PROBABILISTIC QUANTIFICATION OF HAZARDS: A METHODOLOGY USING SMALL ENSEMBLES OF PHYSICS-BASED SIMULATIONS AND STATISTICAL SURROGATES

M. J. Bayarri,[1] J. O. Berger,[2,3] E. S. Calder,[4] A. K. Patra,[5] E. B. Pitman,[6,]* E. T. Spiller,[7] & Robert L. Wolpert[2]

[1] *Departament d Estadistica y Investigacio Operativa, Universitat de València, 46100 Burjassot, Valencia, Spain; Deceased August 19, 2014*

[2] *Department of Statistical Science, Duke University, Durham, North Carolina 27708-0251, USA*

[3] *Department of Statistics, King Abdulaziz University, Jeddah, Saudi Arabia*

[4] *Earth and Planetary Science, The University of Edinburgh, Edinburgh, UK*

[5] *Department of Mechanical and Aerospace Engineering, University at Buffalo, Buffalo, New York 14260, USA*

[6] *Department of Mathematics, University at Buffalo Buffalo, New York 14260, USA*

[7] *Department of Mathematics, Statistics, and Computer Science, Marquette University, Milwaukee, Wisconsiin 53201, USA*

This paper presents a novel approach to assessing the hazard threat to a locale due to a large volcanic avalanche. The methodology combines: (i) mathematical modeling of volcanic mass flows; (ii) field data of avalanche frequency, volume, and runout; (iii) large-scale numerical simulations of flow events; (iv) use of statistical methods to minimize computational costs, and to capture unlikely events; (v) calculation of the probability of a catastrophic flow event over the next T years at a location of interest; and (vi) innovative computational methodology to implement these methods. This unified presentation collects elements that have been separately developed, and incorporates new contributions to the process. The field data and numerical simulations used here are subject to uncertainty from many sources, uncertainties that must be properly accounted for in assessing the hazard. The methodology presented here will be demonstrated with data from the Soufrière Hills Volcano on the island of Montserrat, where there is a relatively complete record of volcanic mass flows from the past 15 years. This methodology can be transferred to other volcanic sites with similar characteristics and where sparse historical data have prevented such high-quality analysis. More generally, the core of this methodology is widely applicable and can be used for other hazard scenarios, such as floods or ash plumes.

KEY WORDS: *volcanic hazards, hazard assessment, computer modeling, statistics, uncertainty*

---

*Correspond to E. B. Pitman, E-mail: abani@buffalo.edu

1. OVERVIEW

Mudslides in California; the Eyjafjallajökull ash plume; pyroclastic flows at Mt. Merapi; floods in Queensland. These natural hazards can devastate large regions, wreaking untold damage and threatening life. What is more, these hazards are largely unpredictable. In spite of this seeming randomness, geo-scientists and mathematicians are working together with civil protection agencies, to outline a theory quantifying the hazard risk to local population.

This paper reviews one approach to assessing hazard risk, due to a large volcanic avalanche. We focus the discussion on volcanic avalanches because of a treasure trove of available data from the Soufrière Hills Volcano (SHV) on the island of Montserrat in the West Indies, but the methodology we present can be applied to other hazards. We discuss a method of finding predictive capability in the unpredictability of volcanic eruptions. We discuss how to account for uncertain inputs to mathematical models of volcanic flows. We discuss the computational effort required to make quantitative predictions of hazards. Throughout, our investigation is informed by a combination of mathematical and computational modeling, statistical analysis, and geological theory.

Our specific goal is to develop and automate a method to calculate the probability of at least one catastrophic event occurring at a specific location (such as the center of a town) over the next T years. If this method could be speeded up, a map of an entire region could be created, providing civil protection officials with quantitative information to help make decisions about risk and evacuation. To make this kind of prediction we must make assumptions about the recurrence of *potentially* catastrophic events occurring with some frequency. Of course assessments of the potential of a catastrophic event cannot rely entirely on field data, since these events are scarce. One alternative is to solicit expert opinion, a traditional way to proceed in many assessments. The catastrophic events we consider are highly nonlinear and are so complex that expert opinion is subject to significant error—an error that is very difficult to quantify. Another alternative is to employ advanced computing together with mathematical modeling to simulate catastrophic events. In this work we incorporate geological insight and mathematical analysis into a computational methodology, to determine the probability of a catastrophe. That said, there are many inputs into the mathematical models and computational codes, any of which may be sources of uncertainty. Moreover, the computational simulations are expensive and often there is not time to run sufficiently many simulations near every point in the space of input parameters, especially near those parameters that generate the rare but devastating flows that cause so much damage.

More specifically, using a large but limited number of computer simulations that explore a physically relevant set of inputs (4 separate input parameters, and other uncertain quantities), we construct a Gaussian Process emulator, a surrogate that approximates, with a computable error, the simulator outputs at untested values of the inputs. Using this emulator, we determine a distinguished locus in the space of *inputs* that characterizes catastrophic flow events. By adaptively resampling near this locus we substantially reduce the expected error in a computation of the probability of a catastrophic event over the next T years. With modest computing resources and careful analysis, one can extend this computation to create a map showing the probability of catastrophe throughout a region.

Beyond the uncertain input parameters, there are other uncertainties associated with geophysical flows, including the elevation map that defines the region of interest, and the apparent dependency of energy dissipation on flow volume. More fundamentally, the model equations themselves represent the mass flows under study only approximately. We discuss these uncertainties and ways of mitigating their effect on our calculations.

As mentioned, the methods that will be discussed here are quite general, and may be useful in a variety of computer modeling and risk assessment settings. Our analysis focuses on volcanic avalanches and related geophysical mass flows, but similar ideas can be applied to the threat from volcanic ash plumes, as evidenced by analysis of the eruption of Eyjafjallajökull in 2010 [1, 2]. An even more direct extension of our approach is to overland flow and flooding, consequent to hurricanes rain, or snow melt (see [3]). More generally, there is interest in the geophysics community in promoting the application of statistical tools, to advance quantitative hazard and risk assessment [4–7].

Before continuing let us clarify some of the expressions and definitions used here. This paper is concerned with the probability of hazardous events occurring at a location or in a region. A related notion is of risk. Risk involves the probability of a hazardous event and the consequent costs (financial and in terms of lives lost) of that event. We will not address risk, although risk is a principal concern of civil protection agencies. Throughout this paper, our concern is with the probability of a "catastrophic event"—defined here as a flow of thickness 1 m or more. In the center of a town at some distance from a volcano, a 1 m flow might well be a catastrophe. Close to the volcano a 1 m thick flow

might not be too unusual and might not cause serious disruption. In any event, the exact thickness of a "catastrophic" flow is ultimately left for the user to specify.

2. INTRODUCTION

The focus of this paper is assessing the risk of extreme natural hazards, such as volcano avalanches producing a large pyroclastic flow or debris avalanche. The geophysical processes activated in a pyroclastic flow or avalanche range from centimeter-sized particles suspended in a mixture of water, rock, soil, and other materials, flowing over terrain that varies in composition and roughness, with speeds of 10–20 m/s or more and producing kilometer-sized runouts [8–11]. During flow, the composition of the deforming mass often changes, as larger particles migrate to the top of the flowing mass, interstitial fluid liquefies the solid particles, and new solid and fluid material is entrained. Needless to say, a detailed description of the physics activated at every length and time scale is all but impossible. Nonetheless, there is a tremendous need to understand these flows—how to describe the physics of the flowing material and how to quantify the uncertainties in the inputs and in the results—to aid civil protection authorities who are charged with hazard assessment and risk mitigation [12, 13].

Most flow events result in only mild damage to an area very close to the volcano. On rare occasions, however, a very large event may occur with catastrophic consequences. We are interested in the probability of at least one catastrophic event occurring in the near future. More specifically, in this exposition we ask *What is the probability of a volcanic event that results in a flow of a depth of 1 m reaching a designated location (say, the center of town) sometime during the next T years?*

As a testbed for our methodology, we use data from the SHV. The SHV was dormant for some 300 years, until it erupted in 1995. The SHV has been erupting frequently since then, with periods of vigorous dome growth, seismic activity, and mass flow events interrupted by pauses in the rate of growth of its lava dome. There is a relatively complete record of volcanic mass flows of given size and runout length, data that provide the backbone for the statistical approach used here. Although this choice may appear to limit the applicability of our approach, in fact it does not. The methodology outlined here will carry over directly to similar dome-forming volcanoes (e.g., Mount St. Helens, or Merapi), even if the record of flow events is less complete [14]. In addition, rapid developments of remote sensing techniques provide detailed field observations at many sites [15, 16], including other types of volcanoes. With suitable modifications, the approach can be applied to floods and to atmospheric plumes also.

2.1 Approaches to Hazard Assessment

At present, hazard assessments are often produced by eliciting expert opinion, which is primarily based on the record of historical flows at a specific volcano, and an assumption that future events are likely to be of approximately the same magnitude and frequency as past events, and affect roughly the same areas. Certainly the study of deposits provides important information about the volcano and its explosive capabilities. Experts studying deposits provide rich insight into potential eruption scenarios. But a hazard map that relies only on historical eruptions suffers certain shortcomings. For example, when examining deposits, smaller flows may be overwhelmed by larger ones, so the frequency of all events may be under-represented. Several large eruptions in the past may appear as a single event if they occurred within a short time period. More importantly, the historical record is sparse, and a scientist cannot make reliable predictions of frequency based solely on deposits and expert opinion.

Historically, hazard maps have tended to be lines drawn on a map, without regard to a quantification of the hazard threat. A not-so-apocryphal story is of a hazard map showing one side of a street to be "safe" while an evacuation is ordered for the other side.

Sheridan and Malin [17] proposed an energy cone model to estimate the furthest extent of a volcanic event. The energy cone compares the elevation H of the starting location of erupting mass, and the total dissipation μ acting on the moving mass, to give a runout length L from the relation $H/L = \mu$. From this runout one can obtain an estimate of at-risk areas. The energy cone concept has been applied to the hazard at Vulcano, Lipari, and Vesuvius [17], Soufrière Hills Volcano on Montserrat [18], and Campi Flegrei, Italy [19], among other locations.

A weakness of the energy line models is that it assumes a straight-line flow trajectory for flows, irrespective of topography. Using personal computers of the time, in the early 1990s Sheridan and Macías [20, 21] developed the FLOW2D code. FLOW2D launches a large number of representative point mass particles, and assumes that the shear resistance of a particle depends on basal friction and viscosity. Particle runout is then calculated, giving areas that might be impacted by a flow. FLOW2D incorporated some relevant physics, such as particle momentum. However, the topography of the region needed to be tabulated for each flow path, a computationally intensive exercise.

Later Kover devised FLOW3D [22, 23], which is a more refined kinematic model of flowing particles over a Triangulated Irregular Network of elevations. The physics of the model accounts for the acceleration of the point mass particles due to gravity, and frictional and viscous resistances to motion. Unfortunately FLOW3D does not track momentum of the particles, which are therefore constrained to flow downhill.

In 1989 Savage and Hutter [24, 25] developed a "thin layer" model of granular avalanches, a system of partial differential equations (PDEs) for the mass and momentum balance of a flowing mass. Iverson and colleagues further developed this idea [26, 27], and other improvements followed [28–30]. To solve this system of equations over real to-pography, Patra and colleagues [31] developed the TITAN2D simulation environment. TITAN2D encodes a Godunov method for hyperbolic balance laws, together with an adaptive mesh technique that focuses computational work on areas of rapid changes in the flow; see Section 3.

These computational models for flow provide insight into the possible paths a mass flow might take, thus providing some indication of areas at risk. A more informative result might give a graduated level of risk—higher in some regions, lower in others. The additional information would better indicate the threat to population, and nearby regions of relatively lower risk.

2.2 A Methodology for Quantifying Hazards

Our goal is to develop a quantitative measure of the hazard threat to a region, providing those charged with public safety more than a yes–no alternative when deciding upon a course of action in response to volcanic activity. The methodology that will be developed combines:

- mathematical modeling of volcanic mass flows;

- field data of avalanche frequency, volume, and runout;

- large-scale numerical simulation of flow events;

- statistical methods to minimize computational costs, and to capture unlikely events;

- calculation of the probability of a catastrophic flow event over the next T years at a location of interest;

- innovative computational methodology to implement these methods.

The accuracy of all our assumptions, and of field data and numerical simulations, must be accounted for in assessing the hazard.

One set of inputs to our computational model contains physical parameters that describe the rheology of the flow-ing material (see Sections 3.1 and 3.2). Any rheological model is at best an approximate description at a macroscopic scale of complex phenomena occurring at all smaller scales. Field data show a relationship between the volume of a flowing mass and its rheology, but the foundation for this relationship is not clear (see Section 3.3.1). A second computational input is a description of the topography of the area of interest, which is likely known to within certain errors, and may be viewed as a stochastic field (see Section 3.3.2).

2.3 Uncertainty Analysis

Having just identified several sources of uncertainty and error in our modeling and simulations, it may be useful to pause for a moment to clarify our approach toward uncertainty.

First let us define the kinds of uncertainty one might encounter in studying a physical system. Epistemic uncertainty involves parameters and inputs which, in principle, could be known but which in practice are not. This imperfect knowledge may be due to not measuring a quantity sufficiently accurately, or because mathematical model equations neglect certain effects that are thought to be small. Aleatoric uncertainty involves unknowns that are truly random and differ each time we run an experiment.

It is also true that the phrase "uncertainty analysis" can mean different things in different settings. Epistemic uncertainty in data, parameters, or other inputs can be propagated through a system of model equations, either numerically or analytically, to give an approximation of the distribution of outputs. Sensitivity analysis might be useful in characterizing the effect of this uncertainty. Classically, asymptotic methods provided insight into the behavior of solutions to equations with random coefficients [32]. More recently a general structure for studying the propagation of uncertainty was developed using a separation of random and deterministic variables together with functional approximations; these ideas can be found in many references, including [33, 34].

In our study of volcanic hazards both epistemic and aleatoric uncertainty may arise. Our approach is to use knowledge of the fundamental geology to inform the decisions we make in order to solve—at least approximately—the problem at hand, using whatever mathematical techniques are available to obtain answers, provided we can make a statement about the accuracy of the answer.

