Hazard Assessment for Pyroclastic Flows

Robert L Wolpert
Duke University
Department of Statistical Science
with
James O Berger (Duke),
E. Bruce Pitman, Eliza S. Calder & Abani K. Patra (U Buffalo),
M.J. Bayarri (U Valencia), & Elaine T. Spiller (Marquette U).

Games & Decisions in Reliability & Risk:
2011 May 21, Laggo Maggiore, IT

Pyroclastic Flows at Soufrière Hills Volcano
**Where is Montserrat?**

![Map of Montserrat](image1)

**What's there?**

![Map of Montserrat with Plymouth](image2)

**Montserrat PFs in One 13-month Period**

![Map of Montserrat with Pyroclastic Flows](image3)

**Plymouth, the former capital**

![Photo of Plymouth](image4)
Modelling Hazard

- Goal: quantify the hazard from Pyroclastic Flows (PFs)—
- For specific locations “x” on Montserrat, we would like to find the probability of a “catastrophic event” (inundation to ≥1m) within $T$ years— for $T = 1$, $T = 5$, $T = 10$, $T = 20$ years
- Reflecting all that is uncertain, including:
  - How often will PFs of various volumes occur?
  - What initial direction will they go?
  - How will the flow evolve?
  - How are things changing, over time?
- This is not what is needed for short-term crisis and event management— here we consider only long-term hazard.

Pieces of the Puzzle

Our team (geologist/volcanologist, applied mathematicians, statisticians, stochastic processes) breaks the problem down into two parts:

- Try to model the volume, frequency, and initial direction of PFs at SHV, based on 15 years’ data;
- Tools used: Probability theory, extreme value statistics.
- For specific volume $V$ and initial direction $\theta$ (and other uncertain things like the basal friction angle $\phi$), try to model the evolution of the PF, and max depth $M_x$ at location $x$;
- Tools used: Titan2D PDE solver, emulator (see below).

Titan2D

TITAN2D (U Buffalo) computes solution to the PDEs with:

- Stochastic inputs whose randomness is the basis of the hazard uncertainty:
  - $V =$ initial volume (initial flow magnitude, in m$^3$),
  - $\theta =$ initial angle (initial flow direction, in radians).
- Deterministic inputs:
  - $\phi =$ basal friction (deg) important & very uncertain
  - $\psi =$ internal friction (deg) less important, ignored
  - $v_0 =$ initial velocity (m/s) less important, set to zero
- Output: flow height and depth-averaged velocity at each of thousands of grid points at every time step. We will focus on the maximum flow height at a few selected grid points.
- Each run takes about 1 hour on 16 processors
Hazard Assessment I: What’s a Catastrophe?

- Let $M_x(z)$ be the computer model prediction with arbitrary input parameter vector $z \in Z$ of whatever characterizes a catastrophe at $x \in X$.

**SHV:** $z = (V, \theta, \phi) \in Z = (0, \infty) \times [0,2\pi) \times (0,90)^\circ$, where:

$$M_x(V, \theta, \phi) = \max \text{PF height at location } x \text{ for PF of characteristics } z = (V, \theta, \phi).$$

- Catastrophe occurs for $z$ such that $M_x(z) \in Y_C$.

**SHV:** Catastrophe if $M_x(z) \geq 1\text{ m}$ (Other options...)

- Determine ‘catastrophic region’ $Z_c$ in the input space:

$$Z_c = \{z \in Z : M_x(z) \in Y_C\}$$

**SHV:** $Z_c = \{z : V > \Psi(\theta, \phi)\}$ where...

---

Which inputs are catastrophic at SHV?

By continuity and monotonicity,

$$Z_c = \{z \in Z : M_x(z) \in Y_C\} = \{(V, \theta, \phi) \in Z : V > \Psi(\theta, \phi)\}$$

for the critical contour $\Psi$, where

$$\Psi(\theta, \phi) \equiv \{\text{value of } V \text{ such that } M_x(V, \theta, \phi) = 1\text{ m}\}$$

Now let’s turn to the problem of evaluating $\Psi$.

---

Emulation

To find $\Psi$, we’d like to evaluate $M_x(V_i, \theta_j, \phi_k)$ by running Titan2D for selected locations $x \in X$ and at each of perhaps:

- $\sim 100$ Volumes $V_i$;
- $\sim 100$ Initiation Angles $\theta_j$;
- $\sim 100$ Basal Friction Angles $\phi_k$;

Which would entail maybe $100 \times 100 \times 100 = 1000000$ runs of Titan2D... but we don’t have $1000000$ hours.

Our Solution:

Build an **Emulator** for our PDE flow model.

---

Emulators

An **Emulator** is a (very fast):

- statistical model (based on Gaussian Processes) for our computer model (based on PDE solver) of the volcano.

