can be simply derived and expressed as the balance equations of mass (Eq. 1) and momentum (Eqs. 2 and 3):

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu_x) + \frac{\partial}{\partial y}(hu_y) = 0 \tag{1}$$

$$\frac{\partial}{\partial t}(hu_x) + \frac{\partial}{\partial x}\left(hu_x^2 + \frac{1}{2}gh^2\right) + \frac{\partial}{\partial y}(hu_xu_y) = \frac{\tau_x}{\rho} \tag{2}$$

$$\frac{\partial}{\partial t}(hu_y) + \frac{\partial}{\partial x}(hu_xu_y) + \frac{\partial}{\partial y}\left(hu_y^2 + \frac{1}{2}gh^2\right) = \frac{\tau_y}{\rho} \tag{3}$$

where u_x and u_y are the components of the velocity vector **u**, g is the gravitational acceleration along z, ρ the average flow density, and τ_x and τ_y are the components of the resistive stress, usually neglected for fluids such as water.

For thin flows over a non-planar surface, the formulation is less simple since longitudinal driving forces and non-hydrostatic terms can be non-negligible. Two main approaches are used in the scientific literature to formulate the mathematical problem in such case. Savage and Hutter (1989) first introduced the use of a local, boundary-fitted coordinate system (Fig. 1 B). Its formulation is intuitive, but entails complex geometric transformations involving curvature terms (Iverson and Denlinger, 2001). The second uses universal Cartesian coordinates (Fig. 1 A) and does not require a geometrical transformation but implies non-hydrostatic (and non-hyperbolic) terms in the depth-averaged equations (e.g., Denlinger and Iverson 2004; Juez et al. 2013). The transformation between the two coordinate systems can be difficult in the case of a realistic topography but the two approaches are generally considered equivalent for simple geometries (i.e., for gentle slopes and small topographic variations), even though no rigorous comparison between the resulting models has been performed yet.
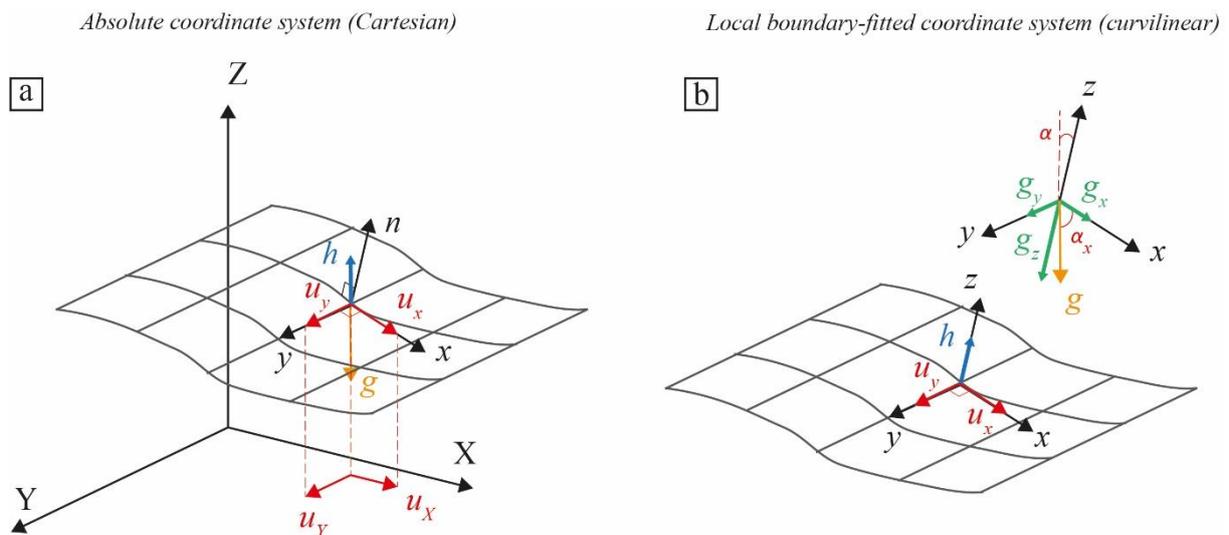
Fig. 1: The two different coordinate systems used for depth-averaged models: a. absolute coordinate system using the cartesian reference frame (X,Y,Z); b. local boundary-fitted coordinate system (x,y,z), tangent to the topography.

2.3. Rheology

The resistive stresses τ_x and τ_y (or shear stresses) are non negligeable in geophysical flows and rheological laws must be used to describe them. Due to the small number of parameters needed to reproduce first order dynamics of gravity-driven flows (Gruber and Bartelt, 2007; Hungr 2008; Fisher

et al. 2012; Lucas et al. 2014; McDougall 2017), two rheological laws are commonly used in the literature: (i) The Coulomb rheology that links the normal stress σ applied by the flow on the ground to its tangential stress (friction) τ by a friction coefficient μ, after a certain threshold C (i.e. cohesive stresses):

$$\tau = \mu\,\sigma + C \tag{4}$$

and (ii) the Voellmy-Salm rheology, developed initially for snow avalanches (Voellmy 1956; Salm et al. 1990), based on the Coulomb rheology with a velocity-dependent dissipative term added to account for particle collisions and interparticle frictions:

$$\tau = \mu\,\sigma + \frac{\rho g u^2}{\xi} \tag{5}$$

where ξ is the empirical Voellmy coefficient, ρ the flow density, g the gravity and u the flow velocity. In these two rheologies, the friction and Voellmy coefficients are constant during the flow emplacement and friction stresses only varies with the normal stress (i.e., flow thickness and flow velocity). The choice of a correct value for each of these rheological parameters is then crucial. The friction coefficient is usually estimated using the ratio H/L (vertical drop over horizontal length of the flow). Recent studies have shown that the coefficient decreases as the flow volume or runout distance increases (Charbonnier and Gertisser 2012; Lucas et al. 2014). The Voellmy coefficient is known to vary with topographic roughness, i.e., the higher the roughness, the higher the dissipative term, so the lower the Voellmy coefficient (Gruber and Bartelt, 2007; Fisher et al. 2012). In the case of CPCs, different studies have pointed out that a constant friction coefficient (i.e., Coulomb rheology) is too simplistic to accurately represent their complex behavior (Kelfoun et al. 2009; Kelfoun 2011; Moretti et al. 2012; Gueugneau et al. 2019). Other rheologies have been tested: some authors found that using a constant resistive stress allow to better model CPCs emplacement dynamics (Kelfoun, 2011; Gueugneau et al, 2017, 2019; Ogburn and Calder, 2017) while recent studies pointed out that the Voellmy rheology, with a velocity-dependent dissipative term, can also model first-order CPC dynamics (Kelfoun 2011; Salvatici et al. 2016; de'Micheli Vitturi et al. 2019; Patra et al. 2020). Even though all four selected models have the capability of using more complex rheologies to further explore the physics of PCs (which is beyond the scope of this paper), the Voellmy-Salm rheology was chosen for this benchmarking exercise.

2.4. Models in local boundary-fitted coordinates

In local coordinates, the same approach used to derive Eqs. (1)-(3) can be adopted, in which (x, y) represent, in each point of the domain, the coordinate directions tangent to the 2D surface,

while (h) is the flow thickness in the direction normal to the tangent plane (Fig. 1 B). The depth-averaged balance equation can be written in this reference frame as:

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu_x) + \frac{\partial}{\partial y}(hu_y) = 0 \tag{6}$$

$$\frac{\partial}{\partial t}(hu_x) + \frac{\partial}{\partial x}(hu_x^2) + \frac{\partial}{\partial y}(hu_x u_y) = S_x \tag{7}$$

$$\frac{\partial}{\partial t}(hu_y) + \frac{\partial}{\partial x}(hu_x u_y) + \frac{\partial}{\partial y}(hu_y^2) = S_y \tag{8}$$

The terms on the right side of the momentum balance equations (7) and (8) are source terms which correspond to the sum of all forces applied to the fluid, and can be expressed as:

$$S_i = \underbrace{g_i h}_{[1]} - \underbrace{\frac{1}{2}\frac{\partial}{\partial i}(g_z h^2)}_{[2]} - \underbrace{\tau_i/\rho}_{[3]} \tag{9}$$

with [1] the gravity acceleration component expressed along x and y axis, [2] the hydrostatic pressure acceleration component, and [3] the resistive stress component. The gravity vector $\vec{g}$ is expressed with its three components in the local coordinate system $\vec{g} = (g_x, g_y, g_z)$, as shown in Fig. 1B. The resistive stress τ, initially depth-averaged by Savage and Hutter (1991) in local coordinates for granular flows with a Coulomb rheology on an arbitrary topography, can be written as follows using the Voellmy rheology:

$$\tau_i = \frac{u_i}{\|\boldsymbol{u}\|}\left(\mu\rho h\left(g\,cos(\theta) + \frac{\|\boldsymbol{u}\|^2}{r}\right) + \frac{\|g\|}{\xi}\|\boldsymbol{u}\|^2\right) \tag{10}$$

where ρ is the flow density, h its thickness, and $\|\boldsymbol{u}\| = \sqrt{u_x^2 + u_y^2}$ is the norm of the flow velocity in the topography-linked coordinate system. The topography is implemented by two different elements: (i) the ground slope angle θ to ensure that the normal stresses stay normal to the local topography, and (ii) centrifugal acceleration effects caused by terrain curvature that can be approximated by the term $\frac{\|\boldsymbol{u}\|^2}{r}$, with r the local curvature radius in the direction of the flow, according to the scale analysis of Savage and Hutter (1989). The computation of r is not straightforward, and several approximations have been suggested, as described further. The exact expression of the curvature involves the topography curvature tensor and not only a scalar r (Mangeney-Castelnau et al. 2005; Mangeney et al. 2007; Peruzzetto et al. 2021). Two of the models described below and tested in this benchmark are based on the local-coordinate formulation.

VolcFlow (Kelfoun and Druitt 2005) was developed to simulate volcanic mass flows like debris avalanches, lahars, lava flows and CPCs but can also be applied to simulate other geophysical flows such as landslides and tsunamis (Giachetti et al, 2011). It relies on a finite-difference method that solves the hyperbolic part of partial differential equations (6) to (8) using a first or second order upwind

11

scheme. The model enables the choice of different rheologies from viscous rheologies like Newtonian and Bingham, plastic rheologies, or granular rheologies such as Coulomb or Voellmy. Hence, using the Voellmy rheology, the source terms can be written as:

$$S_x = g_x\boxed{?} - \frac{1}{2}\frac{\partial}{\partial x}\left(g_z h^2\right) - \frac{u_x}{\|\boldsymbol{u}\|}\left(h\mu\left(g_z + \frac{\|\boldsymbol{u}\|^2}{r}\right) + \varepsilon\|\boldsymbol{u}\|^2\right) \tag{11}$$

$$S_y = g_y\boxed{?} - \frac{1}{2}\frac{\partial}{\partial y}\left(g_z h^2\right) - \frac{u_y}{\|\boldsymbol{u}\|}\left(h\mu\left(g_z + \frac{\|\boldsymbol{u}\|^2}{r}\right) + \varepsilon\|\boldsymbol{u}\|^2\right) \tag{12}$$

where $\varepsilon = \frac{\|\boldsymbol{g}\|}{\xi}$, and g_x, g_y, g_z are the gravity components calculated following: $g_z = g\cos(\alpha)$, $g_x = g\sin(\alpha_x)$, and $g_y = g\sin(\alpha_y)$, with α the ground slope angle, α_x and α_y being the slope angles in the xz and yz planes respectively (see Fig. 1 B). In this code, the curvature radius r is calculated using the approximated formulation:

$$r = \sin(\alpha)\,\gamma_x + \cos(\alpha)\,\gamma_y \tag{13}$$

where γ_x and γ_x are the topography curvature in the direction x and y, calculated following:

$$\gamma_x = \frac{\partial^2 z}{\partial x^2} + \left(1 + \left(\frac{\partial z}{\partial x}\right)^2\right)^{-3/2} \tag{14}$$

$$\gamma_y = \frac{\partial^2 z}{\partial y^2} + \left(1 + \left(\frac{\partial z}{\partial y}\right)^2\right)^{-3/2} \tag{15}$$

The code, written in Matlab, has been verified in Kelfoun and Druitt (2005) and Kelfoun (2017), and confirmed/validated for CPCs in Kelfoun et al. (2009), Charbonnier and Gertisser (2012), Gueugneau et al. (2017) using an alternate version including pore pressure, and more recently in Kelfoun et al. (2017), Gueugneau et al. (2019, 2020) and Charbonnier et al. (2020) using the two-layer version of the code (Kelfoun 2017).

TITAN2D (Patra et al. 2005) was developed initially to model geophysical mass flows. It has been extensively used in volcanology to simulate CPCs. TITAN2D relies on a finite-volume method (i.e. fluxes are based on the centroid of each cells) and solves hyperbolic partial differential equations using a first or second order Gudonov scheme, for which a local grid refinement (Adaptive Mesh Refinement – AMR) is used to increase the accuracy of the simulation, while reducing the computational cost. The friction forces are also expressed following the Savage and Hutter approach with the Coulomb rheology but can also integrates the Pouliquen and Voellmy rheologies as well (Simakov et al. 2019). For the latter, source terms can be written as:

$$S_x = g_x\boxed{?} - \frac{1}{2}\frac{\partial}{\partial x}\left(g_z h^2\right) - \frac{u_x}{\|\boldsymbol{u}\|}\left(h\mu\left(g_z + \frac{\|\boldsymbol{u}\|^2}{r_x}\right) + \frac{\|\boldsymbol{g}\|}{\xi}\|\boldsymbol{u}\|^2\right) \tag{17}$$

$$S_y = g_y\boxed{} - \frac{1}{2}\frac{\partial}{\partial y}\left(g_z h^2\right) - \frac{u_y}{\|\boldsymbol{u}\|}\left(h\mu\left(g_z + \frac{\|\boldsymbol{u}\|^2}{r_y}\right) + \frac{\|\boldsymbol{g}\|}{\xi}\|\boldsymbol{u}\|^2\right) \tag{18}$$

where g_x, g_y and g_z are the projections of the gravity vector **g** along x, y and z axis. Here the curvature is also approximated, but in contrast to VolcFlow, two curvature terms are calculated in the x and y directions by a simplified formulation:

$$\frac{1}{r_x} = \frac{\partial^2 z}{\partial x^2} \tag{19}$$

$$\frac{1}{r_y} = \frac{\partial^2 z}{\partial y^2} \tag{20}$$

This code, written in C, has been verified in Patra et al. (2005), and confirmed/validated for CPCs in various studies, for example by Sheridan et al. (2004), Sulpizio et al. (2010), Capra et al. (2011), Charbonnier and Gertisser (2009; 2012), Stefanescu et al. (2012), Tierz et al (2018) and Patra et al. (2020).

2.5. Models in absolute coordinates

Use of an absolute Cartesian system (such as, for example, the Universal Transverse Mercator – UTM – coordinate system) facilitates the incorporation of a georeferenced topography and the formulation of the transport equations. However, treatment of non-hydrostatic terms can be problematic in case of complex geometries since it gives rise to non-hyperbolic terms in the transport equations (Bouchut and Westdickenberg, 2004; Denlinger and Iverson, 2004; Castro-Orgaz et al. 2014).

Two of the models tested in this benchmark are based on the global-coordinate formulation.

IMEX-SfloW2D (de'Michieli Vitturi et al. 2019) was developed to simulate geophysical mass flow over 3D topographies. The model adopts an absolute Cartesian (X, Y, Z) reference frame (Fig. 1 A), so that mass and momentum equations are integrated along the axis Z parallel to the gravity g. The present formulation of the code neglects non-hydrostatic terms associated with steep slopes and rugged topographies. Therefore, its formulation can be easily recast in a local coordinate system for gentle slopes. The model is based on the finite-volume method, and it is discretized in time with an explicit–implicit Runge-Kutta method, in which the hyperbolic part of the governing equations is solved explicitly with a second-order central-upwind scheme. The main novelty of this model is the implicit treatment of the source terms, which is a key feature to properly model flow stopping (when friction becomes dominant). The model integrates the topography as the function $b = b(X, Y)$, which can be imported as a georeferenced digital elevation model.

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial X}\left(hu_X\right) + \frac{\partial}{\partial Y}\left(hu_Y\right) = 0 \tag{21}$$

$$\frac{\partial}{\partial t}(hu_X) + \frac{\partial}{\partial X}(hu_X^2) + \frac{\partial}{\partial Y}(hu_X u_Y) + gh\frac{\partial(h+b)}{\partial X} = S_X \qquad (22)$$

$$\frac{\partial}{\partial t}(hu_Y) + \frac{\partial}{\partial X}(hu_X u_Y) + \frac{\partial}{\partial Y}(hu_Y^2) + gh\frac{\partial(h+b)}{\partial Y} = S_Y \qquad (23)$$

In the first version of the model, the code only integrates the Coulomb and Voellmy rheologies, and does not integrate non-hydrostatic corrections (i.e; curvature effect; will be added in the future). Therefore, source terms are written as:

$$S_X = -\frac{u_X}{\|\boldsymbol{u}\|}\left(h\mu g_n + \frac{g}{\xi}\|\boldsymbol{u}\|^2\right) \qquad (24)$$

$$S_Y = -\frac{u_Y}{\|\boldsymbol{u}\|}\left(h\mu g_n + \frac{g}{\xi}\|\boldsymbol{u}\|^2\right) \qquad (25)$$

where $\|\boldsymbol{u}\| = \sqrt{u_X^2 + u_Y^2}$ and g_n is the component of the gravity acceleration $\mathbf{g} = (0; 0; -g)$ along the surface normal vector $\boldsymbol{n}$, given by:

$$g_n = \mathbf{g}\cdot\mathbf{n} = \frac{g}{\sqrt{1+\left(\frac{\partial b}{\partial x}\right)^2+\left(\frac{\partial b}{\partial y}\right)^2}} \qquad (26)$$

The code, open source and written is FORTRAN90, has been verified and tested against standard unit problems and applied to the 2014 CPCs from Mount Etna in de' Michieli Vitturi et al. (2019) and to the 2008 Chaiten CPCs in Aravena et al (2020).