3. MATHEMATICAL MODELING AND SIMULATION

We now describe the mathematical basis for modeling thin geophysical mass flows and the simulation of these flows, and discuss the uncertainties associated with models and simulations.

3.1 Thin Layer Models

A volcanic avalanche or pyroclastic flow contains sand and soil made of particles of millimeter size, stones in the centimeter range, and large rocks and boulders that are meter sized, all in a liquid or gas matrix. These mass flows typically contain 10^6–10^8 m^3 or more of solids material, leaving deposits that may be tens of meters deep and several kilometers long. This enormous span of scales presents significant challenges to those developing models of the flow process. In this subsection we discuss the important assumptions about depth-averaged models of geophysical mass flow, and refer the reader to papers by Savage and Hutter [24, 25], Gray and colleagues [28, 35], and Iverson and Denlinger [26, 27]. Here we follow the derivation of Patra et al., [31, 36, 37], which may be consulted for further details.

Consider a thin layer of an incompressible granular material of density ρ_0 with characteristic thickness H and characteristic extent L. Allow this mass to flow over a basal surface described, in a coordinate system aligned with the bed, by the surface $z = b(x, y)$. Scale z by H, x, and y by L, the stresses by $\rho_0 gH$ and time as $t^2 = L/g$. The layer is subject to forces in the x, y, and z directions due to gravity g, friction, and material deformation, assumed to be characterized by Mohr-Coulomb plasticity. Taking advantage of the fact that $\epsilon \equiv H/L \ll 1$, we depth-average the conservation laws of mass and momentum to obtain evolution equations for the flow thickness h and momenta hv^x, hv^y (v^x, v^y are the depth-averaged down slope and cross-slope velocities) of the material at any point (x, y) and time t:

$$
\frac{\partial}{\partial t}\begin{pmatrix} h \\ hv^x \\ hv^y \end{pmatrix} + \frac{\partial}{\partial x}\begin{pmatrix} hv^x \\ hv^{x2} + \frac{1}{2}\eta g^z h^2 \\ hv^x v^y \end{pmatrix} + \frac{\partial}{\partial y}\begin{pmatrix} hv^y \\ hv^x v^y \\ hv^{y2} + \frac{1}{2}\eta g^z h^2 \end{pmatrix} \tag{1}
$$

$$
= \begin{pmatrix} 0 \\ hg^x - \left(\dfrac{v^x}{\sqrt{v^{x2} + v^{y2}}}\right) h\tan(\delta_{\text{bed}})\left[g^z + v^{x2}\dfrac{db}{dx}\right] - \operatorname{sgn}\left(\dfrac{\partial}{\partial y}v^x\right)\dfrac{\partial}{\partial y}\left(\dfrac{\eta}{2}\sin(\delta_{\text{int}})h^2 g^z\right) \\ hg^y - \left(\dfrac{v^y}{\sqrt{v^{x2} + v^{y2}}}\right) h\tan(\delta_{\text{bed}})\left[g^z + v^{y2}\dfrac{db}{dy}\right] - \operatorname{sgn}\left(\dfrac{\partial}{\partial x}v^y\right)\dfrac{\partial}{\partial x}\left(\dfrac{\eta}{2}\sin(\delta_{\text{int}})h^2 g^z\right). \end{pmatrix}
$$

In this system, the first equation represents conservation of mass, and results from depth-averaging the continuity equation. The second and third equations describe momentum balance; inertial and pressure terms are on the left-hand side of the equation, and the right-hand side contains gravitational acceleration, dissipation due to basal friction, and internal frictional dissipation. The parameter δ_{int} is the internal friction angle, a measure of particle-particle dissipation, and δ_{bed} is the basal friction angle, measuring the dissipation of particles on the basal surface. The coefficient $\eta = \epsilon K$ where $K(\delta_{\text{int}}, \delta_{\text{bed}})$, the so-called earth pressure coefficient, is a function of the friction angles and describes the ratio of shear stresses to normal stress within the flowing mass. Specifically K is in the active or passive state, depending on whether the down-slope and cross-slope flows are expanding or contracting. We modify Iverson's assumption [27] (see also [24, 38]) and define two values for K

$$K = 2\frac{1 \pm [1 - \cos^2(\delta_{\text{int}})/\cos^2(\delta_{\text{bed}})]^{1/2}}{\cos^2(\delta_{\text{int}})} - 1 \ ,$$

where the $+$ (passive state) applies when flow is converging, that is, if $(\partial/\partial x)v^x + (\partial/\partial y)v^y < 0$, and the $-$ (active state) applies if $(\partial/\partial x)v^x + (\partial/\partial y)v^y > 0$. We use g^x, g^y, g^z to denote the local components of gravity, and note that local topographic slopes are hidden in these terms. (1) are a variant of the original Savage-Hutter equations that incorporate certain refinements due to Iverson.

We emphasize that in writing these thin layer equations we have made the long wave assumption, so only $O(L)$ variations are represented in the system. Shorter wavelengths are not accounted for. Of course this approximation contributes to the epistemic uncertainty of the problem. These equations are hyperbolic, with two nonlinear wave speeds and one linearly degenerate one; the system loses strict hyperbolicity at the "vacuum" state where $h = 0$. Standard techniques for solving hyperbolic systems may be used to compute solutions (for example, [39, 40]).

The system as written does not include the effects of a second fluid phase (gas or water) on the motion of the flowing mass (so these equations do not model wet debris flows). Two phase models have been developed [30, 37], but have not been extensively used for hazard evaluation. Similarly (1) does not include the effects of erosion. For large mass flows, as much as half the final deposit can be material that was eroded and entrained; see [41, 42] for models that incorporate erosion.

3.2 Simulations and Simulation Input

TITAN2D is an environment for computing solutions to Eq. (1) for dry geophysical mass flows [31]. TITAN2D is a parallel, adaptive grid, second-order Godunov shock capturing method. A principal feature of the code is the incorporation of topographical data into simulations and grid structure. There is a preprocessing routine in which digital elevation model (DEM) data are imported, potentially from one of several sources. Because of its wide use in the geoscience community and its easy availability, TITAN2D principally use GRASS, Geographic Resources Analysis Support System, which is open source software for data management, spatial modeling, image processing, and visualization. These geographic data define a two-dimensional spatial box in which the simulation will occur. The raw data provide elevations at specified locations. Typically, the data divide a region into a mesh, perhaps squares varying from 5 m on a side with a vertical accuracy of about 2 m to 50 m squares with 10 m vertical accuracy; coarser elevation meshes are available, as are some with a mesh as fine as 1 m. By using geographic data, and interpolating between data points where necessary, a rectangular, Cartesian grid is created. This grid is then indexed in a manner consistent with the computational solver. The raw elevation data are filtered to remove the highest frequency noise and errors, and local surface normals, tangents, and curvatures are calculated; all these are ingredients in the governing PDEs. Finally, the grid data are written out for use, together with simulation output, in post-computation visualization [31].

Other major inputs for TITAN2D are the initial volume, location, and flow direction of the material. The judgment of experts is especially important in choosing the range of values for these inputs. One wishes to sample from the entire range of possible events, giving appropriate weight to the more likely ones.

TITAN2D does not try to represent the initiation of a flow explicitly. Rather, an appropriate mass is placed at the specified initial location and it is assumed this mass will flow if gravitational energy can overcome friction. TITAN2D also allows the user to define a mass flux emanating from a specified location, but we do not discuss this feature here.

By extending the spatial region of interest modestly, it is easy to confine the flowing mass to the middle of a computational domain, avoiding the need to impose boundary conditions.

The final input factors are constitutive, namely the internal and bed friction angles.

TITAN2D has been tested against several flows for which we have at least partial data, including laboratory-sized flows of a mass moving down an inclined plane, and field simulations ("post-dictions") of historical flows, including Volcan de Colima in Mexico and the Little Tahoma Peak in Washington [43]. TITAN2D has also been compared with LaharZ, a non-PDE-based tool for predicting mass flow deposits [44]. These comparisons give confidence in using TITAN2D for predictive capabilities.

A simple extension of the model just presented allows for varying basal friction. As presented, δ_{bed} appears as a constant scalar input parameter characterizing the basal surface over which material flows. However δ_{bed} could easily be a function of location, $\delta_{bed} = \delta_{bed}(x, y)$. As one example of this extension, consider the Little Tahoma Peak flow of 1964. This rockfall was initiated on a glacier, proceeded over a rocky flank, and ended on soil surface. In [43] the authors used three different basal friction angles for these different regions; indeed, specifying different friction factors was necessary to reproduce the runout and overbanking data that are known for this flow.

An easy to use version of TITAN2D is available to the public through Vhub (vhub.org/resources/titan2d).

3.3 Uncertainties

In risk assessment, it is important to incorporate all major sources of uncertainty into the analysis. In later sections we talk about the dominant uncertainty, namely the aleatoric uncertainty in the volcanic process itself. But accounting for uncertainties in the computer model is also important in uncertainty quantification. Tests using multiple resolutions and the impact of numerical resolution on hazard probability calculations indicate that numerical accuracy is important. In order to focus on other sources of uncertainty, we begin with a fine computational mesh that is then refined adaptively within TITAN2D during a simulation in order to maintain a high degree of accuracy. Furthermore, with this approach, numerical instability or problems of convergence do not contribute significant variability or error. We turn now to discuss the most significant uncertainties in our modeling and simulations.

3.3.1 Effective Friction

Because they are intuitive physically, one might imagine the friction angles to be well-defined quantities. One way of characterizing the internal friction is as the angle made by a conical heap of material when poured onto a flat surface. To find the basal friction, consider a layer of flowing material one or two grains deep, on a horizontal plane made of the basal material; as one slowly raises this plane, the basal friction is the angle the plane makes when the material first begins to slide. Unfortunately these measurement techniques allow for significant variability. For example, in [45] Blijenberg sampled the angle of repose for several natural materials at locations throughout the Alps. His data show a variability of about ± 2–$3°$ around a mean angle (about a ± 5–7% variability), irrespective of the material (sandstone, limestone, etc.).
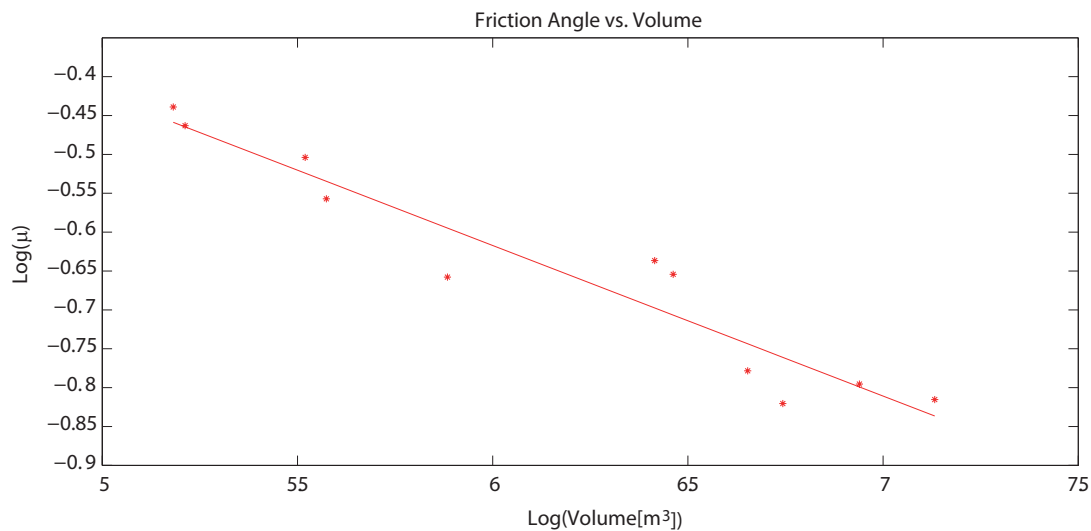
Even more problematic are findings such as [46], which indicates an apparent *volume* dependency of the *effective* basal friction. That is, for a large dataset of terrestrial and extraterrestrial avalanche events, the effective friction—that is, the ratio of the height at which the avalanche originated to the runout length—seems to decrease with increasing volume. This so-called "long runout" problem has been studied extensively (see [12, 47]), but there is no generally accepted explanation of the phenomenon. Using data from Montserrat and other similar volcanoes, we model the effective friction, given the volume of a mass flow.

Table 1 summarizes 11 large mass flows at SHV in which the deposited material was mostly confined to a single valley. The table reports the volume of each flow, together with the drop in elevation H and the horizontal runout L of the flow, and the effective friction angle δ_{bed}^{eff} measured as $\mu = \tan(\delta_{bed}^{eff}) = H/L$. Figure 1 is a plot of this data, showing the $\log_{10}$ of the volume and the $\log_{10}$ of μ, for each of the flows, and an approximate linear fit of the data.

Although suggestive, these 11 data points are insufficient to demonstrate a strong correlation between volume and friction. In Fig. 2 we show similar data from some 100 flow events at SHV, Unzen (Japan), Colima (Mexico),

TABLE 1: Effective Friction illustrated by 11 flows at the Soufrière Hills Volcano

Vertical fall H (km)	Horizontal runout L (km)	Effective friction angle δ_{bed}^{eff} (deg)	Volume V ($\times 10^6$ m^3)
0.59	1.6	20	0.152
0.91	2.9	17.4	0.331
0.86	2.5	19	0.163
0.86	3.1	15.5	0.375
0.86	3.9	12.4	0.766
1.01	6.7	8.6	5.538
0.83	3.6	13.0	2.6
0.9	4.05	12.5	2.9
1.0	5.5	9.46	4.5
0.9	5.6	9.1	8.7
0.92	6.0	8.7	13.563

FIG. 1: Plot showing the (log of the) effective friction μ as a function of the (log of the) flow volume, for each of the 11 flow events listed in Table 1. Each of these flows was largely confined to a single valley on Montserrat. A linear least squares fit of the data shows high correlation.

Augustine (Alaska), Fuego (Guatemala), Asama (Japan), Hakusan (Japan), and Santiaguito (Guatemala).[*] In this figure the axes are on a log scale, and a least-squares fit to the data is shown as a solid line; a one standard deviation bound is indicated by the dashed lines. In all these events, the flowing material is similar—we have no reason, and no ability, to distinguish among the flows by constitutive properties. The natural conclusion is that larger flows generate lower effective dissipation. This conclusion is supported by other geological data; see, for example, [12, 46]. Although the data appear to clearly indicate a "long runout" phenomenon, there is no generally accepted theory of its physical origin.

