Comparative anatomy of geophysical flow models and modeling assumptions using uncertainty quantification

Abani K. Patra[1,3], Andrea Bevilacqua[2], Ali Akhavan-Safaei[1], E. Bruce Pitman[4], Marcus I. Bursik[2], and David Hyman[2]

[1]*Dept. of Mech. and Aero. Eng., University at Buffalo, Buffalo NY 14260*
[2]*Dept. of Geology, University at Buffalo, Buffalo NY 14260*
[3]*Comp. Data Science and Eng., University at Buffalo, Buffalo NY 14260*
[4]*Dept. of Materials Design and Innovation, University at Buffalo, NY, 14260*

{abani,abevilac,aliakhav,pitman,mib,davidhym}@buffalo.edu

Contents

Abstract

Complex systems (e.g., volcanoes, debris flows, climate) commonly have many models advocated by different modelers and incorporating different modeling assumptions. Limited and sparse data on the modeled phenomena does not permit a clean discrimination among models for fitness of purpose, and, heuristic choices are usually made, especially for critical predictions of behavior that has not been experienced. We advocate here for characterizing models and the modeling assumptions they represent using a statistical approach over the full range of applicability of the models. Such a characterization may then be used to decide the appropriateness of a model for use, and, perhaps as needed weighted compositions of models for better predictive power. We use the example of dense granular representations of natural mass flows in volcanic debris avalanches, to illustrate our approach.

1 Introduction

This paper presents a systematic approach to the study of models of complex systems, applied to the case of geophysical mass flows modeling.

A simple though not necessarily comprehensive definition of a model is that:

> *A model is a representation of a postulated relationship among inputs and outputs of a system usually informed by observation and a hypothesis that best explains them.*

The definition captures two of the most important characteristics

- models depend on a *hypothesis*, and,
- models use the *data from observation* to validate and refine the hypothesis.

Errors and uncertainty in the data and limitations in the hypothesis (usually a tractable and computable mathematical construct articulating beliefs like proportionality, linearity etc.) are immediate challenges that must be overcome to construct useful and credible models.

A model is most useful in predicting the behavior of a system for unobserved inputs and interpretability or explainability of the system's behavior. Since, models require a hypotheses implies that the model is a formulation of a belief about the data. The immediate consequence of this that the model may be very poor about such prediction even when sufficient care is taken to use all the available data and information since the *subjectivity of the belief* can never be completely eliminated. Secondly, the data at hand may not provide enough information about the system to characterize its behavior at the desired prediction. What makes this problem even more acute is that we are often interested in modeling outcomes that are not observed and perhaps sometimes not observable.

The consequence of this lack of knowledge and limited data is the multiplicity of beliefs about the complex system being modeled and a profusion of models based on different modeling assumptions and data use. These competing models lead to much debate among scientists. Principles like "Occam's razor" and Bayesian statistics Farrell et al. (2015) provide some guidance but simple robust approaches that allow the testing of models for fitness need to be developed. We present in this paper a simple data driven approach to discriminate among models and the modeling assumptions implicit in each model, given a range of phenomena to be studied. We illustrate the approach by work on granular flow models of large mass flows.

1.1 Models and assumptions

An assumption is a simple concept – any atomic postulate about relationships among quantities under study e.g., a linear stress strain relationship $\sigma = E\epsilon$ or neglecting some quantities in comparison to larger quantities $\theta \approx \sin(\theta)$ for small θ. Models are compositions of many such assumptions. The study of models is thus implicitly a study of these assumptions and their composability and applicability in a particular context. Sometimes a good model contains a useless assumption that may be removed, sometimes a good assumption should be implemented inside a different model - these are usually subjective choices, not data driven. Moreover, the correct assumptions may change through time, making *model choice* more difficult.

The rest of the paper will define our approach and a simple illustration using 3 models for large scale mass flows incorporated in our large scale mass flow simulation framework TITAN2D (Patra et al., 2005, 2006; Yu et al., 2009; Aghakhani et al., 2016). The 4^{th} release of TITAN2D offers multiple rheology options in the same cyber-infrastructure. The availability of 3 distinct models for similar phenomena in the same tool provides us the ability to directly compare inputs, outputs and internal variables in all the 3 models.

So far, TITAN2D achieved many successful applications in the simulation of different geophysical mass flows with specific characteristics (Sheridan et al., 2005; Rupp et al., 2006; Norini et al., 2009; Charbonnier and Gertisser, 2009; Procter et al., 2010; Sheridan et al., 2010; Sulpizio et al., 2010; Capra et al., 2011). Several studies involving TITAN2D were recently directed towards a statistical study of geophysical flows, focusing on uncertainty quantification and propagation (Dalbey et al., 2008; Dalbey, 2009; Stefanescu et al., 2012a,b), or on the more efficient production of hazard maps (Bayarri et al., 2009; Spiller et al., 2014; Bayarri et al., 2015; Ogburn et al., 2016).

1.2 Analysis of Modeling assumptions and models

Let us define $\big(M(A), P_{M(A)}\big)$, where A is a set of assumptions, $M(A)$ is the model which combines those assumptions, and P_M is a probability distribution in the parameter space of M. For the sake of simplicity we assume P_M to be uniformly distributed on selected parameter ranges. While the support of P_M can be restricted to a single value by solving an inverse problem for the optimal reconstruction of a particular flow, this is not possible if we are interested in the general predictive capabilities of the model, where we are interested in the outcomes over a whole range.

Stage 1: Parameter Ranges In this study, we always assume

$$P_M \sim \bigotimes_{i=1}^{N_M} Unif(a_{i,M}, b_{i,M}),$$

where N_M is the number of parameters of M. These parameter ranges will be chosen using information gathered from the literature about the physical meaning of those values together with a preliminary testing for physical consistency of model outcomes and range of inputs/outcomes of interest. If the total friction of the models does not cover a similar span, the statistical comparison is dominated by trivial macroscopic differences, and cannot focus on the rheology details.

Stage 2 Simulations and Data Gathering The simulation algorithms can be represented as:

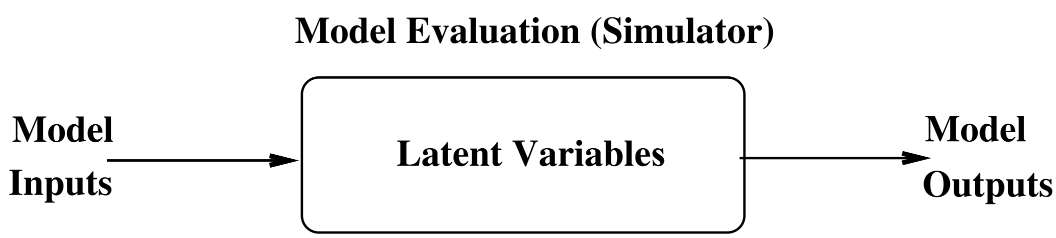

Figure 1: Models and variables

The *model inputs* are the parameters of M, the *latent variables* include quantities in the model evaluation that are ascribable to specific assumptions A_i. These are usually not observed as outputs from the model. For example in momentum balances of complex flow calculations these could be values of different source terms, dissipation terms and inertia terms. Finally, the *model outputs* include explicit outcomes e.g., for flow calculations these could be flow height, lateral extent, area, velocity, acceleration, and derived quantities such as Froude number Fr. In general, for each quantity of interest (QoI), we use a Monte Carlo simulation, sampling the input variables and obtaining a family of graphs plotting their expectation, and their 5^{th} and 95^{th} percentiles. Our sampling technique of the input variables is based on the Latin Hypercube Sampling (LHS) idea, and in particular, on the improved space-filling properties of the orthogonal array-based Latin Hypercubes (see Appendix A).

Stage 3 Results Analysis These and other statistics can now be compared to determine the need for different modeling assumptions and the relative merits of different models. Thus, analysis of the data gathered over the entire range of flows for the state variables and outcomes leads to a quantitative basis for accepting or rejecting particular assumptions or models for specific outcomes.

2 Modeling of geophysical mass flows

Dense large scale granular avalanches are a complex class of flows with physics that has often been poorly captured by models that are computationally tractable. Sparsity of actual flow data (usually only a posteriori deposit information is available), and large uncertainty in the mechanisms of initiation and flow propagation, make the modeling task challenging, and a subject of much continuing interest. Models that appear to represent the physics well in certain flows, may turn out to be poorly behaved in others, due to intrinsic mathematical or numerical issues. Nevertheless, given the large implications on life and property, many models with different modeling assumptions have been proposed. For example in Iverson (1997); Iverson and Denlinger (2001); Denlinger and Iverson (2001); Pitman et al. (2003b); Denlinger and Iverson (2004); Iverson et al. (2004), the depth-averaged model was applied in the simulation of test geophysical flows in large scale experiments. Several studies were specifically devoted to the modeling of volcanic mass flows (Bursik et al., 2005; Kelfoun and Druitt, 2005; Charbonnier and Gertisser, 2009; Kelfoun et al., 2009; Procter et al., 2010; Kelfoun, 2011; Charbonnier et al., 2013). In fact, volcanoes are great sources for a rich variety of geophysical flow types and provide field data from past flow events.

Modeling in this case proceeds by first assuming that the laws of mass and momentum conservation hold for properly defined system boundaries. The scale of these flows – very long and wide with small depth led to the first most generally accepted assumption – shallowness Savage and Hutter (1989). This allows an integration through the depth to obtain simpler and more computationally tractable equations. This is the next of many assumptions that have to be made. Both of these are fundamental assumptions which can be tested in the procedure we established above. Since, there is a general consensus and much evidence in the literature of the validity of these assumptions we defer analysis of these to future work.

The depth-averaged Saint-Venant equations that result are:

$$
\begin{aligned}
\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(h\bar{u}) + \frac{\partial}{\partial y}(h\bar{v}) &= 0 \\
\frac{\partial}{\partial t}(h\bar{u}) + \frac{\partial}{\partial x}\left(h\bar{u}^2 + \frac{1}{2}kg_z h^2\right) + \frac{\partial}{\partial y}(h\bar{u}\bar{v}) &= S_x \\
\frac{\partial}{\partial t}(h\bar{v}) + \frac{\partial}{\partial x}(h\bar{u}\bar{v}) + \frac{\partial}{\partial y}\left(h\bar{v}^2 + \frac{1}{2}kg_z h^2\right) &= S_y
\end{aligned}
\tag{1}
$$

Here the Cartesian coordinate system is aligned such that z is normal to the surface; h is the flow height in the z direction; $h\bar{u}$ and $h\bar{v}$ are respectively the components of momentum in the x and y directions; and k is the coefficient which relates the lateral stress components, $\bar{\sigma}_{xx}$ and $\bar{\sigma}_{yy}$, to the normal stress component, $\bar{\sigma}_{zz}$. The definition of this coefficient depends on the constitutive model of the flowing material we choose. Note that $\frac{1}{2}kg_z h^2$ is the contribution of hydrostatic pressure to the momentum fluxes. S_x and S_y are the sum local stresses: they include the gravitational driving forces, the basal friction force resisting to the motion of the material, and additional forces specific of rheology assumptions.

The final class of assumptions are the assumptions on the rheology of the flows – in particular in this context assumptions used to model different dissipation mechanisms embedded in S_x, S_y that lead to a plethora of models with much controversy on the most suitable model.

2.1 Overview of the models

In the three following sections, we briefly describe *Mohr-Coulomb* (MC), *Pouliquen-Forterre* (PF) and *Voellmy-Salm* (VS) models. Models based on additional heterogeneous assumptions are possible, either more complex (Pitman and Le, 2005; Iverson and George, 2014) or more simple (Dade and Huppert, 1998), but we are focusing on those because of their historical relevance and their comparable degree of complexity.

2.1.1 Mohr-Coulomb

Based on the long history of studies in soil mechanics (Rankine, 1857; Jaeger et al., 1989), the Mohr-Coulomb rheology (MC) was developed and used to represent the behavior of geophysical mass flows Savage and Hutter (1989).

Shear and normal stress are assumed to obey Coulomb friction equation, both within the flow and at its boundaries. In other words,

$$\tau = \sigma \tan\phi, \tag{2}$$

where τ and σ are respectively the shear and normal stresses on failure surfaces, and ϕ is a friction angle. This relationship does not depend on the flow speed.

We can summarize the MC rheology assumptions as:

- *Basal Friction* based on a constant friction angle.
- *Internal Friction* based on a constant friction angle.

- *Earth pressure coefficient* formula depends on the Mohr circle.
- Velocity based *curvature effects* are included into the equations.

Under the assumption of symmetry of the stress tensor with respect to the z axis, the earth pressure coefficient $k = k_{ap}$ can take on only one of three values $\{0, \pm 1\}$. The material yield criterion is represented by the two straight lines at angles $\pm\phi$ (the internal friction angle) relative to horizontal direction. Similarly, the normal and shear stress at the bed are represented by the line $\tau = -\sigma \tan(\delta)$ where δ is the bed friction angle.

MC equations As a result, we can write down the source terms of the Eqs. (1):

$$\begin{aligned} S_x &= g_x h - \frac{\bar{u}}{\|\underset{\sim}{\bar{\mathbf{u}}}\|}\left[h\left(g_z + \frac{\bar{u}^2}{r_x}\right)\tan(\phi_{bed})\right] - hk_{ap}\,\mathrm{sgn}\left(\frac{\partial \bar{u}}{\partial y}\right)\frac{\partial(g_z h)}{\partial y}\sin(\phi_{int}) \\ S_y &= g_y h - \frac{\bar{v}}{\|\underset{\sim}{\bar{\mathbf{u}}}\|}\left[h\left(g_z + \frac{\bar{v}^2}{r_y}\right)\tan(\phi_{bed})\right] - hk_{ap}\,\mathrm{sgn}\left(\frac{\partial \bar{v}}{\partial x}\right)\frac{\partial(g_z h)}{\partial x}\sin(\phi_{int}) \end{aligned} \tag{3}$$

Where, $\underset{\sim}{\bar{\mathbf{u}}} = (\bar{u}, \bar{v})$, is the depth-averaged velocity vector, r_x and r_y denote the radii of curvature of the local basal surface. The inverse of the radii of curvature is usually approximated with the partial derivatives of the basal slope, e.g., $1/r_x = \partial\theta_x/\partial x$, where θ_x is the local bed slope.

In our study, sampled input parameters are ϕ_{bed}, and $\Delta\phi := \phi_{int} - \phi_{bed}$. In particular, the range of ϕ_{bed} depends on the case study, while $\Delta\phi \in [2^\circ, 10^\circ]$ (Dalbey et al., 2008).

2.1.2 Pouliquen-Forterre

The scaling properties for granular flows down rough inclined planes led to the development of the Pouliquen-Forterre rheology (PF), assuming a variable frictional behavior as a function of flow regime (i.e. Froude Number, Fr) and flow depth (Pouliquen, 1999; Forterre and Pouliquen, 2002; Pouliquen and Forterre, 2002; Forterre and Pouliquen, 2003).

PF rheology assumptions can be summarized as:

- *Basal Friction* is based on an interpolation of two different friction angles, based on the flow regime and depth.
- *Internal Friction* is neglected.
- *Earth pressure coefficient* is equal to one.
- Normal stress is modified by a *hydrostatic pressure force* related to the flow height gradient.
- Velocity based *curvature effects* are included into the equations.

Two critical slope inclination angles are defined as functions of the flow thickness, namely $\phi_{start}(h)$ and $\phi_{stop}(h)$. The function $\phi_{stop}(h)$ gives the slope angle at which a steady uniform flow leaves a deposit of thickness h, while $\phi_{start}(h)$ is the angle at which a layer of thickness h is mobilized. They define two different basal friction coefficients.

$$\mu_{start}(h) = \tan(\phi_{start}(h)) \tag{4}$$

$$\mu_{stop}(h) = \tan(\phi_{stop}(h)) \tag{5}$$

An empirical friction law $\mu_b(\|\underset{\sim}{\bar{\mathbf{u}}}\|, h)$ is then defined in the whole range of velocity and thickness. The expression changes depending on two flow regimes, according to a parameter β and the Froude number $Fr = \|\underset{\sim}{\bar{\mathbf{u}}}\|/\sqrt{hg_z}$.

Dynamic friction regime - $Fr \geq \beta$

$$\mu(h, Fr) = \mu_{stop}(h\beta/Fr) \tag{6}$$

Intermediate friction regime - $0 \leq Fr < \beta$

$$\mu(h, Fr) = \left(\frac{Fr}{\beta}\right)^{\gamma} [\mu_{stop}(h) - \mu_{start}(h)] + \mu_{start}(h), \tag{7}$$

where γ is the power of extrapolation, assumed equal to 10^{-3} in the sequel (Pouliquen and Forterre, 2002).

The functions μ_{stop} and μ_{start} are defined by:

$$\mu_{stop}(h) = \tan\phi_1 + \frac{\tan\phi_2 - \tan\phi_1}{1 + h/\mathcal{L}} \tag{8}$$

and

$$\mu_{start}(h) = \tan\phi_3 + \frac{\tan\phi_2 - \tan\phi_1}{1 + h/\mathcal{L}} \tag{9}$$

The critical angles ϕ_1, ϕ_2 and ϕ_3 and the parameters $\mathcal{L}, \beta$ are the parameters of the model.

In particular, $\mathcal{L}$ is the characteristic depth of the flow over which a transition between the angles ϕ_1 to ϕ_2 occurs, in the μ_{stop} formula. In practice, if $h \ll \mathcal{L}$, then $\mu_{stop}(h) \approx \tan\phi_2$, and if $h \gg \mathcal{L}$, then $\mu_{stop}(h) \approx \tan\phi_1$. The effect of the topographic local curvatures is also taken into account.

PF equations The depth-averaged Eqs. (1) source terms thus take the following form:

$$\begin{aligned} S_x &= g_x h - \frac{\bar{u}}{\|\underset{\sim}{\bar{\mathbf{u}}}\|}\left[h\left(g_z + \frac{\bar{u}^2}{r_x}\right)\ \mu_b(\|\underset{\sim}{\bar{\mathbf{u}}}\|, h)\right] + g_z h\frac{\partial h}{\partial x} \\ S_y &= g_y h - \frac{\bar{v}}{\|\underset{\sim}{\bar{\mathbf{u}}}\|}\left[h\left(g_z + \frac{\bar{v}^2}{r_y}\right)\ \mu_b(\|\underset{\sim}{\bar{\mathbf{u}}}\|, h)\right] + g_z h\frac{\partial h}{\partial y} \end{aligned} \tag{10}$$

In our study, sampled input parameters are ϕ_1, $\Delta\phi_{12} := \phi_2 - \phi_1$, and β. In particular, the range of ϕ_1 depends on the case study, whereas $\Delta\phi_{12} \in [10°, 15°]$, and $\beta \in [0.1, 0.85]$. Moreover, $\phi_3 = \phi_1 + 1°$, and $\mathcal{L}$ is equal to 1 dm and 1 mm in the two case studies, respectively (Pouliquen and Forterre, 2002; Forterre and Pouliquen, 2003).

2.1.3 Voellmy-Salm

The theoretical analysis of dense snow avalanches led to the VS rheology (VS) (Voellmy, 1955; Salm et al., 1990; Salm, 1993; Bartelt et al., 1999). Dense snow or debris avalanches consist of mobilized, rapidly flowing ice-snow mixed to debris-rock granules (Bartelt and McArdell, 2009). The VS rheology assumes a velocity dependent resisting term in addition to the traditional basal friction, ideally capable of including an approximation of the turbulence-generated dissipation. Many experimental and theoretical studies were developed in this framework (Gruber and Bartelt, 2007; Kern et al., 2009; Christen et al., 2010; Fischer et al., 2012).

The following relation between shear and normal stresses holds:

$$\tau = \mu\sigma + \frac{\rho\|\underline{\mathbf{g}}\|}{\xi}\|\underset{\sim}{\bar{\mathbf{u}}}\|^2, \tag{11}$$

where, σ denotes the normal stress at the bottom of the fluid layer and $\underline{\mathbf{g}} = (g_x, g_y, g_z)$ represents the gravity vector. The VS rheology adds a velocity dependent *turbulent* friction to the traditional velocity independent basal friction term which is proportional to the normal stress at the flow bottom. The two parameters of the model are the bed friction coefficient μ and the turbulent friction coefficient ξ.

We can summarize VS rheology assumptions as:

- *Basal Friction* is based on a constant coefficient, similarly to the MC rheology.
- *Internal Friction* is neglected.
- *Earth pressure coefficient* is equal to one.
- Additional *turbulent friction* is based on the local velocity by a quadratic expression.

- Velocity based *curvature effects* are included into the equations, following an alternative formulation.

The effect of the topographic local curvatures is again taken into account by adding the terms containing the local radii of curvature r_x and r_y. In this case the formula is considering the modulus of velocity instead than the scalar component Pudasaini and Hutter (2003); Fischer et al. (2012).