 SHALTOP was developed to simulate landslides and debris avalanches. The numerical method used to solve the hyperbolic equation system relies on a finite volume formulation (second-order upwind scheme) coupled with the apparent topography approach of Bouchut et al. (2003) to deal with friction (Mangeney et al. 2007). The detailed derivation of SHALTOP mass and momentum equations is given in Bouchut et al. (2003) for flows on 1D topographies, and in Bouchut and Westdickenberg (2004) for flows on complex topographies. A discussion on these derivations is provided in Peruzzetto et al. (2021). The integration of mass and momentum equations is performed in the direction normal to the topography, which requires the use of an appropriate frame linked to the topography (see Fig 1 A and Peruzzetto et al. 2021). However, the final equations for the depth-averaged velocity $\vec{V}$ are given in the fixed Cartesian coordinate system $(X, Y, Z) = (\mathbf{X}, Z)$, with $\mathbf{X} = (X, Y)$. In the following, the notation $\overrightarrow{}$ is used for 3D vectors, and the bold notation is used for 2D vectors. The slope θ of the topography $b = b(X, Y) = b(\mathbf{X})$ is given by:

$$\cos(\theta) = (1+ \|\nabla_\mathbf{X} b\|^2)^{-1/2} \qquad (27)$$

Then, $\mathbf{s} = \cos(\theta)\,\nabla_\mathbf{X} b$ such that the unit vector tangent to the topography is

$$\vec{n} = (-\mathbf{s}, \cos(\theta)) \qquad (28)$$

With these notations, $\vec{V}$ is parameterized with $\mathbf{u} = (u, u_t)$ following:

$$\vec{V} = (\cos(\theta)\,\mathbf{u}, \mathbf{s}^t\mathbf{u}) \qquad (29)$$

where the first two components are given by $\cos(\theta)\,\mathbf{u}$, and the last component of $\vec{V}$ is deduced from the constraint that $\vec{V}$ is tangent to the topography (that is, $\vec{V}^t\,\vec{n} = 0$). $\mathbf{s}^t\mathbf{u}$ is the scalar product of $\mathbf{s}$ and $\mathbf{u}$ (^{t} is the transpose operator). With these notations, the norm of physical depth-average velocity is given by:

$$\|\vec{V}\| = ((\cos(\theta)\,\|\mathbf{u}\|)^2 + (\mathbf{s}^t\mathbf{u})^2)^{1/2} \tag{30}$$

The formal derivation of the SHALTOP equations involves the topography curvature. The topography curvature tensor is given by:

$$\mathcal{H} = \cos^3(\theta)\begin{pmatrix} \dfrac{\partial^2 b}{\partial X^2} & \dfrac{\partial^2 b}{\partial X \partial Y} \\ \dfrac{\partial^2 b}{\partial X \partial Y} & \dfrac{\partial^2 b}{\partial Y^2} \end{pmatrix} \tag{31}$$

With these equations, SHALTOP solves the mass and momentum equation for $\mathbf{u}$ and h, the thickness of the material layer in the direction normal to the topography, following:

$$\frac{\partial}{\partial t}\left(\frac{h}{\cos(\theta)}\right) + \nabla_\mathbf{X} \cdot (h\mathbf{u}) = 0 \tag{32}$$

$$\frac{\partial (h\mathbf{u})}{\partial t} + \cos(\theta)\,\nabla_\mathbf{X} \cdot \left(h\,(\mathbf{u} \otimes \mathbf{u})\right) = \mathbf{S} \tag{33}$$

Hence, SHALTOP's source term $\mathbf{S}$ is written:

$$\mathbf{S} = \mathbf{F}_g + \mathbf{F}_\gamma + \mathbf{F}_{frictions} \tag{34}$$

with $\mathbf{F}_g$ are the gravity and lateral pressure forces (corresponding to the terms [1] and [2] in the equation (9)), and $\mathbf{F}_\gamma$ is the curvature force following:

$$\mathbf{F}_g = -\frac{h}{\cos(\theta)}\,(Id - \mathbf{s}\mathbf{s}^t)\,\nabla_\mathbf{X}\,\left(g(h\cos(\theta) + b)\right) \tag{35}$$

$$\mathbf{F}_\gamma = -\frac{h}{\cos(\theta)}\,(\mathbf{u}^t\mathcal{H}\mathbf{u})\mathbf{s} + \frac{h}{\cos(\theta)}\,(\mathbf{s}^t\mathcal{H}\mathbf{u})\mathbf{u} \tag{36}$$

With the Voellmy rheology, the friction forces of the flow are:

$$\mathbf{F}_{friction} = -\frac{h\mu g\cos(\theta)\mathbf{u}}{\|\vec{V}\|}\left(1 + \frac{\mathbf{u}^t\mathcal{H}\mathbf{u}}{g\cos(\theta)}\right) - g\frac{\mathbf{u}\|\vec{V}\|}{\xi} \tag{37}$$

Note that curvature effects influence two terms in SHALTOP: the friction forces $\mathbf{F}_{frictions}$ (as in the previous codes), but also the curvature force $\mathbf{F}_\gamma$. SHALTOP also allows for the selection of other granular rheologies like Coulomb or Pouliquen, or viscous fluid rheologies like Newton or Bingham. SHALTOP, written in FORTRAN90, has been used successfully to reproduce both granular flows at the laboratory scale (Mangeney-Castelnau et al. 2005; Mangeney et al. 2007) as well as real landslides deposits (i.e. Lucas et al. 2014; Brunet et al. 2017; Peruzzetto et al. 2019) and dynamics inferred from seismic recordings (Favreau et al. 2010; Moretti et al., 2015, 2020; Yamada et al. 2016). It has also been proven efficient to study CPCs induced by partial dome collapse (Levy et al. 2015).

III. Building a synthetic benchmarking procedure

Because such overspill events represent one of the deadliest and most unpredictable characteristics of CPCs, the source and boundary conditions used in this benchmarking procedure consider both the volume flux in the valley and synthetic channel topographies as modular input parameters.

3.1. Synthetic topographies

To investigate the role of various channel morphology on CPC models, five synthetic topographies were built for this benchmark (Fig. 2, Table 2): four of them contain a channel with a significant topographic feature (as described earlier) while the last one is an inclined plane, used as a control case for our four benchmarked models. These synthetic topographies are restricted to a rectangular domain of 5000 m long and 1500 m wide, enough to contain a valley and its surroundings, and with a scale similar to areas affected by small-volume CPCs, like block-and-ash flows (BAFs; Brown and Andrews, 2015). These topographies are generated numerically as Digital Elevation Models (DEMs), with regular grids of 1000 x 300 cells of 5 m spatial resolution. The synthetic longitudinal profile was simplified into a constant slope of 20°, obtained by averaging the H/L (vertical drop over horizontal length) ratio of 80 BAFs as found in the database FlowDat (Ogburn et al. 2012) and measured at Merapi, Unzen, Soufriere Hills, and Colima volcanoes. A single pseudo-sinusoidal valley (in cross-section) 80-120 m wide and 60 m deep is dug at the center of the domain. The synthetic topographies are:

- **(1) Inclined plane case (Fig 2a)**: rectangular and planar surface of 5000 by 1500 m with a 20° slope.
- **(2) Bend case (Fig 2b):** designed as a channel with two opposed 45° angles bends starting at 1500 m from the source, and distant of 500 m from each other. The channel depth (60 m) is not modified along the bends.
- **(3) Break in slope case (Fig 2c):** designed as a straight channel, but with a well-defined change of slope angle at 2500 m from the source, decreasing from 20° proximally to 10° distally.
- **(4) Obstacle case (Fig 2d):** designed as a straight channel composed of an obstacle located at 2500 m from the source, of 50 m long and 40 m high, corresponding to roughly two-third of the channel depth (60 m).
- **(5) Valley constriction case (Fig 2e):** designed as a straight channel with a sharp narrowing of its width from 1500 to 3000 m from the source, switching from a 120 m to 50 m wide channel cross section.
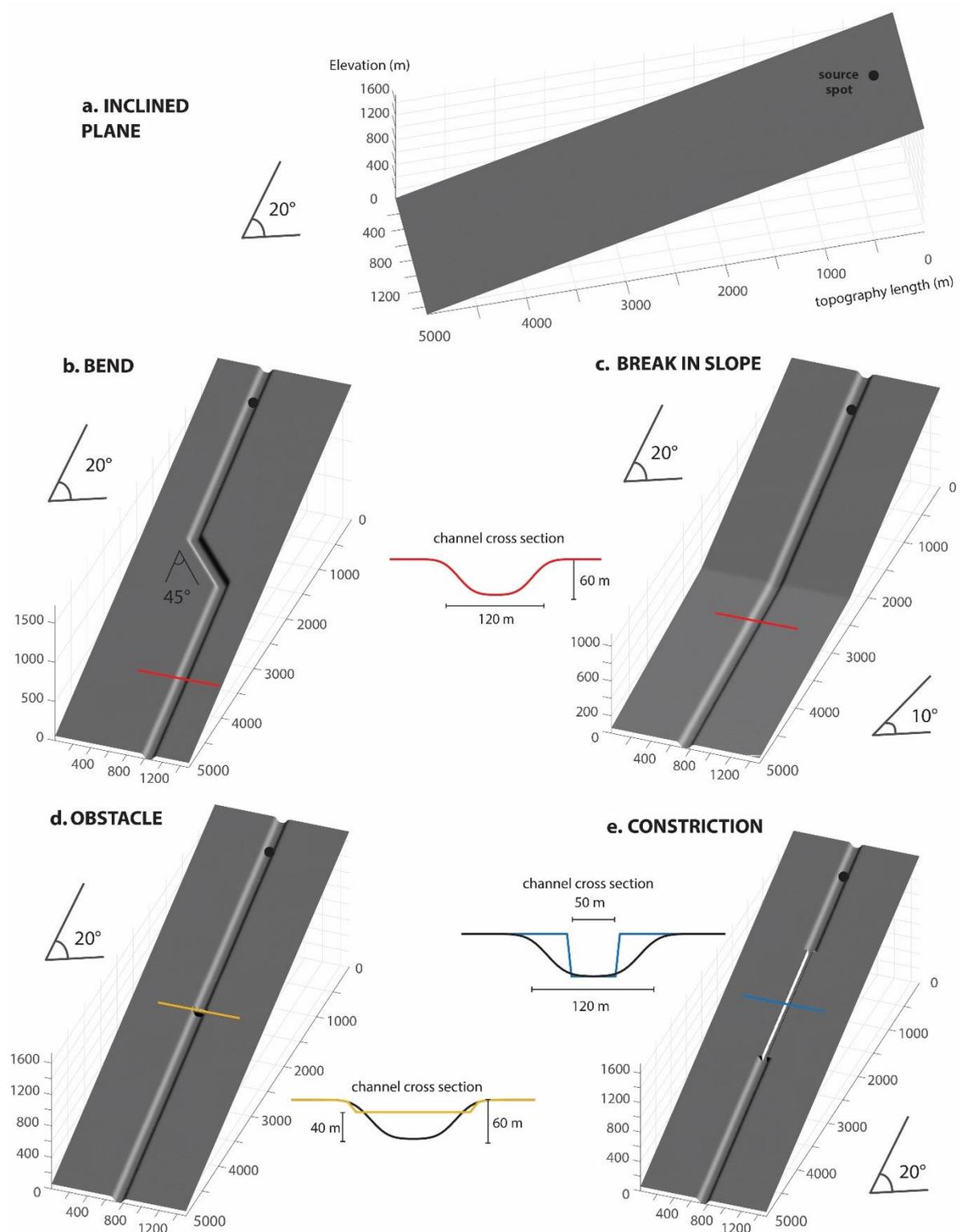
Fig. 2: Representation of the five synthetic topographies with their key features: a. the inclined plane case; b. the bend case; c. the break in slope case; d. the obstacle case and e. the constriction case. The channel morphology is shown on the side of each topography, along with its dimensions. See text for explanations.

Topographical parameters	Value selected
Number of columns	300
Number of rows	1000
Maximal x coordinate	1500 m
Maximal y coordinate	5000 m
Cell size	5 m
Maximal channel width	120 m
Maximal channel depth	60 m
Slope angle	20° (10° for break in slope)

Table 2: Summary of the selected dimensions for the four synthetic topographies.

3.2. Modeled CPC volume flux

To accurately evaluate CPC models in their capacity to overspill from a volcanic valley, different volumetric rates are set as fixed source conditions in our benchmarks to generate flows with different initial volume fluxes. A total volume of $V = 1 \times 10^6$ m^3 is selected, corresponding to the mean volume of 80 BAFs selected from the FlowDat database (Ogburn et al., 2012), with values ranging from 10^5 m^3 to 10^7 m^3. To input a volumetric rate in the models, the total volume is discretized into sub-volumes supplied at each time step, during a specific duration Δt. A decreasing volumetric rate at the source was chosen for all four models: the volume per time step decreases linearly from an initial volume V_{ini} to 0 during the duration Δt. Three different scenarios are defined, i.e. high, medium, and low, in which the total volume $V = 1 \times 10^6$ m^3 is supplied at three different rates, represented in Fig. 3, and summarized in Table 3.

Scenario	Initial volume V_{ini} (m^3 s^{-1})	Supply duration (s)
S1: low volumetric rate	20,000	100
S2: moderate volumetric rate	30,303	66
S2: high volumetric rate	60,606	33

Table 3: Description of the source conditions for each of the scenarios selected for the synthetic benchmarks.
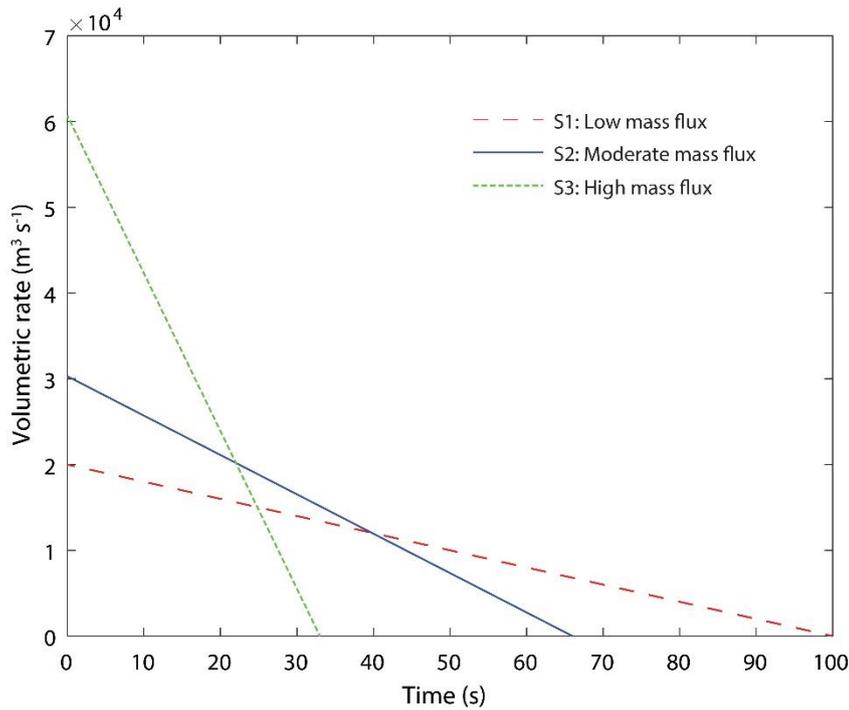
Fig. 3: Initial volumetric/mass fluxes set as fixed input parameters in all four models for the three scenarios considered in this benchmark (see also Table 3).

3.3. Procedure and inputs/outputs parameters

Each of the four selected models was evaluated on all five synthetic topographies. Thus, for each topographic case, the three volumetric rate scenarios (i.e. low, medium and high volume flux) were simulated independently, leading to a total number of 15 simulations per model, or 60 simulations for the entire benchmark exercise. For each simulation, the same input parameter values were used and kept constant to ensure each model run were performed under the exact same conditions. While a representative flow density is not required in these models (simplified in their system of equations), rheological parameter values, such as the friction coefficient μ and the Voellmy drag coefficient ξ, need to be defined. Since the Voellmy law uses empirical parameters, representative values of the two empirical coefficients were taken from a compilation of previous studies: Kelfoun et al. (2011) used the values 0.08 − 0.19; 10 m s^{-2} with VolcFlow for the friction coefficient μ and the Voellmy drag coefficient ξ, respectively, whereas de'Michieli Vitturi et al. (2019) used values of 0.1 − 0.4; 500 m s^{-2} with IMEX_SfloW2D, Salvatici et al. (2016) used values of 0.19; 1000 m s^{-2} with DAN3D, and Patra et al. (2020) used 0.5; 120 m s^{-2} with TITAN2D. Combining values from these previous studies, the average couple 0.2; 750 m s^{-2} was selected for this benchmark.

The source (as represented by black dots in Fig. 2) is approximated here as a circular spot with a 25 m radius sets at the center of the valley (center of the domain for the inclined plane) and at 500

m from the domain top boundary (to avoid back flow issues). Simple boundary conditions are considered, with free inlet/outlet for flows at the borders and absence of surface roughness. All the input parameters and boundary conditions are summarized in Table 4.