Given the importance of the basal friction in TITAN2D computations, how might one connect flow volume and dissipation? In [48], Lopes introduced a hierarchical model that provides a distribution for δ_{bed} depending on the volume. In calculations below, each selection of an initial flow volume is followed by a draw from this posterior distribution of δ_{bed}, which is then used in the simulation or emulation evaluation. Samples from this volume-friction distribution are plotted in Fig. 3.

---

[*]Thanks to Sarah Ogburn for compiling these data.
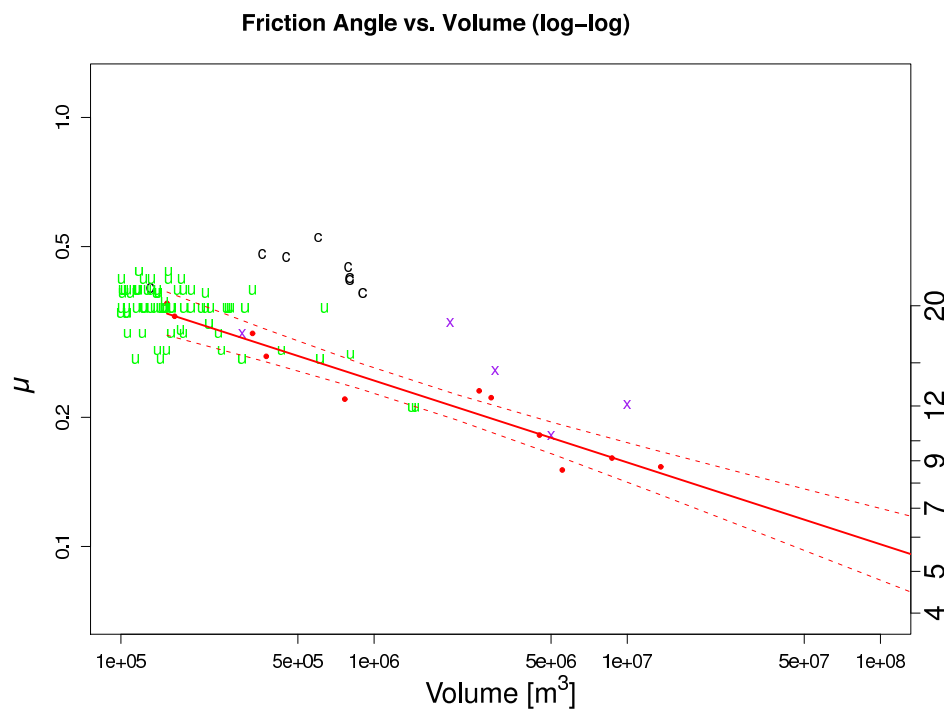
Friction Angle vs. Volume (log–log)

FIG. 2: Figure showing the effective friction versus volume on a log scale, using data from some 100 flow events from volcanoes across the world. Data from different volcanos are marked by a different symbol. A linear fit to the data is shown as the solid line, and one standard deviation bounds are indicated by dashed lines. The left axis shows the effective friction coefficient, and the axis on the right shows the effective friction angle $\delta_{\text{bed}}^{\text{eff}}$.
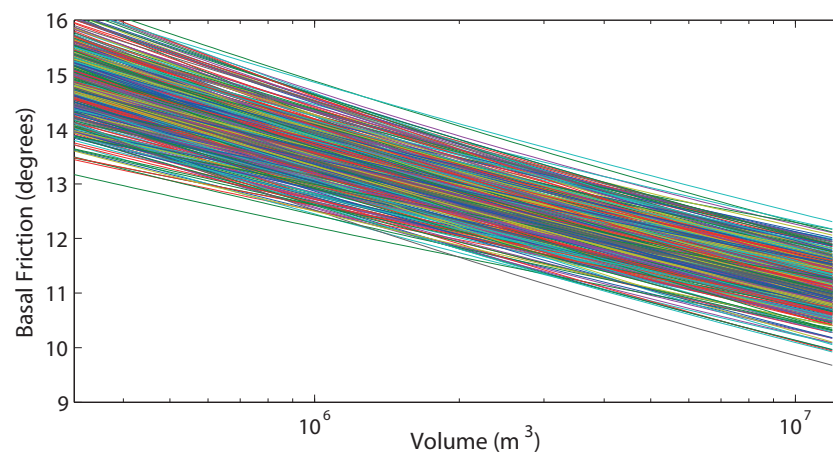

FIG. 3: Two thousand samples drawn from a distribution modeling the volume dependence of the effective friction angle on volume. (A similar figure appeared in [49].)

3.3.2 Uncertainty in DEMs

Another source of uncertainty is the DEM of the topography over which a geophysical flow is computed. The model equations (1) make use of slopes and curvatures of the elevation data. Computation of derivatives amplifies any error that may be present in the elevation measurements. Recent work [50] shows that the variability in DEMs obtained

by different methods (different elevation data sources and different post-processing of the raw elevation data) has significant effects on inundation areas computed by the TITAN2D tool for small and medium sized flows. Figure 4 illustrates this effect. Two different DEMs, each providing a 30 m mesh, were used in TITAN2D simulations with the same initial conditions to compute a modest-sized flow at Mammoth Mountain in California. The regions overrun by the resulting flows are significantly different.

In TITAN2D and other simulation software, the DEM is usually not treated as a model with possible errors, but rather as a "true" representation of the earth's surface [51]. Of course, such an assumption ignores the limitations of the instruments used to determine the DEM. The quality of any elevation model is dependent upon the methods of data acquisition, the nature of the input data, and the methods employed in generating the DEMs, among other factors [52]. At each stage, the error in this data acquisition and DEM generation process is quantifiable. Ultimately the error in the procedure is precisely the difference between reality and our representation of that reality. But we cannot measure the error at every spatial location, for every geographic feature [53–55]. Because "reality" cannot be known precisely, the error must be characterized by methods commonly used to treat epistemic uncertainty.

We can probe the uncertainty in the DEM by examining a measure of the sensitivity of the analysis result to variations in the parameters that are inputs to the DEM and its construction. Naive approaches to treating DEM uncertainty can quickly lead to the use of thousands or millions of random variables. Recognizing the strong relationship between elevation error and terrain, influenced by the spatial autocorrelation of the error (defined as the difference between two independently derived DEMs), we advocate for a procedure introduced by [53–55] that provides a reasonable representation of DEM uncertainty using only a few random variables. A characterization of the error is provided by constructing a random field defined as the difference of two DEMs, one of which is considered to be an accurate baseline DEM. Additional DEMs are then created by sampling the error field and adding this sample to the baseline DEM.

3.3.3 Other Uncertainties

Another uncertain input is the precise location of the initial flowing mass. As shown in [56], if a flow is initiated near the peak of a mountain, small changes in that initiation location can produce flows that travel in very different directions. When using TITAN2D, the user can specify the initial direction and speed of the flow. We treat the initial location, direction, and speed as aleatoric uncertainties, and sample appropriately.

The mathematical model itself makes many simplifying assumptions in order to make the physics tractable. Errors in this physical modeling may be accounted for by a hierarchical model of models. Unfortunately, the data needed to investigate model validity are sparse. One might consider using several different models to help mitigate the
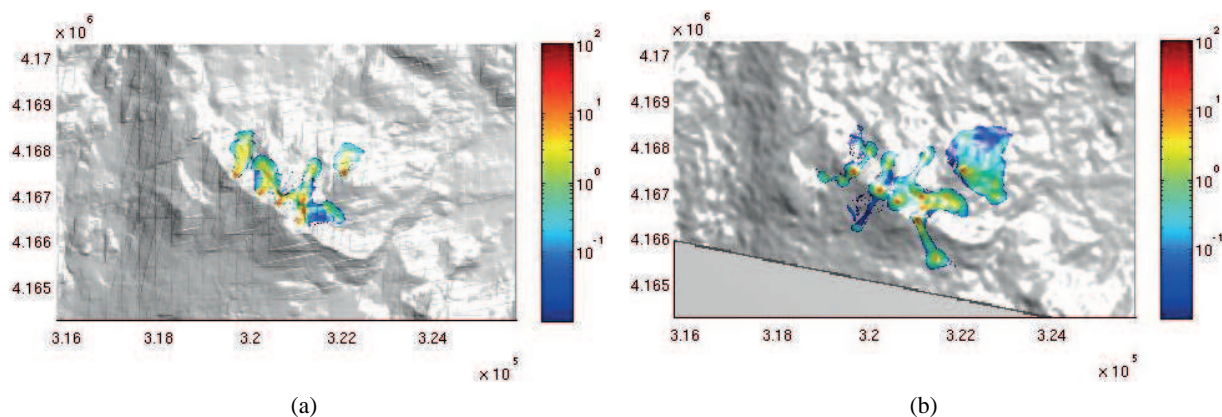

(a)   (b)

FIG. 4: Final flow depth for low-volume flows in UTM easting, northing coordinates, for the same initial inputs but different DEMs, illustrating the sensitivity of flow to DEM accuracy. (a) DEM generated by terrain observation with Topographic Synthetic Aperture Radar (TOPSAR30m) (with spline interpolation) (b) DEM generated by Advanced Spaceborne Thermal Emission and Reflection (ASTER).

shortcomings of one particular model. For example, LAHARZ, FLOW3D, and TITAN2D all provide some notion of the area inundated by mass flows. A Bayesian model average of predictions made by these three models, and perhaps others, might provide a more accurate hazard warning than any one model [57]; results from a model averaging will be presented in future work.

4. STATISTICS, EMULATION, AND THE HAZARD PROBABILITY

Our ultimate goal is to compute a measure of hazard, showing the probability of a catastrophic inundation occurring in the next T years at a specified location, and then to extend this single point calculation to all points in a sensitive region of interest.

We argue that a careful orchestrated mix of computer simulations, probability models, and statistical methods is required to address uncertainty and calculate the hazard probability. No single tool used in isolation will be entirely satisfactory.

We quantify *all* uncertainties (aleatoric or epistemic, whether in the inputs, the initial conditions, or the model) with probabilities. Then a calculation that combines all of these uncertainties is possible by using the laws of probability. In addition, when uncertainties are updated by data, this update can be propagated naturally to the various components of the analysis. We make heavy use of a methodology called objective Bayes analysis [58] which is particularly well suited for efficiently combining disparate sources of uncertainty; with the objective Bayes methodology, we are able to appeal to several sources from which to "borrow" information, to alleviate the difficulties associated with sparse data.

Given a DEM of a region around a volcano (for purposes of this exposition, we assume the DEM to be "true"), and assumed values of input parameters and initial conditions—the initial volume (V) and the direction (angle ϕ) of the initial mass flow—TITAN2D can compute the thickness of a pyroclastic flow at any location at any time during that flow [31]. The large epistemic uncertainties in input parameters and the aleatoric nature of initial conditions must be accounted for in the calculation of the hazard (see Sections 4.4.1 and 4.3). In addition, to estimate the threat of a catastrophic event a probability model for the frequency and severity of eruption events is also required (see Section 4.4.1). Unknown hyper-parameters in all these models are usually given an objective prior distribution which is then updated with data from previous eruptions at SHV (see [59]). To simplify the subsequent discussion, from now on we will refer to the DEM and initial conditions, as well as model and computational input parameters, simply as "inputs", since they are all needed to perform a TITAN2D run. We will sometimes use the term "simulator" to refer to the computer model, reminding the reader that other mass flow models and computer codes could be used instead of TITAN2D.

We use the relatively abundant data at the SVH to fit these input distributions (see Section 4.3 and 4.4.1). Because of limits on computational budgets, it is usually not possible to perform simulations by drawing from the (joint) distribution of the inputs as a set of random inputs and employing a Monte Carlo method to determine the probability of a catastrophic inundation at a given location—both because the simulator is expensive to run, and because a very large number of runs is needed to estimate the extremely small probabilities of the very rare, very large events that cause so much damage.

In Section 4.1 we describe our approach to the hazard assessment problem. Roughly speaking, we do not attempt to directly estimate the probability of a catastrophic inundation, but rather we characterize the critical set of inputs which would produce a catastrophic event, and then use the input distribution to compute the probability of that set (see Section 4.5).

However, we cannot solve this inverse problem (determining the set of inputs from the catastrophic outputs) either, because it would require too many simulator runs. That is, as is common when dealing with complex computer models, the analysis requires results from the simulator (outputs) using inputs at which we have not run the simulator. The usual way to deal with this difficulty is to use the available simulator runs to define an approximation to the computer output at the untried inputs. In Section 4.1 we advocate use of a particular statistical approximation, an *emulator*, which is relatively easy to implement, and produces not only the required approximation but also a measure of the uncertainty in that approximation.

The fact that the most catastrophic events are rare complicates the process of fitting the emulator. We proceed by using an adaptive design. One could adaptively add input points at which to run the simulator "manually"—that is, using human judgment to add additional input points—if one is trying to assess hazards at a small number of locations. However, when the goal is to produce a hazard map, emulators will have to be evaluated at thousands of locations, and manually adapting each emulator is not feasible. Instead we develop a strategy to automatically adapt the design. The process of adapting and automating the design is presented in Section 4.2.

4.1 Emulators

To approximate the outputs of a computer simulation we use statistical models, called *emulators* [60, 61]; specifically we use the very popular GaSPs, or *Gaussian Separable Processes* emulator [49, 59]. A GaSP emulator is constructed to reproduce the simulator output at all design inputs, and to interpolate between these points. Statistical emulators have the advantage of producing not only the required approximations to the simulator at input points where it has not been evaluated, but also quantitative estimates of the likely approximation error. In fact, the GaSP emulator generates full probability distributions, so that uncertainty in the approximated outputs can be combined with all other uncertainties and propagated to all aspects of the analysis. This is an elegant and coherent way to address uncertainty quantification.