VS equations Therefore, the final source terms take the following form:

$$\begin{aligned} S_x &= g_x h - \frac{\bar{u}}{\|\underset{\sim}{\bar{\mathbf{u}}}\|} \left[h \left(g_z + \frac{\|\underset{\sim}{\bar{\mathbf{u}}}\|^2}{r_x} \right) \mu + \frac{\|\underset{\sim}{\mathbf{g}}\|}{\xi} \|\underset{\sim}{\bar{\mathbf{u}}}\|^2 \right], \\ S_y &= g_y h - \frac{\bar{v}}{\|\underset{\sim}{\bar{\mathbf{u}}}\|} \left[h \left(g_z + \frac{\|\underset{\sim}{\bar{\mathbf{u}}}\|^2}{r_y} \right) \mu + \frac{\|\underset{\sim}{\mathbf{g}}\|}{\xi} \|\underset{\sim}{\bar{\mathbf{u}}}\|^2 \right]. \end{aligned} \tag{12}$$

In our study, sampled input parameters are μ, and ξ, on ranges depending on the case study. In particular, ξ uniform sampling is accomplished in log-scale. In fact, values of ξ between 250 and 4,000 m/s^2 have been described for snow avalanches (Salm, 1993; Bartelt et al., 1999; Gruber and Bartelt, 2007).

2.2 Latent variables

For analysis of modeling assumptions we need to record and classify the results of different modeling assumptions. In our case study, we focus on the right-hand side terms in the momentum equation, we call them RHS forces. These terms are explored in detail in the next sections.

$$\boldsymbol{RHS_1} = [g_x h, g_y h], \tag{13}$$

it is the gravitational force term, it has the same formulation in all models.

The formula of **basal friction force $\boldsymbol{RHS_2}$** depends on the model:

$$\begin{aligned} \boldsymbol{RHS_2} &= -\, h g_z \tan(\phi_{bed}) \left[\frac{\bar{u}}{\|\underset{\sim}{\bar{\mathbf{u}}}\|}, \frac{\bar{v}}{\|\underset{\sim}{\bar{\mathbf{u}}}\|} \right], \text{ in MC model.} \\ \boldsymbol{RHS_2} &= -\, h g_z \; \mu_b(\|\underset{\sim}{\bar{\mathbf{u}}}\|, h) \left[\frac{\bar{u}}{\|\underset{\sim}{\bar{\mathbf{u}}}\|}, \frac{\bar{v}}{\|\underset{\sim}{\bar{\mathbf{u}}}\|} \right], \text{ in PF model.} \\ \boldsymbol{RHS_2} &= -\, h g_z \mu \left[\frac{\bar{u}}{\|\underset{\sim}{\bar{\mathbf{u}}}\|}, \frac{\bar{v}}{\|\underset{\sim}{\bar{\mathbf{u}}}\|} \right], \text{ in VS model.} \end{aligned} \tag{14}$$

The formula of the force related to the **topography curvature**, $\boldsymbol{RHS_3}$, also depends on the model:

$$\begin{aligned} \boldsymbol{RHS_3} &= -\, h \tan(\phi_{bed}) \left[\frac{\bar{u}^3}{r_x \|\underset{\sim}{\bar{\mathbf{u}}}\|}, \frac{\bar{v}^3}{r_y \|\underset{\sim}{\bar{\mathbf{u}}}\|} \right], \text{ in MC model.} \\ \boldsymbol{RHS_3} &= -\, h \; \mu_b(\|\underset{\sim}{\bar{\mathbf{u}}}\|, h) \left[\frac{\bar{u}^3}{r_x \|\underset{\sim}{\bar{\mathbf{u}}}\|}, \frac{\bar{v}^3}{r_y \|\underset{\sim}{\bar{\mathbf{u}}}\|} \right], \text{ in PF model.} \\ \boldsymbol{RHS_3} &= -\, h \mu \left[\frac{\bar{u} \|\underset{\sim}{\bar{\mathbf{u}}}\|}{r_x}, \frac{\bar{v} \|\underset{\sim}{\bar{\mathbf{u}}}\|}{r_y} \right], \text{ in VS model.} \end{aligned} \tag{15}$$

All the three models have an additional force term, having a different formula and meaning in the three models:

$$\begin{aligned}
\boldsymbol{RHS_4} =& - hk_{ap}\sin(\phi_{int})\left[\operatorname{sgn}(\frac{\partial\bar{u}}{\partial y})\frac{\partial(g_z h)}{\partial y},\ \operatorname{sgn}(\frac{\partial\bar{v}}{\partial x})\frac{\partial(g_z h)}{\partial x}\right],\ \text{in MC model.}\\
\boldsymbol{RHS_4} =& g_z h\left[\frac{\partial h}{\partial x},\frac{\partial h}{\partial y}\right],\ \text{in PF model.}\\
\boldsymbol{RHS_4} =& -\frac{\|\underset{\sim}{\mathbf{g}}\|}{\xi}\|\underset{\sim}{\bar{\mathbf{u}}}\|^2\left[\frac{\bar{u}}{\|\underset{\sim}{\bar{\mathbf{u}}}\|},\frac{\bar{v}}{\|\underset{\sim}{\bar{\mathbf{u}}}\|}\right],\ \text{in VS model.}
\end{aligned} \tag{16}$$

These latent variables can be analyzed locally and globally for discriminating among the different modeling assumption.

2.3 Monte Carlo Process and Statistical Analysis

For our study, the flow range is defined by establishing boundaries for inputs like flow volume V and different rheology coefficients characterizing the different models. The Latin Hypercube Sampling is performed over $[0,1]^3$ for the MC and VS input parameters, and $[0,1]^4$ for PF input parameters. Those dimensionless samples are linearly mapped to fill the required intervals.

Following the simulations, we generate data for each sample run and each outcome and latent variable $f(\underline{\mathbf{x}},t)$ calculated as a function of time on the elements of the computational grid. This analysis generates tremendous volume of data which must then be analyzed using statistical methods for summative impact. The latent variables in this case are the mass and force terms in the conservation laws defined above. In more detail, local sampling is performed by considering the elements of the adapting mesh which are found to contain the sample points. Instead, spatial integrals are defined by $F(t)=\int_{\mathbb{R}^k} f(\underline{\mathbf{x}},t)d\underline{\mathbf{x}}$. In the most of the cases $k=2$, and $d\underline{\mathbf{x}}$ is given by the area of the mesh elements. Sometimes $k=3$, e.g. concerning speed, and $d\underline{\mathbf{x}}$ is the element of volume corresponding to the mesh elements.

We devise many statistical measures for analyzing the data. For instance, let $(F_i(x,t))_{i=1,\dots,4}$ be an array of force components, where $x\in\mathbf{R}^2$ is a spatial location, and $t\in T$ is a time instant. The degree of contribution of those force terms can be significantly variable in space and time, and we define the *dominance factors* $(p_j)_{j=1,\dots,k}$, i.e., the probability of each F_j to be the dominant force at (x,t). Those probabilities provide insight into the dominance of a particular source or dissipation (identified with a particular modeling assumption) term on the model dynamics. The dominance factors can be adopted to define a statistical decomposition of the contributions of the forces, as detailed in Appendix B.

3 Small scale flow on inclined plane and flat runway

The first case study assumes very simple boundary conditions, and corresponds to an experiment fully described in Webb (2004); Bursik et al. (2005); Webb and Bursik (2016). It is a classical flow down an inclined plane set-up, including a change in slope to an horizontal plane (Fig. 2). Modeling flow of granular material down an inclined plane was explored in detail by several studies, both theoretically and experimentally (Ruyer-Quil and Manneville, 2000; Silbert et al., 2001; Pitman et al., 2003a; Da Cruz et al., 2005).

In our setting, four locations are selected among the center line of the flow to accomplish local testing. These are: the initial pile location $L_1=(-0.7,0)$ m, the middle of the inclined plane $L_2=(-0.35,0)$ m, the change in slope $L_3=(0,0)$ m, the middle of the flat plane $L_4=(0.15,0)$ m (see Section 3).

3.1 Preliminary consistency testing of the input ranges

In this same case study, Dalbey et al. (2008) assumed $\phi_{bed}=[15°,30°]$, while Webb and Bursik (2016) performed a series of laboratory experiments and found $\phi_{bed}=[18.2°,34.4°]$. We relied on those published parameter choices to decide a comprehensive parameter range. Figure 2b displays the maps of max flow height and max velocities observed in the extreme cases tested. Simulation options are - max_time = 2 s, height/radius = 1.34, length_scale = 1 m, number_of_cells_across_axis = 10, order = first, geoflow_tiny = 1e4 (Patra et al., 2005; Aghakhani et al., 2016). Initial pile geometry is cylindrical.

- **Material Volume:** $[449.0\ ,\ 607.0]$ cm^3, i.e. average of 528.0 cm^3 and uncertainty of $\pm15\%$.

- **Rheology models' parameter space:**
  The parameter ranges adopted in this case study are:

  MC - $\phi_{bed} \in [18°, 30°]$.

  PF - $\phi_1 \in [10°, 22°]$.

  VS - $\mu \in [0.22, 0.45]$, $\log(\xi) \in [3, 4]$.

Even if maximum and mininum runout are both matching, the shape and lateral extent of the flow is remarkably different in the maximum runout case, according to the three models. In particular, MC model can produce the largest lateral extent, and VS model displays an accentuated bow-like shape, due to the increased friction in the lateral margins.

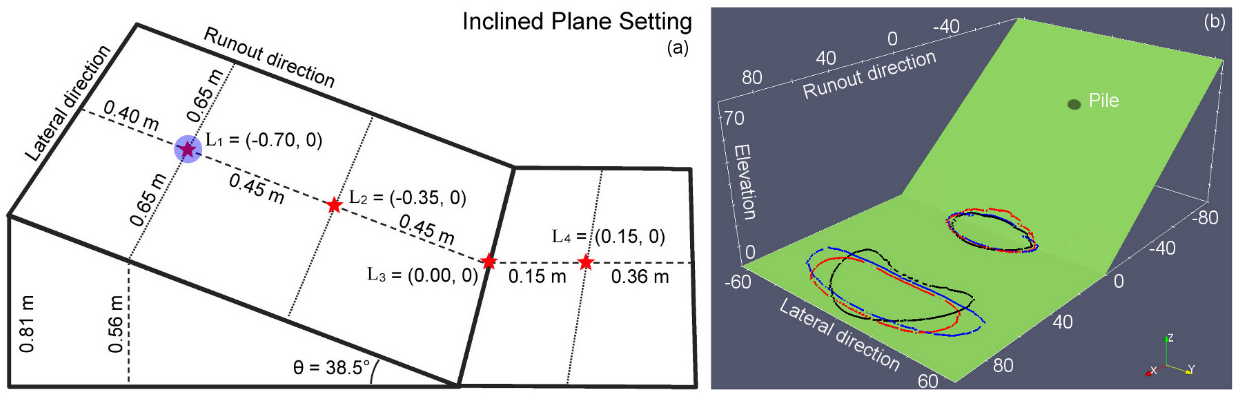

Figure 2: (a) Inclined plane overview, including local samples sites (red stars). Pile location is marked by a blue dot. (b) Contours of $h = 1.0$ mm at last simulated snapshot ($t = 1.5$ sec) for simulated flows with *minimum runout* obtained from ***min volume*** – ***max resistance***, and *maximum runout* obtained from ***max volume*** – ***min resistance***. — : MC, — : PF, — : VS.

3.2 Observable outputs

Observable outputs include the flow height and acceleration as a function of time, measured in the four locations $L_1, \ldots, L_4$ displayed in Fig. 2a. In addition, lateral extension, flow area, and spatially averaged speed and Fr are displayed. Uncertainty quantification (UQ) is always performed, accordingly to the parameter ranges described in Section 3.1. Locally measured Froude Number is included in Supporting Information S1.

3.2.1 Flow height

Figure 3 shows the flow height, $h(L, t)$, at the points $(L_i)_{i=1,\ldots,4}$, for the three rheology models. It clearly shows the differences in the statistics of the flow outcomes induced by the different choices of rheology at different locations in the plane. Availability of data allows us to subject the data to tests of reasonability both for the means and extremal values. Given a particular type of flow and collected data we can clearly distinguish model skill in capturing not only that flow but also possible flows. Past work Webb (2004) allows us to conclude that MC rheology is adequate for modeling simple dry granular flows.

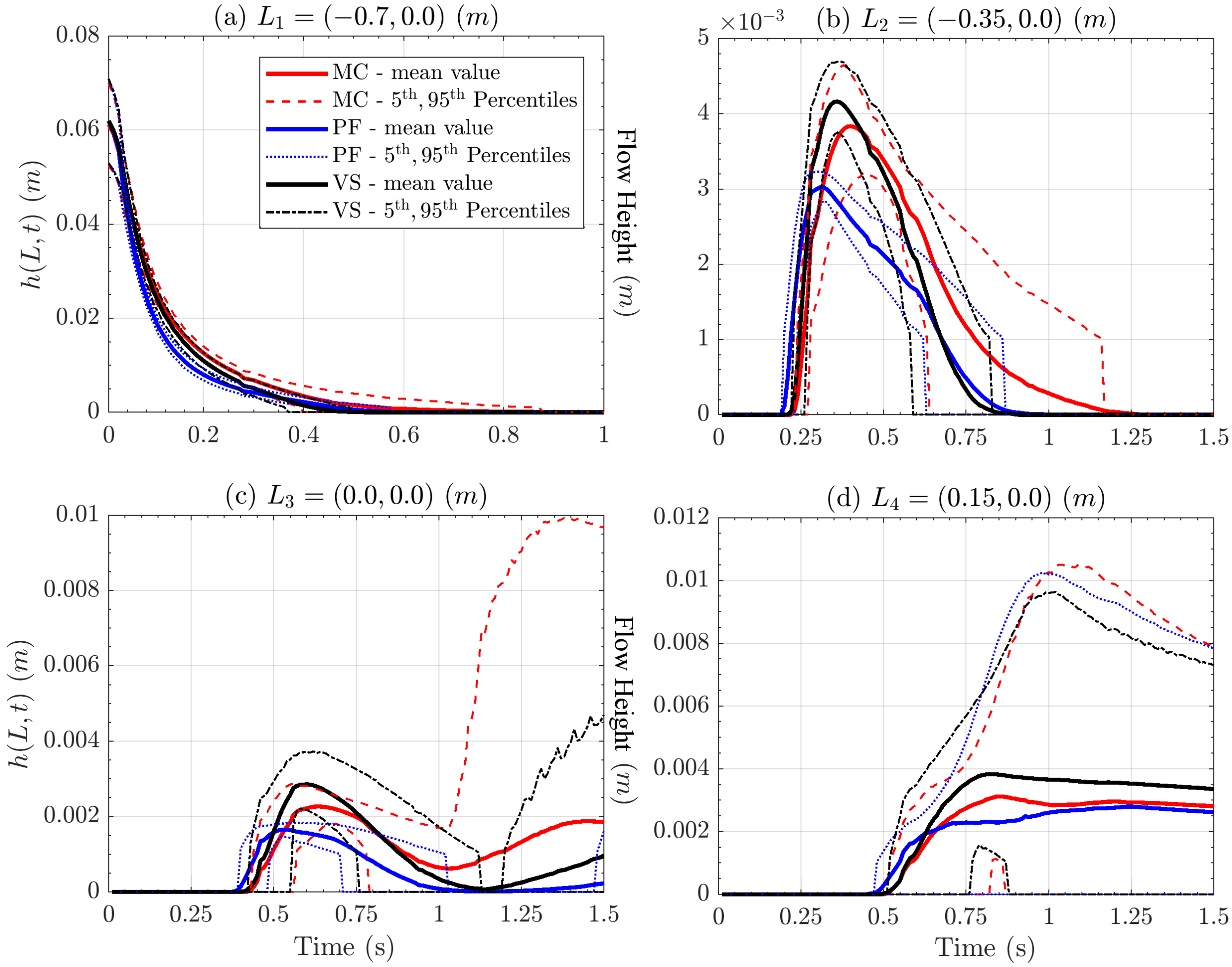

Figure 3: Records of flow height at four spatial locations of interest. Bold line is mean value, dashed/dotted lines are 5^{th} and 95^{th} percentile bounds. Different rheology models are displayed with different colors. Plots are at different scale, for simplifying lecture.

We note the effect of cutting the flow when height is < 1 mm, which is at the scale of the smallest granular size. In fact, continuity assumption would not be valid below this scale. The 5^{th} and 95^{th} percentile plots are vertically cut to zero when they decrease over that threshold. The mean plot is not cut to zero but it is dulled by this cutoff. In plot 3a, related to point L_1 placed on the initial pile, the initial values of $\sim 6 \pm 1$ cm are the equal and express the assigned pile height. The flow height decreases slightly faster in PF model, and slower in MC, compared to VS. Differences are more significant in plot 3b, related to point L_2, placed in the middle of the slope. Maximum flow height on average is greater in VS, 4.1 ± 0.2 mm, but more uncertain in MC, 3.9 ± 0.4 mm, and generally smaller in PF model, 3.0 ± 0.1 mm. After the peak, PF decreases significantly slower than the other models. These height values are about 15 times smaller than initial pile height. None of the models leaves a significant material deposit in L_1 or L_2, and hence the 95^{th} percentile of the height is null at the ending-time. In contrast, a deposit is left at points L_3 and L_4, i.e. plot 3c placed at the change in slope, and plot 3d in the middle of the flat runout. At L_3 MC's deposit, 2 mm with uncertainty [-2,+8] mm, is higher than the other models'. The plot profile is bimodal, showing a first peak at ~ 0.6 s, and then a reduction until 1 s, before the final accumulation. At L_4,deposit it is not significantly different between the three models. It measures ~ 3 mm on average, slightly more than this in VS, with uncertainty [-3,+7] mm.

3.2.2 Flow acceleration

Figure 4 shows the flow speed, $\|\underline{\mathbf{a}}\|(L,t)$, at the points $(L_i)_{i=1,\dots,4}$, for the three rheology models.

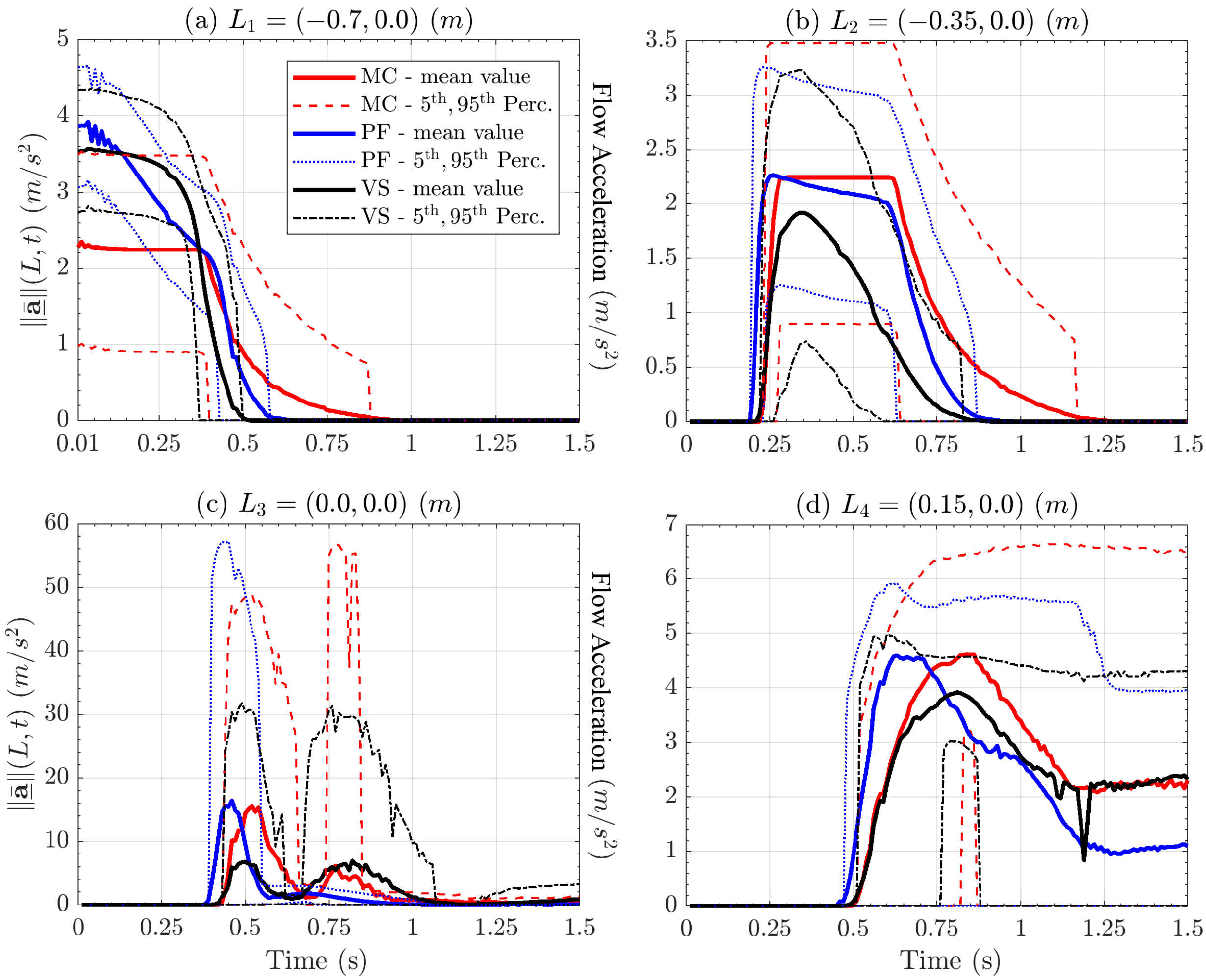

Figure 4: Records of flow acceleration magnitude (computed from LHS). Bold line is mean value, dashed lines are 5^{th} and 95^{th} percentile bounds. Different rheology models are displayed with different colors. Plots are at different scale.