Input parameters	Common value
Volume	10^6 m^3
Gravity	9.81 kg s^{-2}
Friction coefficient μ	0.2
Voellmy drag coefficient ξ	750 m s^{-2}
Coordinates (x,y) of the center of the source	750, 500 (550, 500 for bend case)
Source radius	25 m
End time of simulations	100 s
Boundary conditions	
Bottom surface roughness	0 m
Top, Bottom, East, West boundaries	free inlet-outlet

Table 4: Summary of the selected input parameters and boundary conditions for each simulation.

The ability of our simulations to reproduce natural cases cannot be quantified because we do not have a reference case for each scenario. However, results from each simulation can be compared and the differences observed can be quantified. For consistency and to facilitate post-processing analyses, all simulations are manually stopped at 100 s and the following simulation outputs are used to evaluate the performance of each model:

- The maximum inundated area of the flow at t = 100 s
- The maximum runout of the flow at t = 100 s
- The evolution of the flow thickness through time at two locations of interest along the channel from t = 0 s to t = 100 s
- The evolution of the center of mass velocity through time from t = 0 s to t = 100 s
- The evolution of the front velocity through time from t = 0 s to t = 100 s

As all simulations are artificially stopped after 100 s (i.e., flows are still in movement), no stopping criterion was needed to be implemented. The computational setting (i.e. the real computational time on a desktop PC, size and format of the output files, etc.) is given in Table 5.

IV.   Benchmarking results

4.1. Inter-comparison procedure

Results of the model inter-comparison are presented for each topographic case in Figs. 4-8. The maps of maximum flow extent (a-d) are generated by extracting information from each pixel inundated by the flow at the end of the simulation (after 100 s). As some depth-averaged models tend to produce unrealistic, thin flow edges (thickness $< 10^{-6}$ m), a threshold of minimum flow depth is fixed here at 10^{-3} m. To quantify the differences between model results, two ratios are calculated and displayed in a table below each model's map (Figs. 4-8): the maximum area ratio $A_{X/R}$ and the maximum runout ratio $R_{X/R}$. Both are calculated by comparing the outputs of a reference model (area A_R and runout R_R) to those of the other models (A_X and R_X) following:

$$A_{X/R} = \frac{A_X - A_R}{A_R} \times 100 \tag{38}$$

$$R_{X/R} = \frac{R_X - R_R}{R_R} \times 100 \tag{39}$$

In order to reduce the amount of data displayed in each figure, percentages are only shown using the averaged area and flow runouts obtained for the three scenarios together. However, we invite the reader to refer to the supplementary material in which the complete data analysis for each model is provided (Supplementary Tables 1-4).

4.2. Inclined plane case

Simulations performed in the control case 'inclined plane' (Fig. 4) show a similar lobate shape and aspect ratios for the 4 models. An increase of roughly 5 to 25 % (depending on the model considered) in the maximum extent area and 5 % in the runout can be seen between the medium and the high volume flux scenarios. Flow simulation results with VolcFlow, SHALTOP and IMEX_Sflo2D show a good consistency: areas covered by the flows show only 17% maximum differences, and their runout only 10%. The velocity and thickness curves of these three models (Fig. 4) are almost superimposed, even though the IMEX_SfloW2D simulated flow is slightly faster than the other two, and consequently slightly more widespread and thinner. While velocities measured at the center of mass are very similar in the 3 codes, the front velocities are significantly different during the first 25 s (when the force balance dominates; see Fig. 13 in Mangeney-Castelnau et al. 2003), with much higher initial velocities obtained with SHALTOP and IMEX_SfloW2D. After ~25 s, SHALTOP and VolcFlow give very similar front velocities. However, TITAN2D simulation results strongly differ from those obtained with the three other codes: the simulated flows are much faster (up to two times in the high volume

flux scenario), causing a larger flow extent (up to 70%), and runout (41% to 45%), with lower thicknesses.

4.3. Bend case

Results of the topographic bend case (Fig. 5) highlight significant differences between models, especially for the maximum extent of the simulated flows. The four models overflow at the bend location, but at different scales. Differences are also visible between the 3 scenarios: the TITAN2D simulations easily overflow with the 3 different input volume fluxes, whereas the VolcFlow and IMEX_SfloW2D simulations only overflow in the high and moderate volume fluxes scenarios, while the SHALTOP simulations only do it in the high volume flux scenario. If we arbitrarily consider the VolcFlow simulation as a reference for a comparison with an averaged maximum flow extent of 1.97×10^5 m² (see supplementary materials Table 3), the TITAN2D simulations cover an averaged surface up to 617% larger (7 times, 14.1×10^5 m²) partially due to the presence of a large overspill at the source in the high volume flux scenario that produces an overbank flow outside of the channel, traveling 1400 m downstream (Fig. 5), whereas the SHALTOP and IMEX simulated flows inundate an averaged area of only 16% (2.29×10^5 m²) and 52% larger (2.99×10^5 m²) than the VolcFlow ones, respectively.

The evolution of simulated thicknesses and velocities with time (graphs in Fig. 5) follows almost the same general pattern, especially for VolcFlow, SHALTOP and IMEX, with center of mass velocity curves that are almost superimposed (Fig. 5). Simulated flows accelerate until they reach the bend, and then decelerate until the end of the simulation. Similarly, flow thicknesses also increase sharply to reach values of 9 to 12 m at location 1, and then follow a linear decrease until the end of the simulation, as the mass in the channel is drained and accumulates at the front to build a lobe (not seen in locations 1 or 2). However, simulated flows reach location 1 at different times, with the IMEX ones always arriving first, then the SHALTOP ones always 5 s later and finally the VolcFlow ones 8 s later. This trend is coherent with the observed variations in velocities (from the front or center of mass). Note that, on the contrary, SHALTOP flows were slower than the VolcFlow ones in the inclined plane case. TITAN2D simulated flows show the same general pattern than those from the three other codes, but its center of mass velocities is marked by a sudden acceleration during the first few seconds of the simulations in all three scenarios, leading to velocities 1.5 to 2 times higher. Consequently, TITAN2D generates large overflows (2 to 4 m thick) in all three scenarios and its peak in flow thicknesses at location 1 occurs 20 to 25 s earlier than those from the three other models.

In summary, all simulated flows with the four models show some interactions with the synthetic bend and two trends emerge: (i) TITAN2D flows travel generally faster than the other ones and produce major overflows in all three scenarios, and (ii) VolcFlow, IMEX_SfloW2D and SHALTOP

flows all travel slower than the TITAN2D ones and produce overspills with a limited extent only in the high or moderate volume flux scenarios.
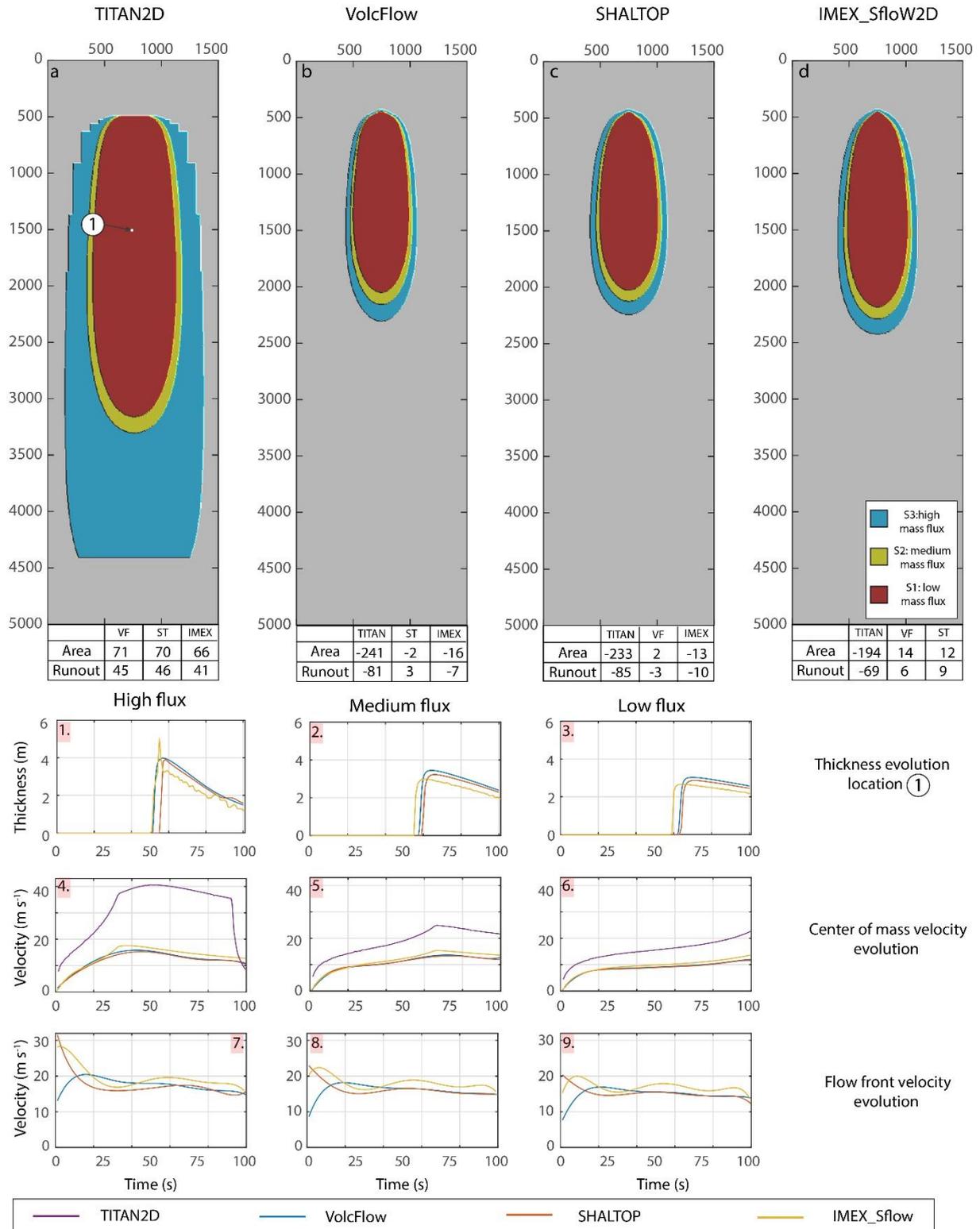
Fig. 4: Results of the model inter-comparison for the inclined plane case. Maps a. to d. show the inundated areas from each model (red colormap for the low volumetric rate, green colormap for the medium volumetric rate, and blue colormap for the high volumetric rate). Graphs 1 to 9 corresponds to the time-varying parameters for each scenario and model, with the flow thickness (1 to 3) at location 1 (shown in the top left map), the center of mass velocity (4 to 6) and the front velocity (7 to 9).

23
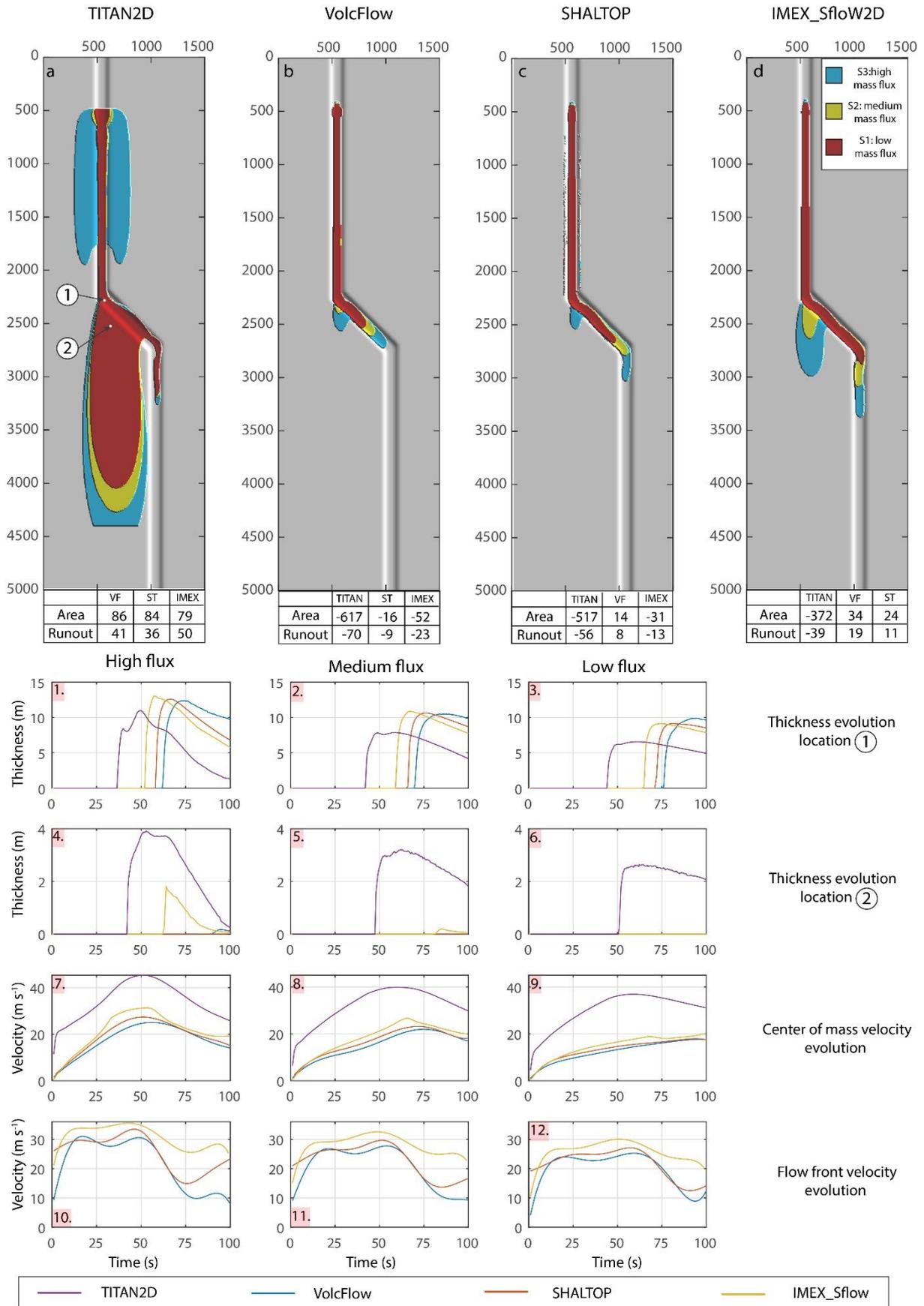
Fig. 5: Results of the model inter-comparison for the bend case. See Fig. 4 caption for details.

4.4. Break in slope case

Results of the break in slope case (Fig. 6) show more consistency between models than in the previous case. As an example, the runout differences between TITAN2D and the other codes are now only between 21–28% (41–50% previously), depending on the scenario considered. Differences obtained between scenarios are also limited to the simulation runouts (5-8%). No overflow occurs and flows stay channelized, except for TITAN2D in the high volumetric rate scenario where overflows occur at the source. When the slope angle is divided by two (from 20° to 10°) at 2500 m from the source, the flow front stops and a frontal lobe starts to form (i.e., graphs 4 to 6, Fig. 6), attested by a sudden drop of both front and center of mass velocities, associated with an increase of the flow thickness in the three scenarios at location 2. We note that the 10° slope is close to the friction coefficient of the simulation flows (11°), potentially explaining formation of a frontal lobe. TITAN2D simulated flows show similar high initial accelerations as those already observed in the bend case, which shifts the center of mass velocity curves up by 20 m s^{-1} compared to those from other models. As a result, VolcFlow, SHALTOP and IMEX_SfloW2D simulations start to build a frontal lobe deposit immediately after the break in slope (front velocity drops; flow thickness at location 2 increases), whereas for TITAN2D simulations such a frontal deposit does not build immediately after the break in slope but at least 1000 m further downstream (i.e., flow thickness at location 2 does not increase). Although no overflow is observed here, the break in slope did modify the flow dynamics in all four models. Similar to the bend case, TITAN2D simulated flows show again higher velocities, longer runouts compared to those from VolcFlow, SHALTOP and IMEX_SfloW2D.
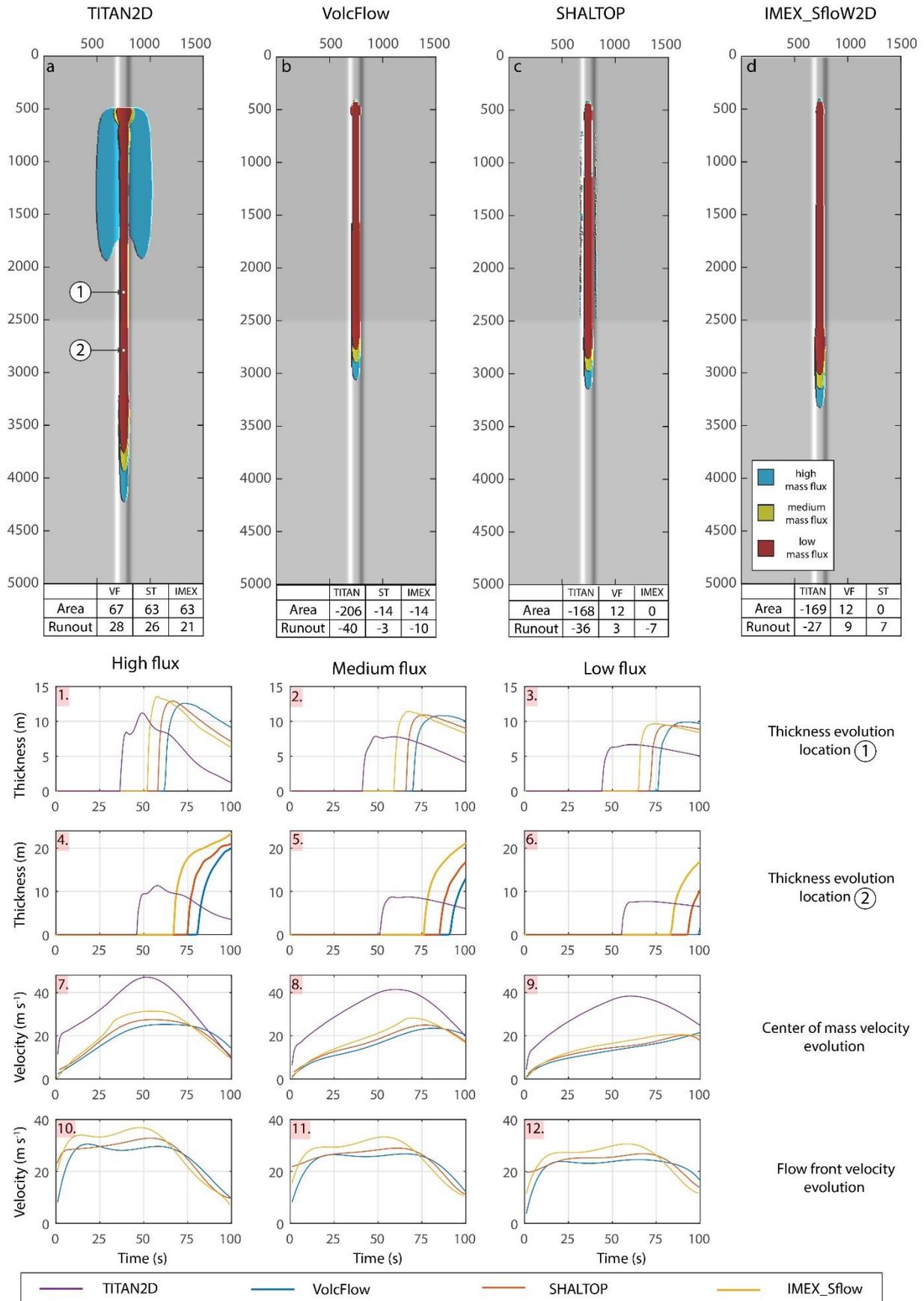
Fig. 6: Results of the model inter-comparison for the break in slope case. See Fig. 4 caption for details.