To simplify the discussion to follow, we adopt notation that is more standard in the statistics literature. For the remainder of this paper we denote inputs to the computer model generically by $\mathbf{x} \in \mathcal{X}$, and the computer simulation outputs of interest by $y^M(\mathbf{x})$. For present purposes we limit consideration to problems in which a scalar output quantity determines whether or not a particular flow event is "catastrophic" so y^M takes real values but $\mathcal{X}$ may be multidimensional (see [62] and references therein for a discussion). As a very specific scenario, we specify a location such as the center of town or a hospital at which we monitor the mass flow thickness, and track the maximum flow depth at that location throughout the entire flow event.

In our volcano flow application, the judgment of domain scientists and numerical experiments with TITAN2D suggest that the most important inputs are $\mathbf{x} = (V, \phi, \delta_{\text{bed}}, \delta_{\text{int}})$, the flow volume and the direction to which the flow is initially released (as an angle measured from due east), and the basal and internal friction angles. As mentioned, the DEM can be the deciding factor for certain flows, but to avoid the additional concerns regarding the impact of DEM uncertainty, TITAN2D simulation are run on an exceptionally accurate DEM at SHV [63]. For such an accurate DEM, uncertainty was deemed negligible in comparison with all the other uncertainties present.

The output of interest $y^M(\mathbf{x}) = \max_{0 \le t \le T} h(*, t)$ is the maximum height of the flow at the specified location $(*)$ from an eruption with input characteristics $\mathbf{x}$. Let $\mathcal{Y}_C$ be the critical set of output values regarded as "catastrophic." That is, the input $\mathbf{x}$ results in a catastrophic event if $y^M(\mathbf{x}) \in \mathcal{Y}_C$. So for our discussion, a flow is catastrophic if $y^M(\mathbf{x}) \ge 1\text{m}$. Thus the hazard quantification is expressed as $\mathrm{P}\big(y^M(\mathbf{x}) \in \mathcal{Y}_C\big) = \mathrm{P}\big(y^M(\mathbf{x}) \ge 1\big)$. Our challenge, then, is to solve the inverse problem—determine the set of inputs that lead to catastrophic outputs.

A naive Monte Carlo evaluation would select inputs $\mathbf{x}$ from their (joint) distribution and compute the proportion of times that the output lies in $\mathcal{Y}_C$. Consider the computational cost of this approach. For a typical DEM, it can take between 20 and 60 min for TITAN2D to simulate a single flow event on 16 computational cores. To run a million simulations (a modest number for estimating the probabilities of rare events) on a 64-processor cluster would take about a year to complete.

There are other shortcomings of a naive Monte Carlo approach. One is the difficulty of approximating rare events. Accurately capturing these low probability events cannot be accelerated by using fast emulators of probability distributions; emulators are typically quite good in the middle of target distributions (in essence interpolating the data) but are not good at approximating near the tails (in essence, extrapolating), the very region of interest. Another difficulty is that any change in the input distribution (for example, from new evidence about friction angles or the volcano's level of activity) would require that the entire simulation and approximation process be restarted from scratch.

Instead of trying to determine the probability of $\mathcal{Y}_C$ in the output space, our approach is to determine the probability of the set of inputs which would give rise to a catastrophic event—that is, $\mathcal{X}_C$. In other words, we use the computational model to identify the critical set of inputs $\mathcal{X}_C = \{\mathbf{x} \in \mathcal{X} : y^M(\mathbf{x}) \in \mathcal{Y}_C\}$, and then compute the probability of $\mathcal{X}_C$ under the input distribution. In light of the previous paragraph, once $\mathcal{X}_C$ is determined, the input

distribution can be updated or changed and the hazard assessment—determining $P(\mathcal{X}_C)$—would not require a new collection of simulator runs. Furthermore, intuition suggests, and both physical and computer experiments largely support, that the maximum flow depth at a specific location varies monotonically with the initial flow volume [i.e., if a given $(V_1, \phi, \delta_{\text{bed}}, \delta_{\text{int}})$ results in a catastrophic event, so will $(V_2, \phi, \delta_{\text{bed}}, \delta_{\text{int}})$ where $V_2 > V_1$]. This means that all we need to determine is the boundary of $\mathcal{X}_C$, called Ψ, since larger initial volumes will produce larger maximum flow thicknesses. Thus we do not have to simulate extremely rare events, but only events near this critical boundary. Because a precise determination of Ψ would require a large number of simulator runs, we instead use the GaSP emulator to determine the boundary of $\mathcal{X}_C$ at least approximately. Note that emulators of the physical model do not suffer from the same "tail extrapolation" difficulty as emulators of probability distributions. That is, we can (and do) run TITAN2D at volumes larger than those in the historical record. As far as Eq. (1) are concerned, these low probability events are in a regime that do not challenge approximations or assumptions used to derive them. Effectively, we are interpolating over physically valid (if unlikely) ranges of initial conditions and parameters to define the boundary of $\mathcal{X}_C$.

We now present an overview of the GaSP construction process. The first step in this construction is to determine an initial set of design points, denoted $\mathcal{D}$, at which to run the simulator. The choice of the initial design is critical in the automation process and will be discussed further in Section 4.2. TITAN2D is run at these inputs producing outputs $\mathbf{y} = \{y^M(\mathbf{x})\}$. Simulations show that the internal friction δ_{int} plays only a small role in the flow runout, so it was averaged out of the input distribution and not considered further. Thus we are effectively left with the fitting of a GaSP 3-input emulator $y^M(\mathbf{x}) = y^M(V, \phi, \delta_{\text{bed}})$ based on the (initial) simulator runs $\mathbf{y}$. [Aside: Because the emulator passes through all the design runs $\mathbf{y}$ (with conditional variance 0 at these points) no notational effort is made to distinguish between the simulator and emulator outputs, and both are denoted by the function $y^M(\cdot)$.] Unless we actually run the simulator at a given input $\mathbf{x}_0$, $y^M(\mathbf{x}_0)$ is effectively unknown at the time of the analysis.

GaSP emulators with power-exponential covariance functions produce relatively smooth response surfaces interpolating the simulators outputs. They are particularly convenient because updating relies on calculations using the conventional multivariate normal. In addition, the specific form of the correlation function (a separable power exponential function) makes extension to higher dimensions easier than with other correlation structures [60, Section 2.3].

However, the introduction of GaSP emulators introduces new—and uncertain—hyperparameters. We generically denote these by $\boldsymbol{\mu}$ (for those hyperparameters entering the mean function), $\boldsymbol{\rho}$ (for those entering the correlation function), and σ^2 (for the variance of the process). That is, before observing the simulator runs $\mathbf{y}$

$$y^M(\mathbf{x}) = y^M(V, \phi, \delta_{\text{bed}}) \sim \text{GaSP}(\boldsymbol{\mu}, \sigma^2, \boldsymbol{\rho}).$$

As mentioned, a distinct advantage of GaSP emulators is the ease in updating. Indeed, for any fixed values of $(\boldsymbol{\mu}, \sigma^2, \boldsymbol{\rho})$, the "updated" emulator (the posterior distribution of $y^M(\mathbf{x})$ given the observed $\mathbf{y}$) is another GaSP with hyperparameters updated by the well-known Kriging equations [64, Section 1.5].

Determining the GaSP hyperparameters using a fully Bayesian fitting is a difficult computational challenge. In contrast, the traditional method of fitting GaSPs (sometimes called *plug-in* GaSP) simply replaces the unknown hyperparameters $(\boldsymbol{\mu}, \sigma^2, \boldsymbol{\rho})$ by their maximum likelihood estimates (MLEs) based on $\mathbf{y}$ ([59] and references therein). We take a middle way, using a modified Bayesian approach that is of approximately the same level of computational difficulty as the traditional plug-in method, but which has a number of important improvements [49, 59]:

1. The correlation function in the GaSP has been generalized to ensure that the emulated height surface is a periodic function of initiation angle ϕ.

2. The uncertainty in the important parameters $(\boldsymbol{\mu}, \sigma^2)$ is ignored when using plug-in estimates, resulting in an underestimate of the total uncertainty in the GaSP. Instead we treat these in a full (objective) Bayesian way and integrate them out from the predictive posterior—which then becomes a Student t process instead of a Gaussian process.

3. Treating the hyperparameters in the correlation structure, $\boldsymbol{\rho}$, as fixed results in considerable simplifications in the computation. In addition, uncertainty in $\boldsymbol{\rho}$ seems to influence the resulting emulator very little. This finding is due to the typically large confounding that is present between $\boldsymbol{\rho}$ and σ. Hence, we deal with $\boldsymbol{\rho}$ in a plug-in

manner. However, we have often found difficulties when using MLEs and we do not recommend their use for fitting GaSPs. For instance, if a likelihood surface is highly irregular, an optimization routine may find values of ρ corresponding to local maxima quite far from from the global maximum. Using an incorrect value of ρ can be much more detrimental to fitting an emulator than ignoring the uncertainty in an accurate value [48]. Use of the mode of objective Bayes posteriors for ρ are slightly more complicated to compute, but result in a numerically more tractable optimization problem (see [65]).

Adaptation of the design space $\mathcal{D}$ is sometimes needed to fit the GaSP to simulator runs that contain many 0's (i.e., no flow passed over the specified location), and to have an accurate emulator near Ψ. Details of the GaSP construction and adaptive approach can be found in [59] and [49].

The emulator constructed in [59] was not entirely satisfactory, for a number of related reasons. As mentioned earlier, the basal friction angle δ_{bed} was found to be extremely important in determining catastrophic inundation at most locations of interest. The large uncertainty about δ_{bed} (see Section 3.3.1) coupled with the lack of physical understanding of how to better characterize the energy dissipation of large volume flows can lead to significant error. The dependence of the effective δ_{bed} on the initial volume, modeled as a stochastic process, provides a way of capturing the uncertainty in this parameter.

We now describe how to incorporate the effective friction into our hazard calculation. We account for the volume dependence by defining $\delta_{bed} = g(V)$ for some function g, and reconceiving the emulator as a 2-input emulator $y^M(V, \phi) = y^M(V, \phi, \delta_{bed} = g(V))$. Note that this remains monotonically increasing in V, since the basal friction decreases as V increases and decreasing friction will create longer flow runouts. To take into account the uncertainty in g, it suffices to generate g's from the posterior distribution (see Section 3.3.1 and illustrated in Fig. 3). This produces a sample of GaSP predictors $y^M(V, \phi)$ (one GaSP per posterior draw g) for the simulator at untried inputs (V, ϕ), capturing the uncertainty both in g and in the emulator [49]. As discussed in Section 3.3.1, the posterior for g is fitted with data not only from SVH but also from other similar volcanoes via an objective Bayes hierarchical analysis. More detail on these ideas can be found in [48].

Thus what remains is a 2-inputs emulator $y^M(V, \phi)$. Our goal remains to find the critical contour Ψ which characterizes the set of inputs leading to catastrophic events. This contour is most conveniently represented by finding, for each angle $\phi \in [0, 2\pi)$, the minimum volume V that causes catastrophic damage. Thus $\Psi = \Psi(\phi) = \inf\{V : y^M(V, \phi) \geq 1 \text{ m}\}$, and the critical region $\mathcal{X}_C$ is the set of values (V, ϕ) above that contour.

The Belham Valley runs west from the flank of the SHV, and borders a still populated region of Montserrat. When those monitoring the SHV believe the Belham might be inundated, public safety warnings are issued. Therefore there is interest in determining the hazard threat at several locations in the Valley. For one selected location, Fig. 5 shows the mean surface of the emulated maximum height GaSP as a function of the initial volume and initiation angle, evaluated at the posterior median of the distribution of g. The black dots identify $\boldsymbol{y}$, the actual TITAN2D outputs at the design points in $\mathcal{D}$. The red curve is an estimate of the critical contour Ψ.

Hazard assessments are now given by the probabilities of the critical region $\mathcal{X}_C = \{(V, \phi) : V \geq \Psi(\phi)\}$. One can then use the input distribution to compute the probability of this set for each given function g; when g is sampled from its posterior distribution, one obtains an ensemble of probabilities providing the required quantification as well as measures of error.

4.2 Automated Design and Emulation

We now discuss the approach we have used to automate the adaptive design process used to construct hazard maps. In automating the adaptive design, there are two competing aspects that need to be balanced:

1. The spatial output of the simulator varies dramatically as inputs vary. This makes intuitive sense—it requires a much larger volume flow to inundate a site further down the valley from the volcano than it would a site up valley, closer to the cone. Likewise, a more mobile flow exercised at exactly the same volume and direction as a less mobile flow could result in a single site getting inundated versus that same site seeing no flow at all. Effectively, the "important" simulator runs for a specific site—those that result in flow height near the threshold
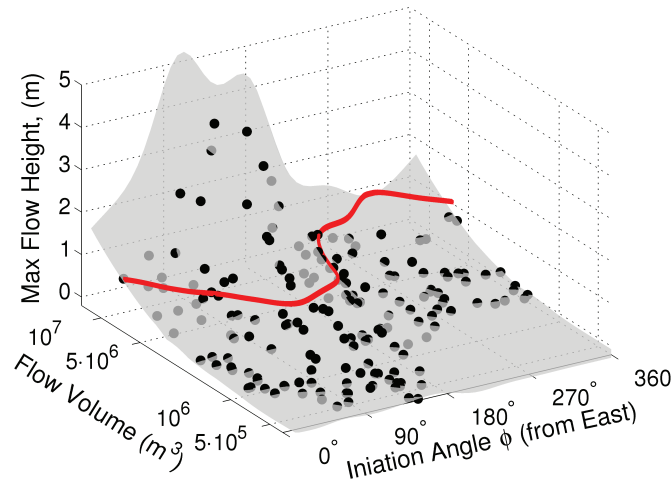
FIG. 5: Max-height surface for a selected location in the Belham Valley. The surface is the mean of the GaSP emulators evaluated at the median volume-basal friction curve. Dots represent the max-height simulation output at design points. The red curve represents the catastrophic contour, $\Psi(\phi)$. (A similar figure appeared in [49].)

output at that site—are confined to a small region of input space. Thus finding a specific site's catastrophic contour requires an emulator fit to a design that is *local* in input space.