Acceleration is the link between force terms and observable motion. We calculated it from the LHS of the dynamical equation, but using the RHS terms produces very similar results, although not identical due to the numerical approximations in solving the equations. In plot 4a, related to point L_1, MC and VS show a plateau before ~ 0.4 sec, at ~ 2.5 m/s^2 and ~ 3.5 m/s^2, respectively, while PF linearly decreases between those same values. In plot 4b, related to point L_2, are MC and PF to show a plateau, at ~ 2.2 m/s^2, while VS has a more bell-shaped profile. UQ tells us that PF has a smaller uncertainty than the other models. In plot 4c, related to point L_3, all the models show a bimodal profile, with peaks at ~ 0.5 sec and 0.8 sec. This is more accentuated in MC and VS, whereas the second peak is almost absent from PF's profile. The second peak is motivated by an increase of velocity in the down-slope direction after its reduction due to lateral spreading of material. At the first peak, acceleration values are significant, with average peaks in MC and PF both at ~ 15 m/s^2, and 95^{th} percentile plot reaching ~ 50 m/s^2 and ~ 55 m/s^2, respectively. VS shows about halved acceleration peak values. At the second peak, average acceleration values are similar in MC and VS, at $\sim 5\ m/s^2$. In contrast, 95^{th} percentile plot is $> 50\ m/s^2$ for MC, while $\sim 30\ m/s^2$ in VS. In plot 4d, related to point L_4, the acceleration has a first peak at $\sim 4\ m/s^2$, and a final asymptote at $\sim 2\ m/s^2$ for MC and VS, $\sim 1\ m/s^2$ for PF. These values generally mean flow deceleration, and uncertainty is more relevant in MC and PF than in VS.

3.2.3 Flow extents and spatial integrals

Figure 5 shows the spatial average of speed and Froude Number, in the three rheology models. Moreover, it shows the (maximum) lateral extent and inundated area of flow, as a function of time. Spatial averages and global quantities as maximum lateral extent and inundated area have smoother plots than local measurements. Most of the details observed in local measurements are not easy to discern.

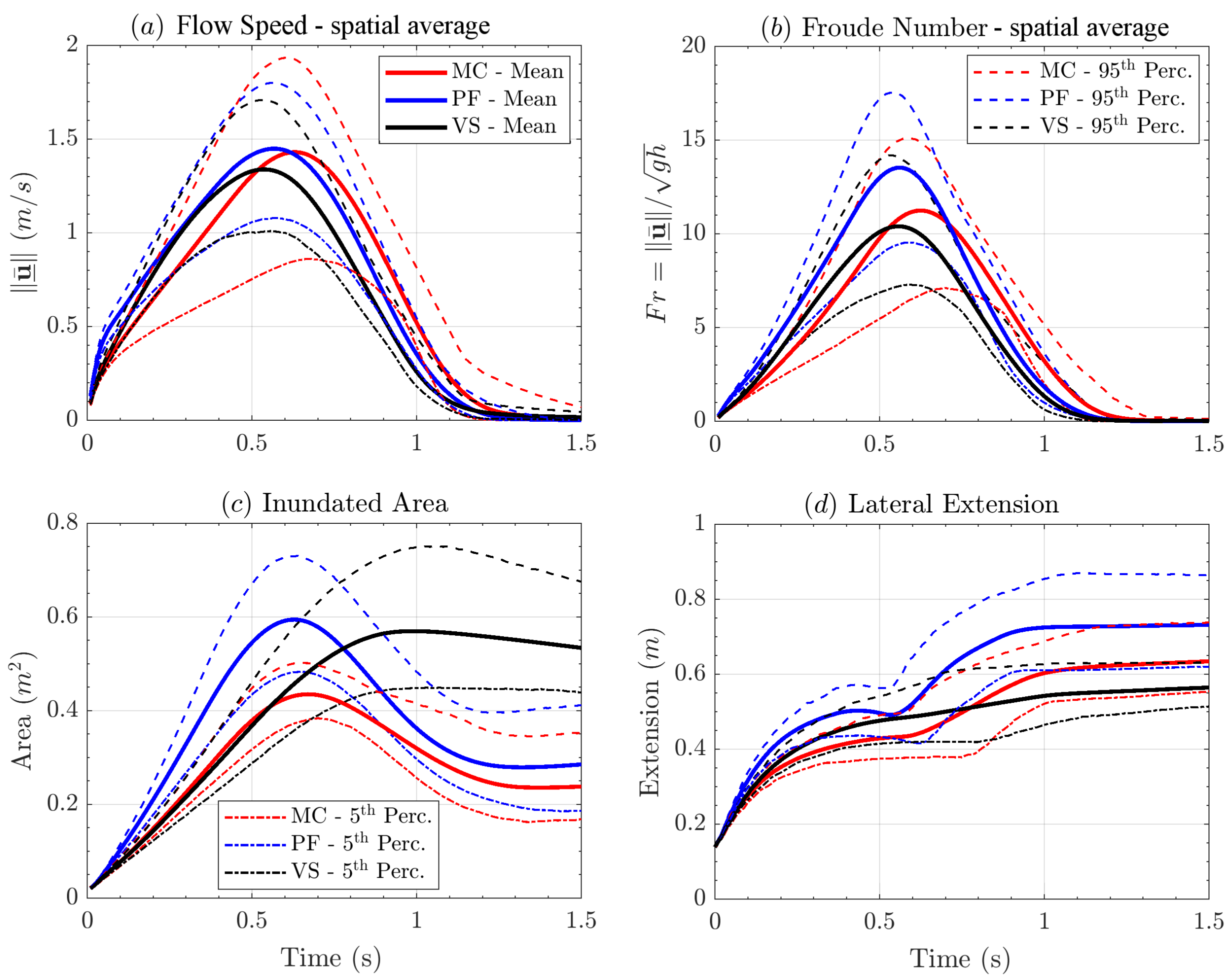

Figure 5: Comparison between spatial averages of (a) flow speed, and (b) Froude Number in addition to the flow (c) lateral extent, and (d) inundated area, as a function of time.

In plot 5a the speed shows a bell-shaped profile in all the models, with an average peak at $\sim 1.4m/s$ and uncertainty range, given by 5^{th} and 95^{th} percentiles, of $\pm 0.4m/s$ for PF and VS. VS is slightly slower, reaching $\sim 1.3m/s$ in the average. MC shows a larger uncertainty range, of $\pm 0.6m/s$. The maximum speed is reached first by VS and PF at $\sim 0.55s$, and last by MC at $\sim 0.65s$. In plot 5b, also Froude Numbers have a bell-shaped profile. Fr peaks are temporally aligned with speed peaks, and are ~ 10 in VS, ~ 11 in MC, ~ 13.5 in PF, on average. Uncertainty range is about ± 4 in all models. In plot 5c inundated area shows similar max values in PF and VS, at $\sim 0.6m^2$ on average, and uncertainty of

$\pm 0.15m^2$. MC is lower, at $\sim 0.45m^2$ on average, and less uncertain, $\pm 0.10m^2$. VS does not decrease significantly after reaching the peak, whereas the other models contract their area, to approximately half of the maximum extent. In plot 5d the lateral extent starts equal to the pile diameter $15cm$, and then rises in two stages in MC and PF, the second and greater rise starting at $\sim 0.6s$, corresponding to the time of arrival at the change in slope L_3 (see Fig. 3c). VS rises without showing two phases. After the first phase, average lateral extent is at $\sim 50cm$ in PF and VS, while it is $\sim 43cm$ in MC. Uncertainty range is $\pm 7cm$ for all models at that time. Final extent is $\sim 75cm$ in PF, $\sim 65cm$ in MC, $\sim 55cm$ in VS. Uncertainty range is $\pm 5cm$ in VS, but rises to $\pm 10cm$ in MC and PF.

3.3 Statistical analysis of latent variables

In this section we focus on the force terms and related powers, as well as their statistical analysis. The decomposition of force terms follows the definitions in section 2.2. In particular, all the estimates assume a material density of the flow $\rho = 805kg/m^3$. This a fixed scaling factor, and the plots aspect is not affected by its value.

3.3.1 Power terms

Figure 6 shows the spatial integral of powers in the three rheology models. The spatial integration is performed on half spatial domain, due to the symmetry with respect to the flow central axis. Corresponding plots of the force terms are included in Supporting Information S2. The scalar product with velocity imposes a bell-shaped profile, as observed in Fig. 5a. In plot 6a the power of $\boldsymbol{RHS_1}$ represents the effect of the gravity in all the models. It starts from zero and rises up to $\sim 1.5W$ at $\sim 0.55s$, then decreases to zero after the material crosses the change in slope. Uncertainty range of $\pm 0.5W$ on the peak values. MC decreases slower, and has a more significant uncertainty, after the change in slope. PF decreases faster. In plot 6b the power of $\boldsymbol{RHS_2}$ represents the friction at the base of the flow. It is negative and peaks to $\sim 1.1 \pm 0.2W$ in MC, $\sim 1.0 \pm 0.2W$ PF, $\sim 0.7 \pm 0.3W$ in VS. A similar bell-shaped profile is shared by the three models. In contrast, the force plots in Figure S2 show an initial ripple before $\sim 0.1s$ and a plateau until $\sim 1s$. Basal friction power becomes negligible at the end, whereas the basal friction force in MC can be still significantly large. In plot 6c the power of $\boldsymbol{RHS_3}$ is related to the curvature effects, and is not null only at the change in slope. It is always dissipative, i.e. opposing flow velocity, indeed it is equivalent to the friction due to the additional weigh generated by centrifugal forces. It is weaker than $-0.1W$ on average, ten times smaller than the previous powers, although MC lower percentile reaches $\sim -0.25W$. VS displays a bimodal profile, with a second and weaker peak at $\sim 0.75s$. In plot 6d the power of $\boldsymbol{RHS_4}$ is related to the additional forces of the models, differently characterized. This term is really relevant in VS, although also in PF has a very short lasting positive peak up to $0.3W$ before to become null at $\sim 0.1s$. This power in VS is a speed dependent term, always dissipative. It is bell shaped and null before $\sim 0.1s$ and after $\sim 1s$. At the time of change in slope it is $\sim -0.7W$, $\pm 0.3W$. In plot 6e the power of the total force $\sum_{i=1}^{4} \boldsymbol{RHS_i}$ summarizes the energy accumulation and dissipation, following a sinusoid profile. The profile is characterized by a positively valued stage before $\sim 0.55s$ before the change in slope, and by a negatively valued stage after that, with bell-shaped profile, and a negative peak at $\sim 0.75s$. In the first stage MC increases more linearly, while PF and VS have a concave shape. In the second stage MC is affected by a larger uncertainty and the decrease occurs later in time of $\sim 0.1s$. PF and VS are remarkably similar, but VS wanes slightly faster.

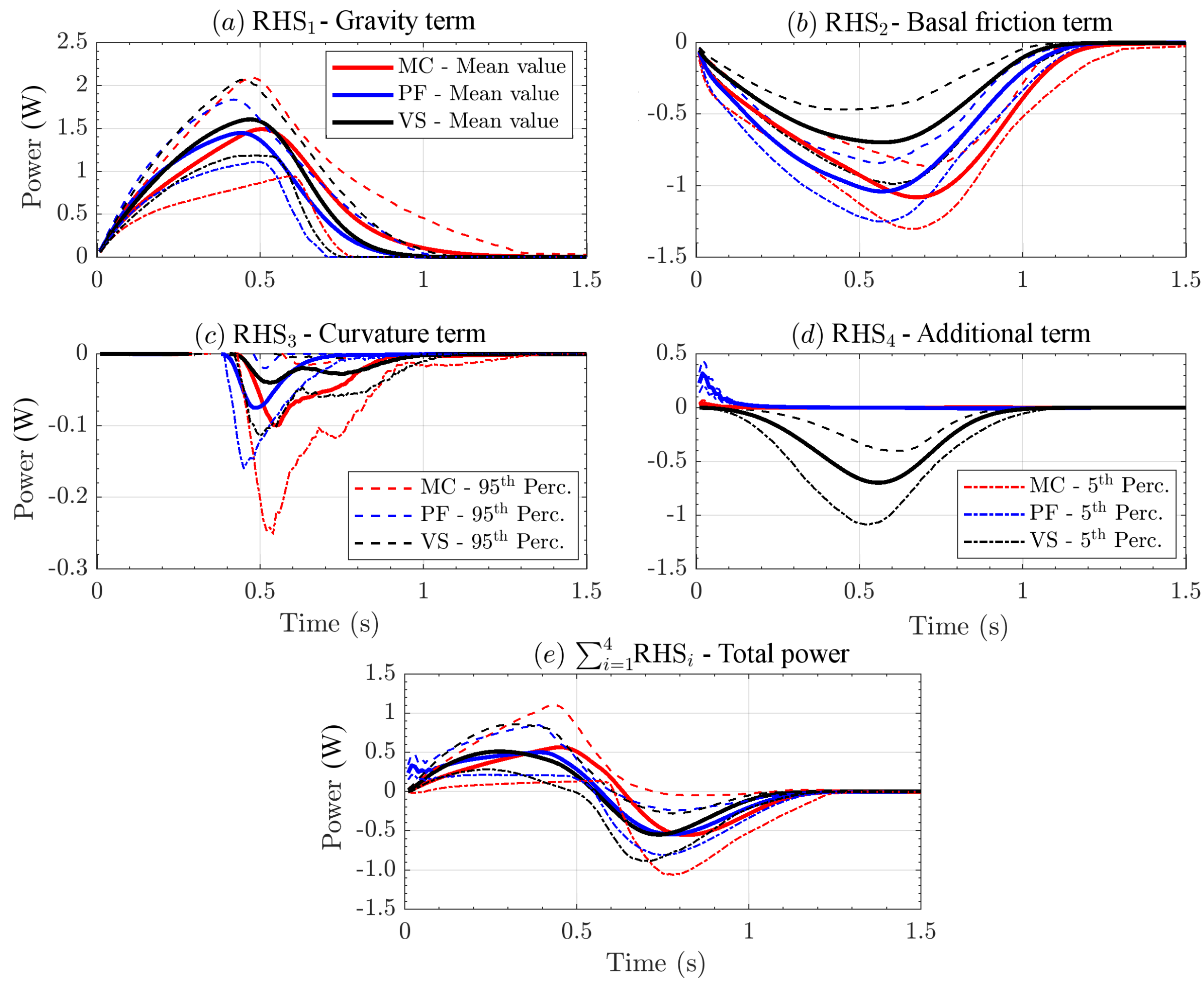

Figure 6: Spatial sum of the RHS powers. Bold line is mean value, dashed lines are 5th and 95th percentile bounds. Model comparison on the mean value is also displayed.

3.3.2 Force dominance factors

Figure 7 shows the Dominance Factors $(P_i)_{i=1,\ldots,4}$ in the three rheology models, focusing on the RHS terms in the slope direction. The dominance factor is the probability of a force term to be the greatest one. Because these values are probabilities, they always belongs to $[0,1]$. The plots include also the probability of no-flow being observed at the considered point. The different models are plotted separately: 7a,d,g,j assume MC; 7b,e,h,k assume PF; 7c,f,i,l assume VS. The plots 7a,b,c are related to point L_1, placed on the initial pile. Only $\boldsymbol{RHS_1}$ can be the dominant force, and no-flow probability is $(1-P_1)$. Same thing in the plots 7d,e,f related to point L_2, placed in the middle of the slope. Then, plots 7g,h,i are related to point L_3, placed at the change in slope. In L_3, $\boldsymbol{RHS_3}$ can be the dominant term for a short time, with a peak probability of $\sim 30\%$. Plots 7j,k,l are related to point L_4, placed in the middle of the flat runout. In L_4 only $\boldsymbol{RHS_2}$ can be the dominant term, except in PF where there is $\sim 10\%$ that $\boldsymbol{RHS_4}$ is the dominant term at then ending-time.

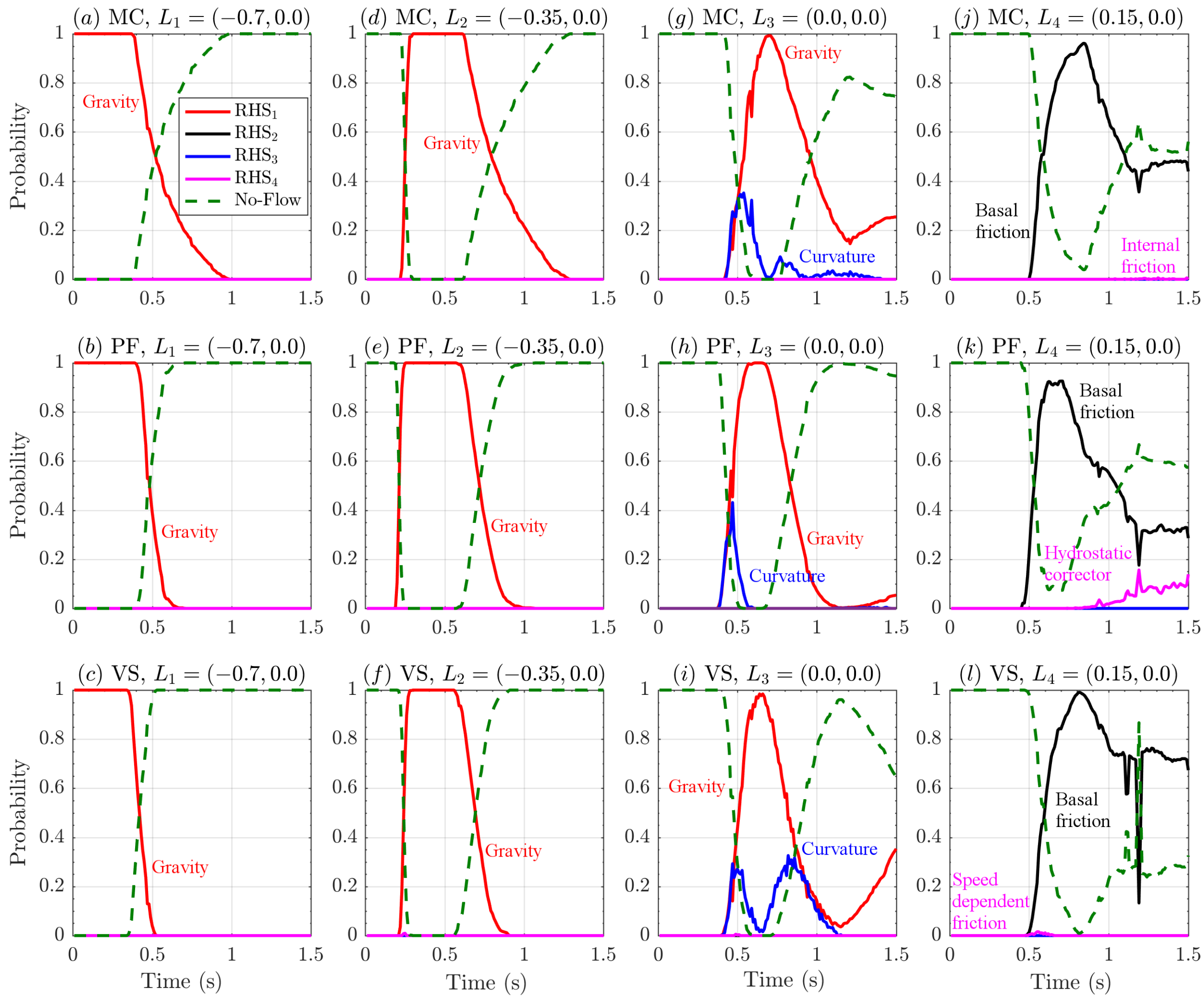

Figure 7: Records of dominance factors of **RHS** forces, in the slope direction, in four spatial locations of interest. Different rheology models are displayed with different colors. No-flow probability is also displayed with a green dashed line.