4.5. <u>Obstacle case</u>

In the obstacle case (Fig. 7), no overflow is observed but the differences in runout between simulated flows from each different model are more important than in the two previous cases. Here, SHALTOP simulated flows show the shortest runouts, as they stop at the foot of the obstacle in the low and moderate volume flux scenarios and travel only 250 m after passing over the obstacle in the high volume flux scenario. As a result, TITAN2D simulated flows have an averaged runout 81% longer than those from SHALTOP, with 17% and 38% longer runouts for those from VolcFlow and IMEX_SfloW2D, respectively. The TITAN2D flows also do not show any sign of flow accumulation prior to or after the obstacle, since both thicknesses and velocities remain constant or even decrease toward the end of the simulation (graphs 1 to 6 in Fig. 7). As a result, they reach the bottom edge of the DEM with runouts exceeding 4000 m in the 3 scenarios (note that for TITAN2D flows, cropping due to the AMR does not allow them to reach > 4500m). VolcFlow, SHALTOP and IMEX_SfloW2D simulated flows show similar patterns but with larger differences than in the previous cases. Their flow thicknesses evolution at location 1 show two phases (graphs 1 to 3 in Fig. 7): (i) one peak when the flow front reaches the probed location 1 (roughly 8 to 10 m thick, similar to the TITAN2D flows), (ii) a second peak a few tens of seconds later that reaches more than 40 m thick (similar to the obstacle height), corresponding to the accumulation of mass at the foot of the obstacle that fills the channel until they reach the top of the obstacle. Surprisingly, IMEX and VolcFlow flows reach location 2 (graphs 4 to 6 in Fig. 7) before the mass accumulation peak (second peak location 1), so that their flow fronts already passed the 40 m high obstacle when the mass accumulation phase starts (note that the enlargement of the flow around the obstacle is not a sign of an overflowing around it but simply due to the filling of the channel before the obstacle). This is also confirmed by their front velocities that do not decrease significantly at the passage of the obstacle. It is not the case for the SHALTOP flows that reach location 2 a few tens of seconds after the mass accumulation peak (or never reach it, like in the case of the lower volume flux scenarios).

To summarize, the presence of a large bump/obstacle obstructing the channel, perpendicular to the greatest slope gradient, seems to affect the four models considered here differently: (i) TITAN2D flows do not really interact with the obstacle and do not show any sign of mass accumulation in its vicinity, (ii) VolcFlow and IMEX_SfloW2D flows seem to moderately interact with it and exhibit a delayed mass accumulation after their fronts already passed the obstacle, and (iii) SHALTOP flows highly interact with the obstacle and are unable to cross it if the mass accumulating at its foot does not exceed its top height.
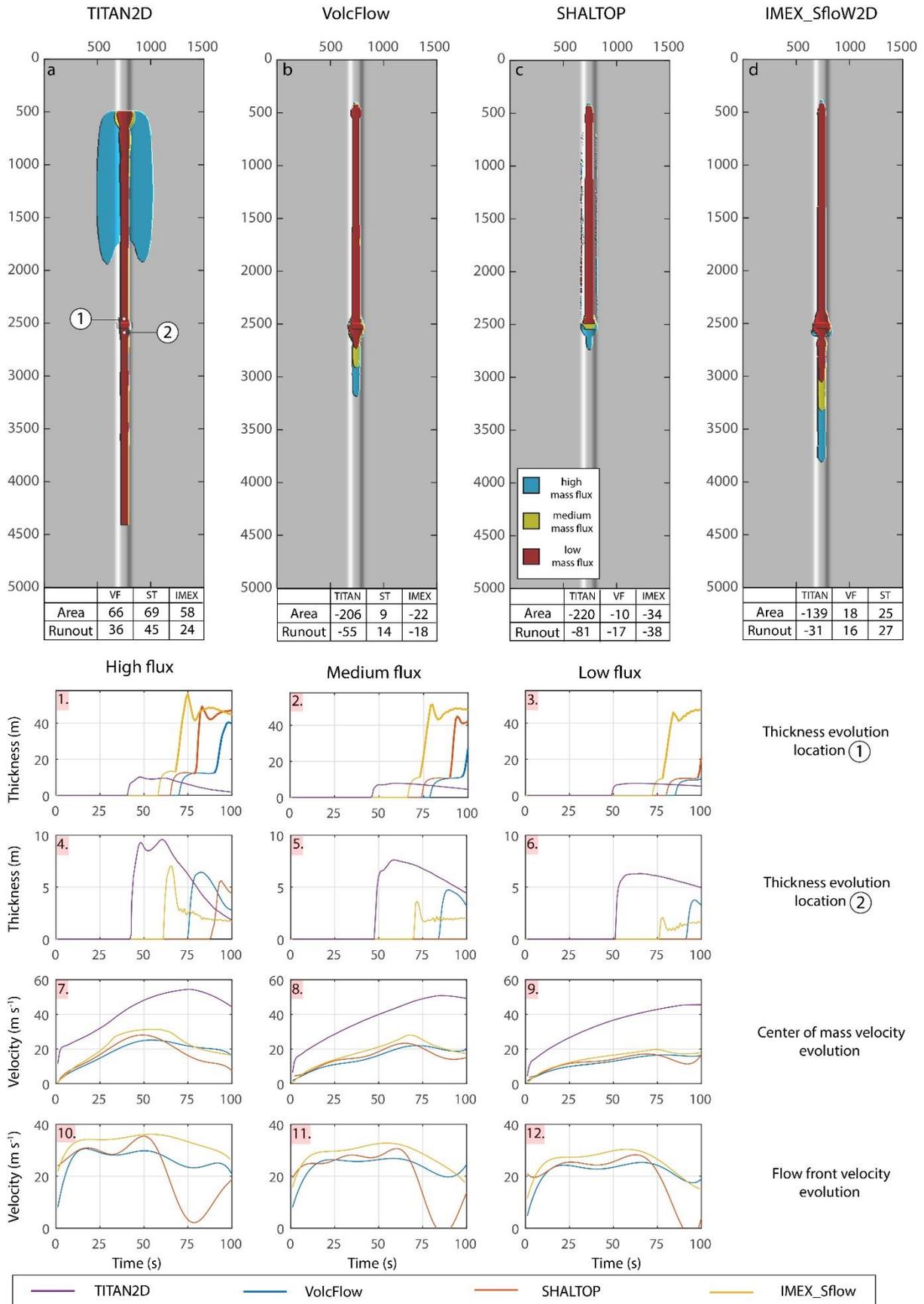
Fig. 7: Results of the model inter-comparison for the obstacle case. See Fig. 4 caption for details.

4.6. Constriction case

In the constriction case (Fig. 8), TITAN2D simulations are not shown because they all ended unexpectedly with an error generated at the constriction point. VolcFlow is the only model that generates a small overflow in all three scenarios after the simulated flow encounters the constriction. The area inundated by this overflow is limited (5.85×10^3 m² maximum, corresponding to only 3 % of the maximum extent area, see supplementary material table 2) as the overbank flow does not spread much laterally away from the channel (< 40 m). However, such overflows of the Volcflow simulations cause a loss of flow momentum, as shown by a drop in the both front and center of mass velocities stronger than in the other models (see graphs 1 to 3 in Fig. 8), reducing the flow propagation into the constriction. In consequence, VolcFlow simulated flow runouts hardly exceed 2000 m, even in the high volume flux scenario, which is 46% and 56% shorter than those of SHALTOP and IMEX_SfloW2D, respectively. SHALTOP simulated flows do not show any overflow at the constriction point (location 1 in Fig. 8) but exhibit a sudden mass accumulation when flows enter the constriction, as shown by the small deceleration of both front and center of mass velocities (graphs 7 to 12 in Fig. 8). Interestingly, this change in flow dynamics at the constriction point is followed by a strong re-acceleration (from 30 to 51 m/s in the next 20 s for the high volume flux scenario) followed by a strong deceleration at the end of the constriction when the channel width increases (graphs 4 to 6 in Fig. 8). These complex flow dynamics result in long SHALTOP simulated flow runouts, exceeding 3000 m for the high volume flux scenario. We also note that a gap in each SHALTOP simulated flow is visible inside the constriction part, and such gaps are interpreted here as mass flow separations, such processes being commonly observed with depth-averaged models simulating flow over complex topographies (Levy et al. 2015). Finally, while IMEX_SfloW2D simulated flows do not accumulate mass at the constriction point, its velocities slowly decrease, causing an increase of the flow thickness at the exit (graphs 4 to 6 in Fig. 8) and maximum flow runouts similar to the SHALTOP ones.

In conclusion, results of the model inter-comparison in the constriction case exhibit the most complicated results among the four topographical cases. Overall, the reduction of the cross-sectional area of the channel (by modifying only its width) seems to drastically modify the simulated flow dynamics with all three models, as well as to generate an overflow with VolcFlow.

4.7. Computational performances and usability

In order to give a representative computational time scale and performance comparison for all four models in our benchmark cases, all simulations with TITAN2D, SHALTOP and VolcFlow were performed on the same computer (i.e., a desktop PC equipped with a quad-core (8 threads) i7-4770K 3.5 GHz CPU, 16 GB of RAM, and a 1TB SSD), while IMEX_SfloW2D simulations were performed on a
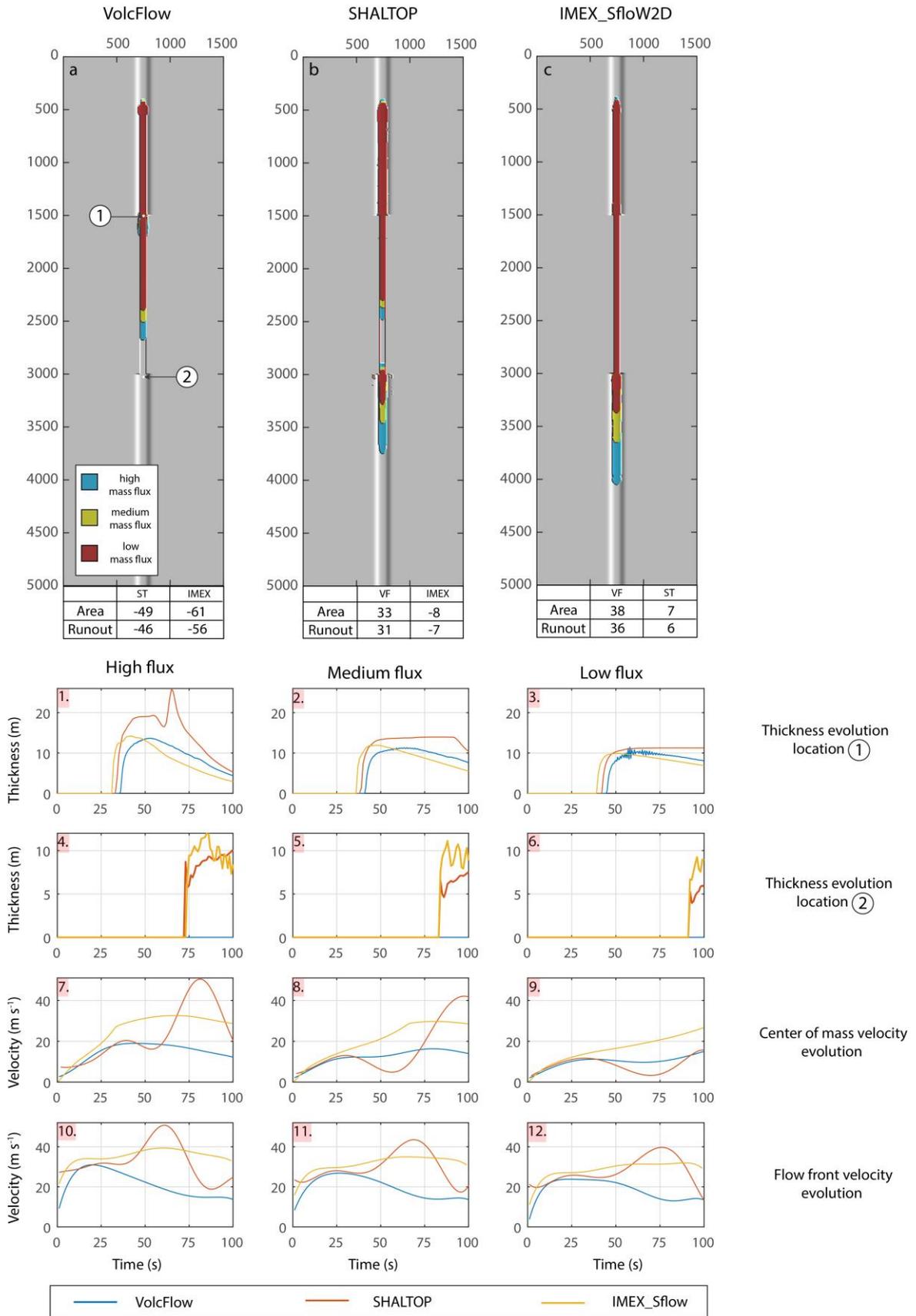
Fig. 8: Results of the model inter-comparison for the constriction case. See Fig. 4 caption for details.

laptop computer with similar specifications than the first PC. All simulations were run as scripts and no visual representation were activated so that all the computer resources were fully dedicated to the modeling tasks. Simulation time steps are adjusted automatically by each code, and data are saved every second (except for TITAN2D for which every time step is saved automatically). For consistency, the computational time is given for simulations performed with a second order scheme, except for TITAN2D. In addition to the computational time, output data size and format, as well as the visualization method used to analyze each model output are given in Table 5. For the data size, only the data related to the flow thicknesses and velocities are recorded to minimize the calculation time. A short and qualitative comparison of the usability of each model, summarizing the key advantages and drawbacks of each code for a first-time user, is listed in Table 5. The benchmarks proposed here can be completed in the future by other models that might be used in CPC hazard and risk assessment. All the material needed for completing these synthetic benchmarks, including the procedure, the DEMs used and the results from all four models are available upon request at:

https://vhub.org/groups/benchmarking_models

Models	Computational time (min)	Output Data size and format	Visualization	Usability
VolcFlow (Matlab, on Windows or Linux)	1.5 - 8	1.5 GB .dat + exportable in any format (Matlab)	Directly on matlab + Post simulation processing with Matlab (.dat)	+ Fast and easy to use, visualize and analyze, real time visualization possible - Matlab dependent, not open source
TITAN2D (Online or Linux)	51 – 97	1 – 2.5 GB Raw data + ASCII files for outputs at last time step (+ various file formats for outputs at user-defined time steps)	Post simulation processing with Paraview (.h5), GRASS-GIS or Tecplot (.tec)	+ Directly exportable on a GIS software (Georef. data), available online with a GUI - complex and large data files
SHALTOP (Linux)	39 - 47	0.5 GB Raw data (text) and .dat	Post simulation processing with Matlab/Python	+ Easy to use, visualize and analyze (simple raw data) - slower, cannot use simple DEM grid for the topography, needs to be processed

		1 GB	Post simulation	+ Easy to visualize and analyze
IMEX_SfloW2D	1.5	Raw data (text) and	processing with	(simple raw data), very fast, open
(Linux)		ASCII files	Matlab/Python	source
				- Not user friendly (no GUI)

Table 5: Comparison of performances and outputs of the four codes used in the synthetic benchmarks. The detailed description of the conditions for which these data were obtained is given in the text.