2. A probabilistic hazard map that covers a significant area will, by default, include sites with localized designs that, taken together, span a very broad swath of input space. At the same time, we want to ensure that our automated approach accounts for physically plausible yet rare scenarios. For example, we perform some runs of the simulator at volumes larger than any recorded. Specifically we seek a design that samples the support of tail events for a variety of aleatoric models we might consider. To account for these needs the design must be *global*.

So how are we to proceed? We begin by exercising TITAN2D at design points in a large region of the four-dimensional space $V, \phi, \delta_{bed}, \delta_{int}$, given as $\mathcal{X} = [10^5 m^3, 10^{9.5} m^3] \times [0, 2\pi] \times [6°, 18°] \times [18°, 35°]$ of the input space. A set of 2048 points are chosen, according to a Latin hypercube design [60, 61, 66, 67]. The output—the maximum flow height for all the locations of interest—for each run is saved. *Once these 2048 runs have been completed, no further TITAN2D runs are necessary in the subsequent hazard analyses.*

The strategy is to choose site specific *subdesigns* from the original 2048 runs and fit *subemulators* to each. To do so, cull from the full design points that (i) are not likely to be physically relevant based on model-data calibration, and (ii) result in no flow or "too much" flow (e.g., those which exceed, say, twice the critical height) at the site in question. This culling results in a relatively small (roughly size $N/10$), but nonstandard subdesign (not space filling on a hyperrectangle, owing to many excluded zero-height runs). On the plus side, the small size of the resulting subdesign leads to easy matrix inversion when fitting the GaSP. On the negative side, since we have not included any design points that result in no flow reaching the site of interest, the GaSP is not constrained to go to zero where it ought. To remedy this shortcoming, we re-introduce a small number of zero-height runs, choosing those that are closest to subdesign points in input space (measured via a normalization of the distance in $\mathcal{X}$). Using this strategy, fitting subemulators drawn from the original 2048 TITAN2D runs at as many sites as desired, to fill in a map (on the order of thousands of sites typically), and finding their corresponding catastrophic contours, is trivially parallelizable. If, at any site, the error bounds for the catastrophic contour are too large, additional points may be added to the subdesign, choosing from the original 2048 design points those close to the site-specific Ψ contour. Details of the strategy outlined here can be found in [49]. That paper also describes a strategy for monitoring the quality of the automated subemulators.

Figure 5 shows a 2D (volume and initiation angle) surface obtained by evaluating a 3D subemulator at the median basal friction–volume curve. From the same subemulator, Fig. 6 highlights the selected subdesign points and the resulting catastrophic contour.

4.3 The Initiation Angle

The simplest assumption to make in describing the initiation angle is that the angle is uniformly distributed, and independent of event volumes. There is no physical reason to suspect otherwise, and this is the assumption made in [59].

However, careful examination of the data suggests this assumption might not be correct, perhaps because, at Montserrat, there are only a handful of valleys down which all flows ultimately travel. Consider a polar representation of the data. For each initiation angle ϕ plot the volume of a flow (on a log scale) event as the radius. The resulting Fig. 7 clearly suggests that the angle distribution is not uniform. The standard statistical model for a nonuniform angle is the von Mises distribution [68, 69] with pdf

$$f(\phi \mid \kappa, \vartheta) = \frac{e^{\kappa \cos(\phi - \vartheta)}}{2\pi I_0(\kappa)},$$

centered at an angle ϑ (which here will be close to 0) and with a concentration parameter, κ, controlling how much of the mass is located near ϑ. The resulting joint density for V and ϕ is

$$V, \phi \sim \alpha \epsilon^\alpha V^{-\alpha-1} \mathbf{1}_{V>\epsilon} f(\phi \mid \kappa, \vartheta); \tag{2}$$

it is also possible that κ could depend on V. Detailed hazard estimates exploiting this relationship will be reported elsewhere.

4.4 The Probability Calculation

As mentioned, SHV is one of the most well-studied volcanoes on earth. Among the data is an extensive collection of (virtually) every mass flow from the volcano since the eruption phase began, together with an estimate of the volume of the flow and the valley(s) in which it traveled. From these we determine a volume-frequency relationship for mass flows in Section 4.4.1—that is, we quantify approximately how often flows of specified sizes might be expected to occur.
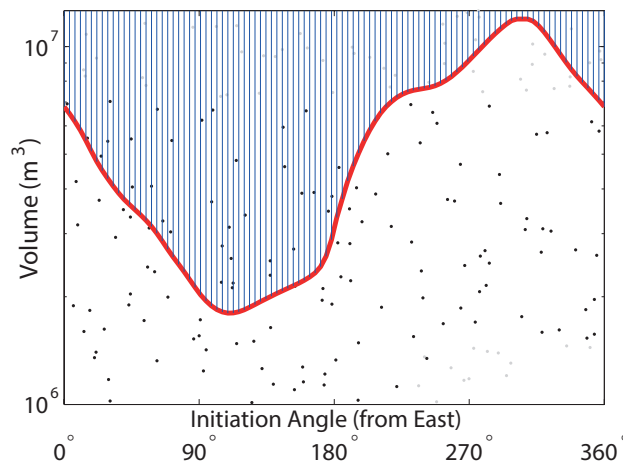

FIG. 6: Catastrophic contour separates design space into a region which leads to inundation (shaded blue), $V \geq \Psi_{\hat{M}}(\theta)$, from initial conditions which do not lead to inundation, $V < \Psi_{\hat{M}}(\theta)$. Dots represent design points where TITAN2D was exercised (that is, some of the original 2048 runs), and the darker dots were those selected for the subdesign through the automation process. (A similar figure appeared in [49].)
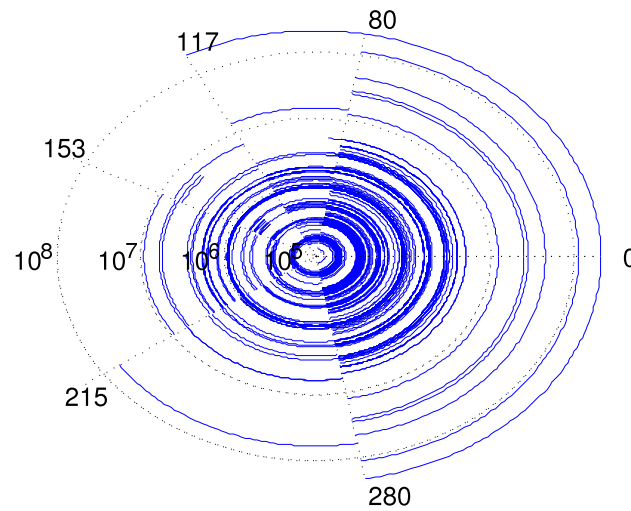
FIG. 7: Initiation angle dependence at Soufrière Hills Volcano. Each curve corresponds to an event where the angle sector represents the valley the flow went down (in degrees from East) and the radius represents the (log of the) volume of the flow event (in m^3).

4.4.1 Frequency and Severity of Extreme Events

A key ingredient in the probabilistic model we wish to construct is a suitable input distribution for the frequency of pyroclastic flows of a given magnitude. Figure 8 shows an empirical plot of the number of pyroclastic flows of volume $V_j \geq v$ vs. v during the period from March, 1996 through July, 2008 for the Soufrière Hills Volcano, on a log-log scale, for volumes larger than some minimum $v \geq \epsilon$ (here $\epsilon = 5 \times 10^4$ m^3). Three principal sources of information were used to assemble this inventory: (1) the record of daily field observations made by the Montserrat Volcano Observatory scientists and reported in daily reports or weekly reports [70], supplemented with (2) published papers describing specific events at SHV, and (3) internal open-file reports of the Observatory. The vertical banding at smaller volumes is an artifact, suggesting some binning of smaller reported flows.

The approximately linear fall-off on the log scale suggests that the probability distribution of flow volumes satisfies the linear equation

$$\log P[V \geq v \mid V \geq \epsilon] \approx -\alpha \log(v) + c \tag{3}$$

for some constants $\alpha > 0$ and $c \in \mathbb{R}$, and hence the distribution of the $\{V_j\}$ is approximately Pareto, with

$$P[V \geq v] \approx (v/\epsilon)^{-\alpha}, \qquad v \geq \epsilon. \tag{4}$$

This is just what one would expect for a peaks-over-threshold model in the (heavy-tailed) domain of the Fréchet extreme value distribution [71, Section 4.2]. Estimating the value of α can be challenging. The analysis in [59] gives an estimate of $\alpha \approx 0.64$. Because $\alpha < 1$, the flow volumes do not have finite means nor finite variance—over time one should expect the average of the observed flow volumes to continue increasing. The total accumulated flow over any fixed time interval will have approximately an α-stable distribution, not a normal distribution.

To give some sense of the flows at the Soufrière Hills Volcano, during the 1996–2008 period there were eruptions generating

- flows exceeding 10^4 m^3 daily, with a 0–1.6 km runout

- flows exceeding 10^5 m^3 fortnightly, with a 2–3 km runout

- flows exceeding 10^6 m^3 quarterly, with a 4–6 km runout

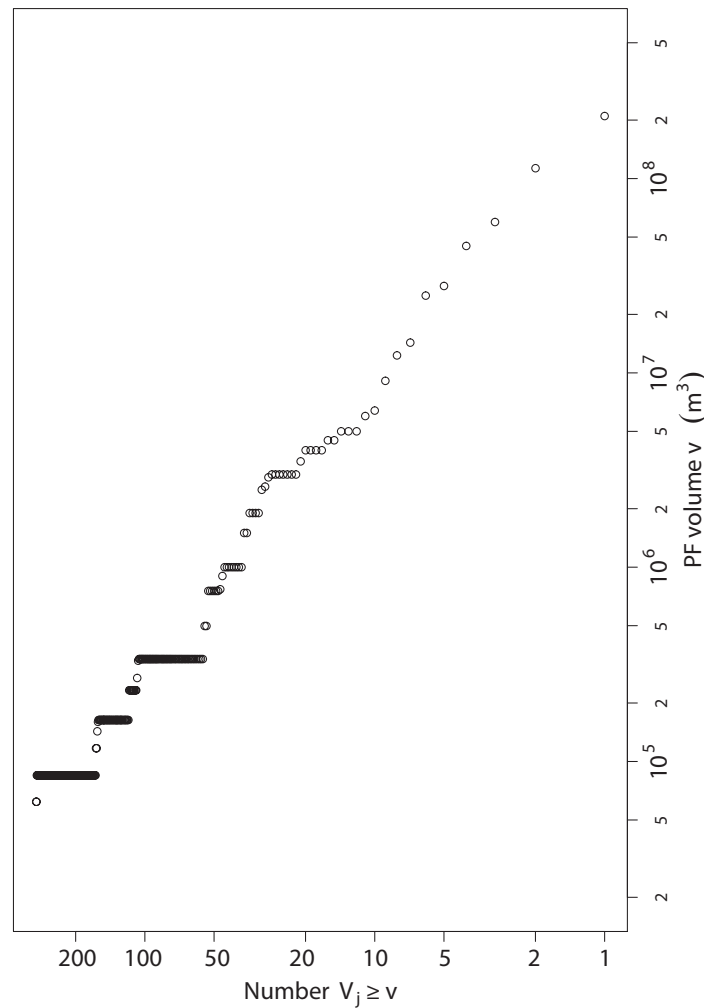- flows exceeding 10^7 m^3 yearly, with a 6+ km runout

FIG. 8: Frequency-vs.-magnitude plot for pyroclastic flows at Soufrière Hills Volcano for the period 1996–2008. Each marker represents a specified number of flows of a given volume. At small flow sizes, several observations are lumped at approximately the same size, and the binning is evident. The approximately linear fall-off suggests a Pareto distribution.

- flows exceeding 10^8 m^3 twice, of unknown runout (flows went into the ocean)

These frequencies and runouts are, of course, only approximately true, but this list provides some intuition about the eruption history.

The "fat tails" of a Pareto distribution with an exponent $\alpha < 1$ suggest that there is a significant chance of observing a flow considerably larger than any that have yet been observed, making naive hazard estimates based on recent experience overly optimistic. But this model can only be an approximation to reality—for example, it also predicts a non-negligible chance of unphysically large events, such as a $10^{9.5}$ m^3 flow, which is so large that it could cover the entire bottom half of the island in a deposit 1 m deep. Clearly the volume-frequency distribution must fall off more rapidly than the Pareto for sufficiently large volumes. However, the data are too sparse to provide good estimates for the rate of such a fall-off. In spite of this shortcoming, the model fits well in the range of 10^6–10^8 m^3 that is important for our hazard analysis. That is, errors in the probability calculations owing to the tail being incorrectly

modeled for very large flows are very small. As a consequence, a more complex model is not needed, at least not at this time.

4.4.2 Switch Points in the Distribution

Our approximation of flow frequency assumes a flow frequency that is stationary in time. As the lava dome at SHV grows, however, avalanches occur more frequently and are usually larger. Following a dome collapse event, avalanche frequency diminishes for a time. Dome growth itself is a surrogate for lava extrusion, and an increase in lava being pumped out of the conduit suggests more vigorous activity.

A careful examination of the data shows five periods when the SHV dome has grown measurably, separated by periods of quiescence with little or no dome growth. That is, on short time scales (i.e., on scales of several months) SHV may be inactive, with few or no flow events. A triggering event then turns the internal geologic mechanisms "on," and the volcano enters another active phase lasting for several months or a few years.

Modeling these switch points in SHV activity can predict the chances for the volcano turning on and off, and can be included in a more detailed analysis of flow frequency, and consequent hazard analysis; this is likely not important for long-term hazard prediction but would clearly be crucial for short-term prediction.

4.5 The Hazard Calculation

The final step in computing hazard risk is to combine all the ingredients of the probability calculation. For a given model of scenario variability, $\pi(V, \phi)$, the key to efficient hazard calculation is that $\Psi(\phi)$ replaces the indicator function in probability calculations. That is,

$$\mathrm{P}(\text{catastrophe} \mid \text{PF occurs}, \Psi) = \int_{[0,\infty)\times[0,2\pi)} \mathbf{1}_{y^M(V,\phi)\in\mathcal{Y}_C}\,\pi(V,\phi)dV\,d\phi = \int_{[\Psi(\phi),\infty)\times[0,2\pi)} \pi(V,\phi)dV\,d\phi.$$

The integral over V can be done in closed form, with the integral over ϕ being done numerically or by Monte Carlo. Making the reasonable further assumption that pyroclastic flows occur as a Poisson process with with rate λ/year [59], it follows that the probability of hazard over T years is

$$\mathrm{P}(\text{catastrophe within } T \text{ years} \mid \Psi, \lambda, \alpha) = 1 - \exp\left(-T\lambda\,\mathrm{P}(\text{catastrophe} \mid \text{PF occurs})\right),$$

where we also indicate the dependence of this probability on the Pareto parameter α.