Dominance factors provide an informative description of the main dynamics of the flow, but they do not tell anything on the not-dominant forces. The Contribution Coefficients can complete the statistical description of the latent variables. They are obtained dividing the force terms by the dominant force. This is a tool to compare the different force terms, scaling the plots by the dominant dynamics - in this case in $[-1, 1]$. It also represent the degree of relevance of the assumptions behind the force terms, changing as a function of time. See Appendix B for more details. Figure 8 shows the Contribution Coefficients $(C_i)_{i=1,...,4}$, for the three rheology models. The different models are plotted separately: 8a,d,g,j assume MC; 8b,e,h,k assume PF; 8c,f,i,l assume VS.

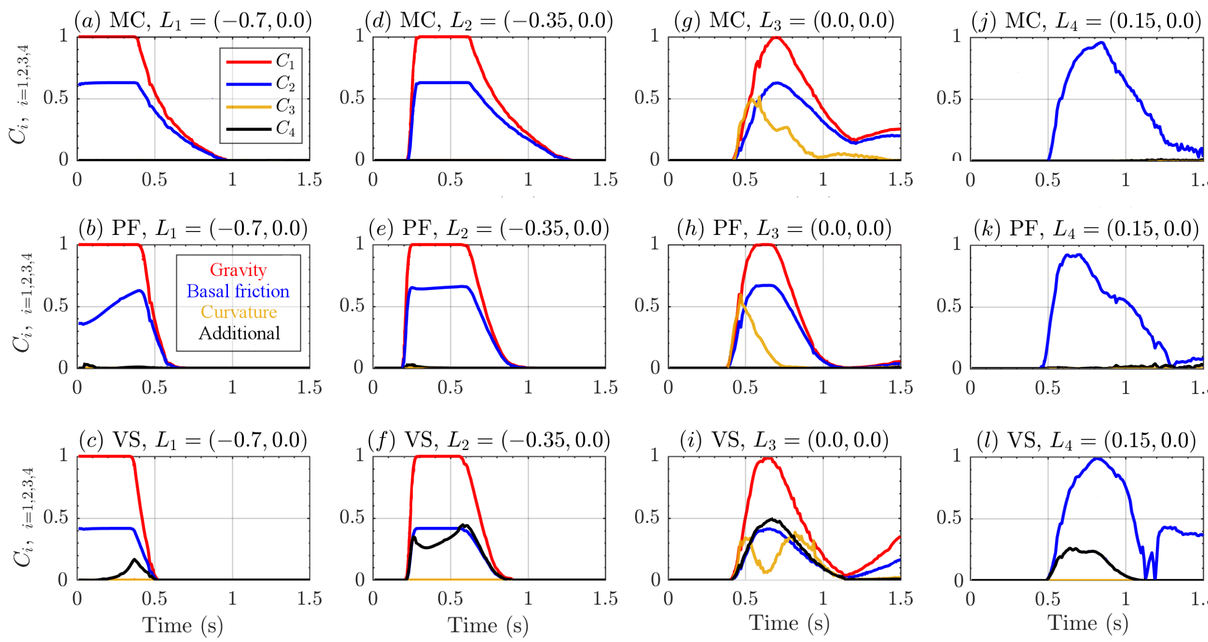

Figure 8: Records of contribution coefficients of **RHS** forces, in the slope direction, in four spatial locations of interest. Different rheology models are displayed with different colors. Dominant function Φ is calculated based on the l^∞ norm

The plots 8a,b,c are related to point L_1, placed on the initial pile. C_1 and C_2 play the major roles, with a minor contribution C_4 in VS. Contributions profiles are flat plateaus that start to wane after $0.4s$, with the rise of the probability of no-flow at the point L_1. Uncertainty is related to the no-flow times, and to the value of C_2. The contribution C_4 can be significant, as it is shown by the 95$^{\text{th}}$ percentile values, but appears after $0.25s$. The plots 8d,e,f are related to point L_2, placed in the middle of the slope. Again the major contributions are C_1 and C_2, with trapezoidal profile preceded and followed by no-flow. In VS, C_4 becomes as a significant as C_2, but it is bimodal instead than trapezoidal. The plots 8g,h,i are related to point L_3, placed at the change in slope. The contributions C_1 and C_2 are still the largest, but their profiles are bell-shaped. In VS, C_4 is almost identical to C_2. In all the models, C_3 is also significant, with a peak similar to C_2, but has a different profiles - triangular for MC and PF, bimodal for VS. In MC the decrease occurs in two stages. The uncertainty tells us that C_3 can also be the dominant force, for a shorter time. Due to the presence of deposit, all the contributions are small (particularly small in PF), but not zero at the ending time. The plots 8j,k,l are related to point L_4, placed in the middle of the flat runout. Only C_2 has a major role, with a bell shaped profile faster to wax than to wane. Contribution C_4 has a minor role in VS and PF. Uncertainty affects this force, which can be shortly the dominant term, in PF.

4 Large scale flows on the SW slope of Volcán de Colima

The second case study is a block and ash flow down the slope of Volcán de Colima (MX) - an andesitic stratovolcano that rises to 3,860 m above sea level, situated in the western portion of the Trans-Mexican Volcanic Belt (Fig. 9).

Volcán de Colima has historically been the most active volcano in México (la Cruz-Reyna, 1993; Zobin et al., 2002; González et al., 2002). The modeling of pyroclastic flows generated by explosive eruptions and lava dome collapses of Volcán de Colima is a well studied problem (Martin Del Pozzo et al., 1995; Sheridan and Macías, 1995; Saucedo et al., 2002, 2004, 2005; Sarocchi et al., 2011; Capra et al., 2015). The presence of a change in slope followed by multiple ravines characterize this framework. The volcano has been already used as a case study in several research involving the Titan2D code (Rupp, 2004; Rupp et al., 2006; Dalbey et al., 2008; Yu et al., 2009; Sulpizio et al., 2010; Capra et al., 2011; Aghakhani et al., 2016). During July 10^{th}-11^{th}, 2015, the volcano underwent its most intense eruptive phase since its Subplinian-Plinian 1913 AD eruption (Saucedo et al., 2010; Zobin et al., 2015; Reyes-Dávila et al., 2016; Capra et al., 2016; Macorps et al., 2018).

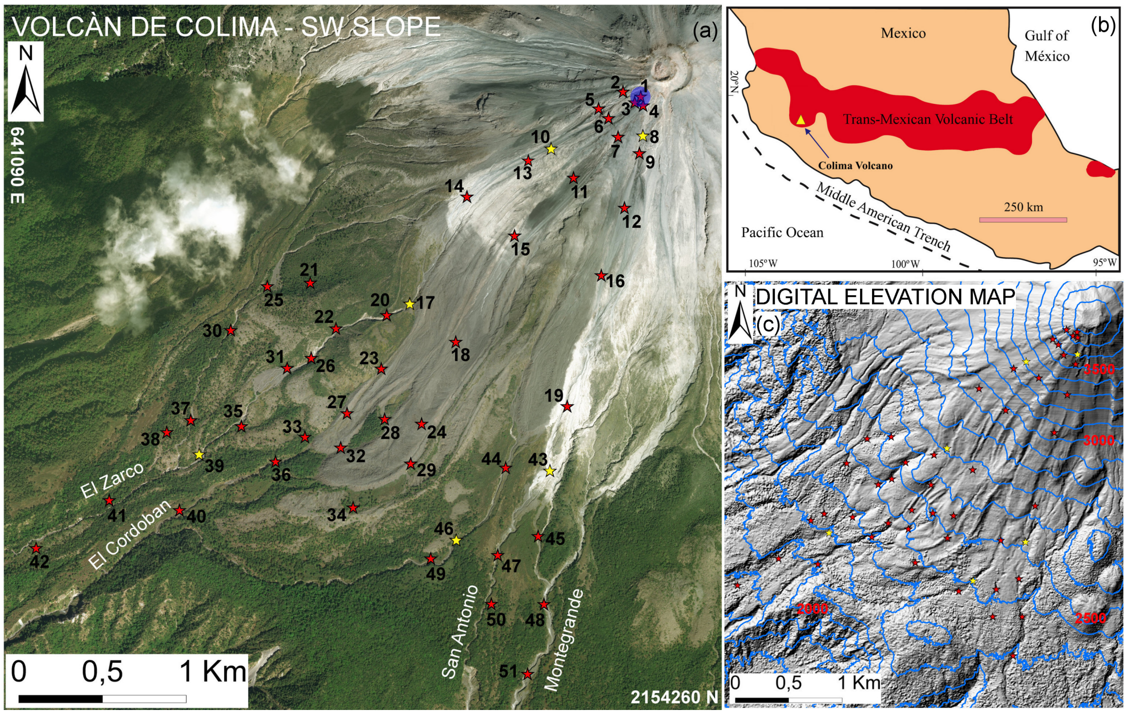

Figure 9: (a) Volcán de Colima (México) overview, including 51 numbered local sample sites (stars) and four labeled major ravines channeling the flow. Pile location is marked by a blue dot. Reported coordinates are in UTM zone 13N. Background is a satellite photo. (b) Regional geology map. (c) Digital elevation map. Six points that are adopted as preferred locations are highlighted in yellow. Elevation isolines are included in blue, elevation values in red.

We assume the flow to be generated by the gravitational collapse of a material pile placed close to the summit area. The volcano already produced either pyroclastic flows generated by lava dome collapse, called Merapi style flows, or by an eruptive column partial collapse, i.e. Soufriere style flows (Macorps et al., 2018). A dome collapse occurs when there is a significant amount of recently-extruded highly-viscous lava piled up in an unstable configuration around a vent. Further extrusion and/or externals forces can cause the still hot dome of viscous lava to collapse, disintegrate, and avalanche downhill (Bursik et al., 2005; Wolpert et al., 2016; Hyman and Bursik, 2018). Eruptive column collapse can occur during explosive eruptions, when the eruption column can no longer sustain the weight of material due to loss of pressure, and hence it partially collapses down. The hot, dense blocks in this "block and ash" flow (BAF) will typically range from centimeters to a few meters in size. The matrix is composed of fine ash from the comminuted blocks. Computations were performed on a DEM with 5m-pixel resolution, obtained from LiDAR data acquired in 2005 (Davila et al., 2007; Sulpizio et al., 2010). We placed 51 locations along the flow inundated area to accomplish local testing. Six of them are then adopted as preferred locations, being representative of different flow regimes.

4.1 Preliminary consistency testing of the input ranges

In this case study, Dalbey et al. (2008) assumed $\phi_{bed} = [15°, 35°]$, while (Capra et al., 2011) adopted $\phi_{bed} = 30°$ in the simulation of a BAF in that same setting. Moreover, Spiller et al. (2014); Bayarri et al. (2015); Ogburn et al. (2016) found a statistical correlation between flow size and effective basal friction inferred from field observation of geophysical flows. The size of the BAF of this study would have $\phi_{bed} = [13°, 18°]$ according to their estimates. Figure 10 displays the maps of max flow height and max velocities observed in the extreme cases tested. Simulation options are - max_time = 7200 s (2 hours), height/radius = 0.55, length_scale = 4e3 m, number_of_cells_across_axis = 50, order = first, geoflow_tiny = 1e4 (Patra et al., 2005; Aghakhani et al., 2016). Initial pile geometry is paraboloid.

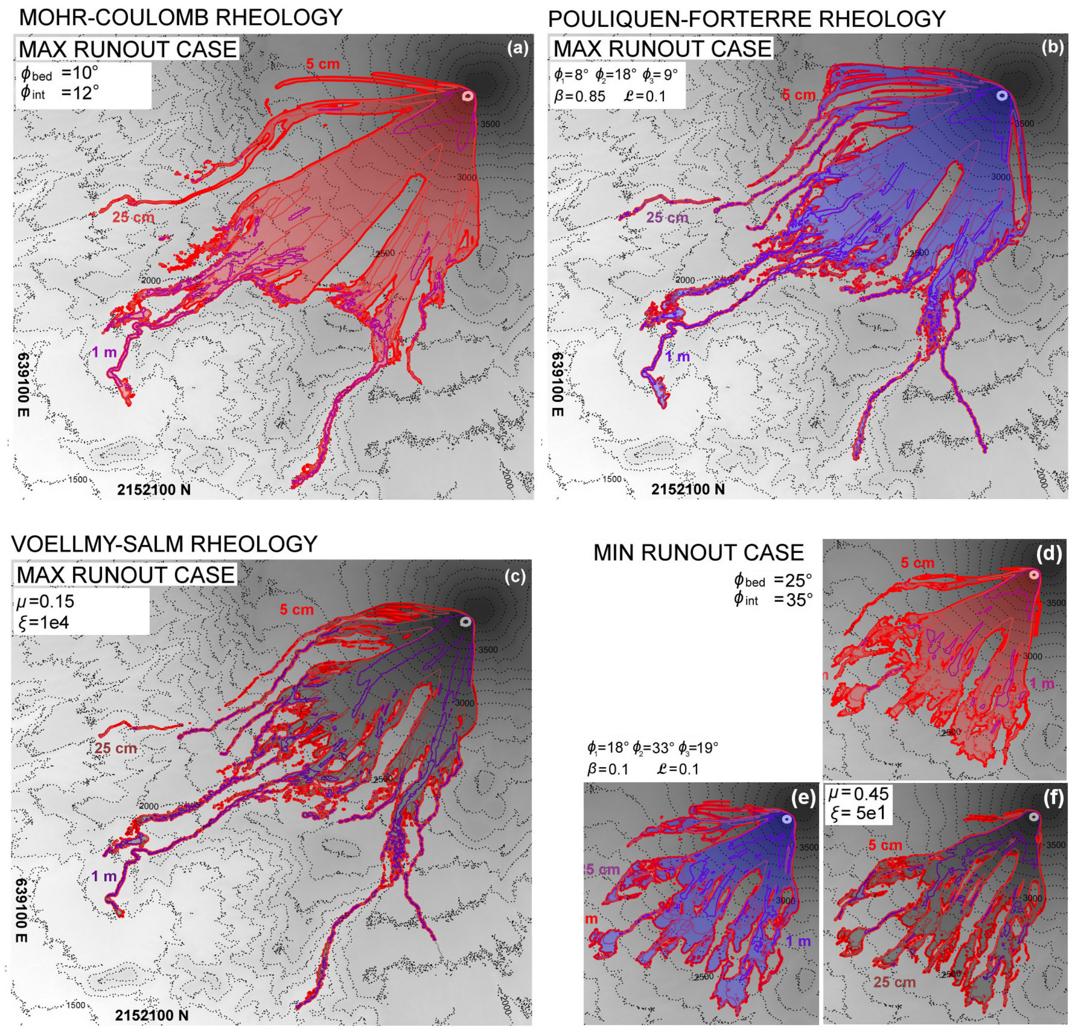

Figure 10: Volcán de Colima BAF. Comparison between *max flow height* maps of simulated flow, assuming Mohr-Coulomb (a),(d), Pouliquen-Forterre (b),(e), and Voellmy-Salm (c),(f) rheology. Extreme cases producing maximum and minimum runout, i.e. (a),(b),(c) ***max volume – min resistance*** and (d),(e),(f) ***min volume – max resistance***.

- **Material Volume:** $[2.08, 3.12] \times 10^5\ m^3$, i.e. average of $2.6 \times 10^5\ m^3$ and uncertainty of ±20%.
- **Rheology models' parameter space:**
  The parameter ranges adopted in this case study are:

  MC - $\phi_{bed} \in [10°, 25°]$.

  PF - $\phi_1 \in [8°, 18°]$.

VS - $\mu \in [0.15, 0.45]$, $\log(\xi) \in [1.7, 4]$.

The models possess different features, but the maximum runout after channelization in the ravines is matching. In particular, VS lateral spreading is significant lower and the material reaches higher thickness, whereas PF model seems to stop more gradually than MC with a more complex inundated area boundary lines. These features are the macroscopic consequence of the rheology differences.

4.2 Flow height in 51 locations - Mohr-Coulomb model

In this case study, the number of spatial locations is significantly high. We placed 51 points to span the entire inundated area, in search of different flow regimes, as displayed in Fig. 9. First we show the average flow height in all the locations, then, based on that, we select six points which we find representative of interesting flow regimes. Figure 11 assumes MC model, while the results related to PF and VS models are detailed on the six selected points.

Figure 11 shows the mean flow height, $h(L, t)$, at the 51 spatial locations of interest, according to MC.

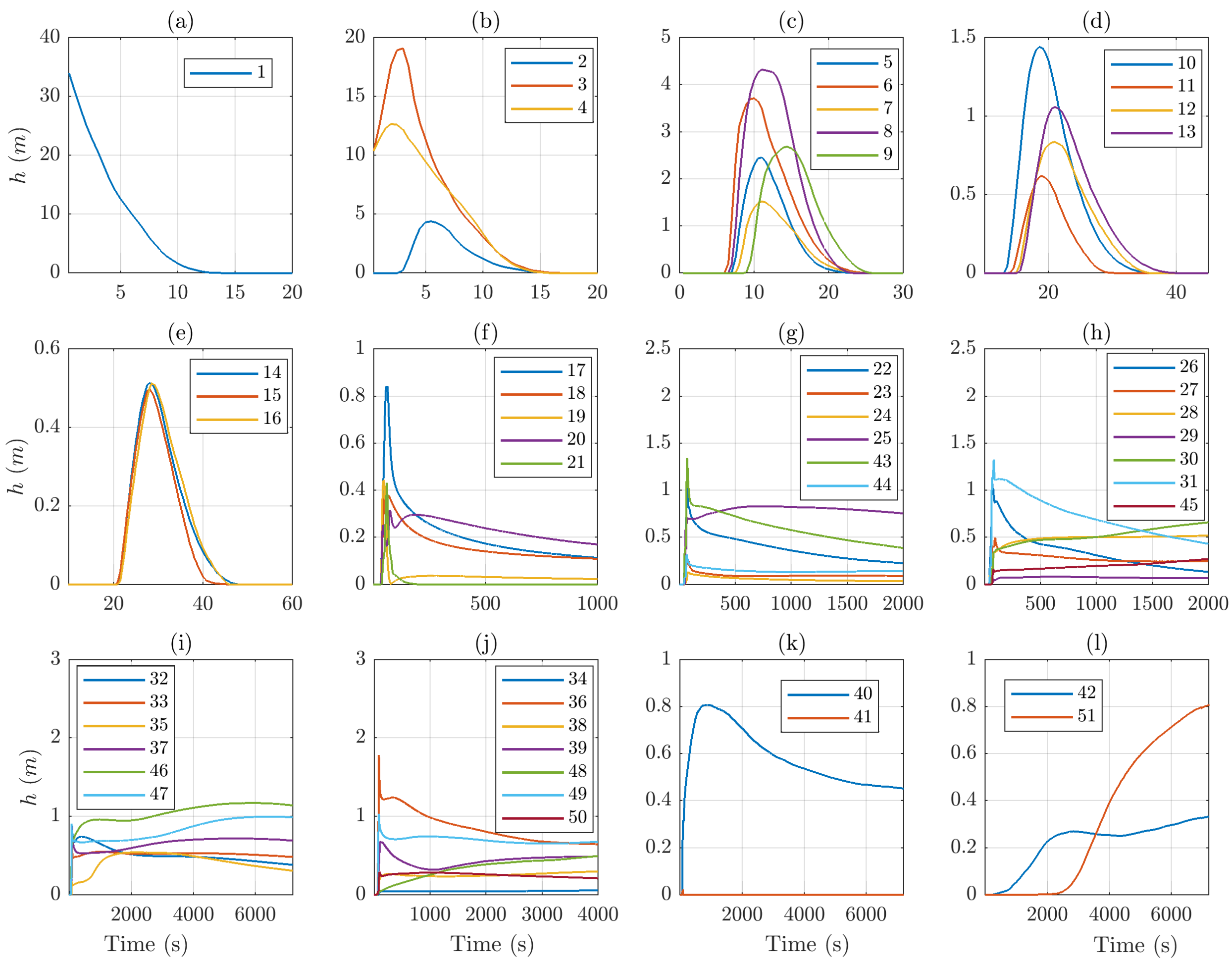

Figure 11: MC model, records of average flow height, $h(L, t)$, in 51 spatial locations of interest (Fig. 9).

Different plots have different scales on either time and space axes. In plot 11a, the only location is set on the center of the initial pile, and the profile is similar to what observed in point L_1 of the inclined plane case study, in Fig.3a. In particular, the height decreases from the initial value to zero in $\sim 15s$.