V. Discussion and perspectives for hazard and risks

Results of the synthetic benchmarks performed with four depth-averaged models highlight their abilities to simulate the interaction of CPCs with various channel morphologies, but some discrepancies between the simulation results are noticeable. While all simulations were based upon the same (i) source conditions, (ii) digital topographies, and (iii) flow rheologies, output parameters obtained with the four different codes show important variability. It is worth mentioning that the magnitude of these differences is associated to a specific value of each of the two rheological parameters and cannot be generalized. Thus, only the causes of these differences will be discussed here. Two groups of models can be distinguished: 1) VolcFlow, IMEX and SHALTOP simulations that highly interact with topographic features and give similar (but not identical) flow velocity, thickness and aerial distribution, and 2) TITAN2D simulations with limited topographic interaction, higher velocities, greater inundated area and longer runouts than the three other codes.

5.1. The curvature effects

Results of the four channelized cases (Figs 5-8) display a similar trend for the first model group: IMEX_SfloW2D always reaches the channel topographic feature (location 1) first as its front and center of mass velocities are the highest, followed by SHALTOP with moderate velocities and then VolcFlow in third with the lowest velocities of the three. Interestingly, when simulations are performed on the inclined plane (Fig. 4), there are much less variability and VolcFlow is slightly faster than SHALTOP. The presence of a terrain curvature (i.e. the channel) and rapid topographic changes seem to impact each model differently, with SHALTOP interacting much more with the topography.

With depth-averaged models, Patra et al. (2020) and Peruzzetto et al. (2021) have demonstrated that curvature effects have a limited influence on the flow dynamics when the Voellmy rheology is used, as the velocity-dependent stresses represent a significant contribution of the total resistive stresses with such slope (20°). However, in our four benchmark cases (Figs. 5 to 8), the channel is narrow and curved, and the simulated flows encounter sudden changes in channel morphology

and/or slope during propagation. Such topographic changes increase the centrifugal acceleration and modify the associated resistive term in the basal friction force (see Eq. 10). The four models used in this study do not simulate these effects the same way: IMEX_SfloW2D does not consider any curvature effect, VolcFlow and TITAN2D consider an approximation of the terrain curvature in the directions parallel to the flow (friction term Eqs. 13-15 and 19-20 respectively), and SHALTOP considers the full curvature tensor in the friction term (Eq. 31) plus the curvature force in Eq. 36). Hence, the different formulation of the terrain curvature might be responsible for a non-negligible part of the observed differences in our modeling results. Nevertheless, since the TITAN2D simulations also show important discrepancies with the ones from the three other codes in the case of the inclined plane (that does not contain any terrain curvature), the analysis of TITAN2D results will be treated separately in the section 5.3.

Peruzzetto et al. (2021) demonstrated that expressing the terrain curvature for a channelized flow with an approximated formulation like VolcFlow (Eqs. 13-15) breaks the rotational invariance of the equations and generally leads to higher resistive stresses, in comparison to models using the full curvature tensor, as SHALTOP. Hence, this could explain why VolcFlow simulated flows have the slowest front velocity, the shortest runouts and the smallest inundated areas. To verify this hypothesis, additional simulations were performed only with VolcFlow and SHALTOP and presented in Fig. 9: (i) simulations without any curvature effects, similar to IMEX_SfloW2D (i.e. curvature in the friction force is null for both models, and in SHALTOP), (ii) simulations with the approximated curvature formulation Eq. (13) in the friction force (the curvature force $\mathbf{F}_\gamma$ in SHALTOP is still kept null). To better highlight the differences, simulations were performed only for the two benchmarks with the sharpest curvature variations, i.e. the bend case and obstacle case, using the highest volume flux scenario (S3).

Results from the simulations performed with the 'no curvature' condition (Fig. 9 left side) show that for the bend case, removing curvature effects increases runout and inundated area compared to simulations with the 'exact' curvature condition in Fig. 5 (also reported in Fig. 9 as white dashed lines). In SHALTOP specifically, this could be explained by the fact that the curvature force maintains the flow in the central channel by reducing the bouncing effect, and thus reduces overflow (see Figs 7b, 7c; Peruzzetto et al. 2021). The differences between these two models (VolcFlow, SHALTOP) and IMEX-SfloW2D are then reduced by 30% for the inundated area and 50% for the flow runout, respectively. Simulation results obtained in the obstacle case with the 'no curvature' condition does not seem to significantly modify the runout of the two models. However, it seems to affect the overspill of SHALTOP's simulated flow at the obstacle location by increasing its inundated area by 19%. In the latter case, the presence of a straight channel parallel to the flow propagation direction most likely reduces
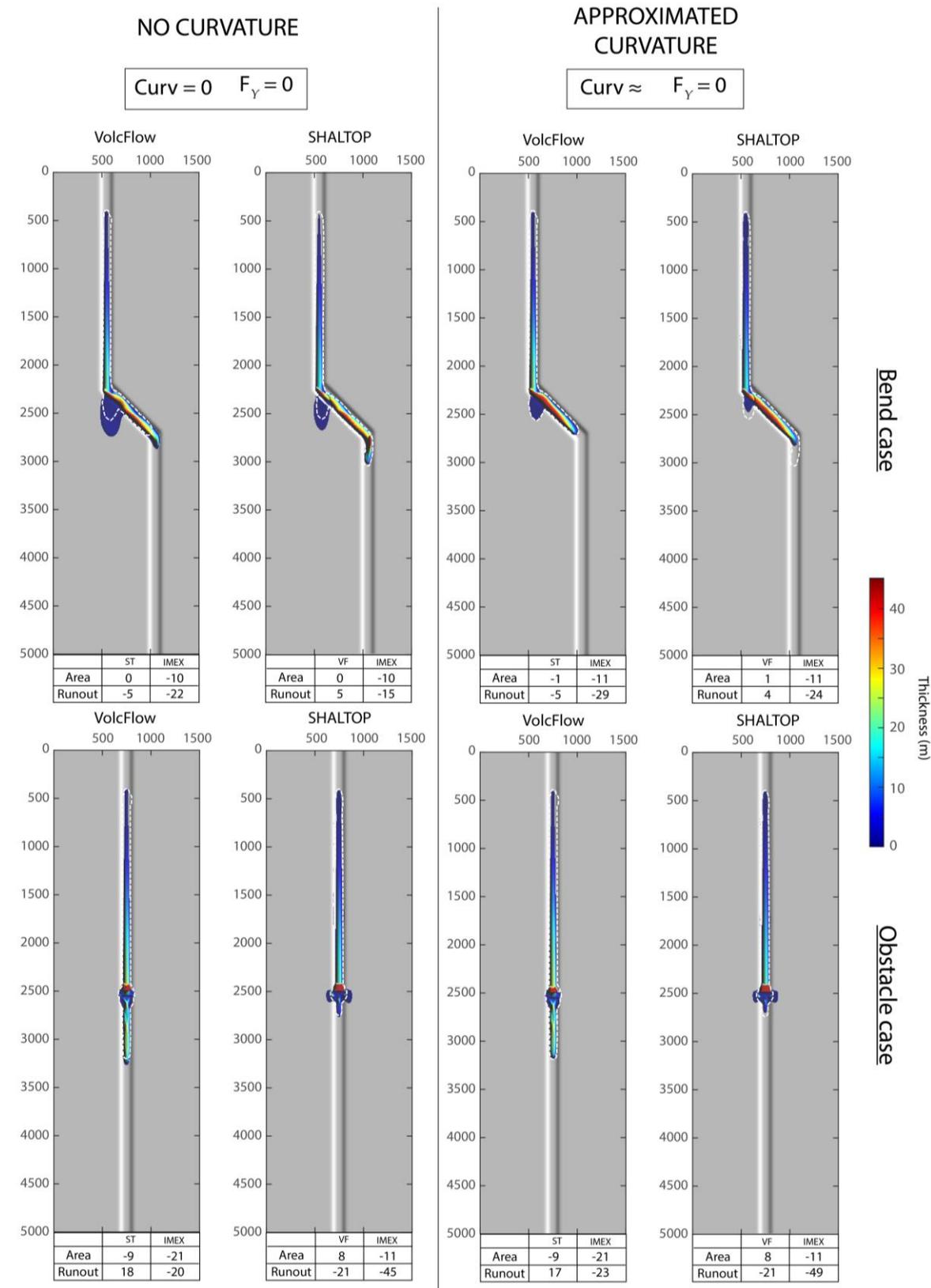
Fig. 9: Results of the complementary benchmark between VolcFlow, and SHALTOP for the bend case (top half) and the obstacle case (bottom half). The left part of the figure shows the results of simulations without any curvature effects implemented in the models, whereas the right part shows the results of simulations with an approximated curvature implemented in VolcFlow and SHALTOP. The colormap of each simulation refers to the thickness distribution of the flow after 100s of simulation time. White dash lines show the flow outlines with the curvature effects, as shown in Figs. 5 and 7.

the role of curvature effects during flow emplacement, even though the curvature force increases both the front and center of mass velocities, as discussed above (Peruzzetto et al., 2021).

Using the 'approximated curvature' condition (i.e., in the friction force) in the bend case (Fig. 9 right side), the SHALTOP simulated flow has a shorter runout than simulations with the 'exact' curvature condition (white dashed lines), providing results closer to the VolcFlow simulations (only 1-5% differences). Such a condition also does not really modify the SHALTOP results in the obstacle case. This can be related to the orientation of the channel. In the obstacle case, the channel is aligned with the y axis, such that the approximated expression of the curvature in Eq. (13) in the friction term is consistent with the 'exact' expression. This shows that the overflow observed at the obstacle in SHALTOP simulations results from the lack of the curvature force and not from the lack of curvature in the friction term. In the bend case, the channel is rotated by a 45° angle. At the bottom of the channel, the curvature in the flow direction is zero, but γ_x and γ_y in Eqs. (14, 15) are positive. In turn, the approximated curvature in Eq. (8) is also positive, which artificially increases friction. Consequently, the approximation of the curvature can indeed reduce the runout and inundated area of the simulated flow in a channel not aligned with the referential axis, which is consistent with the results of Peruzzetto et al. (2021).

In summary, results of our benchmark confirmed the conclusions of Patra et al. (2020) and Peruzzetto et al. (2021) that the impact of the curvature effects with a Voellmy rheology on channelized flow propagation is limited on smooth topographies but can be non-negligible in case of sudden topographic changes. Even though our simulations were performed on simplified topographies, results highlight the importance of considering such centrifugal acceleration effects when dealing with CPCs. Such curvature conditions for CPCs should be non-approximated and invariant in rotation, as these flows are usually emplaced in sharp and tortuous valleys.

5.2. Differences in flow velocity calculation

Even when curvature effects are disabled, SHALTOP flows do not interact the same way with the obstacle as those of VolcFlow and IMEX_SfloW2D (see Figs. 7 and 9). Difference in the mathematical models could be one explanation. Indeed, in VolcFlow the projection of the gravity terms is not performed in an orthonormal system and in IMEX_SfloW2D, the shallow approximation and the depth-averaging are performed in the vertical direction, which significantly increases the runout distance (see figure 6a in Delgado-Sanchez et al. 2020). This results in differences in how the source terms are solved in each code. Because of the different integration methods used for the momentum and mass balance equations (Eqs. 1-3) in each code, the resulting flow velocity is not equivalent for all four models: (i) VolcFlow and TITAN2D use a two-components velocity in the plane x,y tangent to the

slope (see Eqs. 11-12, 17-18 and Fig. 1 B), (ii) IMEX_SfloW2D uses a two-components velocity in the horizontal plane X,Y perpendicular to $\vec{g}$ (see Eq. 24-25, and Fig 1 A), and (iii) SHALTOP uses the physical flow velocity, which has three components in the X,Y,Z coordinate system (see Eq. 30). As detailed in Savage and Hutter (1989, 1991), in the depth-averaged approach the resistive stresses τ_x and τ_y are tangent to the slope. This implies that the flow velocity must remain tangent to the topography, colinear to τ_x and τ_y. For cases (i) and (ii) the tangentiality is easily achieved over planar or smooth topography but not over complex, rough topographies where non-hydrostatic forces like the 'curvature force' arise (Bouchut and Westdickenberg 2004; Iverson and Denlinger 2004; Peruzzetto et al. 2021). A result of this is that resistive stresses calculated in cases (i) and (ii) can be underestimated when flows encounter rapid topographic changes. In case (iii) the use of the physical velocity enables the calculation of the 'curvature force' $\mathbf{F}_\gamma$, which ensures that the flow velocity stays tangent to the topography. This could potentially explain why SHALTOP simulations barely cross the obstacle in Figs. 7 and 9: when the flows reach the obstacle, the curvature force $\mathbf{F}_\gamma$ becomes high, and the frictions increase. This causes the flows to abruptly decelerate (see front velocities graphs 10 to 12 in Fig. 7), while the curvature force helps the flows to stay confined in the channel and accumulate at the foot of the obstacle before eventually overtopping it. In summary, the use of a two-components slope-tangent velocity to calculate the friction terms in depth-averaged models seems coherent for smooth or planar topographies (i.e. the break in slope and inclined plane cases in our benchmarks), but can lead to important discrepancies with models using a three-components velocity over more complex topographies (i.e. the obstacle and bend cases). In particular, the curvature effects in SHALTOP (case (iii)) makes it much more sensitive to topography variations.

5.3. Numerical framework

While terrain curvature and the flow velocity treatment explain some of the differences observed in the benchmark results, some discrepancies remain unexplained. For instance, IMEX_SfloW2D flows are always 5 to 15% faster than VolcFlow and SHALTOP flows (center of mass velocity), even in the inclined plane case. These limited residual differences can be attributed to the different numerical framework used by each code. However, terrain curvature and the flow velocity calculation do not explain why TITAN2D flows behave so differently than the others. The most remarkable characteristic of TITAN2D flows is the fast acceleration during the first computational steps, shifting up velocity curves by 20 m s⁻¹ or more compared to those from the other three models, regardless of the scenario chosen. Without such a large initial acceleration, TITAN2D simulation results could be comparable to those from the three other codes, at least for the evolution of flow velocities (see graphs 7 to 12 in Figs. 5-8). After verification of the TITAN2D source code used for this study (4.1.0), the cause of this rapid acceleration was identified and attributed to an error in the numerical implementation of the

Voellmy-Salm rheology. A new version of the source code (4.2.0) is currently under development and was tested here for the bend case with the low volume flux scenario (see Supplementary Material Fig. S1). The new simulated flow velocity and resulting inundated area seem to be closer to those obtained with the three other codes, but some instabilities remain (see velocity curve) and must be improved. The corrected version of the code will be released soon after undergoing further confirmation and validation tests.

5.4. Model performance and usability for hazard assessment

Performance and usability of models play an important role in the user's model choice for a particular case study. This choice also strongly depends on the type of the hazard assessment performed. Regarding the choice of the four selected models for CPC hazard assessment, time is an important variable, and a balance must be stricken in terms of the total computational time required to couple the physical model and the uncertainty quantification (UQ) technique chosen (e.g. Marzocchi and Bebbington 2012; Calder et al. 2015; Bevilacqua et al. 2019; Tierz et al. 2021). For rapid crisis management, when a few simulations must be run in a limited amount of time for risk mitigation, the computational time (Table 5) is the most crucial metric. It is worth remarking, however, that such a comparison might be incomplete, because the models might have a different convergence rate to the solution (at decreasing grid size). Nevertheless, our benchmark results show that the four selected models have reasonable computational times (minutes to < two hours) when used with moderately large topographies like our synthetic ones (grids of 300,000 cells, see Table 2) with a standard computer configuration (see table 5 and Supplementary Material for detail of the computational setup). Models that give results within a few minutes such as VolcFlow and IMEX_SfloW2D seem more suitable for this specific task. Note that the values of computational time given here are dependent on the computational resources used and can significantly change from one computer configuration to another. For other hazard assessment purposes, the quality of the assessment is dependent on both the diversity of models and the UQ technique selected. For example, using an ensemble run of simulations from a single PC model only but with a sophisticated UQ solution (i.e., dominance factors or expected contributions) is not enough to fully assess the epistemic uncertainty of the system. However, a probabilistic assessment using an ensemble of PC models coupled with a standard UQ technique (i.e., inversions or emulators) will allow a modeler to capture the values and variability in some relevant variables for PC hazard assessment (e.g., Bayarri et al. 2009; Stefanescu et al. 2012; Spiller et al. 2014; Tierz et al. 2016; Patra et al. 2020). The strong variability obtained in the benchmarking results presented in this study highlights the importance of using an ensemble of different models for the same phenomena to directly compare outputs and internal variables in all the

models while controlling other factors like numerical solution procedures, input ranges and computer hardware.

Model accessibility is also an important aspect in a user's decision to choose a particular code for their purposes. VolcFlow and TITAN2D are freely available on their respective websites and can be used through a graphical user interface (GUI) and/or cyberinfrastructure (i.e., TITAN2D on VHub), which does increase their accessibility (note that VolcFlow needs Matlab, which requires a paid license). Moreover, the availability of proper documentation (i.e., user guide and website) allows any user to run these models without any prior training. In contrast, SHALTOP and IMEX_SfloW2D lack some of these resources, even though the last one is available through GitHub. Accessibility improvement for these two codes should be considered in the future.

Model performance and usability metrics should also include both pre- and post-processing analyses (Table 5) that can drastically increase the total time needed to display a final simulation result. A harmonization of the input data implementation and a standardization of output formats (i.e., georeferenced ASCII files for the final state, and a compressed binary format for kinematic data) could help potential users to process data more efficiently and speed up the hazard assessment process, while also significantly decreasing the time spent for future similar inter-comparison of these models.