To summarize the findings we have discussed, the data suggest an aleatory model in which:

1. The flow volumes are iid from a Pareto distribution $\mathrm{Pa}(\alpha, \epsilon)$ for an exponent $\alpha \approx 0.64$ and a minimum flow ϵ taken to be 5×10^4 m^3;

2. The initiation angles likely have a von Mises distribution, but a uniform angular distribution results in a simple illustrative calculation;

3. The event arrival times follow a stationary Poisson process with a rate $\lambda \approx 22$/yr, although on short time scales nonstationarity is evident.

How should uncertainties be incorporated? Some of the aleatoric uncertainties are captured through the distribution of inputs, $\pi(V, \phi)$, in our illustration. If new data or expert opinion are considered, or if models are refined, one can quickly calculate hazard probabilities associated with different aleatory variability models [e.g., different $\pi(V, \phi)$]) without the need for additional expensive computer simulations. This allows one to systematically compare the probability of hazard under different models of variability. Likewise, other sources of variability such as the probability of the volcano "switching off," could be included in a similar fashion.

In order to account for epistemic uncertainty, we suggest a Bayesian approach. To incorporate epistemic uncertainty from the volume-basal friction relationship, we can find a catastrophic contour corresponding to each of

the posterior samples in Fig. 3 (instead of just the median). This strategy produces M inundation contours, $\Psi_j(\phi)$, $j = 1, \ldots, M$. The spread of these contours reflects uncertainty in the modeled relationship and the resulting contours are plotted in Fig. 9. The inundation probability would then become

$$\text{P(catastrophe within } T \text{ years} \mid \lambda, \alpha) \approx \frac{1}{M} \sum_{j=1}^{M} \text{P(catastrophe within } T \text{ years} \mid \Psi_j, \lambda, \alpha) \, .$$

Finally, we need to account for the epistemic uncertainty in the parameters λ and α. Instead of using these plug-in estimates for the parameters, we incorporate their epistemic uncertainty by objective Bayesian analysis, obtaining the posterior distribution $\pi(\alpha, \lambda \mid \text{data})$ for these unknown parameters [59]. The final hazard probability calculation then becomes

$$\text{P(catastrophe within } T \text{ years)} = \int \text{P(catastrophe within } T \text{ years} \mid \lambda, \alpha) \pi(\alpha, \lambda \mid \text{data}) d\alpha d\lambda \, ,$$

which can easily be computed by numerical integration and importance sampling.

The hazard maps in Fig. 10 were generated with the emulator-based hazard computations presented in this section (run in parallel for various locations on the map). Note, once the subemulators are constructed, producing these hazard maps takes roughly 5 min on a cluster. Figure 10 illuminates the power of this approach. That is, we can easily and quickly compare the effects that the choice of scenario model has on the inundation probability. Furthermore, we can do this while simultaneously accounting for potential effects of epistemic uncertainty on these calculations.

5. ALTERNATIVE METHODS

Depending on the goals of a particular analysis, other methods may shed light on the challenges of uncertainty. Here we briefly discuss two other approaches, and their virtues and shortcomings. We report on implementation of these ideas in practical settings.

5.1 PC and PCQ

Polynomial chaos, or PC, is a class of methods designed to propagate uncertainty from inputs and parameters to outputs. PC originated with Wiener [72], and was generalized by Xiu and Karniadakis, and Ghanem and Spanos
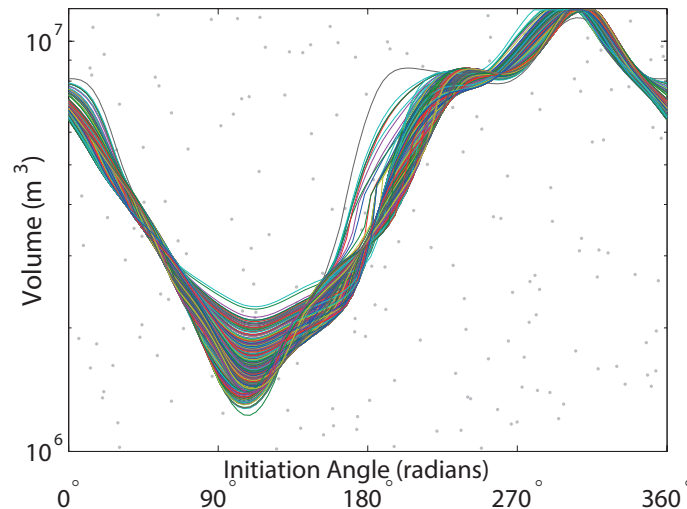

FIG. 9: Each contour above is analogous to the one in Fig. 6, but where the 3D emulator is evaluated at a volume–friction curve sampled from Fig. 3 and reduced to a 2D surface (flow height vs. volume and initiation angle). From each resulting surface, a catastrophic contour $\Psi_j(\theta)$ is plotted above. (A similar figure appeared in [49].)
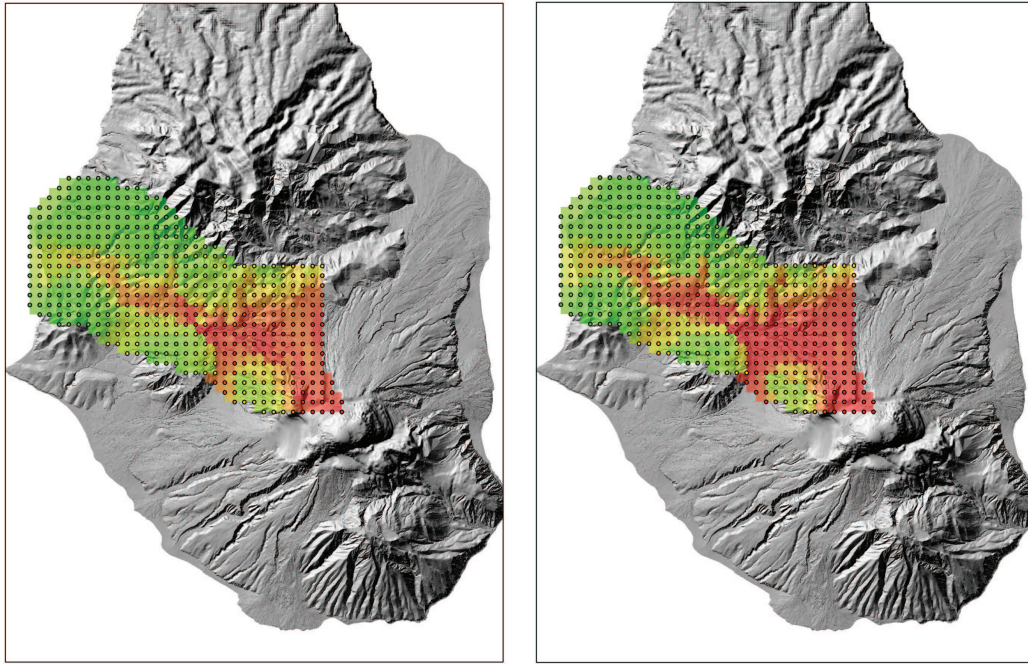
FIG. 10: Two probabilistic hazard maps for the Belham Valley on the island of Montserrat. The color map, from red to green, represents probabilities from 1 to 0 of inundation from a pyroclastic flow over the next 2.5 years (under the assumptions above which let to choices of Pareto and Poisson models for volume and frequency respectively). Probabilities were calculated at the map points marked with circles. The two sets of calculations use they same set of 2048 TITAN-2D runs, with $O(10^5)$ Monte Carlo samples for each map point. The difference between the two plots is the choice of the scenario variability model for the initiation angle: uniform distribution (left) and point mass distribution to the north/northwest (right). (A similar figure appeared in *SIAM News*, vol. 45, 2012.)

[33, 34, 73, 74]. Roughly speaking, PC separates random variables (labeled ξ) from deterministic ones (such as space z and time t), yielding a relation

$$u(z, t, \xi) = \sum_{i=1}^{\infty} u_i(z, t)\psi_i(\xi).$$

Here the ψ_i are known basis functions (often chosen as an orthonormal basis) associated with the assumed distribution of u (e.g., Hermite polynomials if the distribution is taken to be normal).

Incorporating any spatial differentiation into the right-hand side, a system of differential equations may be written

$$\frac{\partial u}{\partial t} = r(u; \beta)$$

where u is a vector of N_{eqns} equations, and β is a vector of uncertain parameters. Expanding, we have $u = \sum u_i\psi_i$ and $\beta = \sum \beta_i\psi_i$. Substituting back into the differential equations

$$\frac{\partial u_i\psi_i}{\partial t} = r(u_i\psi_i; \beta_i\psi_i).$$

In the original development of PC, a Galerkin projection was performed to arrive at a system of equations for the u_i's

$$\frac{\partial u_m}{\partial t} = \frac{\langle r(u_j\psi_j; \beta_k\psi_k)\psi_m \rangle}{\langle \psi_m^2 \rangle}.$$

Here we use the angle brackets to denote averaging over the random variables.

The PC equations for the u_m's are usually highly nonlinear, and numerical codes to integrate these equations are not simply related to codes for the original system of differential equations. In addition, the Galerkin projection is impossible to perform directly, if the nonlinearities in r are nonpolynomial. Of course a Taylor expansion can be performed, at the expense of additional error.

A different approach is provided by stochastic collocation and nonintrusive stochastic projection methods [75] dubbed here polynomial chaos quadrature (PCQ). Roughly speaking, the Galerkin projection is replaced by a quadrature, accurate up to integration error:

$$\frac{\partial u_m}{\partial t} = \frac{\sum_q w_q r(u_j \psi_j(\xi_q); \beta_k \psi_k(\xi_q)) \psi_m(\xi_q)}{\sum_q w_q \psi_m^2(\xi_q)}.$$

Integrating from time 0 to T yields

$$u_m(T) = \frac{1}{\sum_q w_q \psi_m^2(\xi_q)} \sum_q w_q \psi_m(\xi_q) \left(u(0) + \int_0^T r(u_j \psi_j(\xi_q); \beta_k \psi_k(\xi_q)) dt \right).$$

Denote the integral by $R(u(0); \alpha) \equiv u(0) + \int_0^T r(u_j \psi_j(\xi_q); \alpha_k \psi_k(\xi_q)) dt$. Now combining all the u_m's and taking the Nth moment, we have

$$
\begin{aligned}
\langle u(T)^N \rangle &= \langle (u_m(T) \psi_m(\xi))^N \rangle \\
&= \left\langle \left(\frac{\sum_q w_q \psi_m(\xi_q) R(u(0, \xi_q); \beta(xi_q))}{\sum_q w_q \psi_m(\xi_q)} \psi_m(\xi) \right)^N \right\rangle \\
&= \langle R(u(0, \xi); \beta(\xi))^N \rangle \\
&= \sum_q w_q R\big(u(0, \xi_q); \beta(\xi_q)\big)^N.
\end{aligned}
$$

This formula suggest that PCQ can be viewed as a "smart" way to perform sampling, not unlike a Monte Carlo approach, but one with better convergence properties. Specifically, if the sample points ξ_q are chosen as quadrature points for the polynomial basis functions (e.g., Gauss or Clenshaw-Curtis quadrature points), high integration accuracy can be obtained.

In hazard mapping applications there appear two principal drawbacks to using PC and PCQ methods (see [76]): (i) when the number of uncertain variables gets large, PCQ calculations suffer from the "curse of dimensionality," and the number of random variables grows exponentially; (ii) PCQ is, ultimately, a spectral method and cannot capture higher frequency behavior (in realistic time), nor easily sample from tail events from any probability distribution. The difficulty in sampling was established in an *ad hoc* fashion [76] by restricting the volume inputs to only very large flows. Such a restriction is difficult to do in a systematic fashion and the predictive power of the methodology is entirely dependent on the choice of that cutoff. Figure 11 shows a PCQ-based analysis from [76] illustrating the impact of this choice, in a hazard calculation at Colima Volcano in Mexico. A realistic choice of flow volume ranges of $>10^4$ m^3 yields a hazard analysis declaring all areas safe! A more restrictive choice of volume to $10^7 - 10^8$ m^3 provides a more reasonable hazard map.