In plots 11b,c,d,e, the locations are are set at less than ~ 1 km radius from the initial pile (projected distance, without considering slope). Their profiles are similar to point L_2 in Fig.3b. The height profile is bell-shaped, starting from zero and then waning back to zero in ~ 20 s. All the dynamics occurs during the first minute. Plot (b) shows transitional features, and focuses on points at the boundary of the initial pile. In plots 11f,g,h,i,j, points are set where the slope reduces, and the flow can channelize and leave a deposit. Projected distance from the initial pile is $\sim 2-3$ km. The profiles are sometimes similar to L_3 of Fig.3c, other times to L_4 of Fig.3d, in a few cases showing intermediate aspects. In general is either observed an initial short-lasting bulge followed by a slow decrease lasting minutes and asymptotically tending to a positive height, or a steady increase of material height tending to a positive height. In both cases it is sometimes observed a bimodal profile in the first 5 minutes. Finally, plots 11k,l focus on three points set at about the runout distance of the flow, in the most important ravines, at $\sim 4-5$ km projected distance from the initial pile. Profiles are similar to what observed in point L_4 of Fig.3d. Plots of the Froude Number at the 51 points are included in Supporting Information S3 - strongly bimodal profiles and sharp changes are observed due to the interplay between flow height and flow speed.

4.3 Observable outputs - UQ in 6 selected locations

The six selected locations are $[L_8, L_{10}, L_{17}, L_{39}, L_{43}, L_{46}]$, displayed in Figure 12). First two points, L_8 and L_{10} are both significantly close to the initiation pile. Points L_{17} and L_{43} are placed where the slope is reducing and the ravine channels start, and L_{39} and L_{46} are placed in the channels, further down-slope. Moreover, L_8, L_{43}, and L_{46} are placed at the western side of the inundated area, whereas L_{10}, L_{17}, and L_{39} are placed at the eastern side.

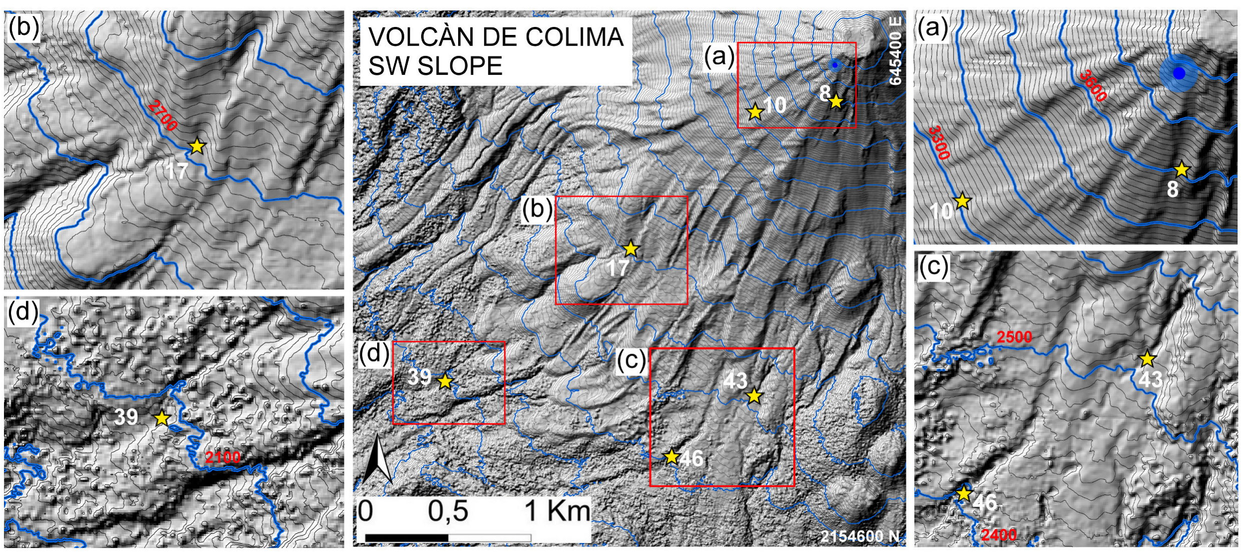

Figure 12: Volcán de Colima (México) overview, including six selected sample sites (stars). In (a), (b), (c), (d) are enlarged the topographic features in proximity of those locations. Pile location is marked by a blue dot. Reported coordinates are in UTM zone 13N. Elevation isolines are included, at intervals of $10m$ in black, and $100m$ in bold blue. Elevation values in red.

Observable outputs include the flow height and acceleration as a function of time. In addition, flow area, and spatially averaged speed and Fr are displayed. Locally measured Froude Number and flow acceleration are included in Supporting Information S4 and S5.

4.3.1 Flow height

Figure 13 shows the flow height, $h(L,t)$, at the points $(L_i)_{i=8,10,17,39,43,46}$, for the three rheology models.

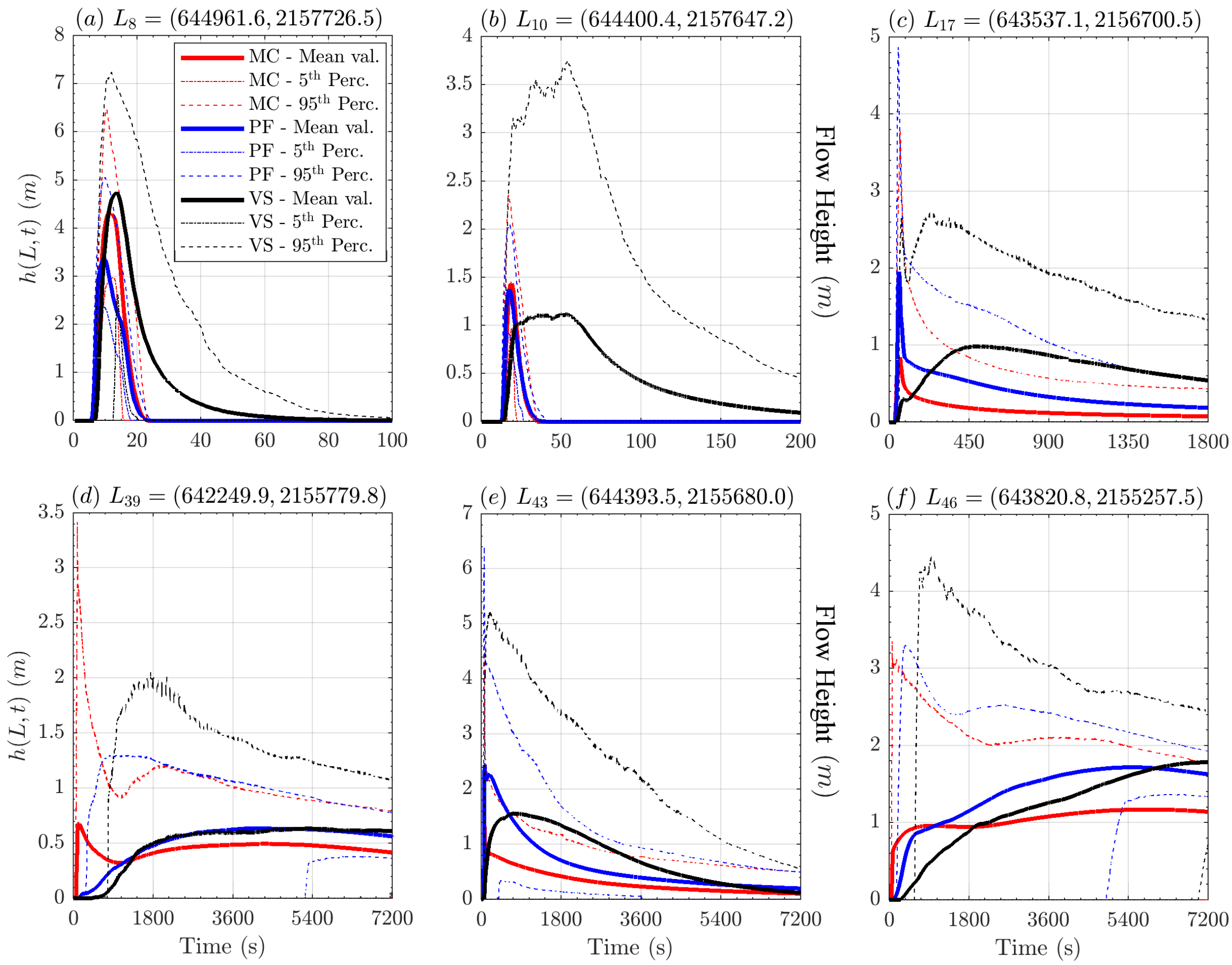

Figure 13: Records of flow height at six selected locations. Bold line is mean value, dashed/dotted lines are 5^{th} and 95^{th} percentile bounds. Different rheology models are displayed with different colors. Plots are at different scale, for simplifying lecture.

In plots 13a,b, we show the flow height in points L_8 and L_{10}, $\sim 200m$ and $\sim 500m$ from the initial pile, respectively. L_8 is on the east side, and L_{10} on the west side of the flow. Models MC and PF display similar profiles, positive for less than $15s$ and bell-shaped. VS requires a significantly longer time to decrease, particularly in point L_{10}, where the average flow height is still positive after $\sim 200s$. Peak average values in L_8 are $3.4m$ in PF, $4.3m$ in MC, $4.7m$ in VS. Uncertainty is $\sim \pm 2m$, halved on the lower side in MC, and PF. In L_{10}, models MC and PF are almost indistinguishable, with peak height at $1.4m$ and uncertainty $\pm 0.5m$. Model VS, in contrast, has a maximum height of $1.1m$ lasting for $50s$, and 95^{th} percentile reaching $3.7m$. In plots 13c,e, we show the flow height in points L_{17} and L_{43}, both at $\sim 2km$ from the initial pile, on the west and east side of the flow, respectively. All the three models show in both the points a fast spike during the first minute, followed by a slow decrease, still showing a positive average height after 30 minutes. Again, VS is significantly different from MC and PF, and has a secondary rise peaking at $\sim 450s$, which is not observed in the other models. This produces higher values for the most of the temporal duration, but converges to similar deposit thickness after more than 1 hour. Maximum values are $1m$ for MC, $2m$ for PF, and $1.5m$ for VS, in both locations. The 5^{th}

percentile is zero in all the models, meaning that the parameter range does not always allow the flow to reach those locations. The 95^{th} percentile is always above $5m$ in the models, except in VS, point L_{17}. In plots 13d,f, we show the flow height in points L_{39} and L_{46}, both placed at more than $3km$ from the initial pile, on the west and east side of the flow, respectively. The three models all show an increasing profile, except in MC, in point L_{39}, which display an initial spike and a decrease before to rise again. A similar decreasing profile can be also observed in the 95^{th} percentiles of all the models. It is significant that the 5^{th} percentile of PF becomes positive after $\sim 5400s$, meaning that the flows almost surely have reached that location. Deposit thickness is $\sim 0.5m$ in all the models in point L_{39}, and $1.7m$ in VS, $1.6m$ in PF, $1.2m$ in MC, in L_{46}.

4.3.2 Flow area and spatial integrals

Figure 14 shows the spatial average of speed and Froude Number, in the three rheology models. It also shows the inundated area of flow, as a function of time. Analysis of model suitability can be conducted here given recorded deposits. In past work Patra et al. (2005), we have tuned MC rheology to match deposits for known block and ash flows but *a priori* predictive ability was limited by inability to tune without knowledge of flow character.

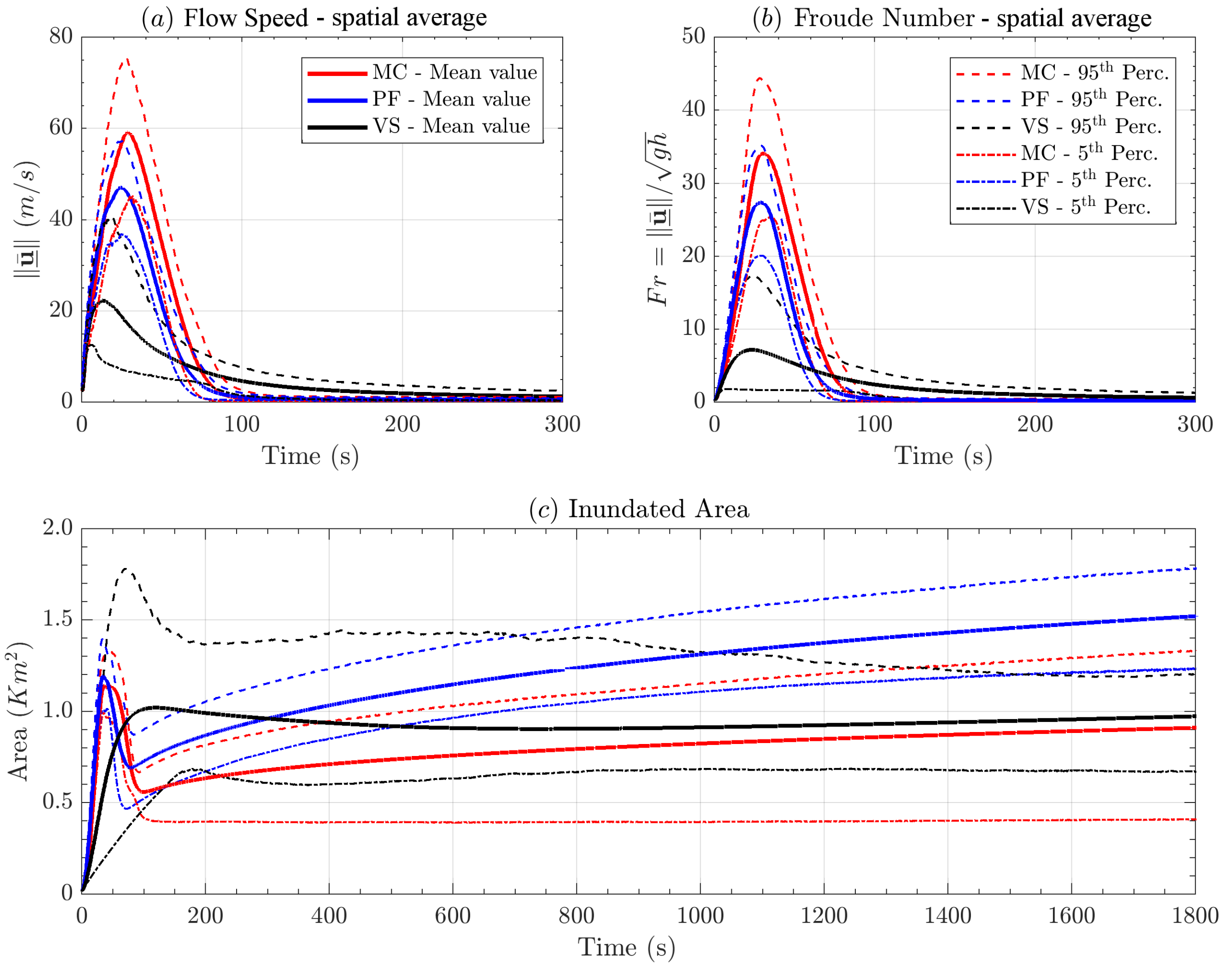

Figure 14: Comparison between spatial averages of (a) flow speed, and (b) Froude Number, in addition to the (c) inundated area, as a function of time.

Like in Fig.5,spatial averages and inundated area have smoother plots than local measurements, and most of the details observed in local measurements are not easy to discern. In plot 14a the speed shows a bell-shaped profile in all the models, but whereas the values were significantly similar in the inclined plane experiment, in this case the maximum speed is $\sim 60m/s$ in MC, $\sim 50m/s$ in PF, $\sim 20m/s$ in VS, on average. Uncertainty is $\pm 18m/s$ in MC, similar, but skewed on the larger values in VS, $\pm 10m/s$ in PF. In plot 14b, the Froude profile is very similar to the speed, but the difference between VS and the other models is accentuated. Maximum values are ~ 50 in MC, ~ 38 in PF, ~ 5 in VS, whereas uncertainty is ± 10 in MC, ± 7 in PF, and skewed $[-5, +10]$ in VS. In plot 14c, inundated area has a first peak in MC and PF, both at $\sim 1.15km^2$, followed by a decrease to $0.55km^2$ and $0.7km^2$,respectively, and then a slower increase up to a flat plateau at $0.9km^2$ and $1.5km^2$, respectively. Uncertainty is $\sim \pm 0.2km^2$ in both MC and PF until $\sim 100s$, and then it increases at $\pm 0.3km^2$ and $[-0.5, +0.4]km^2$, respectively. In MC the increase of uncertainty is concentrated at $\sim 100s$, while it is more gradual in PF. VS has a different profile. The initial peak is significant only in the 95^{th} percentile values, and occurs significantly later, i.e. at $\sim 100s$ against $\sim 50s$ in MC and PF. It is of $\sim 1km^2$ on the average, but up to $\sim 1.8km^2$ in the 95^{th} percentile. The decrease after the peak is very slow and the average inundated area is never below $0.85km^2$, and eventually reaches back to $\sim 1km^2$. Uncertainty is $[-0.3, +0.2]km^2$.

4.4 Statistical analysis of latent variables

All the estimates in this section assume the material density of the flow $\rho = 1800kg/m^3$. This a fixed scaling factor, and the plots aspect is not affected by its value.

4.4.1 Power terms

Figure 15 shows the spatial sum of the powers, for the three rheology models. Corresponding plots of the force terms are included in Supporting Information S6. The scalar product with velocity imposes the bell-shaped profile already observed in Fig. 6a. In plot 15a the power of $\boldsymbol{RHS_1}$ represents the effect of the gravity in all the models. It starts from zero and rises up to $\sim 1.4e11W$ in MC, $\sim 1.2e11W$ in PF, $\sim 6.5e10W$ in VS. Uncertainty is $\pm 4e10W$ in MC, $\pm 3e10W$ in PF, $[-4e10, +5e10]W$ in VS. The decrease of gravitational forces is related to the slope reduction, and this decrease is more gradual in VS than in the other models. In Figure S5, at $\sim 10s$ the forces in all the models show a slowing down of the decrease, which is permanent in VS, and lasting $\sim 10s$ in MC and PF. This is not visible in the power plots. In plot 15b the power of $\boldsymbol{RHS_2}$ represents the effect of the basal friction in all the models. It is negative and peaks to $\sim -6.5e10W$ in MC, $\sim -5e10W$ in PF, $\sim -2e10W$ in VS. In VS this dissipative power is significantly more flat than in the other models. MC and PF show negligible powers after $\sim 100s$, VS after $\sim 200s$. Uncertainty is $\pm 2e10W$ in MC, $\pm 1.5e10W$ in PF, $[-2e10, +1e10]W$ in VS. In Figure S2, the basal friction force in MC shows a bimodal profile with a peak at $\sim 5s$ and a second one at $\sim 60s$. This is not visible in powers. In PF, the plot starts from stronger values than in the other models, but it is also the faster to wane. In plot 15c the power of $\boldsymbol{RHS_3}$ represents the effect of the curvature of terrain. It has a similar profile to the corresponding force, peaking to $\sim -7e10W$ in MC, $\sim -4.5e10W$ in PF, $\sim 5e9W$ in VS. The rise of this dissipative power has a similar profile, but a different duration between the models. The decrease is more gradual in VS. Uncertainty on the peak value is $[-4.5e10, +3.5e10]W$ in MC, $[-2.5e10, +2e10]W$ in PF, $[-1e10, +5e9]W$ in VS. The three models all show a bell-shaped profile, MC and PF waning to zero at $90s$, VS at $\sim 30s$. In plot 15d the power of $\boldsymbol{RHS_4}$ has a different meaning in the three models. In MC it is the internal friction term, and it only has almost negligible ripple visible in the first second. In PF it is a depth averaged correction in the hydrostatic pressure, and has an almost negligible effect only during the first second of simulation, at $5e9W$. It becomes null at $\sim 10s$. In VS, instead, it is a speed dependent term, and has a very relevant effect. The plot shows a bell-shaped profile, with a peak of $\sim -3.5e10W$, $[-2e10, +1e10]W$. After that, this dissipative power gradually decreases, and becomes negligible at $200s$. In plot 15e the power of the total force $\sum_{i=1}^{4} \boldsymbol{RHS_i}$ summarizes the energy accumulation and dissipation, following, like in the inclined plane case study, a sinusoid profile. First is observed a steep increase due to the domination of the gravity upon the friction, then a negative part when the friction stops the dynamics. As displayed in Fig. 6e, MC profile is delayed compared to the others, and is affected by a larger uncertainty.