5.5. Volcanological implications

Results from these benchmarks highlight the ability of the four selected depth-averaged models to simulate first-order CPC dynamics: (i) flow velocities and flow thickness distribution inside the synthetic channels are similar to those from natural CPCs like block-and-ash flows (Calder et al. 1999; Brown and Andrews 2015), (ii) simulated flows stay confined within the synthetic valleys and overflow only at specific locations. Overflows occurred in the bend case with all models, and for the constriction cases with VolcFlow only (TITAN2D overbanks near the source are not consider here due to errors found in the code). Even though no overflow occurred in the break in slope and obstacle cases (Figs. 6 and 7), some processes associated with deposition, linked to both a sudden decrease in flow velocities and increase of flow thickness, were observed both before the obstacle and after the break in slope, drastically reducing the channel capacity and promoting late flow overspills. Hence, the first order dynamics of CPC overspill processes seem to be successfully reproduced by the models during these synthetic benchmarks.

Simulation results support field observations that a sudden change in channel geometries (shape, slope, dimensions, sinuosity), combined with a high volume flux, are keys to generate overflows. To illustrate the relationship between flow overspill processes and channel geometry, the overbank width of VolcFlow simulated flows measured along the synthetic channels for both the 'bend' and the
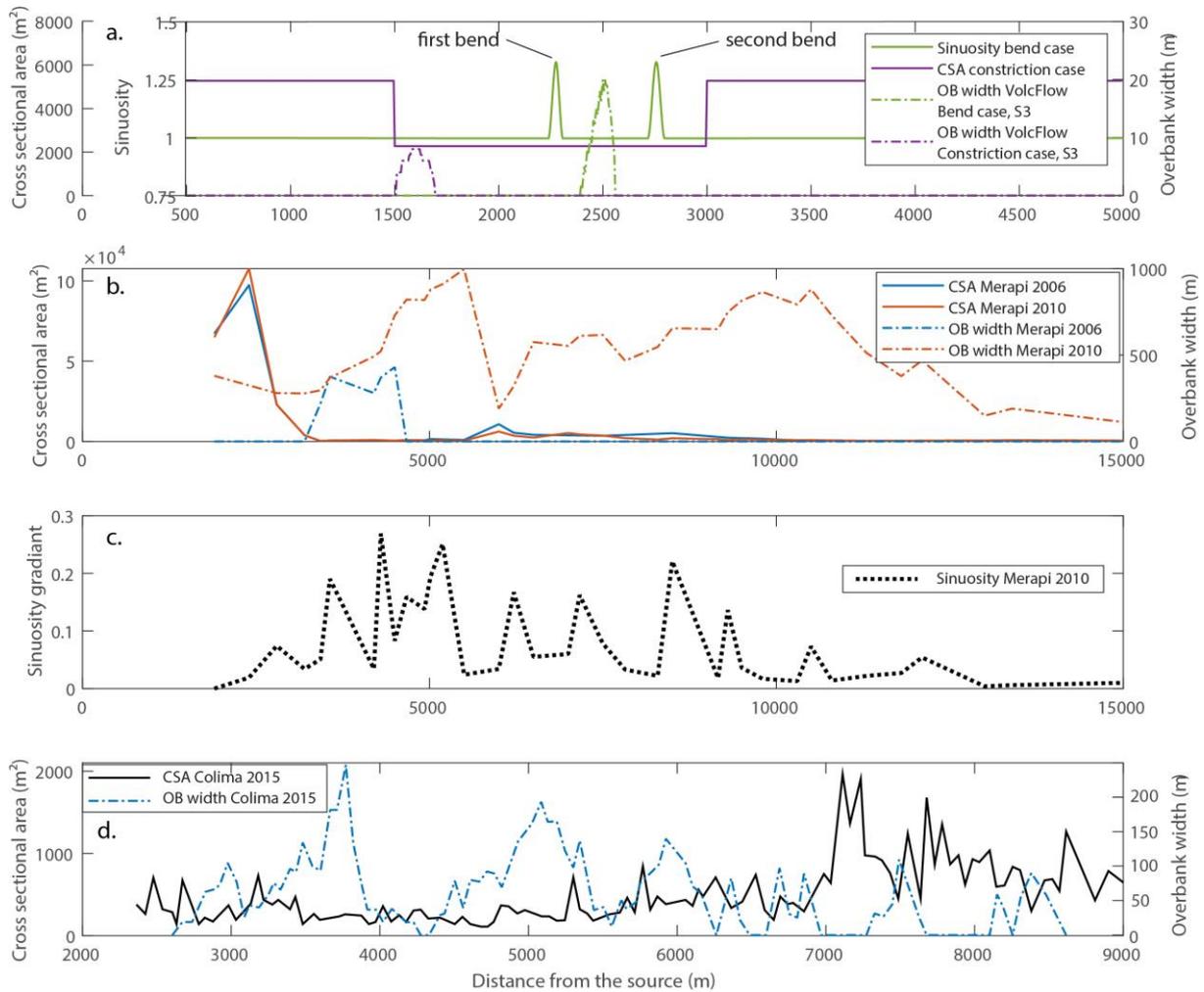
Fig. 10: Morphometric data extracted from the benchmark results compared to a set of field data compiled from three recent eruptions associated with block-and-ash flows (BAFs): a. Channel sinuosity and cross-sectional area measured along the synthetic channel in the 'bend' and the 'constriction' benchmark cases respectively, as well as the corresponding overbank width extracted from two VolcFlow simulations. b. cross-sectional area measured along the Gendol channel at Merapi, and the corresponding overbank width measured before and after the emplacement of the 2006 and 2010 BAFs, respectively. b. Sinuosity gradient measured along the Gendol river at Merapi in 2010 before the eruption. d. cross-sectional area measured along the Montegrande channel at Colima in 2005, and the corresponding overbank width measured after the emplacement of the 2015 BAFs (from Macorps et al. 2018).

'constriction' cases with the high volume flux scenario (Fig. 10a) are compared to the channel cross-sectional area and the channel sinuosity extracted from the synthetic topographies of the 'bend' and the 'constriction' cases, respectively. Results show that overspill processes occur either after a drop in the channel capacity (constriction) or a peak in channel sinuosity (bend). The same observations were made for several past natural BAFs: the June 14, 2006 BAFs (Charbonnier and Gertisser 2008; 2011; Lube et al. 2011) and the November 5, 2010 BAFs at Merapi (Charbonnier et al. 2013; Cronin et al. 2013) and the July 11, 2015 BAFs at Colima (Macorps et al, 2018). To highlight the similarities between

these natural cases and the benchmark results, the same set of data (channel cross-sectional area, channel sinuosity, overbank width) as for the synthetic topographies were extracted from these three natural case studies (Fig. 10 b-d). Similar correlations are obtained where a sharp decrease of channel capacity, or a sudden increase of channel sinuosity, is linked to an increase of the overbank width (Fig. 10 b-d).

In addition, the flow discharge rate seems to be as important as the channel geometry for triggering overspill processes. Fig. 11 shows that such a relationship was indeed obtained in our synthetic benchmarks by the VolcFlow, IMEX_SfloW2D and TITAN2D simulations in the 'bend' and 'constriction' cases: the cumulative overbank area measured from these simulated flows increases proportionally with the averaged flow discharge rate at the source. This also corroborates with field observations at Merapi for both the 2006 and 2010 BAFs: using Charbonnier and Gertisser (2008) data, the discharge rate of the June 14, 2006 BAF can be estimated at 2.5×10^3 m^3 s^{-1}, whereas the averaged discharge rate of the November 5, 2010 BAF was estimated by Kelfoun et al. (2017) at 43×10^3 m^3 s^{-1}. While the channel capacity of the Gendol river did not change significantly between 2006 and 2010 (see Fig.10b), a much higher volumetric rate in 2010 (by twenty times) allowed the occurrence of significantly larger overflows than in 2006 (cumulated overbank areas increased by almost twenty times, see Fig. 10b). Similar trends are observed in the synthetic benchmarks for both VolcFlow and TITAN2D simulated flows, although no predictive pattern is found.
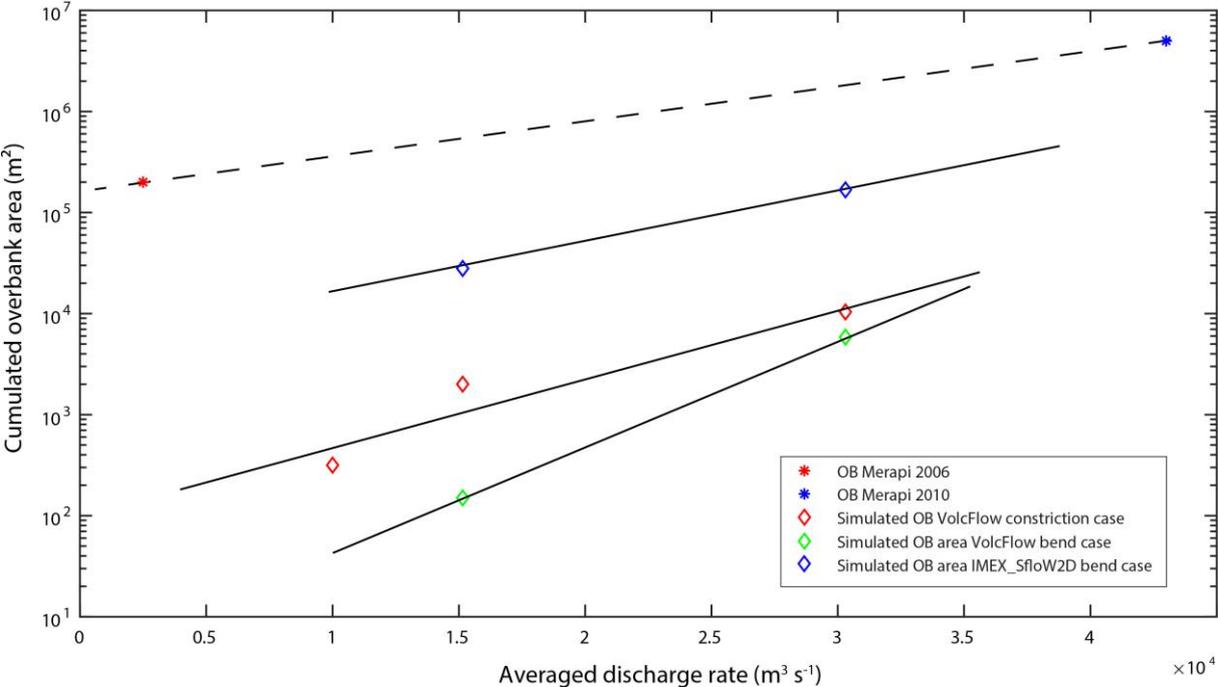

Fig. 11: Cumulated overbank area versus averaged discharge rate of at the source extracted from VolcFlow and IMEX_SfloW2D simulation results in the 'constriction' and 'bend' benchmark cases, compared to those estimated from field observations after the 2006 and 2010 BAFs at Merapi.

5.6. Outcomes and perspectives

The main outcome of this CPC model benchmarking exercise is that, except for TITAN2D, all selected models coherently reproduce similar emplacement dynamics (depth-averaged velocity, thickness) and interactions with topographic artefacts as those observed with natural CPCs, even though a first-order rheological law was used (Voellmy). With TITAN2D, large discrepancies in flow simulation outputs compared to the ones obtained with the other three models allowed us to detect issues with the Voellmy-Salm implementation in the code. Our benchmarks also highlighted noticeable differences both between simulations performed with different models for the same benchmark case, and those performed with the same model between different benchmark cases. In contrast to the previous inter-comparison exercise of Ogburn and Calder (2017), the use of similar source conditions and flow rheology in all our simulations enable us to better interpret those differences: most discrepancies arise from: (i) the various velocity resulting from the gravity projection or from the direction of the shallow approximation and curvature formulations implemented in the source terms of each model, and (ii) the diverse numerical frameworks implemented in each model, an inherent characteristic of each code that cannot be altered by the benchmark procedure itself (in contrast to input parameters). As the computational performance of such models increase year after year, allowing the use of more accurate DEMs with higher spatial and vertical resolution and rougher channel topographies, the divergence of such model benchmarking results is likely to become more important in the future and the most accurate (i.e. realistic) models must be prioritized. It is important to stress that this exercise was based on a synthetic dataset, model accuracy cannot be quantified but simply estimated. The need of a reference framework to better quantify such CPC model inter-comparison and benchmarking efforts is crucial. To respond to that specific need, a new large-scale experimental facility has been developed at USF (Tampa, FL, USA) to serve as a reference framework for future experimental CPC benchmarks. Finally, due to the large variability in the results obtained here, the use of a single model for hazard and risk assessment purposes is not recommended and a compilation of output data from different models with a robust UQ solution, without significantly increasing the pre- and post-processing time, seems to be more appropriate. For such UQ modeling purposes, one would need to quantify the aleatory uncertainty via one or more models (and/or, e.g. the structural -epistemic- uncertainty of each model), and then, for instance, use an ensemble of models to further characterize the epistemic uncertainty (e.g. Tierz et al., 2016; 2018). This method is already used in routine in other geosciences, for example hurricane tracking in meteorology is assessed by compiling simulations of several different codes using a well-established workflow in near real-time.

VI. CONCLUSIONS

We present the results of the first synthetic benchmarks for CPC numerical models. Except for TITAN2D, for which issues with the Voellmy-Salm implementation were detected thanks to the benchmarks, the three other codes, VolcFlow, SHALTOP and IMEX_SfloW2D, satisfactorily produce similar first-order CPC dynamics over four simplified volcanic-like valleys with various topographic changes (channel bend and constriction, break-in-slope and obstacle). Results, specific to the Voellmy rheology, are coherent among the various models and were obtained in a short amount of time, from a few minutes to an hour depending on the model and topographic setting, which demonstrated the usefulness and ability of such model benchmarking efforts to correctly evaluate and select appropriate models for hazard assessment purposes. The benchmarks also highlight some discrepancies in the results obtained both between models and benchmark cases, especially regarding their ability to reproduce CPC overspill processes, which have been proven to be the most deadly and unpredictable phenomenon associated with CPC during recent eruptions at Merapi, Colima and Fuego volcanoes. It is shown that incorporating curvature effects into such models is a key component to accurately simulate CPC overspill processes related to topography variations, even on a mean slope of 20°. Except for TITAN2D, these discrepancies in the simulation results can be partly compensated by implementing the same curvature and velocity formulations in all models. The remaining differences come from the various numerical frameworks used, intrinsic to each code and which cannot be altered by the benchmarking procedure itself.

Nevertheless, even with the lack of a reference framework, results obtained here show that such a synthetic benchmark procedure seems to be adapted to qualitatively evaluate the performance of CPC numerical models over natural-like volcanic terrains and should be complemented in the future with new large-scale experimental designs, more realistic volcanic digital topographies, and a more diverse set of numerical models with robust UQ techniques.

References

Andrews BJ, Manga M (2011) Effects of topography on pyroclastic density current runout and formation of coignimbrites. Geology 39:1099–1102. https://doi.org/10.1130/G32226.1

Andrews BJ, Manga M (2012) Experimental study of turbulence, sedimentation, and coignimbrite mass partitioning in dilute pyroclastic density currents. J Volcanol Geotherm Res 225–226:30–44. https://doi.org/10.1016/j.jvolgeores.2012.02.011

Andrews BJ (2014) Dispersal and air entrainment in unconfined dilute pyroclastic density currents. Bull Volcanol 76:852. https://doi.org/10.1007/s00445-014-0852-4

Aravena A, Cioni R, Bevilacqua A, Michieli Vitturi M, Esposti Ongaro T, Neri A (2020) Tree-Branching-Based Enhancement of Kinetic Energy Models for Reproducing Channelization Processes of Pyroclastic Density Currents. J Geophys Res Solid Earth 125:1–20. https://doi.org/10.1029/2019JB019271

Bayarri MJ, Berger JO, Calder ES, Dalbey K, Lunagomez S, Patra AK, et al. (2009) Using statistical and computer models to quantify volcanic hazards. Technometrics 51:402–413

Benage MC, Dufek J, Mothes PA (2016) Quantifying entrainment in pyroclastic density currents from the Tungurahua eruption, Ecuador: Integrating field proxies with numerical simulations. Geophys Res Lett 43:6932–6941. https://doi.org/10.1002/2016GL069527

Bernabeu N, Saramito P, Smutek C (2014) Numerical modeling of non-Newtonian viscoplastic flows: Part II. Viscoplastic fluids and general tridimensional topographies. Int J Numer Anal Model 11:213–228

Bevilacqua A, Patra AK, Bursik MI, Pitman EB, Macías JL, Saucedo R, et al. (2019) Probabilistic forecasting of plausible debris flows from Nevado de colima (Mexico) using data from the atenquique debris flow, 1955. Nat Hazards Earth Syst Sci 19:791–820. https://doi.org/10.5194/nhess-19-791-2019

Bouchut F, Mangeney-Castelnau A, Perthame B, Vilotte J-P (2003) A new model of Saint Venant and Savage–Hutter type for gravity driven shallow water flows. Comptes Rendus Math 336:531–536. https://doi.org/10.1016/S1631-073X(03)00117-1

Bouchut F, Westdickenberg M (2004) Gravity driven shallow water models for arbitrary topography. Commun Math Sci 2:359–389. https://doi.org/10.4310/cms.2004.v2.n3.a2

Bourdier JL, Abdurachman E (2001) Decoupling of small-volume pyroclastic flows and related hazards at Merapi volcano, Indonesia. Bull Volcanol 63:309–325. https://doi.org/10.1007/s004450100133