5.2 Bayes Linear Models

A different approach to emulator construction can be found in [77], where a Bayes linear method (BLM) [78] is used in conjunction with a localization procedure (see also [79]). The BLM is based on an updating of a prior distribution (called beliefs) in light of new data. Given the prior and new data, the mean and variance of the distribution are adjusted by minimizing the L^2 difference between the beliefs and the data. That is, BLM projects beliefs onto the span of the data. Because only means and variances are determined, the BLM is exact only for Gaussian distributions.
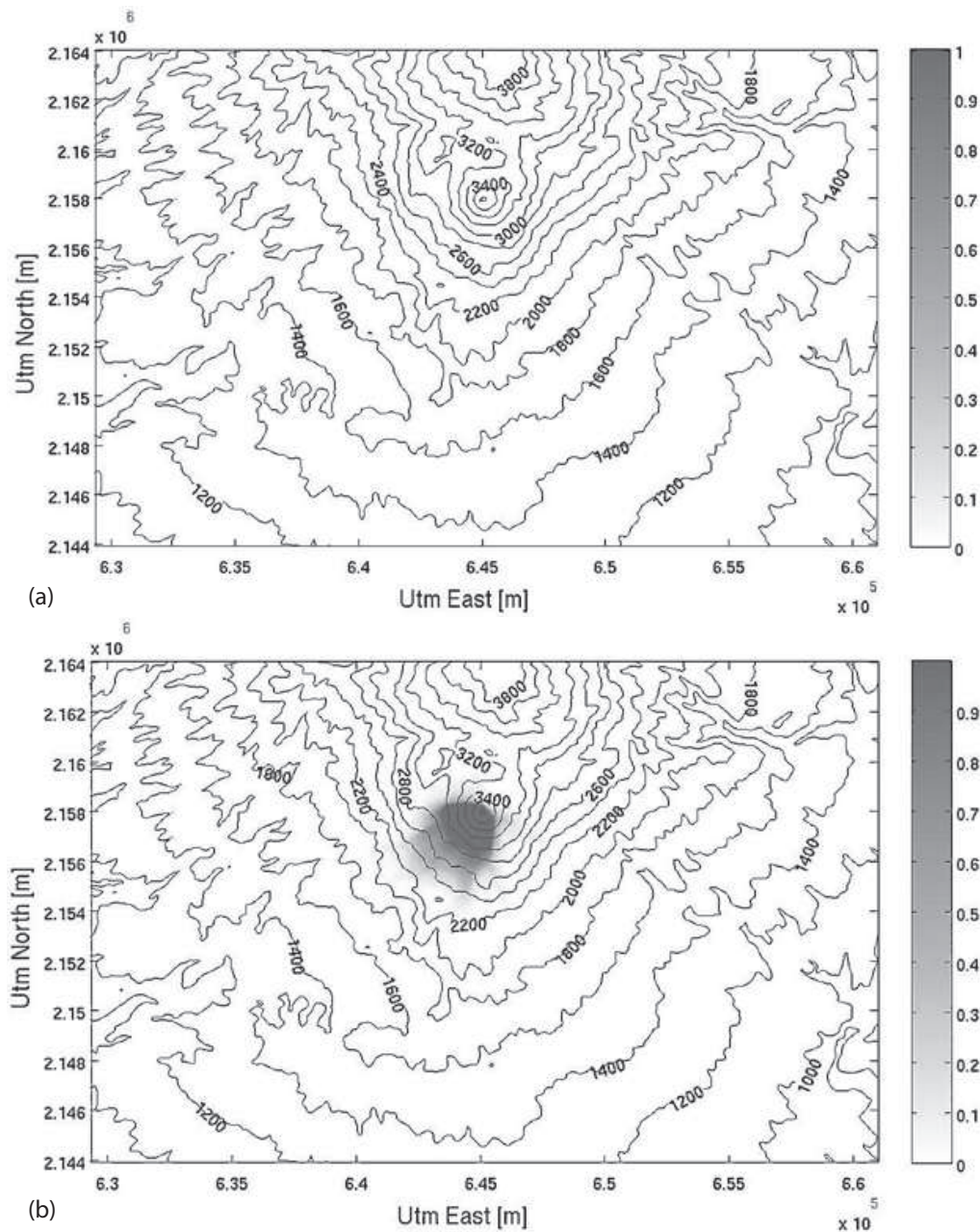
FIG. 11: Probability of maximum flow depth at Colima Volcano being greater than 1m given that a flow occurs, and initial volumes exponentially distributed (a) greater than 10^4 m^3 and (b) from 10^7 m^3 to 10^8 m^3. Bed friction is uniformly distributed between 15^0 and 35^0. Contours indicate elevations, and shading indicates flow probabilities.

In the BLM updating of mean and variance, the covariance matrix must be inverted (as in constructing a GaSP emulator). When there are many variables, this inversion can become computationally expensive. To avoid this difficulty, [77] introduced a localization based on a measure of distance in parameter space. This localization breaks a large problem into many smaller ones, allows use of parallel computing and efficient computation workflows [80]. Related ideas can be seen in the treed Gaussian approaches of Lee and co-workers [81].

6. CONDITIONAL HAZARD MAPS

In Section 4 we computed the probability of a catastrophic flow occurring at a specific location, over the next several years, using an automated construction of subemulators that allows risk-at-a-point calculations in parallel for large numbers of locations. This calculation is applicable in several kinds of studies. For example, it may be useful to find the hazard probability at specific "trigger points," places where, if a flow were to reach there, civil authorities would take action about evacuation.

One might expect hazardous events at some locations to be correlated with those at certain other locations—for example, hazards at different locations in the same river valley may well be correlated. Consider then a different kind of correlated risk that arises naturally in evacuation planning. If the conditional probability of a catastrophe in area B, given a catastrophic event at a region A is large—that is, $P(B|A) > p_{\min}$—then civil authorities should try to avoid using B as an evacuation route or destination for residents of A.

The calculation of this conditional risk at B is not much harder than the calculation of the risk at A described in Section 4, *if* the correlation structure is well understood. However, this correlation is not simple to specify. For flat terrain, a simple distance between A and B could likely capture much of the correlation; for more realistic terrains the correlation will depend critically on the topography—locations A and B separated by a specified distance will be more correlated if they lie in the same valley than if they are separated by a ridge.

Even with an accepted correlation structure, computing conditional hazard maps entails calculations among thousands of spatial points. The computational burdens, in particular the matrix arithmetic, then becomes taxing, and difficult to perform in parallel. Localization procedures as in [82, 83] may be useful in computations. The geographic nature of the correlation structure suggests that it may be useful to *approximate* the covariance structure by one in which distinct valleys are treated as independent, leading to a domain decomposition approach for identifying block structure in the correlations, potentially alleviating the linear algebra roadblock.

7. COMPUTATIONAL APPROACHES

We review here hardware and software approaches designed to surmount computational issues arising from the methodology described above, allowing the rapid calculation of a hazard map. In computing this ensemble, problems involved in data extraction, movement of data over the network, and data replication, all must be addressed. Any computational methodology that may be employed should be flexible enough to improve resolution in the calculated hazard maps, using, for example, a localization to allow access to more data during the inference process (a tessellation based localization may be found in [84], and a tree-based localization in [80]). In our work, we have used a combination of specialized data intensive computing hardware (the IBM Netezza platform) and traditional supercomputing, joined in a workflow designed to map operations to platforms, to maximize net throughput.

Let us explain. For very large data sets, the most efficient procedure moves the computational platform close to where the data reside. This approach is very different from traditional scientific computing, wherein data are moved from a storage area network to the compute engine. As mentioned, data extraction, data movement, and data replication constitute perhaps the most time-consuming phase of many large data analysis problems [85]. In-database computing using hardware like the Netezza can locate certain of the processing activities onto the database management system, eliminating the need for data extraction and replication, greatly improving scalability and performance of certain analytic computations. In particular, the Map-Reduce model is being extensively used in applications requiring analysis of large volumes of data. The Hadoop API based on a Map-Reduce model provides a high level of parallelism, and offers a level of abstraction that significantly simplifies programming. Alternative program interfaces overcome other obstacles that may arise when running map-reduce in a local Hadoop implementation. Map-reduce using MPI (MR-MPI) created by Sandia National Laboratories, is one alternative implementation of map-reduce using the Message Passing Interface [86]. Unlike Hadoop, MR-MPI does not have its own distributed file system, and works with mainly C/C++, although it also has Python bindings together with OINK wrappers. As an example of this approach, in a recent exercise of this workflow a hazard map was generated using a Netezza server and 504 processors on a high-performance cluster in about 12 h, during which more than 1.5 TB of data were transferred over the network to Netezza server.

8. TRANSFERRING THE METHODOLOGY

The hazard risk calculations outlined here have made use of the geological data collected over the past 15 years at SHV. Several significant findings from SHV may be transferrable to other volcanoes, especially other lava dome volcanoes. For example, given basic information about the size of the volcano crater and dome growth rates, data that are available from ground-based measurements, one may be able to infer a likely flow volume-frequency relationship. Such a rate could be partially validated by any historical data that might be available. The hierarchical effective friction model relies on data obtained from volcanic sites worldwide, so is applicable universally. Thus, much of the hazard calculation could be performed at other sites, even if eruption data are not readily available (although uncertainty in the posterior might be larger than at Soufrière Hills Volcano).

A different kind of prediction is to compute the chance that volcanic ash (of a specified concentration) might be present in the atmosphere above a given location. A volcanic ash plume moves via advection and diffusion, and particles fall out at a certain rate [1, 87]. The uncertainty in the advecting windspeed, diffusion parameter, and the initial data can be approximated in a fashion similar to our friction and initial data uncertainties, and a similar calculation performed.

Modeling floods resulting from storm surges onto land require DEMs and estimates of the initial flowing mass similar to volcanic mass flows, and share the associated uncertainties. If the rate and intensity of triggering storms could be determined, a calculation very similar to our risk-at-a-point algorithm may provide quantifiable hazard risk estimates for coastline cities and towns.

9. DISCUSSION

We have suggested a methodology for computing hazard risk associated with the geophysical mass flows consequent to volcanic eruptions. Although this application is specific to a particular natural hazard, the methodology is directly applicable to ash plumes and floods, and may be adapted to other natural hazards.

We call attention to the special features of the mass flow problem which we exploit in order to make our analysis tractable. First and most importantly, the approach we have taken allows us to separate the stochastic and deterministic parts of the problem. Deterministic computer models were used to identify the region of catastrophic inputs. The determination of this region is an inverse problem—what inputs cause a catastrophic output. We have circumvented the difficulties usually associated with the computation of an inverse problem through use of a statistical procedure allowing us to compute the posterior distribution P(catastrophe | data). Because of the special character of the geological problem being studied, we were able to restrict effort to a region around the boundary curve $\Psi(\phi)$, a one-dimensional curve in two-dimensional parameter (V, ϕ) space. That is, we exchanged a stochastic inverse problem for a forward calculation of a deterministic model, with an adaptive design of inputs to speed up computations.

In computing a hazard map, we computed the probability of "catastrophe" for a representative set of locations, avoiding a large matrix inversion and instead computing a small matrix problem at each location independently, an easily parallelized operation.

More generally, our progress in volcanic hazard prediction rests on a combination of statistical and mathematical expertise coupled with a deep understanding of geological science, a paradigm for many problems in applied mathematics.

ACKNOWLEDGMENTS

REFERENCES

1. Bursik, M., Jones, M., Carn, S., Dean, K., Patra, A., Pavolonis, M., Pitman, E., Singh, T., Singla, P., Webley, P., Bjornsson, H., and Ripepe, M., Estimation and propagation of volcanic source parameter uncertainty in an ash transport and dispersal model: Application to the Eyjafjallajokull plume, *Bull. Volcanology*, 74(10):2321–2338, 2012.

2. Patra, A. K., Bursik, M., Dehn, J., Jones, M. D., Madankan, R., Morton, D. J., Pavolonis, M., Pitman, E. B., Pouget, S., Singh, T., Singla, P., Stefanescu, E. R., and Webley, P. W., Challenges in developing dddas based methodology for volcanic ash hazard analysis—Effect of numerical weather prediction variability and parameter estimation, In *ICCS*, pp. 1871–1880, 2013.

3. Freer, J., Beven, K., Neal, J., Schumann, G., Hall, J., and Bates, P., Flood risk and uncertainty, In *Risk and Uncertainty Assessment for Natural Hazards*, Rougier, J., Sparks, R. S. J., and Hill, L. (eds.), Cambridge University Press, Cambridge, pp. 190–233, 2013.

4. Furlan, C., Extreme value methods for modelling historical series of large volcanic magnitudes, *Stat. Model.*, 10(2):113–132, 2010.

5. Sparks, R. S. J., Aspinall, W. P., Crosweller, H. S., and Hincks, T. K., Risk and uncertainty assessment of volcanic hazards, In *Risk and Uncertainty Assessment for Natural Hazards*, Rougier, J., Sparks, R. S. J., and Hill, L. (eds.), Cambridge University Press, Cambridge, pp. 190–2337, 2013.

6. Aspinall, W. P., Hincks, T. K., Mahony, S., Pooley, J., Nadim, E. S. F., and Sparks, R. S. J., *Volcano Hazard and Exposure in Track II Countries and Risk Mitigation Measures—GFDRR Volcano Risk Study*, 2011.

7. Neri, A., Aspinall, W., Cioni, R., Bertagnini, A., Baxter, P., Zuccaro, G., Andronico, D., Barsotti, S., Cole, P., Ongaro, T., Hincks, T. K., Macedonio, G., Papale, P., Rosi, M., Santacroce, R., and Woo, G., Developing an event tree for probabilistic hazard and risk assessment at Vesuvius, *J. Volcanology Geothermal Res.*, 178:397–415, 2008.

8. National Research Council, The eruption of Nevado del Ruiz Volcano, The Committee on Natural Disasters, Colombia, South America, Nov. 13, 1991.

9. National Research Council, Mount Ranier: Active Cascade Volcano, The U. S. Geodynamics Committee, 1994.

10. Tilling, R. I., *Volcanic Hazards*, Vol. I, American Geophysical Union Short Course, 1989.

11. Vallance, J. W., *Postglacial Lahars and Potential Hazards in the White Salmon River System on the Southwest Flank of Mount Adams, Washington*, U.S. Geological Survey, Reston, VA, Bulletin 2161, 1999.

12. Legros, F., The mobility of long runout landslides, *Eng. Geol.*, 63:301–331, 2002.

13. Wasowski, J., Keefer, D., and Jibson, R. W., Special issue from the symposium on landslide hazards in seismically active regions, *Eng. Geol.*, 58:v–vi, 2000.

14. Voight, B., Sparks, R. S. J., Miller, A. D., Stewart, R. C., Hoblitt, R. P., Clarke, A. B., Ewart, J., Aspinall, W. P., Baptie, B., Calder, E. S., Cole, P. D., Druitt, T. H., Hartford, C., Herd, R. A., Jackson, P., Lejeune, A.-M., Lockhart, A. B., Loughlin, S. C., Luckett, R., Lynch, L., Norton, G. E., Robertson, R., Watson, I. M., Watts, R. B., and Young, S. R., Magma flow instability and cyclic activity at Soufrière Hills Volcano, Montserrat, British West Indies, *Sci.*, 283(5405):1138–1142, 1999.

15. Wadge, G., Macfarlane, D. G., Robertson, D. A., Hale, A. J., Pinkerton, H., Burrell, R. V., Norton, G. E., and James, M. R., Avtis: A novel millimetre-wave ground based instrument for volcano remote sensing, *J. Volcanology Geothermal Res.*, 146:307–318, 2005.