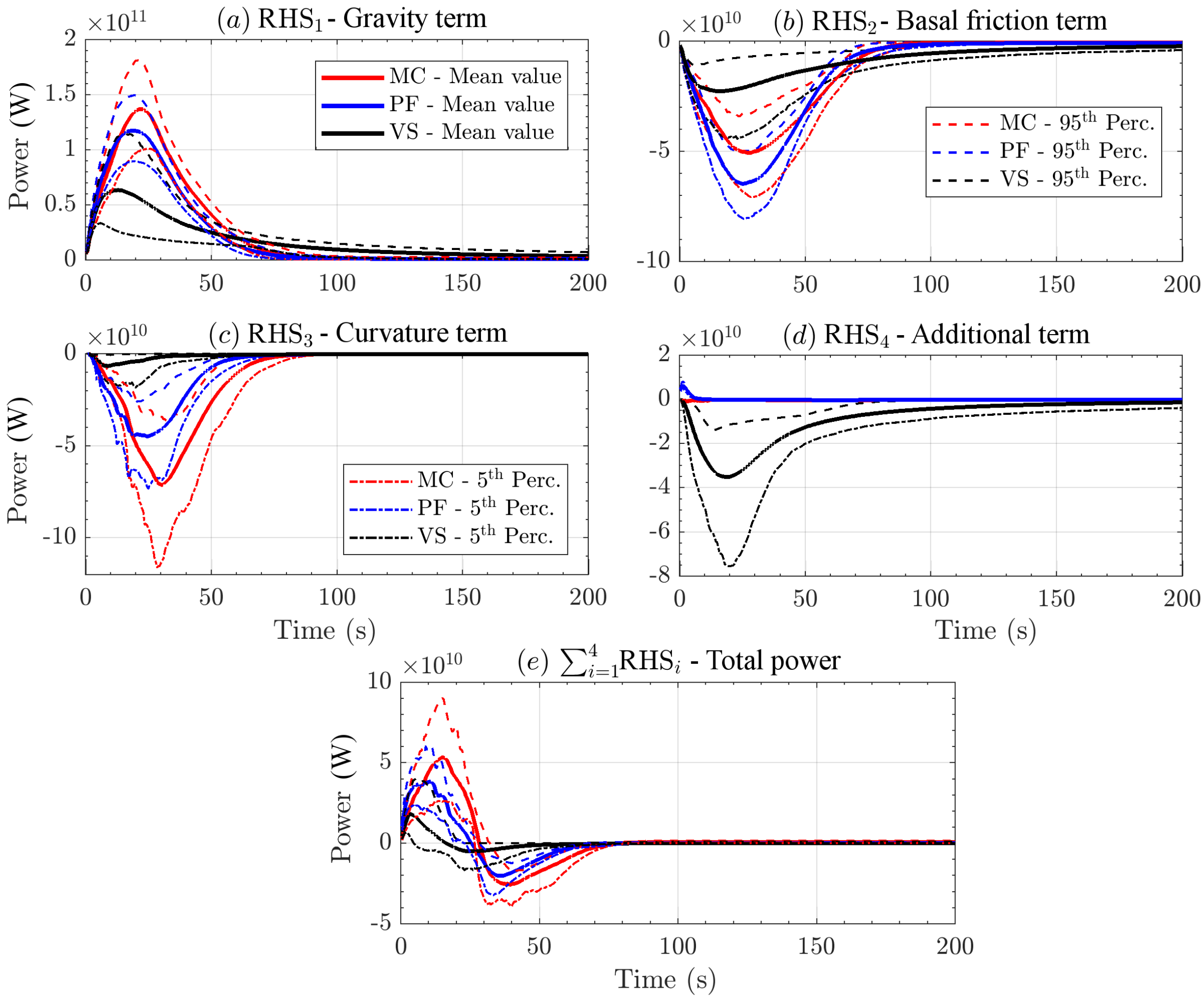

Figure 15: Spatial sum of the RHS powers. Bold line is mean value, dashed lines are 5^{th} and 95^{th} percentile bounds. Model comparison on the mean value is also displayed.

4.4.2 Force dominance factors

Figure 16 shows the Dominance Factors $(P_i)_{i=1,\dots,4}$, in the three rheology models and focusing on the RHS terms moduli, in the three selected points L_8, L_{10}, and L_{17}, closer than 1 km to the initial pile (in horizontal projection). The plots 16a,b,c and 16d,e,f are related to point L_8 and L_{10}, respectively. They are significantly similar. $\boldsymbol{RHS_1}$ related to the gravitational force is the dominant force with a very high chance, $P_1 > 90\%$. In MC and PF there is a small probability, i.e. $P_3 = 5\% - 30\%$ at most, of $\boldsymbol{RHS_3}$ being the dominant force for a short amount of time, i.e. $\sim 5s$. This occurs in the middle of the time interval in which the flow is almost surely inundating the points being observed. In VS it is observed a $P_4 = 5\%$ chance of $\boldsymbol{RHS_4}$ being dominant, for a few seconds, anticipating the minimum of no-flow probability. Plots 16g,h,i, are related to point L_{17}, and the plots are split in two sub-frames, following different temporal scales. In all the models, $\boldsymbol{RHS_2}$ is the most probable dominant force, and its dominance factor has a bell-shaped profile, similar to the complementary of no-flow probability. In all the models, $\boldsymbol{RHS_1}$ has a small chance of being the dominant force. In MC, this is more significant, at most $P_1 = 30\%$, for $\sim 20s$ after the flow arrival, and has again about $P_1 = \%2$ chance to be dominant in $[100, 7200]s$. In PF, the chance is $P_1 = 15\%$ at most, and has two maxima, one short lasting at about $55s$, and the second in $[100, 500]s$. Also in VS, the chance is at most $P_1 = 15\%$, reached at $[300, 500]s$, but its profile is unimodal in time, and becomes lower than $P_1 = 2\%$ after $2000s$. In MC and PF, $\boldsymbol{RHS_3}$ has a chance of $P_3 = 10\%$ of being the dominant force, for a short amount of time $[30, 50]s$ and $[40, 50]s$, respectively. More details about the correspondent force contributions are reported in Supporting Information S7.

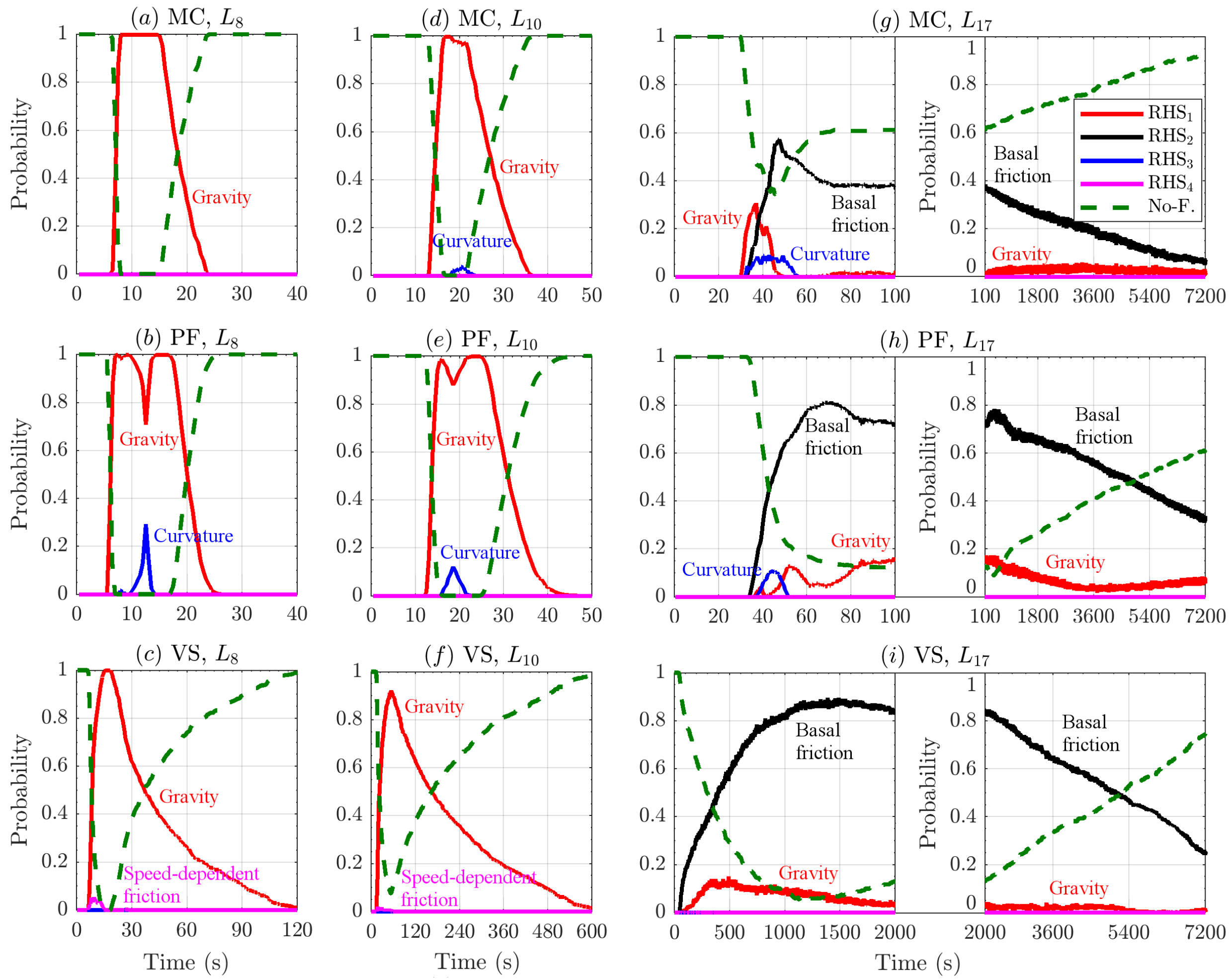

Figure 16: Records of dominance probabilities of **RHS** force moduli, at three spatial locations of interest, in the first km of runout. Bold line is mean value, dashed/dotted lines are 5$^{\text{th}}$ and 95$^{\text{th}}$ percentile bounds. Different rheology models are displayed with different colors. No-flow probability is also displayed.

Figure 17 shows the Dominance Factors $(P_i)_{i=1,...,4}$, in the three rheology models, in the three selected points L_{39}, L_{43}, and L_{46}, over 2 km far away from the initial pile.

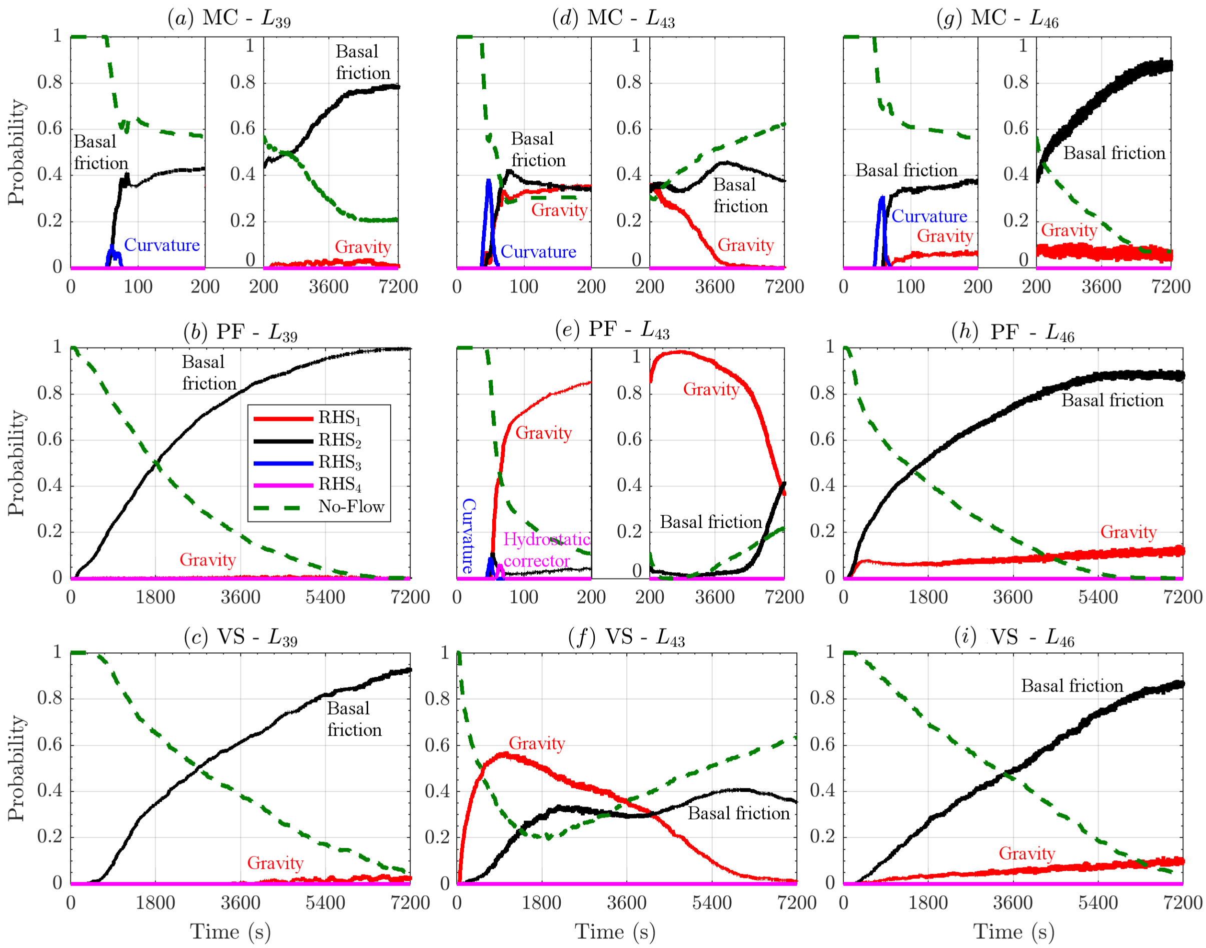

Figure 17: Records of dominance probabilities of **RHS** force moduli, at three spatial locations of interest, after 2 km of runout. Bold line is mean value, dashed/dotted lines are 5th and 95th percentile bounds. Different rheology models are displayed with different colors. No-flow probability is also displayed.

The plots 17a,b,c and 17g,h,i are related to point L_{39} and L_{46}, respectively, and they are significantly similar. Those points are both in the most distal part of the ravines, on the western and eastern side of the inundated area, respectively. They are also similar to the dominance factors at point L_{17}, shown in plots 16g,h,i, but the no-flow probability never becomes negligible in this case. The probability of $\boldsymbol{RHS_2}$ being the dominant force is $P_2 > 90\%$ till the end of the temporal window. Moreover, $\boldsymbol{RHS_1}$ does not show any initial peak in P_2, and generally increases slowly, reaching at most $P_2 = 10\%$ chance of being the dominant force in L_{46}, after more than $3600s$, in all the models. Plots 17d,e,f, are related to point L_{43}, are remarkably different. In MC, either $\boldsymbol{RHS_1}$ and $\boldsymbol{RHS_2}$ can be dominant forces, both with chances of $\sim 35\%$ in the first $200s$. Then, $\boldsymbol{RHS_2}$ increases its chance, and becomes the only dominant force after $3600s$. The no-flow probability is never below 30%. $\boldsymbol{RHS_3}$ has a chance P_3 up to 35% for a short amount of time, i.e. $[40, 60]s$. In PF, $\boldsymbol{RHS_1}$ is the dominant force with chance above $P_1 = 90\%$ until $3600s$. $\boldsymbol{RHS_2}$ chance rises only in the very last amount of time, reaching $P_2 = 40\%$ at $7200s$, at the same level of P_1. The no-flow probability is very low during the most of the temporal window, rising at 20% only at $7200s$. Both $\boldsymbol{RHS_3}$ and $\boldsymbol{RHS_4}$ have a short peak of dominance factor, at about 10% for a few seconds, at $[50, 60]s$. Finally, in VS the no-flow probability is never below 20%, and the profile is similar to MC, although $\boldsymbol{RHS_1}$ is more relevant, up to $4000s$. The other force are almost never dominant.

5 Discussion on the comparative anatomy of geophysical flow models

In this section we summarize the general features which differ between the models, with the purpose of depicting a thorough identikit of them, and breaking down the effects of the modeling assumptions listed in section 2.1. The main focus is on the average plots, but the uncertainty percentiles are also considered, and often they exalt the differences between the models. Moreover, we describe a new method that uses the statistical analysis of the models to evaluate their similarity to a specific observation.

5.1 Characteristic features and their motivations

Even in the process of the preliminary alignment of the parameter ranges, the differences features in the models already affected the shape of the inundated area, both in the small flow on the inclined plane, and in the large scale geophysical flow. In Fig. 2b the max runouts display a larger lateral extent in PF, and a bent shape in VS where the lateral wings remain behind the central section of the flow. In Fig. 10, the three models look clearly different both in the min-runout and in the max-runout cases. Even if the maximum runout is matched, MC displays a further distal spread before entering the ravines, PF shows a larger angle of lateral spread at the initiation pile, VS is less laterally extended and generally looks significantly channelized, and displays several not-inundated spots due to minor topographical coulées.

5.1.1 Flow height and acceleration

Flow height gives additional insight on the model features. In Fig. 3, MC looks more distally stretched, but starts to deposit material earlier and closer to the initial pile. PF height is generally shorter, and displays a small temporal anticipation in its arrival at the sample points. These features are probably due to the hydrostatic correction term, which pushes forward the material at the very beginning, during the pile collapse. Also, the interpolation between the minimum ϕ_1 and maximum ϕ_2 friction angles, as a function of flow height and speed, plays a role in the linear cut in the flow height profile when the flow thins on the slope. VS tends to be higher than the other models, because of the reduced lateral spreading of material. In Fig. 13 the large scale flow height plots allow us to classify the points according to their similarity with the fours points of the small scale flow. There is a significant feature which was not observed in the small scale flow - VS is temporally stretched, and material arrives later and stays longer in the sample points. This is clearly a consequence of the speed dependent term.

Flow acceleration brings our analysis from the mechanics to the dynamics of the flow. In Fig. 4, at the point L_2 on the slope, MC has a flat plateau, while PF is linearly decreasing, and VS is bell shaped. Those differences are a consequence of the assumptions behind the models - double angle in PF, and speed dependent term in VS. At the slope change point L_3, VS and MC display a bimodal profile in the acceleration. This is not a statistical effect, but it is physically observed in single simulations. The first maximum is when the head of the flow hitting the ground, while the second maximum is when the accumulating material in the tail arrives. In VS the maxima are equal, because its tail is not laterally spread and hits the ground compactly. In contrast, PF does not show such a second peak, due to the accentuated lateral spreading in the tail. In point L_4 on the flat runway, the average acceleration in PF is lower than in the other models at the end of the simulation time, probably because the material stops more suddenly, and the probability of no deposit is greater.

5.1.2 Flow extents, average speed, and spatial sum of power terms

In Fig. 5 the average speed and Froude Number of the different models are significantly similar, and this points to the fact that the differing features are mostly localized in space. VS is confirmed to be significantly slower than the other models, after the initial collapse. Moreover, it is the only model which presents a significant amount of long lasting and slowly moving material. Inundated area in PF has a greater maximum value, because of the accentuated lateral spread. In VS the inundated area almost does not decrease from its peak, because of the strictly increasing lateral spread. Vice versa, lateral spread in MC and PF has a temporary stop when the bulk of the flow hits the ground. This is a consequence of the interplay of accumulating material and the push of new material, which is stronger in the middle than in the lateral wings. In PF lateral extent is even decreasing slightly for a short time. The hydrostatic correction term may generate the pull reducing the lateral extent. In Fig. 14 average speed and Froude Number display that VS slower speed is not a local feature. Inundated area is again significantly larger in PF.

Power terms have several features in common with the corresponding forces, and provide the decomposition of the acceleration sources. Main dissimilarity between forces and powers is that gravitational and basal friction powers have a profile starting from zero when the flow initiates, because the flow speed starts from zero. In Fig. 6 the difference between models is particularly relevant in term $\boldsymbol{RHS_4}$. Speed dependent power in VS is at least one order of magnitude larger than the maximum values of the corresponding terms in MC and PF. Those are decreasing to zero after a short time from the initiation. Hydrostatic correction in PF is clearly positive in the speed direction, and hence contributing to push the flow ahead, at the beginning. The effects of internal friction in MC are almost negligible, and initially positive, then negative. This is motivated by an initial compression of the material during the pile collapse, followed by its stretching. It is worth remarking that $\boldsymbol{RHS_2}$ and $\boldsymbol{RHS_3}$ are both smaller in VS, due to the lower basal friction angles involved. In Fig. 15, the differences are accentuated, because of the topography complexity. Gravity term is larger in VS, because a portion of the flow lingers on the higher slopes for a long time. Basal friction has a higher peak in PF, due to the effect of the interpolation of the friction angle to the larger value ϕ_2 in several portions of the flow.