- The emulator can predict (in seconds) what Titan2D would return (in hours), with an estimate of its accuracy;

- Based on Gaussian Stochastic Process (GaSP) Model. We:
  - Pick a few hundred LHC “design points” $(V_j, \theta_j, \phi_j)$;
  - Run Titan2D at each of them to find Output $M_x(V_j, \theta_j, \phi_j)$;
  - Train the GaSP to return a statistical estimate
    $$E[M_x(V, \theta, \phi) \mid \{M_x(V_j, \theta_j, \phi_j)\}]$$
    of model output $M$ for site $x$ at untried points $(V, \theta, \phi)$.
Our immediate goal is to find a “threshold function” for each location of concern $x$ in Montserrat:

$$\Psi(\theta, \phi) = \text{Smallest volume } V \text{ that would inundate } x \text{ if flow begins in direction } \theta \text{ with friction angle } \phi$$

for each possible direction $\theta$ (0–360 in degrees, or 0–$2\pi$ radians, with $0 = \text{due East}$ and $\pi/2 = \text{due North}$) and basal friction angle $\phi$.

**Simplification**: Take $\phi_i \equiv \hat{\phi}(V_i)$ (empirical; see below). Then we can quantify the hazard at $x$ for $T$ years as

$$\Pr \{ V_i \geq \Psi(\theta_i) \text{ for some PF } (V_i, \theta_i) \text{ in time } (t_0, t_0 + T) \}$$

which in turn we study with the probability models.

---

**Design points $V, \theta$**

2048 Design Points total; 40 used for determining $\Psi$ at $x$=

- **Level 4 Trigger Point**: Dyers River Valley (head of Belham valley).

Black $\iff M_x(V, \theta, \phi) = 0$, Red $\iff M_x(V, \theta, \phi) > 0$.

---

**Polar View of Design, Subdesign, & $\Psi$**

---

**50 Simulations of $\Psi(\theta)$ for Level 4 Trigger Point, Dyers RV**

$\psi(\theta)$ with sampled effective friction
Some Details about our Emulator...

Emulators are very fast approximations for (some of) the outputs $M_x(z)$ of (slow) computer models. They are used for many purposes (design, optimization, inference, sensitivity analyses, . . .) for expensive computer models.

- Begin with a Maxi-Min Latin Hypercube statistical design to select some number $N$ of design points $z_i$ in the large region $\mathcal{Z} = [10^5, 10^{9.5}] \times [0, 2\pi] \text{rad} \times [5, 25] \text{deg}$.
- Run the slow computer model $M_x(z_i)$ at these $N$ preliminary points.
- For fitting an emulator to find $\mathcal{Z}_c$, keep only design points $\mathcal{D}$ in a region ‘close’ to the boundary $\partial \mathcal{Z}_c$.

Too many details? Skip ahead 5 or 6 frames...

Handling the Unknown Hyperparameters

$$\theta = (\rho_V, \rho_\theta, \rho_\phi, \alpha_V, \alpha_\theta, \alpha_\phi, \sigma_z^2, \beta, m)$$

- Deal with the crucial parameters ($\sigma_z^2, \beta, m$) via a fully Bayesian analysis (here an extension of Kriging) using objective priors: $\pi(\beta) \propto 1$, $\pi(m) \propto 1$, and $\pi(\sigma_z^2) \propto 1/\sigma_z^2$.
- Compute the marginal posterior mode, $\hat{\xi}$, of $\xi = (\rho_V, \rho_\theta, \rho_\phi, \alpha_V, \alpha_\theta, \alpha_\phi)$ using the above priors and the reference prior for $\xi$; then $\hat{R} \equiv R(\hat{\xi})$ is completely specified (big simplification—no matrix decomp inside MCMC loop).
  - A fully Bayesian analysis, accounting for uncertainty in $\hat{\xi}$, is difficult and rarely affects the final answer significantly because of confounding of variables.

Gaussian Process Emulators in the Region $\mathcal{Z}_c$

- Since we are interested in regions where the flow is small (1m), fit an emulator to $\tilde{M}_x(z) \equiv \log(1 + M_x(z))$. Let $\tilde{y}$ be the transformed vector of computer model runs $\tilde{M}_x(z)$ for $z \in \mathcal{D}$.
- Model the unknown $\tilde{M}(z)$ as a Gaussian process $\tilde{M}(z) = \beta + mV + Z(V, \theta, \phi)$ (note that we expect a monotonic trend in $V$, but not $\theta; \phi$ discussed later), where $Z(V, \theta, \phi)$ is a stationary GP with
  - Mean 0, Variance $\sigma_2^2$;
  - Product exponential correlation: i.e., $(\forall z_i = (V_i, \theta_i, \phi_i) \in \mathcal{D})$, the correlation matrix $R$ is:
    $$R_{ij} = \exp \left\{ - \frac{|V_i - V_j|}{\rho_V} - \frac{1}{\rho_\theta} - \frac{1}{\rho_\phi} \right\}$$
    with range parameters $\rho_\bullet$, smoothness parameters $\alpha_\bullet$.