16. Hort, M., Vöge, M., Seyfried, R., and Ratdomopurbo, A., In situ observation of dome instabilities at Merapi volcano, Indonesia: A new tool for volcanic hazard mitigation, *J. Volcanology Geothermal Res.*, 153(3-4):301–312, 2006.

17. Sheridan, M. F. and Malin, M. C., Application of computer-assisted mapping to volcanic hazard evaluation of surge eruption: Vulcano, Lipari, Vesuvius, *J. Volcanology Geothermal Res.*, 17(1-4):187–202, 1983.

18. Wadge, G. and Isaacs, M. C., Mapping the volcanic hazards from Soufrière Hills Volcano, Montserrat, West Indies using an image processor, *J. Geol. Soc. London*, 145(4):541–551, 1998.

19. Alberico, I., Lirer, L., Petrosino, P., and Scandone, R., A methodology for the evaluation of long-term volcanic risk from pyroclastic flows in Campi Flegrei (Italy), *J. Volcanology Geothermal Res.*, 116:63–78, 2002.

20. Sheridan, M. F. and Macías, J. L., PC software for 2-dimensional gravity driven flows: Application to Colima and El Chichon volcanoes, In *Second International Meeting on Volcanology*, Colima, México, 1992.

21. Sheridan, M. F. and Macías, J. L., Estimation of risk probability for gravity-driven pyroclastic flows at Volcán Colima, México,

J. Volcanology Geothermal Res., 66(1-4):251–256, 1995.

22. Kover, T. P., Application of a digital terrain model for the modeling of volcanic flows: A tool for volcanic hazard determination, Master's thesis, State University of New York at Buffalo, 1995.

23. Sheridan, M. F. and Kover, T. P., FLOW-3D: A computer code for simulating rapid, open-channel volcanic flows, *Proceedings of the US-Japan Science and Technology Workshop on the Technology of Disaster Prevention against Local Severe Storms*, Norman, OK, 1996.

24. Savage, S. B. and Hutter, K., The motion of a finite mass of granular material down a rough incline, *J. Fluid Mech.*, 199:177–215, 1989.

25. Savage, S. B. and Hutter, K., The dynamics of avalanches of granular materials from initiation to runout. Part I: Analysis, *Acta Mech.*, 86(1):201–223, 1991.

26. Iverson, R. M., The physics of debris flows, *Rev. Geoph.*, 35:245–296, 1997.

27. Iverson, R. M. and Denlinger, R. P., Flow of variably fluidized granular material across three-dimensional terrain 1. Coulomb mixture theory, *J. Geoph. Res.*, 106:537–552, 2001.

28. Gray, J. M. N. T., Granular flow in partially filled slowly rotating drums, *J. Fluid Mech.*, 441:1–29, 2001.

29. Pitman, E. B., Nichita, C. C., Patra, A. K., Bauer, A. C., Sheridan, M. F., and Bursik, M. I., Computing granular avalanches and landslides, *Phys. Fluids*, 15(12):3638–3646, 2003.

30. Pelanti, M., Bouchut, F., Mangeney, A., and Vilotte, J.-P., Numerical modeling of two-phase gravitational granular flows with bottom topography, In *Hyperbolic Problems: Theory, Numerics, Applications Proceedings of the 11th International Conference on Hyperbolic Problems, Ecole Normale Supéreure*, Benzoni-Gavage, S. and Serre, D. (eds.), pp. 825–832, Springer-Verlag, 2008.

31. Patra, A. K., Bauer, A. C., Nichita, C. C., Pitman, E. B., Sheridan, M. F., and Bursik, M. I., Parallel adaptive numerical simulation of dry avalanches over natural terrain, *J. Volcanology Geothermal Res.*, 139(1-2):1–21, 2005.

32. Papanicolaou, G., Asymptotic analysis of stochastic equations, In *Studies in Probability Theory*, Rosenblatt, M. (ed.), MAA Studies, Mathematical Association of America, pp. 111–179, 1978.

33. Xiu, D. and Karniadakis, G. E., Modeling uncertainty in flow simulations via generalized polynomial chaos, *J. Comput. Phys.*, 187(1):137–167, 2003.

34. Ghanem, R. G. and Spanos, P. D., *Stochastic Finite Elements: A Spectral Approach*, Dover Publications, Springer-Verlag, New York, 2004.

35. Gray, J. M. N. T., Tai, Y.-C., and Noelle, S., Shock waves, dead zones and particle free regions in rapid granular free-surface flows, *J. Fluid Mech.*, 491:161–181, 2003.

36. Patra, A. K., Nichita, C. C., Bauer, A. C., Pitman, E. B., Bursik, M. I., and Sheridan, M. F., Parallel adaptive discontinuous Galerkin approximation of the debris flow equations, *Comput. Geosci.*, 32(7):912–926, 2006.

37. Pitman, E. B. and Lê, L., A two-fluid model for avalanche and debris flows, *Philos. Trans. R. Soc. London, Series A*, 363:1573–1601, 2005.

38. Terzaghi, K., The shearing resistance of saturated soils and the angle between planes of shear, In *Proceedings of the First International Conference on Soil Mechanics and Foundation Engineering*, MIT Press, Cambridge, MA, pp. 54–56, 1936.

39. LeVeque, R. J., *Numerical Methods for Conservation Laws*, Lectures in Mathematics, ETH Zürich, Birkhauser-Verlag, Basel, CH, 2nd ed., 1992.

40. Toro, E. F., *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction*, Springer-Verlag, Berlin, DE, 3rd ed., 2009.

41. Pitman, E. B., Nichita, C. C., Patra, A. K., Bauer, A. C., Bursik, M. I., and Webber, A. C., A model of granular flows over an erodible surface, *Discrete Continuous Dyn. Syst., Series B*, 3(4):589–599, 2003.

42. Lê, L. and Pitman, E. B., A model for granular flows over an erodible surface, *SIAM J. Appl. Math.*, 70(5):1407–1427, 2009.

43. Sheridan, M. F., Stinton, A. J., Patra, A. K., Pitman, E. B., Bauer, A. C., and Nichita, C. C., Evaluating TITAN2D mass-flow model using the 1963 Little Tahoma Peak avalanches, Mount Rainier, Washington, *J. Volcanology Geothermal Res.*, 139(1-2):89–102, 2005.

44. Sheridan, M. F., Patra, A. K., Dalbey, K., and Hubbard, B., Probabilistic digital hazard maps for avalanches and massive pyroclastic flows using TITAN2D, *Geol. Soc. Am. Special Papers*, 464:281–291, 2010.

45. Blijenberg, H. M., In-situ strength tests of coarse, cohesionless debris on scree slopes, *Eng. Geol.*, 39(3):137–146, 1995.

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

47. Bursik, M. I., Patra, A. K., Pitman, E. B., Nichita, C. C., Macías, J. L., Saucedo, R., and Girina, O., Advances in studies of dense volcanic granular flows, *Rep. Prog. Physics*, 68:271–301, 2005.

48. Lopes, D., Development and implementation of Bayesian computer model emulators, PhD thesis, Duke University, 2012.

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

50. Stefanescu, E. R., Bursik, M. I., Cordoba, G., Dalbey, K., Jones, M. D., Patra, A. K., Pieri, D. C., Pitman, E. B., and Sheridan, M. F., Digital elevation model (DEM) uncertainty and hazard analysis using a geophysical flow model, *Proc. R. Soc. London, Series A: Math., Phys. Eng. Sci.*, 468(2142):1543–1563, 2012.

51. Wechsler, S. P. and Kroll, C. N., Quantifying DEM uncertainty and its effects on topographic parameters, *Photogrammetric Engineering & Remote Sensing*, 72(9):1081–1090, 2006.

52. Shearer, J. W., The accuracy of digital terrain models, In *Terrain Modelling in Surveying and Civil Engineering*, Petrie, G. and Kennie, T. J. M. (eds.), pp. 315–336, McGraw-Hill, NY, 1990.

53. Goodchild, M. F., Buttenfield, B. P., and Wood, J., Introduction to visualizing data quality, In *Visualization in Geographic Information Systems*, Hearshaw, H. M. and Unwin, D. J. (eds.), John Wiley & Sons, Hoboken, NJ, USA, pp. 141–149, 1994.

54. Hunter, G. J. and Goodchild, M. F., Modeling the uncertainty of slope and aspect estimates derived from spatial databases, *Geographical Anal.*, 29(1):35–49, 1997.

55. Hunter, G. J., Caetano, M., and Goodchild, M. F., A methodology for reporting uncertainty in spatial database products, *J. Urban Regional Information Syst. Assoc.*, 7(2):11–21, 1995.

56. Rupp, B., Bursik, M. I., Patra, A. K., Pitman, E. B., Bauer, A. C., Nichita, C. C., Webber, A., Saucedo, R., and Macías, J. L., Simulation of pyroclastic flows of Colima Volcano, Mexico by the TITAN2D program, 2003.

57. Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T., Bayesian model averaging: A tutorial, *Stat. Sci.*, 14:382–417, 1999.

58. Berger, J. O., The case for objective Bayesian analysis, *Bayesian Anal.*, 1(3):385–402, 1999.

59. 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., Using statistical and computer models to quantify volcanic hazards, *Technometrics*, 51(4):402–413, 2009.

60. Santner, T. J., Williams, B. J., and Notz, W. I., *The Design and Analysis of Computer Experiments*, Springer Series in Statistics, Springer-Verlag, New York, NY, 2003.

61. Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P., Design and analysis of computer experiments, *Stat. Sci.*, 4(4):409–435, 1989.

62. Bayarri, M. J., Berger, J. O., Garcia-Donato, G., Liu, F., Palomo, J., Paulo, R., Sacks, J., Walsh, D., Cafeo, J., and Parthasarathy, R., Computer model validation with functional output, *Ann. Stat.*, 35:1874–1906, 2007.

63. Ogburn, S. E., Computational modeling of pyroclastic flows for hazard mapping, PhD thesis, University at Buffalo, 2014.

64. Stein, M. L., *Interpolation of Spatial Data*, Springer-Verlag, New York, 1999.

65. Paulo, R., Default priors for Gaussian processes, *Ann. Stat.*, 33(2):556–582, 2005.

66. Tang, B., Orthogonal array-based Latin hypercubes, *J. Am. Stat. Assoc.*, 88(424):1392–1397, 1993.

67. Koehler, J. R. and Owen, A. B., Computer experiments, In *Handbook of Statistics*, Ghosh, S. and Rao, C. R. (eds.), Vol. 13, pp. 261–308, Elsevier Science, Amsterdam, NL, 1996.

68. Mises, R., Über die "Ganzzahligkeit" der Atomgewicht und verwandte Fragen, *Phys. Z.*, 19:490–500, 1918.

69. Mardia, K. V. and Jupp, P. E., *Directional Statistics*, John Wiley & Sons, New York, 1999.

70. Montserrat Volcano Observatory Website, http://www.mvo.ms/, 2014.

71. Coles, S. G., *An Introduction to Statistical Modeling of Extreme Values*, Springer-Verlag, New York, 2001.

72. Wiener, N., The homogeneous chaos, *Am. J. Math.*, 60(4):897–936, 1938.

73. Xiu, D. and Karniadakis, G. E., The Wiener-Askey polynomial chaos for stochastic differential equations, *SIAM J. Sci. Comput.*, 24(2):619–644, 2002.

74. Xiu, D. and Karniadakis, G. E., A new stochastic approach to transient heat conduction modeling with uncertainty, *Int. J. Heat Mass Transfer*, 46(24):4681–4693, 2003.

75. Xiu, D., Efficient collocational approach for parametric uncertainty analysis, *Commun. Comput. Phys.*, 2(2):293–309, 2007.

76. Dalbey, K., Patra, A. K., Pitman, E. B., Bursik, M. I., and Sheridan, M. F., Input uncertainty propagation methods and hazard mapping of geophysical mass flows, *J. Geoph. Res.*, 113(B05203):1–16, 2008.

77. Dalbey, K., Jones, M. D., Pitman, E. B., Calder, E. S., Bursik, M. I., and Patra, A. K., Hazard risk analysis using computer models of physical phenomena and surrogate statistical models, *Int. J. Uncertainty Quantification*, 2012, to appear.

78. Goldstein, M. and Wooff, D. A., *Bayes Linear Statistics: Theory and Methods*, Wiley Series in Probability and Statistics, John Wiley & Sons, New York, 2007.

79. Stefanescu, E. R., Bursik, M. I., Dalbey, K., Jones, M. D., Patra, A. K., and Pitman, E. B., DEM uncertainty and hazard analysis using a geophysical flow model, In *iEMSs 2010 Conference on Environmental Modeling and Software*, Swayne, D. A., Yang, W., Voinov, A., Rizzoli, A., and Filatova, T. (eds.), pp. 418–426, 2010.

80. Shivaswamy, R., Patra, A. K., and Chaudhari, V., Integrating data and compute intensive workflows for uncertainty quantification in large scale simulation—Application to model based hazard analysis, *Int. J. Comput. Math.*, 91:730–747, 2014.

81. Gramacy, R. B., Lee, H. K. H., and MacReady, W. G., Parameter space exploration with Gaussian process trees, In *ICML-2004: Proceedings of the 21st International Conference on Machine Learning*, Greiner, R. and Schuurmans, D. (eds.), pp. 353–360, ACM Press, 2004.

82. Gramacy, R. B. and Apley, D. W., Local Gaussian process approximation for large computer experiments, Tech. Rep., arXiv:1303.0383, 2013.

83. Kaufman, C., Bingham, D., Habib, S., Heitmann, K., and Frieman, J., Efficient emulators of computer experiments using compactly supported correlation functions, with an application to cosmology, *Ann. Appl. Statist.*, 5:2470–2492, 2011.

84. Dalbey, K., *Predictive simulation and model based hazard maps of geophysical mass flows*, State University of New York at Buffalo, 2009.

85. Stodder, D., *TWDI Monograph Series: Seven Keys to High Performance Data Management for Advanced Analytics*, TWDI Research (1105 Media Inc.), 2011.

86. Plimpton, S. J. and Devine, K. D., Mapreduce in MPI for large-scale graph algorithms, *Paral. Comput.*, 37:610–632, 2011.

87. Stockie, J. M., The mathematics of atmospheric dispersion modeling, *SIAM Rev.*, 53(2):349–372, 2011.