Breard ECP, Lube G, Jones JR, Dufek J, Cronin SJ, Valentine GA, Moebis A (2016) Coupling of turbulent and non-turbulent flow regimes within pyroclastic density currents. Nat Geosci 9:767–771. https://doi.org/10.1038/ngeo2794

Breard ECP, Lube G (2017) Inside pyroclastic density currents – uncovering the enigmatic flow structure and transport behaviour in large-scale experiments. Earth Planet Sci Lett 458:22–36. https://doi.org/10.1016/j.epsl.2016.10.016

Breard ECP, Dufek J, Lube G (2018) Enhanced Mobility in Concentrated Pyroclastic Density Currents: An Examination of a Self-Fluidization Mechanism. Geophys Res Lett 45:654–664. https://doi.org/10.1002/2017GL075759

Breard ECP, Dufek J, Roche O (2019) Continuum Modeling of Pressure-Balanced and Fluidized Granular Flows in 2-D: Comparison With Glass Bead Experiments and Implications for Concentrated Pyroclastic Density Currents. J Geophys Res Solid Earth 124:5557–5583. https://doi.org/10.1029/2018JB016874

Brosch E, Lube G (2020) Spatiotemporal sediment transport and deposition processes in experimental dilute pyroclastic density currents. J Volcanol Geotherm Res 401:106946. https://doi.org/10.1016/j.jvolgeores.2020.106946

Brown RJ, D. M. Andrews G (2015) Deposits of Pyroclastic Density Currents. In: Sigurdsson H (ed) The Encyclopedia of Volcanoes, 2nd edn. Academic Press, Amsterdam, pp 631–648. https://doi.org/10.1016/B978-0-12-385938-9.00036-5

Brown SK, Jenkins SF, Sparks RSJ, Odbert H, Auker MR (2017) Volcanic fatalities database: analysis of volcanic threat with distance and victim classification. J Appl Volcanol 6:15. https://doi.org/10.1186/s13617-017-0067-4

Brunet M, Moretti L, Le Friant A, Mangeney A, Fernández Nieto ED, Bouchut F (2017) Numerical simulation of the 30–45 ka debris avalanche flow of Montagne Pelée volcano, Martinique: from volcano flank collapse to submarine emplacement. Nat Hazards 87:1189–1222. https://doi.org/10.1007/s11069-017-2815-5

Bursik MI, Woods AW (1996) The dynamics and thermodynamics of large ash flows. Bull Volcanol 58:175–193

Calder ES, Cole PD, Dade WB, Druitt TH, Hoblitt RP, Huppert HE, Ritchie L, Sparks RSJ, Young SR (1999) Mobility of pyroclastic flows and surges at the Soufriere Hills Volcano, Montserrat. Geophys Res Lett 26:537–540. https://doi.org/10.1029/1999GL900051

Calder E, Wagner K, Ogburn S (2015) Volcanic hazard maps. In: Loughlin S, Sparks S, Brown S, Jenkins S, and Vye-Brown C (eds) Global Volcanic Hazards and Risk. Cambridge University Press, Cambridge, pp 335-342

Capra L, Manea VC, Manea M, Norini G (2011) The importance of digital elevation model resolution on granular flow simulations: a test case for Colima volcano using TITAN2D computational routine. Nat Hazards 59:665–680. https://doi.org/10.1007/s11069-011-9788-6

Castro-Orgaz O, Hutter K, Giraldez J V., Hager WH (2015) Nonhydrostatic granular flow over 3-D terrain: New Boussinesq-type gravity waves? J Geophys Res Earth Surf 120:1–28. https://doi.org/10.1002/2014JF003279

Cerminara M, Brosch E, Lube G (2018) A theoretical framework and the experimental dataset for benchmarking numerical models of dilute pyroclastic density currents. arXiv preprint. https://arxiv.org/abs/2106.14057. Accessed 13 August 2021

Charbonnier SJ, Gertisser R (2008) Field observations and surface characteristics of pristine block-and-ash flow deposits from the 2006 eruption of Merapi Volcano, Java, Indonesia. J Volcanol Geotherm Res 177:971–982. https://doi.org/10.1016/j.jvolgeores.2008.07.008

Charbonnier SJ, Gertisser R (2009) Numerical simulations of block-and-ash flows using the Titan2D flow model: Examples from the 2006 eruption of Merapi Volcano, Java, Indonesia. Bull Volcanol 71:953–959. https://doi.org/10.1007/s00445-009-0299-1

Charbonnier SJ, Gertisser R (2011) Deposit architecture and dynamics of the 2006 block-and-ash flows of Merapi Volcano, Java, Indonesia. Sedimentology 58:1573–1612. https://doi.org/10.1111/j.1365-3091.2011.01226.x

Charbonnier SJ, Gertisser R (2012) Evaluation of geophysical mass flow models using the 2006 block-and-ash flows of Merapi Volcano, Java, Indonesia: Towards a short-term hazard assessment tool. J Volcanol Geotherm Res 231–232:87–108. https://doi.org/10.1016/j.jvolgeores.2012.02.015

Charbonnier SJ, Germa A, Connor CB, Gertisser R, Preece K, Komorowski JC, Lavigne F, Dixon T, Connor L (2013) Evaluation of the impact of the 2010 pyroclastic density currents at Merapi volcano

from high-resolution satellite imagery, field investigations and numerical simulations. J Volcanol Geotherm Res 261:295–315. https://doi.org/10.1016/j.jvolgeores.2012.12.021

Charbonnier SJ, Varner LA, Escobar-Wolf RP, Rodriguez LA, Chigna G, Chun C, Gonzalez D, Juarez F, Merida R, Calder E (2019) Unravelling the Dynamics and Hazards of the June 3 rd, 2018 Pyroclastic Currents at Fuego volcano (Guatemala): A Multi-Parameter Approach. In: AGU Fall meeting, Abstract #619589, San Francisco, CA

Charbonnier SJ, Thouret J-C, Gueugneau V, Constantinescu R (2020) New Insights Into the 2070calyrBP Pyroclastic Currents at El Misti Volcano (Peru) From Field Investigations, Satellite Imagery and Probabilistic Modeling. Front Earth Sci 8: 557788. https://doi.org/10.3389/feart.2020.557788

Chédeville C, Roche O (2018) Autofluidization of collapsing bed of fine particles: Implications for the emplacement of pyroclastic flows. J Volcanol Geotherm Res 368:91–99. https://doi.org/10.1016/j.jvolgeores.2018.11.007

Christen M, Kowalski J, Bartelt P (2010) RAMMS: Numerical simulation of dense snow avalanches in three-dimensional terrain. Cold Reg Sci Technol 63:1–14. https://doi.org/10.1016/j.coldregions.2010.04.005

Cole PD, Calder ES, Sparks RSJ, Clarke AB, Druitt TH, Young SR, Herd RA, Harford CL, Norton GE (2002) Deposits from dome-collapse and fountain-collapse pyroclastic flows at Soufrière Hills Volcano, Montserrat. Geol Soc London, Mem 21:231–262. https://doi.org/10.1144/GSL.MEM.2002.021.01.11

Cordonnier B, Lev E, Garel F (2015) Benchmarking lava-flow models. Geol Soc London, Spec Publ 426:425–445. https://doi.org/10.1144/SP426.7

Costa A, Suzuki YJ, Cerminara M, Devenish BJ, Ongaro TE, Herzog M, Van Eaton AR, Denby LC, Bursik M, de' Michieli Vitturi M, Engwell S, Neri A, Barsotti S, Folch A, Macedonio G, Girault F, Carazzo G, Tait S, Kaminski E, Mastin LG, Woodhouse MJ, Phillips JC, Hogg AJ, Degruyter W, Bonadonna C (2016) Results of the eruptive column model inter-comparison study. J Volcanol Geotherm Res 326:2–25. https://doi.org/10.1016/j.jvolgeores.2016.01.017

Cronin SJ, Lube G, Dayudi DS, Sumarti S, Subrandiyo S, Surono (2013) Insights into the October–November 2010 Gunung Merapi eruption (Central Java, Indonesia) from the stratigraphy, volume and characteristics of its pyroclastic deposits. J Volcanol Geotherm Res 261:244–259. https://doi.org/10.1016/j.jvolgeores.2013.01.005

Dartevelle S (2004a) Numerical modeling of geophysical granular flows: 1. A comprehensive approach to granular rheologies and geophysical multiphase flows. Geochemistry, Geophys Geosystems 5:Q08003. https://doi.org/10.1029/2003GC000636

Dartevelle S, Rose WI, Stix J, Kelfoun K, Vallance JW (2004b) Numerical modeling of geophysical granular flows: 2. Computer simulations of plinian clouds and pyroclastic flows and surges. Geochemistry, Geophys Geosystems 5: Q08004. https://doi.org/10.1029/2003GC000637

De'Michieli Vitturi M, Esposti Ongaro T, Lari G, Aravena A (2019) IMEX-SfloW2D 1.0: A depth-averaged numerical flow model for pyroclastic avalanches. Geosci Model Dev 12:581–595. https://doi.org/10.5194/gmd-12-581-2019

Dellino P, Büttner R, Dioguardi F, Doronzo DM, La Volpe L, Mele D, Sonder I, Sulpizio R, Zimanowski B (2010) Experimental evidence links volcanic particle characteristics to pyroclastic flow hazard. Earth Planet Sci Lett 295:314–320. https://doi.org/10.1016/j.epsl.2010.04.022

Dellino P, Dioguardi F, Doronzo DM, Mele D (2019) The rate of sedimentation from turbulent suspension: An experimental model with application to pyroclastic density currents and discussion on the grain-size dependence of flow runout. Sedimentology 66:129–145. https://doi.org/10.1111/SED.12485

Denlinger RP, Iverson RM (2004) Granular avalanches across irregular three-dimensional terrain: 1. Theory and computation. J Geophys Res Earth Surf 109:F01014. https://doi.org/10.1029/2003JF000085

Dietterich HR, Lev E, Chen J, Richardson JA, Cashman K V. (2017) Benchmarking computational fluid dynamics models of lava flow simulation for hazard assessment, forecasting, and risk management. J Appl Volcanol 6:9. https://doi.org/10.1186/s13617-017-0061-x

Doronzo DM, Valentine GA, Dellino P, de Tullio MD (2010) Numerical analysis of the effect of topography on deposition from dilute pyroclastic density currents. Earth Planet Sci Lett 300:164–173. https://doi.org/10.1016/j.epsl.2010.10.003

Doyle EE, Hogg AJ, Mader HM, Sparks RSJ (2008) Modeling dense pyroclastic basal flows from collapsing columns. Geophys Res Lett 35:L04305. https://doi.org/10.1029/2007GL032585

Dufek J, Bergantz GW (2007) Suspended load and bed-load transport of particle-laden gravity currents: The role of particle-bed interaction. Theor Comput Fluid Dyn 21:119–145. https://doi.org/10.1007/s00162-007-0041-6

Dufek J, Wexler J, Manga M (2009) Transport capacity of pyroclastic density currents: Experiments and models of substrate-flow interaction. J Geophys Res Solid Earth 114:1–13. https://doi.org/10.1029/2008JB006216

Dufek J, Esposti Ongaro T, Roche O (2015) Pyroclastic Density Currents: Processes and Models. In: Sigurdsson H (ed) The Encyclopedia of Volcanoes, 2nd edn. Academic Press, Amsterdam, pp 617–629. https://doi.org/10.1016/B978-0-12-385938-9.00035-3

Dufek J (2016) The Fluid Mechanics of Pyroclastic Density Currents. Annu Rev Fluid Mech 48:459–485. https://doi.org/10.1146/annurev-fluid-122414-034252

Esposti Ongaro T, Clarke AB, Neri A, Voight B, Widiwijayanti C (2008) Fluid dynamics of the 1997 Boxing Day volcanic blast on Montserrat, West Indies. J Geophys Res 113:B03211. https://doi.org/10.1029/2006JB004898

Esposti Ongaro T, Clarke AB, Voight B, Neri A, Widiwijayanti C (2012) Multiphase flow dynamics of pyroclastic density currents during the May 18, 1980 lateral blast of Mount St. Helens. J Geophys Res Solid Earth 117: B06208. https://doi.org/10.1029/2011JB009081

Esposti Ongaro T, Cerminara M, Charbonnier SJ, Lube G, Valentine GA (2020) A framework for validation and benchmarking of pyroclastic current models. Bull Volcanol 82:51. https://doi.org/10.1007/s00445-020-01388-2

Favreau P, Mangeney A, Lucas A, Crosta G, Bouchut F (2010) Numerical modeling of landquakes. Geophys Res Lett 37:L15305. https://doi.org/10.1029/2010GL043512

Freundt A, Bursik MI (1998) Pyroclastic flow transport mechanisms. In: Freundt A and Rosi M (eds) From magma to tephra: Modelling physical processes of explosive volcanic eruptions. Developments in Volcanology, 4th edn. Elsevier, Amsterdam, pp 173-245

Gertisser R, Charbonnier SJ, Keller J, Quidelleur X (2012) The geological evolution of Merapi volcano, Central Java, Indonesia. Bull Volcanol 74:1213–1233. https://doi.org/10.1007/s00445-012-0591-3

Giachetti T, Paris R, Kelfoun K, Pérez-Torrado FJ (2011) Numerical modelling of the tsunami triggered by the Gumar debris avalanche, Tenerife (Canary Islands): Comparison with field-based data. Mar Geol 284:189–202. https://doi.org/10.1016/j.margeo.2011.03.018

Gruber U, Bartelt P (2007) Snow avalanche hazard modelling of large areas using shallow water numerical methods and GIS. Environ Model Softw 22:1472–1481 . https://doi.org/10.1016/J.ENVSOFT.2007.01.001

Gueugneau V, Kelfoun K, Roche O, Chupin L (2017) Effects of pore pressure in pyroclastic flows: Numerical simulation and experimental validation. Geophys Res Lett 44:2194–2202. https://doi.org/10.1002/2017GL072591

Gueugneau V, Kelfoun K, Druitt T (2019) Investigation of surge-derived pyroclastic flow formation by numerical modelling of the 25 June 1997 dome collapse at Soufrière Hills Volcano, Montserrat. Bull Volcanol 81:25. https://doi.org/10.1007/s00445-019-1284-y

Gueugneau V, Kelfoun K, Charbonnier S, Germa A, Carazzo G (2020) Dynamics and Impacts of the May 8th, 1902 Pyroclastic Current at Mount Pelée (Martinique): New Insights From Numerical Modeling. Front Earth Sci 8:279. https://doi.org/10.3389/feart.2020.00279

Horrillo J, Grilli ST, Nicolsky D, Roeber V, Zhang J (2015) Performance Benchmarking Tsunami Models for NTHMP's Inundation Mapping Activities. Pure Appl Geophys 172:869–884. https://doi.org/10.1007/s00024-014-0891-y

Hungr, O, Morgenstern N, Wong H (2007) Review of benchmarking exercise on landslide debris runout and mobility modelling. In: Ho and Li (ed) 2007 International Forum on Landslide Disaster Management. Geotechnical Division, Hong Kong Institution of Engineers, Hong-Kong, China.