5.1.3 Force dominance factors

The force dominance factors illustrate the predominant force term through time and the probability of no-flow. In Fig. 7 the differences between the models are minor in the inclined plane case study. In general, there is only a single dominant force, and its profile is complementary with the no-flow probability. More details can be found in the force contributions analysis (**??**, Appendix B and Supporting Information S7). Only in the slope change point L_3 the differences are remarkable. Curvature term dominance probability is bimodal in MC and VS, and in MC the second peak is smaller. The profile is similar to the spatial average of the force term. In the flat runway, at point L_4, in PF the hydrostatic correction can be the dominant force with a small chance. In Fig. 16 and 17, the complexity in the dominance factors is greater. At the points L_8 and L_{10}, proximal to the initiation, gravitational force is dominant with a high probability until the no-flow probability become predominant. Curvature force can be dominant in MC and PF, when the no-flow change is zero, and for a short time. Gravitational force is dominant for a much larger time, because of the longer presence of the flow. The speed dependent friction can be dominant with a small probability, at the beginning of the dynamics, when flow is faster. At the points L_{17}, L_{39}, and L_{46}, the basal friction is dominant with high probability. L_{17} is left by the flow, whereas L_{39}, and L_{46} have a deposit at the end with a high chance. In general, in MC the no-flow probability tends to be larger than in the other models, because part of the parameter spaces stops earlier, or completely leaves the site. Again, curvature can have a small chance to be dominant in MC and PF, particularly when the speed is high. Point L_{43} is particularly interesting. It is not close to the initiation, but it is placed right after a local hill. This location shows an intermediate regime between the proximal points and the distal ones. The no-flow probability is increasing at the end, meaning that the material is leaving the site, like in L_{17}. Moreover, the gravity can be dominant like in the proximal points, but also the bed friction can be dominant, and in MC and VS both the two forces have similar chances to be dominant for the most of the time of the simulation. In PF, only the gravitational force is dominant with a high chance, and the no-flow probability is almost null, meaning that there is accelerating material for the most of time. This is probably because point L_{43} is placed in a place right downhill of a place where a significant amount of material stops according to that model.

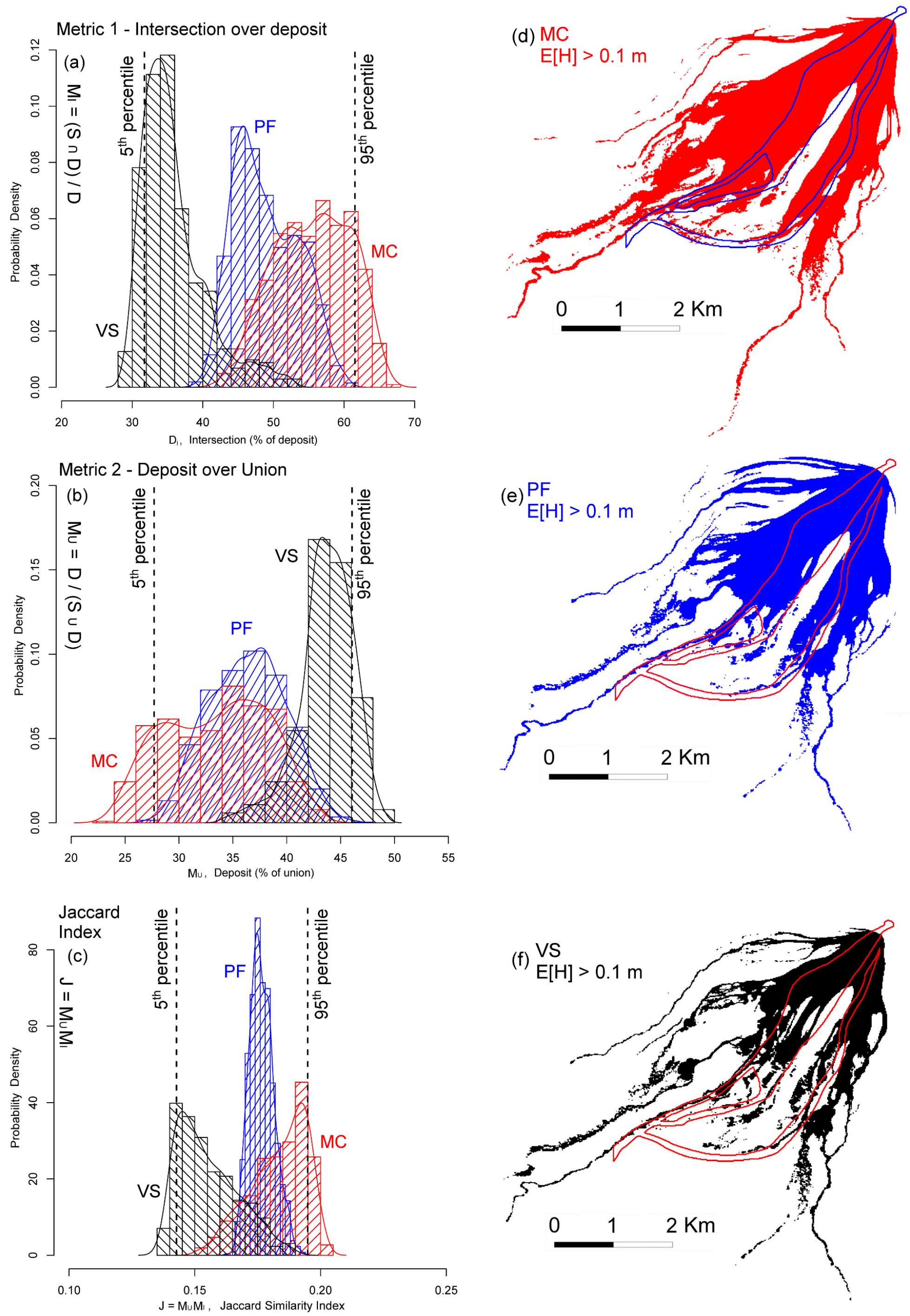

Figure 18: Probability density functions of the similarity indices of the Volcán de Colima inundated regions with the deposit of the 16 April 1991 BAF. (a) is based on $\mathcal{M}_I$, (b) on $\mathcal{M}_U$, (c) on $\mathcal{J}$. Different colors correspond with the models MC (red), PF (blue), VS (black). Data histograms are displayed in the background. Global 5^{th} and 95^{th} percentile values are indicated with dashed lines. Plots (d,e,f) display the average inundated region $\{\mathbf{x} : E[H(\mathbf{x}] > 10cm\}$. The boundary of the real deposit is marked with a colored line.

5.2 Example of model performance calculation

The new procedure developed in this study has also potential applications in the quantification of model performance, i.e. the similarity of the outputs and real data. We remark that the measured performance refers to the couple (M, P_M), and that different parameter ranges produce different performances. Our example concerns the Volcán de Colima case study, and in particular we compare the inundated region of our simulations to the deposit of a real BAF occurred 16 April 1991 (Saucedo et al., 2004). The inundated region is defined as the points in which the maximum flow height H is greater than $10cm$. A similar procedure may be applied to any observable variable produced by the models, if specific data become available.

Let $\mathcal{M} : \mathcal{P}(\mathbb{R}^2) \to [0,1]$ be a similarity index defined on the subsets of the real plane. An equivalent definition may be based on the pseudo-metric $1 - \mathcal{M}$. For example, we define

$$\mathcal{M}_I := \frac{\int_{\mathbb{R}^2} 1_{S\cap D}(\mathbf{x})d\mathbf{x}}{\int_{\mathbb{R}^2} 1_D(\mathbf{x})d\mathbf{x}}, \quad \mathcal{M}_U := \frac{\int_{\mathbb{R}^2} 1_D(\mathbf{x})d\mathbf{x}}{\int_{\mathbb{R}^2} 1_{S\cup D}(\mathbf{x})d\mathbf{x}}, \quad \mathcal{J} := \mathcal{M}_I \cdot \mathcal{M}_U,$$

where $S \subset \mathbb{R}^2$ is the inundated region, and $D \subset \mathbb{R}^2$ is the 16 April 1991 BAF deposit. In particular, $\mathcal{M}_I$ is the area of the intersection of inundated region and deposit over the area of the deposit, $\mathcal{M}_U$ is the area of the deposit over the area of the union of inundated region and deposit, $\mathcal{J}$ is the product of the previous, also called Jaccard Index (Jaccard, 1901).

Figure 18 shows the probability distribution of the similarity indices over the space of our simulations, according to the uniform probability P_M on the parameter ranges defined in this study. Different metrics produce different performances, for example MC overestimates the inundated region, while VS underestimates the inundated portion of the deposit. The global 5^{th} and 95^{th} percentile values $[a, b]$, defined assuming to choose the model randomly with equal chance, are shown in the Figure.

Let $g : [a, b] \to [0, 1]$ be a score function defined over the percentile range of the similarity index. Then the performance score G_g of model (M, P_M) is defined as

$$G_g\,(M, P_M) = \int_{[a,b]} g(x)df_M(x),$$

where f_M is the pdf related to the model. Possible score functions include a step function at the global median, a linear or quadratic function from zero to one, a sigmoid function. Supporting Information S8 reports a Table of alternative performance scores, according to changing similarity indices and score functions.

6 Conclusions

In this study, we have introduced a simple, robust statistically driven method for analyzing complex models. We have used 3 different models arising from different rheology assumptions. The data shows unambiguously the performance of the models across a wide range of possible flow regimes and topographies. We analyze local and global quantities and latent variables. The analysis of latent variables is particularly illustrative of the impact of modeling assumption. Knowledge of which assumptions dominate, and, by how much, at the level of assumptions will allow us to construct efficient models for desired inputs. Such model composition is the subject of ongoing and future work. We have included an example calculating model performance according to the similarity or the models with real data.

Our new method enabled us to break down the effects of the different physical assumptions in the dynamics, providing an improved understanding of what characterizes each model. We considered a Monte Carlo simulation on a range of parameter inputs, and estimated the mean and the uncertainty range of the corresponding outputs. Fitting the results on a specific flow was thus not necessary, and the procedure unfolded the physics of the models in a general framework. We also developed the concept of force dominance factor, in the study of the force terms in the different regions of space and time. This paves the way to future research aimed at designing a new generation of algorithms, which could completely bypass the search for a unique best model. They would go beyond the concept of alternative models, first decomposing them into their characterizing features, and then focusing instead on a more flexible selection of the assumptions which are locally suitable for each specific part of the simulation.

In particular, we presented:

- a short review of the assumptions characterizing three commonly used rheologies of Mohr-Coulomb, Poliquenne-Forterre, Voellmy-Salm. This included a qualitative list of such assumptions, and the breaking down of the different terms in the differential equations.
- a preliminary procedure for selecting the portions of the parameter ranges of the models which produced consistent outputs. This was fundamental because the model comparison cannot focus on the modeling differences if the scales of the stress tensor were on different scales.
- an articulated statistical framework, processing the mean, and the uncertainty range of many mechanical and latent variables in the flow simulation. A new function was featured in the TITAN2D solver, capable of producing local measurements of those quantities at selected sites. Spatial integrals and spatial averages were also performed, illustrating the characteristics of the entire flow under the different models. The new concept of force dominance factors enabled a simplified description of the local dynamics.
- the detailed results of the procedure applied to two different case studies: an inclined plane with a flat runway, and the real DEM on the SW lope of Volcán de Colima (MX). All the differences were quantitatively commented.
- a final discussion, explaining all the observed features in the results on the light of the known physical assumptions of the models, and the evolving flow regime in space and time.
- a model performance estimation method according to the similarity or the models with real data. The results are strongly dependent on the metric and the cost function adopted.

Our statistical analysis based on UQ depicted three very different models, due to the different assumptions underlying to them. Compared to a classical MC model, PF lacks of internal friction and this produces an accentuated lateral spread. This is increased by the hydrostatic correction, which briefly pushes the flow ahead and laterally during the initial collapse. That force can also have some minor effects in the final deposit accumulation. The interpolation of the smaller bed friction angle ϕ_1 with the larger value ϕ_2, suddenly stops the flow if it thins compared to its speed. This mechanism does not allow for large speed peaks. Instead in VS, the speed dependent friction has a great effect in reducing lateral spread and producing channeling features even due to minor ridges in topography. The flow tends to be significantly slower and more stretched in the slope direction. The effects of the different formulation of the curvature term are less impacting than the effects of the lower basal friction and speed.

The procedure may be applied to other models, or modeling assumptions. Moreover, the two case studies considered had the purpose of exploring two very different scale of flows, the first on a controlled setting, the second on a real topography. Additional research concerning other case studies, and different parameter ranges, might reveal other flow regimes, and hence differences in the consequences of the modeling assumptions under new circumstances.

7 Appendix A: Latin Hypercubes and orthogonal arrays

The Latin Hypercube Sampling (LHS) is a well established procedure for defining pseudo-random designs of samples in $\mathbb{R}^d$, with good properties with respect to the uniform probability distribution on an hypercube $[0,1]^d$ (McKay et al., 1979; Owen, 1992a; Stein, 1987; Ranjan and Spencer, 2014; Ai et al., 2016). In particular, compared to a random sampling, a LHS: (i) enhances the capability to fill the d-dimensional space with a finite number of points, (ii) in case $d > 1$, avoids the overlapping of point locations in the one dimensional projections, (iii) reduces the dependence of the number of points necessary on the dimensionality d.

Definition 1 (Latin hypercube sampling) *Let $\Xi = \{\xi_i \ : \ i = 1, \dots, N\}$ be a set of points inside the d-dimensional hypercube $C = [0,1]^d$. Let $[0,1] = \bigcup_{j=1}^N I_j$, where $I_j = [\frac{(j-1)}{N}, \frac{j}{N}]$. Let $\xi_i = \left(\xi_i^1, \dots, \xi_i^d\right)$, and for each $k \in \{1, \dots, d\}$, let $\Xi^k = \{\xi_i^k \ : \ i = 1, \dots, N\}$. Let λ^d be the uniform probability measure supported inside C, called Lebesgue measure. Then Ξ is a latin hypercube w.r.t. $\lambda^d \iff \forall j \in \{1, \dots, N\}$, $\forall k \in \{1, \dots, d\}$, $\left|I_j \cap \Xi^k\right| = 1$.*

The procedure is simple: once the desired number of samples $N \in \mathbb{N}$ is selected, and $[0,1]$ is divided in N equal bins, then each bin will contain one and only one projection of the samples over every coordinate. The LHS definition is trivially generalized over $C = \prod_i^d [a_i, b_i]$, i.e. the cartesian product of d arbitrary intervals. That will be applied in this study, defining LHS over the parameter domain of the flow models.

There are a large number of possible designs, corresponding the number of permutations of the bins in the d-projections, i.e. $d \cdot N!$. If the permutations are randomly sampled there is a high possibility that the design will have good properties. However, this is not assured, and clusters o points or regions of void space may be observed in C. For this reason, we base our design on the orthogonal arrays (OA) (Owen, 1992b; Tang, 1993).

Definition 2 (Orthogonal arrays) *Let $S = \{1, \ldots, s\}$, where $s \geq 2$. Let $Q \in S^{n\times m}$ be a matrix of such integer values. Then Q is called an $OA(n, m, s, r) \iff$ each $n \times r$ submatrix of Q contains all possible $1 \times r$ row vectors with the same frequency $\lambda = n/s^r$, which is called the index of the array. In particular, r is called the strength, n the size, $(m \geq r)$ the constrains, and s the levels of the array.*

Orthogonal arrays are very useful for defining latin hypercubes which are also forced to fill the space (or its r-dimensional subspaces) in a more robust way, at the cost of potentially requiring a larger number of points than a traditional LHS.

Proposition 3 *Let Q be an $OA(n, m, s, r)$. Then let $U \in \mathbb{R}^{n\times m}$ be defined as follows:*

$$\forall k \in \{1, \ldots, s\},\ \forall j \in \{1, \ldots, m\},\ \{Q[\cdot, j]\ :\ Q[i, j] = k\} = \Pi\left(\{(k-1)\lambda s^{r-1}, \ldots, k\lambda s^{r-1}\}\right),$$

where Π is a random permutation of λs^{r-1} elements. Then $\Xi = \{\xi_i = U[i, \cdot]\ :\ i = 1, \ldots, n\}$ is a LHS w.r.t to λ^m over $C = [0, 1]^m$. Moreover, let $[0, 1]^r = \bigcup_{(h_1,\ldots,h_r)=1}^{s} I_{(h_i)}$, where $I_{(h_i)} = \prod_i^r [\frac{(h_i-1)}{s}, \frac{h_i}{s}]$. Then $\forall D = (d_1, \ldots, d_r) \subseteq \{1, \ldots, m\}$, let $\Xi^D = \{(\xi^{d_1}, \ldots, \xi^{d_r})\ :\ i = 1, \ldots, n\}$. We have that

$$\forall k \in \{1, \ldots, s\},\ \forall (h_i\ : i = 1, \ldots, r) \in \{1, \ldots, d\}^r,\ \left|I_{(k_i)} \cap \Xi^D\right| = \lambda.$$

For each column of Q we are replacing the λs^{r-1} elements with entry k by a random permutation of $\left((k-1)\lambda s^{r-1} + h\right)_{h\in 1,\ldots,\lambda s^{r-1}}$. After the replacement procedure is done, the newly obtained matrix U is equivalent to a LHS which inherits from Q the property of fully covering s^r equal r-dimensional hypercubes in every r-dimensional projection. Each hypercube contains λ points. In other words, inside each r-dimensional projection, the design associated to U fills the space like a regular grid at the scale of those s^r hypercubes, but it is still an LHS at a finer scale, i.e. the λs^{r-1} one dimensional bins. A complete proof can be found in Tang (1993) and it is a straightforward verification of the required properties.

However, even in an LHS based on an $OA(n, m, s, r)$, if $r < m$ what happens in the projections with dimension $r' > r$ is not controlled, and randomizing procedures are made more difficult by the additional structure imposed by the OA. Moreover, the total number of points necessary to achieve a full design increases with r, and hence is affected by dimensionality issues.

Dealing with relatively small d, i.e. $d \in \{3, 4\}$, we adopt a LHS U created by a $OA(s^d, d, s, d)$. The strength is equal to the dimension d, hence the design fills the entire space like a d dimensional grid, but it is a LHS as well. In this case there is one point in each hypercube, and $\lambda = 1$. We take $s = 8$ for the 3-dimensional designs over the parameter space of Mohr-Coulomb and Voellmy-Salm models, i.e. 512 points; we took $s = 6$ for the 4-dimensional designs over the more complex parameter space of the Pouliquen-Forterre model, i.e. 1296 points.

8 Appendix B: Force contribution coefficients, and their conditional decomposition

Force contributions represent an additional tool to compare the different force terms, following a less restrictive approach than the Dominance Factors. They are obtained dividing the force terms described in section 2.2 by the dominant force Φ.

In general we focus on the moduli of the forces, or their projection on the slope direction. Hence, in the following the forces are scalar and not vectorial terms. It is important to remark that all the forces depend on the input variables, and they are thus considered as random variables. The definitions are not depending on the location, but all the results will significantly depend of that choice. Next notation will assume to be in a selected location $x = L_k$, where $k \geq 1$.

Definition 4 (Contribution coefficients) *Let $(F_i)_{i\in I}$ be random variables on $(\Omega, \mathcal{F}, P)$, representing the considered force components in location x at time t. Then, for each component i, the contribution coefficient is defined as:*

$$C_i := \mathbb{E}\left[\frac{F_i}{\Phi}\right],$$

where Φ is a dominating function, i.e. $\Phi \geq |F_i|$, $\forall i \in I$.

The *total force*, i.e $\sum_i F_i$ excluding the inertial terms, is not a good candidate for a dominating function. Indeed, the terms often have opposite signs, and their sum can be really small. Another issue is given by the existence of a subset of times Θ characterized by the absence of flow in the selected location x. In Θ the dominant force is null, and cannot be the denominator of a fraction.

In our study Φ is the dominant force - our approach is based on the l^∞ norm:

$$\Phi := \begin{cases} \max_i |F_i|, & \text{if not null;} \\ 1, & \text{otherwise.} \end{cases}$$

In particular, for a particular location x, time t, and parameter sample ω, we have $C_i = 0$ if there is no flow or all the forces are null. The expectation of C_i is reduced by the chance of F_i being small compared to the other terms, or by the chance of having no flow in (x,t). Moreover, $\mathbb{E}[C_i] \in [-1,1]$, $\forall i$.

There is an additional result explaining the meaning of those coefficients through the conditional expectation.

Proposition 5 *Let $(F_i)_{i\in I}$ be random variables on $(\Omega, \mathcal{F}, P)$, representing the considered force components in location x at time t. For each i, let C_i be the contribution coefficient of force F_i, assuming $\Phi = \Phi_2$. Then we have the following expression:*

$$C_i = \sum_j p_j \mathbb{E}\left[\frac{F_i}{|F_j|} \;\Big|\; \Phi = |F_j|\right],$$

where $p_j := P\{\Phi = |F_j|\}$.