The Posterior Mode of $\xi = (\rho_V, \rho_\theta, \rho_\phi, \alpha_V, \alpha_\theta, \alpha_\phi)$

- MLE fitting of $\xi$ has enormous problems; we’ve given up on it.
- A big improvement is finding the marginal MLE of $\xi$ from the marginal likelihood for $\xi$, available by integrating lh wrt the objective prior $\pi(\beta, m, \sigma_z^2) = 1/\sigma_z^2$. The expression is:
  $$L(\xi) \propto |R(\xi)|^{-\frac{1}{2}} |X' R(\xi)^{-1} X|^{-\frac{1}{2}} (S^2(\xi))^{-\left(\frac{q-2}{2}\right)}$$
  where
  - $X = [\mathbf{1}, V]$ is the design matrix for the linear parameters, i.e., $\mathbf{1}$ is the column vector of ones and $V$ is the vector of volumes \{V_i\} in the data set, and $\mu = (\beta, m)$ (of dimension $q = 2$);
  - $S^2(\xi) = [\tilde{y} - X \hat{\mu}]' R(\xi)^{-1} [\tilde{y} - X \hat{\mu}]$;
  - $\hat{\mu} = [X' R(\xi)^{-1} X]^{-1} X' R(\xi)^{-1} \tilde{y}$. 

Too many details? Skip ahead 5 or 6 frames...
An even bigger improvement arises by finding the posterior mode from \( L(\xi)\pi^R(\xi) \), where \( \pi^R(\xi) \) is the reference prior for \( \xi \) (Paulo, 2005 AoS). Note that it is computationally expensive to work with the reference posterior in an MCMC loop, but using it for a single maximization to determine the posterior mode is cheap.

The reference prior for \( \xi \) is \( \pi^R(\xi) \propto |I^*(\xi)|^{1/2} \), where

\[
I^*(\xi) = 
\begin{pmatrix}
(n-q) & \text{tr}W_1 & \cdots & \text{tr}W_p \\
\text{tr}W_1 & \text{tr}W_2 & \cdots & \text{tr}W_p \\
\vdots & \vdots & \ddots & \vdots \\
\text{tr}W_p & \cdots & \cdots & \text{tr}W_p \\
\end{pmatrix}
\]

\[
W_k = \left( \frac{\partial}{\partial \xi_k} \right) \mathbf{R}(\xi)^{-1} \left\{ I_n - X (X' R \xi)^{-1} X \right\}^{-1} X' \mathbf{R}(\xi)^{-1},
\]

with \( q = 2 \) the dimension of \( \mu \) and \( p = 3 \) the dimension of \( \xi \).

The posterior distribution of \( (\sigma^2_z, \beta, m) \), conditional on \( \tilde{y} \) and \( \hat{\xi} \), yields the final emulator (in transformed space) at input \( z^* \):

\[
\tilde{M}_x(z^*) \mid \tilde{y}, \hat{\xi} \sim t(y^*(z^*), s^2(z^*), N - 2),
\]

noncentral \( t \)-distribution with \( N-2 \) degrees of freedom and location & scale parameters:

\[
y^*(z^*) = r'TR^{-1}\tilde{y} + \frac{1}{T R^{-1} 1} \left[ 1 - r'TR^{-1} 1 \right] + \frac{\tilde{V}^{-1} T R^{-1} 1}{V^{-1} T R^{-1} V} \left[ V^* - r'TR^{-1} \tilde{V} \right]
\]

\[
s^2(z^*) = \left[ 1 - r'TR^{-1}r \right] + \frac{1 - r^{-1} r^{-1}}{1 - r^{-1} r^{-1}} \left[ V^* - r'TR^{-1} \tilde{V} \right] \left[ V^{-1} T R^{-1} V \right] + \frac{\left( \tilde{V}^{-1} T R^{-1} \tilde{y} \right)^2}{V^{-1} T R^{-1} V} \left[ V^{-1} T R^{-1} V \right] \frac{\left( \tilde{V}^{-1} T y \right)^2}{V^{-1} T R^{-1} V}
\]

where \( \tilde{V}_i = V_i - V_R, \tilde{V}^* = V^* - V_R, V_R = 1 T R^{-1} V / 1 T R^{-1} 1 \), and \( r^T = (R(z^*, z_1), \ldots, R(z^*, z_N)) \). Tedium, but tractable.