Hungr O (2008) Simplified models of spreading flow of dry granular material. Can Geotech J 45:1156–1168. https://doi.org/10.1139/T08-059

Ishimine Y (2005) Numerical study of pyroclastic surges. J Volcanol Geotherm Res 139:33–57. https://doi.org/10.1016/j.jvolgeores.2004.06.017

Iverson RM, Logan M, Denlinger RP (2004) Granular avalanches across irregular three-dimensional terrain: 2. Experimental tests. J Geophys Res Earth Surf 109:F01015. https://doi.org/10.1029/2003JF000084

George DL , Iverson RM (2014) A depth-averaged debris-flow model that includes the effects of evolving dilatancy. II. Numerical predictions and experimental tests. Proc R Soc A Math Phys Eng Sci 470:20130820–20130820 . https://doi.org/10.1098/rspa.2013.0820

Jenkins S, Komorowski JC, Baxter PJ, Spence R, Picquout A, Lavigne F, Surono (2013) The Merapi 2010 eruption: An interdisciplinary impact assessment methodology for studying pyroclastic density current dynamics. J Volcanol Geotherm Res 261:316–329. https://doi.org/10.1016/j.jvolgeores.2013.02.012

Juez C, Murillo J, García-Navarro P (2013) 2D simulation of granular flow over irregular steep slopes using global and local coordinates. J Comput Phys 255:166–204. https://doi.org/10.1016/j.jcp.2013.08.002

Kelfoun K, Druitt TH (2005) Numerical modeling of the emplacement of Socompa rock avalanche, Chile. J Geophys Res Solid Earth 110:1–13. https://doi.org/10.1029/2005JB003758

Kelfoun K, Samaniego P, Palacios P, Barba D (2009) Testing the suitability of frictional behaviour for pyroclastic flow simulation by comparison with a well-constrained eruption at Tungurahua volcano (Ecuador). Bull Volcanol 71:1057–1075. https://doi.org/10.1007/s00445-009-0286-6

Kelfoun K (2011) Suitability of simple rheological laws for the numerical simulation of dense pyroclastic flows and long-runout volcanic avalanches. J Geophys Res 116:1–14 . https://doi.org/10.1029/2010JB007622

Kelfoun K, Vallejo Vargas S (2015) VolcFlow capabilities and potential development for the simulation of lava flows. Geol Soc London, Spec Publ 426:337-343. https://doi.org/10.1144/SP426.8

Kelfoun K (2017) A two-layer depth-averaged model for both the dilute and the concentrated parts of pyroclastic currents. J Geophys Res Solid Earth 122:1–19 . https://doi.org/10.1002/2017JB014013

Kelfoun K, Gueugneau V, Komorowski J-C, Cholik N, Merciecca C (2017) Simulation of block-and-ash flows and ash-cloud surges of the 2010 eruption of Merapi volcano with a two-layer model. J Geophys Res Solid Earth 122:1–16. https://doi.org/10.1002/2017JB013981

Komorowski JC, Jenkins S, Baxter PJ, Picquout A, Lavigne F, Charbonnier S, Gertisser R, Preece K, Cholik N, Budi-Santoso A, Surono (2013) Paroxysmal dome explosion during the Merapi 2010 eruption: Processes and facies relationships of associated high-energy pyroclastic density currents. J Volcanol Geotherm Res 261:260–294. https://doi.org/10.1016/j.jvolgeores.2013.01.007

Levy C, Mangeney A, Bonilla F, Hibert C, Calder ES, Smith PJ (2015) Friction weakening in granular flows deduced from seismic records at the Soufrière Hills Volcano, Montserrat. J Geophys Res B Solid Earth 120:7536–7557. https://doi.org/10.1002/2015JB012151

Loughlin SC, Calder ES, Clarke A, Cole PD, Luckett R, Mangan MT, Pyle DM, Sparks RSJ, Voight B, Watts RB (2002) Pyroclastic flows and surges generated by the 25 June 1997 dome collapse, Soufriere Hills Volcano, Montserrat. Geol Soc London, Mem 21:191–209. https://doi.org/10.1144/GSL.MEM.2002.021.01.09

Lube G, Cronin SJ, Platz T, Freundt A, Procter JN, Henderson C, Sheridan MF (2007) Flow and deposition of pyroclastic granular flows: A type example from the 1975 Ngauruhoe eruption, New Zealand. J Volcanol Geotherm Res 161:165–186. https://doi.org/10.1016/j.jvolgeores.2006.12.003

Lube G, Cronin SJ, Thouret JC, Surono S (2011) Kinematic characteristics of pyroclastic density currents at merapi and controls on their avulsion from natural and engineered channels. Bull Geol Soc Am 123:1127–1140. https://doi.org/10.1130/B30244.1

Lube G, Breard ECP, Jones J, Fullard L, Dufek J, Cronin SJ, Wang T (2019) Generation of air lubrication within pyroclastic density currents. Nat Geosci 12:381–386. https://doi.org/10.1038/s41561-019-0338-2

Lube G, Breard ECP, Esposti-Ongaro T, Dufek J, Brand B (2020) Multiphase flow behaviour and hazard prediction of pyroclastic density currents. Nat Rev Earth Environ 1:348–365. https://doi.org/10.1038/s43017-020-0064-8

Lucas A, Mangeney A, Ampuero JP (2014) Frictional velocity-weakening in landslides on Earth and on other planetary bodies. Nat Commun 5:1–9. https://doi.org/10.1038/ncomms4417

Macorps E, Charbonnier SJ, Varley NR, Capra L, Atlas Z, Cabré J (2018) Stratigraphy, sedimentology and inferred flow dynamics from the July 2015 block-and-ash flow deposits at Volcán de Colima, Mexico. J Volcanol Geotherm Res 349:99–116. https://doi.org/10.1016/j.jvolgeores.2017.09.025

Malin MC, Sheridan MF (1982) Computer-assisted mapping of pyroclastic surges. Science 217:637–640. https://doi.org/10.1126/science.217.4560.637

Mangeney-Castelnau A, Vilotte JP, Bristeau MO, Perthame B, Bouchut F, Simeoni C, Yerneni S (2003) Numerical modeling of avalanches based on Saint-Venant equations using a kinetic scheme. J Geophys Res 108:2527. https://doi.org/10.1029/2002JB002024

Mangeney-Castelnau A, Bouchut F, Vilotte JP, Lajeunesse E, Aubertin A, Pirulli M (2005) On the use of Saint Venant equations to simulate the spreading of a granular mass. J Geophys Res B Solid Earth 110:1–17. https://doi.org/10.1029/2004JB003161

Mangeney A, Bouchut F, Thomas N, Vilotte JP, Bristeau MO (2007) Numerical modeling of self-channeling granular flows and of their levee-channel deposits. J Geophys Res 112:F02017. https://doi.org/10.1029/2006JF000469

Martinec Z, Klemann V, van der Wal W, Riva REM, Spada G, Sun Y, Melini D, Kachuck SB, Barletta V, Simon K, A G, James TS (2018) A benchmark study of numerical implementations of the sea level equation in GIA modelling. Geophys J Int 215:389–414. https://doi.org/10.1093/gji/ggy280

Marzocchi W, Bebbington MS (2012) Probabilistic eruption forecasting at short and long time scales. Bull Volcanol 74:1777-1805.

McDougall S, Hungr O (2004) A model for the analysis of rapid landslide motion across three-dimensional terrain. Can Geotech J 41:1084–1097. https://doi.org/10.1139/t04-052

McDougall S (2017) 2014 Canadian Geotechnical Colloquium: Landslide runout analysis — current practice and challenges. Can Geotech J 54:605–620. https://doi.org/10.1139/cgj-2016-0104

Moretti L, Allstadt K, Mangeney A, Capdeville Y, Stutzmann E, Bouchut F (2015) Numerical modeling of the Mount Meager landslide constrained by its force history derived from seismic data. J Geophys Res Solid Earth 120:2579–2599 . https://doi.org/10.1002/2014JB011426

Moretti L, Mangeney A, Walter F, Capdeville Y, Bodin T, Stutzmann E, Le Friant A (2020) Constraining landslide characteristics with Bayesian inversion of field and seismic data. Geophys J Int 221:1341–1348. https://doi.org/10.1093/gji/ggaa056

Nakada S, Fujii T (1993) Preliminary report on the activity at Unzen Volcano (Japan), November 1990-November 1991: Dacite lava domes and pyroclastic flows. J Volcanol Geotherm Res 54:319–333. https://doi.org/10.1016/0377-0273(93)90070-8

Neri A, Esposti Ongaro T, Macedonio G, Gidaspow D (2003) Multiparticle simulation of collapsing volcanic columns and pyroclastic flow. J Geophys Res Solid Earth 108:2202. https://doi.org/10.1029/2001JB000508

Neri A, Esposti Ongaro T, Voight B, Widiwijayanti C (2015) Pyroclastic Density Current Hazards and Risk. In: Shroder JF and Papale P (eds) Volcanic Hazards, Risks and Disasters. Elsevier, Amsterdam, pp 109–140

Oberkampf WL, Trucano TG, Roy CJ, Oberkampf WL (2002) Verification and validation in computational fluid dynamics. Prog Aerosp Sci 38:209–272

Ogburn SE (2012) FlowDat: Mass flow database. vHub platform. https://vhub.org/groups/massflowdatabase. Accessed 16 August 2021

Ogburn SE, Calder ES, Cole PD, Stinton AJ (2014) The effect of topography on ash-cloud surge generation and propagation. Geol Soc London, Mem 39:179–194. https://doi.org/10.1144/M39.10

Ogburn SE, Calder ES (2017) The Relative Effectiveness of Empirical and Physical Models for Simulating the Dense Undercurrent of Pyroclastic Flows under Different Emplacement Conditions. Front Earth Sci 5:83. https://doi.org/10.3389/feart.2017.00083

Palladino DM (2017) Simply pyroclastic currents. Bull Volcanol 79:1–5. https://doi.org/10.1007/s00445-017-1139-3

Patra AK, Bauer AC, Nichita CC, Pitman EB, Sheridan MF, Bursik M, Rupp B, Webber A, Stinton AJ, Namikawa LM, Renschler CS (2005) Parallel adaptive numerical simulation of dry avalanches over natural terrain. J Volcanol Geotherm Res 139:1–21. https://doi.org/10.1016/j.jvolgeores.2004.06.014

Patra A, Bevilacqua A, Akhavan-Safaei A, Pitman EB, Bursik M, Hyman D (2020) Comparative Analysis of the Structures and Outcomes of Geophysical Flow Models and Modeling Assumptions Using Uncertainty Quantification. Front Earth Sci 8:275. https://doi.org/10.3389/feart.2020.00275

Peruzzetto M, Komorowski J-C, Le Friant A, Rosas-Carbajal M, Mangeney A, Legendre Y (2019) Modeling of partial dome collapse of La Soufrière of Guadeloupe volcano: implications for hazard assessment and monitoring. Sci Rep 9:13105. https://doi.org/10.1038/s41598-019-49507-0

Peruzzetto M, Mangeney A, Bouchut F, Grandjean G, Levy C, Thiery Y, Lucas A (2021) Topography curvature effects in thin-layer models for gravity-driven flows without bed erosion. J Geophys Res Earth Surf 126:e2020JF005657. https://doi.org/10.1029/2020JF005657

Richardson JA (2017) Modeling the Construction and Evolution of Distributed Volcanic Fields on Earth and Mars. Dissertation, University of South Florida

Roche O, Montserrat S, Niño Y, Tamburrino A (2008) Experimental observations of water-like behavior of initially fluidized, dam break granular flows and their relevance for the propagation of ash-rich pyroclastic flows. J Geophys Res 113:B12203. https://doi.org/10.1029/2008JB005664

Roche O, Montserrat S, Niño Y, Tamburrino a. (2010) Pore fluid pressure and internal kinematics of gravitational laboratory air-particle flows: Insights into the emplacement dynamics of pyroclastic flows. J Geophys Res Solid Earth 115:1–18. https://doi.org/10.1029/2009JB007133

Roche O (2012) Depositional processes and gas pore pressure in pyroclastic flows: an experimental perspective. Bull Volcanol 74:1807–1820. https://doi.org/10.1007/s00445-012-0639-4

Roche O, Philips JC, Kelfoun K (2013) Pyroclastic Density Currents. In: Fagents SA, Gregg T, Lopes R (eds) Modelling Volcanic Processes. Cambridge University Press, Cambridge, UK, pp 203–224

Roche O (2015) Nature and velocity of pyroclastic density currents inferred from models of entrainment of substrate lithic clasts. Earth Planet Sci Lett 418:115–125. https://doi.org/10.1016/j.epsl.2015.03.001

Roche O, Buesch DC, Valentine GA (2016) Slow-moving and far-travelled dense pyroclastic flows during the Peach Spring super-eruption. Nat Commun 7:10890. https://doi.org/10.1038/ncomms10890

Rowley PJ, Roche O, Druitt TH, Cas R (2014) Experimental study of dense pyroclastic density currents using sustained, gas-fluidized granular flows. Bull Volcanol 76:1–13. https://doi.org/10.1007/s00445-014-0855-1

Saint-Venant AJCB de (1871) Théorie du mouvement non-permanent des eaux avec application aux crues des rivières et a l'introduction des marées dans leur lit. Comptes Rendu l'Académie des Sciences 73:147–154

Salm B, Burkard A, Gubler HU (1990) Berechnung von Fliesslawinen: eine Anleitung für Praktiker mit Beispielen. Eidgenössisches Institut für Schnee-und Lawinenforschung. Weissfluhjoch/Davos

Salvatici T, Di Roberto A, Di Traglia F, Bisson M, Morelli S, Fidolini F, Bertagnini A, Pompilio M, Hungr O, Casagli N (2016) From hot rocks to glowing avalanches: Numerical modelling of gravity-induced pyroclastic density currents and hazard maps at the Stromboli volcano (Italy). Geomorphology 273:93–106. https://doi.org/10.1016/j.geomorph.2016.08.011

Savage SB, Hutter K (1989) The motion of a finite mass of granular material down a rough incline. J Fluid Mech 199: 177-215. https://doi.org/10.1017/s0022112089000340

Savage SB, Hutter K (1991) The Dynamics of Avalanches of Antigranulocytes Materials from Initiation to Runout.1. Analysis. Acta Mech 86:201–223. https://doi.org/10.1007/BF01175958

Schilling SP (1998) LAHARZ: GIS programs for automated mapping of lahar-inundation hazard zones. US Geological Survey, Open-File Report 98-638

Sheridan MF, Hubbard B, Carrasco-Nunez G, Siebe C (2004) Pyroclastic Flow Hazard at Volcán Citlaltépetl. Nat Hazards 33:209–221. https://doi.org/10.1023/B

Shimizu HA, Koyaguchi T, Suzuki YJ (2019) The run-out distance of large-scale pyroclastic density currents: A two-layer depth-averaged model. J Volcanol Geotherm Res 381:168–184 . https://doi.org/10.1016/j.jvolgeores.2019.03.013

Smith GM, Williams R, Rowley PJ, Parsons DR (2018) Investigation of variable aeration of monodisperse mixtures: implications for pyroclastic density currents. Bull Volcanol 80:67. https://doi.org/10.1007/s00445-018-1241-1

Spiller ET, Bayarri MJ, Berger JO, Calder ES, Patra AK, Pitman EB, et al. (2014) Automating emulator construction for geophysical hazard maps. SIAM/ASA J Uncertain Quant 2:126–152. https://doi.org/10.1137/120899285

Stefanescu ER, Bursik M, Cordoba G, Dalbey K, Jones MD, Patra AK, et al. (2012) Digital elevation model uncertainty and hazard analysis using a geophysical flow model. Proc R Soc Lond A Math Phys Eng Sci 468:1543–1563. https://doi.org/10.1098/rspa.2011.0711

Sulpizio R, Capra L, Sarocchi D, Saucedo R, Gavilanes-Ruiz JC, Varley NR (2010) Predicting the block-and-ash flow inundation areas at Volcán de Colima (Colima, Mexico) based on the present day (February 2010) status. J Volcanol Geotherm Res 193:49–66. https://doi.org/10.1016/j.jvolgeores.2010.03.007

Sulpizio R, Dellino P, Doronzo DM, Sarocchi D (2014) Pyroclastic density currents: state of the art and perspectives. J Volcanol Geotherm Res 283:36–65. https://doi.org/10.1016/j.jvolgeores.2014.06.014

Sulpizio R, Castioni D, Rodriguez-Sedano LA, Sarocchi D, Lucchi F (2016) The influence of slope-angle ratio on the dynamics of granular flows: insights from laboratory experiments. Bull Volcanol 78:1–11. https://doi.org/10.1007/S00445-016-1069-5

Suzuki YJ, Costa A, Cerminara M, Esposti Ongaro T, Herzog M, Van Eaton AR, Denby LC (2016) Inter-comparison of three-dimensional models of volcanic plumes. J Volcanol Geotherm Res 326:26–42. https://doi.org/10.1016/j.jvolgeores.2016.06.011

Sweeney MR, Valentine GA (2017) Impact zone dynamics of dilute mono- and polydisperse jets and their implications for the initial conditions of pyroclastic density currents. Phys Fluids 29: 093304. https://doi.org/10.1063/1.5004197

Tierz P, Sandri L, Costa A, Zaccarelli L, Di Vito MA, Sulpizio R, Marzocchi W (2016) Suitability of energy cone for probabilistic volcanic hazard assessment: validation tests at Somma-Vesuvius and Campi Flegrei (Italy). Bull Volcanol 78:79. https://doi.org/10.1007/s00445-016-1073-9

Tierz P, Stefanescu ER, Sandri L, Sulpizio R, Valentine GA, Marzocchi W, et al. (2018). Towards quantitative volcanic risk of pyroclastic density currents: probabilistic hazard curves and maps around somma-vesuvius (Italy). J Geophys Res Solid Earth 123:6299–6317. https://doi.org/10.1029/2017JB015383

Tierz P, Bevilacqua A, Mead S, Spiller E, Sandri L (2021) Field Data, Models and Uncertainty in Hazard Assessment of Pyroclastic Density Currents and Lahars: Global Perspectives. Front Earth Sci 9:665725. https://doi.org/10.3389/feart.2021.665725

Valentine GA, Wohletz KH (1989) Numerical models of Plinian eruption columns and pyroclastic flows. J Geophys Res 94: 1867–1887. https://doi.org/10.1029/JB094iB02p01867

Valentine GA, Sweeney MR (2018) Compressible Flow Phenomena at Inception of Lateral Density Currents Fed by Collapsing Gas-Particle Mixtures. J Geophys Res Solid Earth 123:1286–1302. https://doi.org/10.1002/2017JB015129

Valentine GA (2020) Initiation of dilute and concentrated pyroclastic currents from collapsing mixtures and origin of their proximal deposits. Bull Volcanol 82:20. https://doi.org/10.1007/s00445-020-1366-x

Voellmy A (1955) Über die Zerstörungskraft von Lawinen. Schweizerische Bauzeitung 73:159–165

Wadge G, Jackson P, Bower SM, Woods AW, Calder E (1998) Computer simulations of pyroclastic flows from dome collapse. Geophys Res Lett 25:3677–3680. https://doi.org/10.1029/98GL00710

Wang Y, Voskov D, Khait M, Bruhn D (2020) An efficient numerical simulator for geothermal simulation: A benchmark study. Appl Energy 264:114693. https://doi.org/10.1016/j.apenergy.2020.114693

Weit A, Roche O, Dubois T, Manga M (2018) Experimental Measurement of the Solid Particle Concentration in Geophysical Turbulent Gas-Particle Mixtures. J Geophys Res Solid Earth 123:3747–3761. https://doi.org/10.1029/2018JB015530

Widiwijayanti C, Voight B, Hidayat D, Schilling SP (2009) Objective rapid delineation of areas at risk from block-and-ash pyroclastic flows and surges. Bull Volcanol 71:687–703. https://doi.org/10.1007/s00445-008-0254-6

Yamada M, Mangeney A, Matsushi Y, Moretti L (2016) Estimation of dynamic friction of the Akatani landslide from seismic waveform inversion and numerical simulation. Geophys J Int 206:1479–1486. https://doi.org/10.1093/gji/ggw216