Proof. Let Z be a discrete random variable such that, for each $j \in \mathbb{N}$, $(Z = j) \iff (\Phi = |F_j|)$. Then, by the rule of chain expectation:

$$C_i = \mathbb{E}\left[\frac{F_i}{\Phi}\right] = \mathbb{E}\left[\mathbb{E}\left[\frac{F_i}{\Phi} \;\Big|\; Z = j\right]\right] =$$

$$= \mathbb{E}\left[\mathbb{E}\left[\frac{F_i}{|F_j|} \;\Big|\; Z = j\right]\right] = \sum_j P\{Z = j\}\, \mathbb{E}\left[\frac{F_i}{|F_j|} \;\Big|\; Z = j\right].$$

Moreover, by definition, $p_j = P\{Z = j\}$. This completes the proof. $\square$

The last proposition brings to a conditional decomposition of the contribution coefficients, taking advantage of the *dominance factors* $(p_j)_{j=1,\ldots,k}$, i.e. the probability of each F_j to be the dominant force in (x,t).

Definition 6 (Conditional contributions and dominance factors) *Let $(F_i)_{i=1,\ldots,k}$ be random variables on $(\Omega, \mathcal{F}, P)$, representing arbitrary force components in location x at time t. Then, for each pair of components (i,j), the conditional contribution $C_{i,j}$ is defined as:*

$$C_{i,j} := \mathbb{E}\left[\frac{F_i}{|F_j|} \;\Big|\; \Phi = |F_j|\right],$$

where $\Phi = \Phi_2$ is the dominant force. In particular, for each component i, the dominance factor is defined as:

$$p_j := P\{\Phi = |F_j|\}.$$

References

Aghakhani, H., Dalbey, K., Salac, D., and Patra, A. K. (2016). Heuristic and Eulerian interface capturing approaches for shallow water type flow and application to granular flows. *Computer Methods in Applied Mechanics and Engineering*, 304:243–264.

Ai, M., Kong, X., and Li, K. (2016). A general theory for orthogonal array based latin hypercube sampling. *Statistica Sinica*, (26):761–777.

Bartelt, P. and McArdell, B. (2009). Granulometric investigations of snow avalanches. *Journal of Glaciology*, 55(193):829–833(5).

Bartelt, P., Salm, B., and Gruber, U. (1999). Calculating dense-snow avalanche runout using a voellmy-fluid model with active/passive longitudinal straining. *Journal of Glaciology*, 45(150):242–254.

Bayarri, M. J., Berger, J. O., Calder, E. S., Dalbey, K., Lunagomez, S., Patra, A. K., Pitman, E. B., Spiller, E. T., and Wolpert, R. L. (2009). Using Statistical and Computer Models to Quantify Volcanic Hazards. *Technometrics*, 51(4):402–413.

Bayarri, M. J., Berger, J. O., Calder, E. S., Patra, A. K., Pitman, E. B., Spiller, E. T., and Wolpert, R. L. (2015). Probabilistic Quantification of Hazards: A Methodology Using Small Ensembles of Physics-Based Simulations and Statistical Surrogates. *International Journal for Uncertainty Quantification*, 5(4):297–325.

Bursik, M., Patra, A., Pitman, E., Nichita, C., Macias, J., Saucedo, R., and Girina, O. (2005). Advances in studies of dense volcanic granular flows. *Reports on Progress in Physics*, 68(2):271.

Capra, L., Gavilanes-Ruiz, J. C., Bonasia, R., Saucedo-Giron, R., and Sulpizio, R. (2015). Re-assessing volcanic hazard zonation of volcán de colima, méxico. *Natural Hazards*, 76(1):41–61.

Capra, L., Macís, J., Cortés, A., Dávila, N., Saucedo, R., Osorio-Ocampo, S., Arce, J., Gavilanes-Ruiz, J., Corona-Chávez, P., García-Sánchez, L., Sosa-Ceballos, G., and Vázquez, R. (2016). Preliminary report on the july 10–11, 2015 eruption at volcán de colima: Pyroclastic density currents with exceptional runouts and volume. *Journal of Volcanology and Geothermal Research*, 310(Supplement C):39 – 49.

Capra, L., Manea, V. C., Manea, M., and Norini, G. (2011). The importance of digital elevation model resolution on granular flow simulations: a test case for colima volcano using titan2d computational routine. *Natural Hazards*, 59(2):665–680.

Charbonnier, S., Germa, A., Connor, C., Gertisser, R., Preece, K., Komorowski, J.-C., Lavigne, F., Dixon, T., and 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. *Journal of Volcanology and Geothermal Research*, 261(Supplement C):295 – 315. Merapi eruption.

Charbonnier, S. J. and 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. *Bulletin of Volcanology*, 71(8):953–959.

Christen, M., Kowalski, J., and Bartelt, P. (2010). RAMMS: Numerical simulation of dense snow avalanches in three-dimensional terrain. *Cold Regions Science and Technology*, 63:1–14.

Da Cruz, F., Emam, S., Prochnow, M., Roux, J.-N., and Chevoir, F. (2005). Rheophysics of dense granular materials: Discrete simulation of plane shear flows. *Physical Review E*, 72:021309.

Dade, W. B. and Huppert, H. E. (1998). Long-runout rockfalls. *Geology*, 26(9):803–806.

Dalbey, K., Patra, A. K., Pitman, E. B., Bursik, M. I., and Sheridan, M. F. (2008). Input uncertainty propagation methods and hazard mapping of geophysical mass flows. *Journal of Geophysical Research: Solid Earth*, 113:1–16.

Dalbey, K. R. (2009). *Predictive Simulation and Model Based Hazard Maps*. PhD thesis, University at Buffalo.

Davila, N., Capra, L., Gavilanes-Ruiz, J., Varley, N., Norini, G., and Vazquez, A. G. (2007). Recent lahars at volcán de colima (mexico): Drainage variation and spectral classification. *Journal of Volcanology and Geothermal Research*, 165(3):127 – 141.

Denlinger, R. P. and Iverson, R. M. (2001). Flow of variably fluidized granular masses across three-dimensional terrain: 2. Numerical predictions and experimental tests. *Journal of Geophysical Research*, 106(B1):553–566.

Denlinger, R. P. and Iverson, R. M. (2004). Granular avalanches across irregular three-dimensional terrain: 1. Theory and computation. *Journal of Geophysical Research (Earth Surface)*, 109(F1):F01014.

Farrell, K., Tinsley, J., and Faghihi, D. (2015). A bayesian framework for adaptive selection, calibration, and validation of coarse-grained models of atomistic systems. *Journal of Computational Physics*, 295.

Fischer, J., Kowalski, J., and Pudasaini, S. P. (2012). Topographic curvature effects in applied avalanche modeling. *Cold Regions Science and Technology*, 74-75:21–30.

Forterre, Y. and Pouliquen, O. (2002). Stability analysis of rapid granular chute flows: formation of longitudinal vortices. *Journal of Fluid Mechanics*, 467:361–387.

Forterre, Y. and Pouliquen, O. (2003). Long-surface-wave instability in dense granular flows. *Journal of Fluid Mechanics*, 486:21–50.

González, M. B., Ramírez, J. J., and Navarro, C. (2002). Summary of the historical eruptive activity of volcán de colima, mexico 1519–2000. *Journal of Volcanology and Geothermal Research*, 117(1):21 – 46.

Gruber, U. and Bartelt, P. (2007). Snow avalanche hazard modelling of large areas using shallow water numerical models and GIS. *Environmrntal Modelling & Software*, 22:1472–1481.

Hyman, D. and Bursik, M. (2018). Time-Dependence of Passive Degassing at Volcán Popocatépetl, Mexico from Infrared Measurements: Implications for Gas Pressure Distribution and Lava Dome Stability. *Journal of Geophysical Research*, (in preparation).

Iverson, R. M. (1997). The physics of debris flows. *Reviews of Geophysics*, 35:245–296.

Iverson, R. M. and Denlinger, R. P. (2001). Flow of variably fluidized granular masses across three-dimensional terrain: 1. Coulomb mixture theory. *Journal of Geophysical Research*, 106(B1):537.

Iverson, R. M. and George, D. L. (2014). A depth-averaged debris-flow model that includes the effects of evolving dilatancy. i. physical basis. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 470(2170).

Iverson, R. M., Logan, M., and Denlinger, R. P. (2004). Granular avalanches across irregular three-dimensional terrain: 2. experimental tests. *Journal of Geophysical Research: Earth Surface*, 109(F1):n/a–n/a. F01015.

Jaccard, P. (1901). Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines. *Bulletin de la Société Vaudoise des Sciences Naturelles*, 37:241–272.

Jaeger, H. M., Liu, C., and Nagel, S. R. (1989). Relaxation at the Angle of Repose. *Physical Review Letters*, 62(1):40–43.

Kelfoun, K. (2011). Suitability of simple rheological laws for the numerical simulation of dense pyroclastic flows and long-runout volcanic avalanches. *Journal of Geophysical Research*, 116:B08209.

Kelfoun, K. and Druitt, T. H. (2005). Numerical modeling of the emplacement of socompa rock avalanche, chile. *Journal of Geophysical Research: Solid Earth*, 110(B12).

Kelfoun, K., Samaniego, P., Palacios, P., and Barba, D. (2009). Testing the suitability of frictional behaviour for pyroclastic flow simulation by comparison with a well-constrained eruption at Tungurahua volcano (Ecuador). *Bulletin of Volcanology*, 71:1057–1075.

Kern, M., Bartelt, P., Sovilla, B., and Buser, O. (2009). Measured shear rates in large dry and wet snow avalanches. *Journal of Glaciology*, 55(190):327–338.

la Cruz-Reyna, S. D. (1993). Random patterns of occurrence of explosive eruptions at colima volcano, mexico. *Journal of Volcanology and Geothermal Research*, 55(1):51 – 68.

Macorps, E., Charbonnier, S. J., Varley, N. R., Capra, L., Atlas, Z., and Cabré, J. (2018). Stratigraphy, sedimentology and inferred flow dynamics from the july 2015 block-and-ash flow deposits at volcán de colima, mexico. *Journal of Volcanology and Geothermal Research*, 349:99 – 116.

Martin Del Pozzo, A. M., Sheridan, M. F., Barrera, M., Hubp, J. L., and Selem, L. V. (1995). Potential hazards from Colima Volcano, Mexico. *Geofisica Internacional*, 34:363–376.

McKay, M. D., Beckman, R. J., and Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245.

Norini, G., De Beni, E., Andronico, D., Polacci, M., Burton, M., and Zucca, F. (2009). The 16 november 2006 flank collapse of the south-east crater at mount etna, italy: Study of the deposit and hazard assessment. *Journal of Geophysical Research: Solid Earth*, 114(B2):n/a–n/a. B02204.

Ogburn, S. E., Berger, J., Calder, E. S., Lopes, D., Patra, A., Pitman, E. B., Rutarindwa, R., Spiller, E., and Wolpert, R. L. (2016). Pooling strength amongst limited datasets using hierarchical bayesian analysis, with application to pyroclastic density current mobility metrics. *Statistics in Volcanology*, 2:1–26.

Owen, A. B. (1992a). A Central Limit Theorem for Latin Hypercube Sampling. *Journal of the Royal Statistical Society.*, 54(2):541–551.

Owen, A. B. (1992b). Orthogonal arrays for computer experiments, integration and visualization. *Statistica Sinica*, 2(2):439–452.

Patra, A., Nichita, C., Bauer, A., Pitman, E., Bursik, M., and Sheridan, M. (2006). Parallel adaptive discontinuous galerkin approximation for thin layer avalanche modeling. *Computers & Geosciences*, 32(7):912 – 926. Computer Simulation of natural phenomena for Hazard Assessment.

Patra, A. K., Bauer, A. C., Nichita, C. C., Pitman, E. B., Sheridan, M. F., Bursik, M., Rupp, B., Webber, A., Stinton, A. J., Namikawa, L. M., and Renschler, C. S. (2005). Parallel adaptive numerical simulation of dry avalanches over natural terrain. *Journal of Volcanology and Geothermal Research*, 139(1-2):1–21.

Pitman, E., Nichita, C., Patra, A., Bauer, A., Bursik, M., and Webb, A. (2003a). A model of granular flows over an erodible surface. *Discrete and Continuous Dynamical Systems - Series B*, 3(4):589–599.

Pitman, E. B. and Le, L. (2005). A two-fluid model for avalanche and debris flows. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 363(1832):1573–601.

Pitman, E. B., Nichita, C. C., Patra, A., Bauer, A., Sheridan, M., and Bursik, M. (2003b). Computing granular avalanches and landslides. *Physics of Fluids*, 15(12):3638.

Pouliquen, O. (1999). Scaling laws in granular flows down rough inclined planes. *Physics of Fluids*, 11(3):542–548.

Pouliquen, O. and Forterre, Y. (2002). Friction law for dense granular flows: application to the motion of a mass down a rough inclined plane. *Journal of Fluid Mechanics*, 453:133–151.

Procter, J. N., Cronin, S. J., Platz, T., Patra, A., Dalbey, K., Sheridan, M., and Neall, V. (2010). Mapping block-and-ash flow hazards based on titan 2d simulations: a case study from mt. taranaki, nz. *Natural Hazards*, 53(3):483–501.

Pudasaini, S. P. and Hutter, K. (2003). Rapid shear flows of dry granular masses down curved and twisted channels. *Journal of Fluid Mechanics*, 495:193–208.

Ranjan, P. and Spencer, N. (2014). Space-filling latin hypercube designs based on randomization restrictions in factorial experiments. *Statistics and Probability Letters*, 94(Supplement C):239 – 247.

Rankine, W. J. M. (1857). On the Stability of Loose Earth. *Phil. Trans. R. Soc. Lond.*, 147(2):9–27.

Reyes-Dávila, G. A., Arámbula-Mendoza, R., Espinasa-Perena, R., Pankhurst, M. J., Navarro-Ochoa, C., Savov, I., Vargas-Bracamontes, D. M., Cortés-Cortés, A., Gutiérrez-Martínez, C., Valdés-González, C., Domínguez-Reyes, T., González-Amezcua, M., Martínez-Fierros, A., Ramírez-Vázquez, C. A., Cárdenas-González, L., Castaneda-Bastida, E., de los Monteros, D. M. V. E., Nieto-Torres, A., Campion, R., Courtois, L., and Lee, P. D. (2016). Volcan de colima dome collapse of july, 2015 and associated pyroclastic density currents. *Journal of Volcanology and Geothermal Research*, 320(Supplement C):100 – 106.

Rupp, B. (2004). An analysis of granular flows over natural terrain. Master's thesis, University at Buffalo.

Rupp, B., Bursik, M., Namikawa, L., Webb, A., Patra, A. K., Saucedo, R., Macías, J. L., and Renschler, C. (2006). Computational modeling of the 1991 block and ash fows at Colima Volcano, Mèxico. *Geological Society of America Special Papers*, 402(11):223–237.

Ruyer-Quil, C. and Manneville, P. (2000). Improved modeling of flows down inclined planes. *Eur. Phys. J. B*, 15:357–369.

Salm, B. (1993). Flow, flow transition and runout distances of flowing avalanches. *Annals of Glaciology*, 18:221–226.

Salm, B., Burkard, A., and Gubler, H. (1990). Berechnung von Fliesslawinen: eine Anleitung für Praktiker mit Beispielen. *Mitteilungen des Eidgenössische Institutes für Schnee- und Lawinenforschung*, 47.

Sarocchi, D., Sulpizio, R., Macías, J., and Saucedo, R. (2011). The 17 july 1999 block-and-ash flow (baf) at colima volcano: New insights on volcanic granular flows from textural analysis. *Journal of Volcanology and Geothermal Research*, 204(1):40 – 56.

Saucedo, R., Macías, J., Bursik, M., Mora, J., Gavilanes, J., and Cortes, A. (2002). Emplacement of pyroclastic flows during the 1998–1999 eruption of volcán de colima, méxico. *Journal of Volcanology and Geothermal Research*, 117(1):129 – 153.

Saucedo, R., Macías, J., Gavilanes, J., Arce, J., Komorowski, J., Gardner, J., and Valdez-Moreno, G. (2010). Eyewitness, stratigraphy, chemistry, and eruptive dynamics of the 1913 plinian eruption of volcán de colima, méxico. *Journal of Volcanology and Geothermal Research*, 191(3–4):149 – 166.

Saucedo, R., Macías, J., Sheridan, M., Bursik, M., and Komorowski, J. (2005). Modeling of pyroclastic flows of colima volcano, mexico: implications for hazard assessment. *Journal of Volcanology and Geothermal Research*, 139(1):103 – 115. Modeling and Simulation of Geophysical Mass Flows.

Saucedo, R., Macías, J. L., and Bursik, M. (2004). Pyroclastic flow deposits of the 1991 eruption of Volcán de Colima, Mexico. *Bulletin of Volcanology*, 66(4):291–306.

Savage, S. B. and Hutter, K. (1989). The motion of a finite mass of granular material down a rough incline. *Journal of Fluid Mechanics*, 199:177.

Sheridan, M., Stinton, A., Patra, A., Pitman, E., Bauer, A., and Nichita, C. (2005). Evaluating titan2d mass-flow model using the 1963 little tahoma peak avalanches, mount rainier, washington. *Journal of Volcanology and Geothermal Research*, 139(1):89 – 102. Modeling and Simulation of Geophysical Mass Flows.

Sheridan, M. F. and Macías, J. L. (1995). Estimation of risk probability for gravity-driven pyroclastic flows at volcan colima, mexico. *Journal of Volcanology and Geothermal Research*, 66(1):251 – 256. Models of Magnetic Processes and Volcanic Eruptions.

Sheridan, M. F., Patra, A. K., Dalbey, K., and Hubbard, B. (2010). *Probabilistic digital hazard maps for avalanches and massive pyroclastic flows using TITAN2D.*, volume Stratigraphy and geology of volcanic areas, 464, chapter 14, page 281. Geological Society of America.

Silbert, L. E., Ertas, D., Grest, G. S., Halsey, T. C., Levine, D., and Plimpton, S. J. (2001). Granular flow down an inclined plane: Bagnold scaling and rheology. *Physical Review E*, 64(5):051302.

Spiller, E. T., Bayarri, M. J., Berger, J. O., Calder, E. S., Patra, A. K., Pitman, E. B., and Wolpert, R. L. (2014). Automating emulator construction for geophysical hazard maps. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1):126–152.

Stefanescu, E. R., Bursik, M., Cordoba, G., Dalbey, K., Jones, M. D., Patra, A. K., Pieri, D. C., Pitman, E. B., and Sheridan, M. F. (2012a). Digital elevation model uncertainty and hazard analysis using a geophysical flow model. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 468(2142):1543–1563.

Stefanescu, E. R., Bursik, M., and Patra, A. K. (2012b). Effect of digital elevation model on mohr-coulomb geophysical flow model output. *Natural Hazards*, 62(2):635–656.

Stein, M. (1987). Large sample properties of simulations using latin hypercube sampling. *Technometrics*, 29(2):143–151.

Sulpizio, R., Capra, L., Sarocchi, D., Saucedo, R., Gavilanes-Ruiz, J., and Varley, N. (2010). Predicting the block-and-ash flow inundation areas at volcán de colima (colima, mexico) based on the present day (february 2010) status. *Journal of Volcanology and Geothermal Research*, 193(1):49 – 66.

Tang, B. (1993). Orthogonal array-based latin hypercubes. *Journal of the American Statistical Association*, 88(424):1392–1397.

Voellmy, A. (1955). Über die Zerstörungskraft von Lawinen. *Schweiz Bauzeitung*, 73:159–165, 212–217, 246–249, 280–285.

Webb, A. (2004). Granular flow experiments to validate numerical flow model, TITAN2D. Master's thesis, University at Buffalo.

Webb, A. and Bursik, M. (2016). Granular flow experiments for validation of numerical flow models. Technical report, University at Buffalo.

Wolpert, R. L., Ogburn, S. E., and Calder, E. S. (2016). The longevity of lava dome eruptions. *Journal of Geophysical Research: Solid Earth*, 121(2):676–686. 2015JB012435.

Yu, B., Dalbey, K., Webb, A., Bursik, M., Patra, A. K., Pitman, E. B., and Nichita, C. (2009). Numerical issues in computing inundation areas over natural terrains using Savage-Hutter theory. *Natural Hazards*, 50(2):249–267.

Zobin, V., Luhr, J., Taran, Y., Bretón, M., Cortés, A., Cruz-Reyna, S. D. L., Domínguez, T., Galindo, I., Gavilanes, J., Muniz, J., Navarro, C., Ramírez, J., Reyes, G., Ursúa, M., Velasco, J., Alatorre, E., and Santiago, H. (2002). Overview of the 1997–2000 activity of volcán de colima, méxico. *Journal of Volcanology and Geothermal Research*, 117(1):1 – 19.

Zobin, V. M., Arámbula, R., Bretón, M., Reyes, G., Plascencia, I., Navarro, C., Téllez, A., Campos, A., González, M., León, Z., Martínez, A., and Ramírez, C. (2015). Dynamics of the january 2013–june 2014 explosive-effusive episode in the eruption of volcán de colima, méxico: insights from seismic and video monitoring. *Bulletin of Volcanology*, 77(4):31.