**Response Surfaces**

Currently we simply estimate the function \( \phi(V) \), and replace \( \phi \) in the emulator by this function. The emulator thus becomes only a function of \( (V, \theta) \). We are now moving to a full 3-dimensional \( Z_c \).

With this short-cut, \( \Psi(z) \) depends on only one quantity: \( \theta \).

**Figure:** Median of the emulator, transformed back to the original space. Left: Plymouth, Right: Bramble Airport.

Black points: max-height simulation outputs at subdesign points \( D_c \).
Catastrophic event contours $\Psi$
(response surface slices at ht. 1 m)

Adapting the design

- We added new design points near the boundary $\partial Z_c$ of the critical region where:
  - contours $\Psi(z)$ pass between design points $z_i$ with $M_x(z_i) = 0$ and $z_j$ with $M_x(z_j) \gg 1$; or
  - the confidence bands for $\Psi(z)$ are widest.
- The computer model was re-run at these new design points.
- The emulator was then re-fit and critical contour $\Psi$ was re-computed.
- Median contours $\Psi$ did not change much, but confidence bands for $\Psi$ were much narrower, so it was judged that no further adaptation was needed.

(For other adaptive designs see R.B. Gramacy et al. (ICML, 2004), P. Ranjan et al. (Technometrics, 2008), B.J. Williams et al. (Stat Sin, 2000))

Hazard Assessment II: Probability of Catastrophe

For us a PF is catastrophic if its volume $V$ exceeds an uncertain threshold $\Psi(\theta)$ that depends on the initiation angle $\theta$.

For a hazard summary we wish to report, for each $T > 0$,

$$\Pr[\text{Catastrophe at } x \text{ within } T \text{ Years } ] = \Pr[\{(V_i, \theta_i, \tau_i) \in \mathbb{R}^3 : V_i > \Psi(\theta_i), \tau_i \leq T \}]$$

SO, we need a joint model for points $\{(V_i, \theta_i, \tau_i)\} \subset \mathbb{R}^3$.

Let’s do it in that order: first Volumes, then Angles, then Times.
PF Volume vs. Frequency

Linear log-log plots of Magnitude vs. Frequency are a hallmark of the Pareto probability distribution \( Pa(\alpha, \epsilon) \),

\[
P[V > v] = (v/\epsilon)^{-\alpha}, \quad v > \epsilon.
\]

Which is kind of bad news.

The Pareto Distribution

The negative slope seems to be about \( \alpha \approx 0.64 \) or so.

The Pareto distribution with \( \alpha < 1 \) has:

- Heavy tails;
- Infinite mean \( E[V] = \infty \), infinite variance \( E[V^2] = \infty \);
- No Central Limit Theorem for sums (skewed \( \alpha \)-Stable);
- Significant chance that, in the future, we will see volumes larger than any we have seen in the past. Like \( V > 10^9 \text{m}^3 \).
- The Pareto comes up all the time in the Peaks over Threshold (PoT) approach to the Statistics of Extreme Events— related to Fisher/Tippett Three Types Theorem.

How big is bad?

Kinda depends on which direction it goes...
PF Initiation Angles

Our data on angles is quite vague—we only know which of 7 or 8 valleys were reached by a given PF, from which we can infer a sector but not a specific angle $\theta$:

Any nonuniformity for $\theta$? Any dependence of $\theta$ on Volume $V$?

Discussion

Angle/Volume Data (cont)

We need a joint density function for $V$ and $\theta$. Without much evidence against independence, we use product pdf:

$$V, \theta \sim \alpha \epsilon^{\alpha} V^{-\alpha-1} 1\{V > \epsilon\} \pi_{\kappa}(\theta)$$

where $\pi_{\kappa}(\theta)$ is the pdf for the von Mises $vM(\mu, \kappa)$ distribution,

$$\pi_{\kappa}(\theta) = \frac{e^{\kappa \cos(\theta - \mu)}}{2\pi I_0(\kappa)},$$

centered at $\theta \approx \mu$ close to zero (East) with concentration $\kappa$ that might depend on $V$ if the data support that.

PF Times

If we observe $\lambda$ PFs per year of volume $V > \epsilon$, then what is the probability that such a PF will occur in the next 24 hours?

For short-term predictions, it may be important to note this can depend on many things, such as:

- How high is the dome just now?
- Any seismic activity suggesting dome growth and instability?
- Has it rained recently?
- How long since last PF?

But, for long-term predictions all these factors average out and we assume that:

- PF occurrences in disjoint time intervals are independent
- PF rates are constant over time, neither rising nor falling.

PF Times (cont)

Under those assumptions (which we will revisit), the number of PFs in any time interval has a Poisson Distribution, with mean proportional to its length.

For an interval of length $\Delta t = 1/365$, a single day, the expected number of PFs is $\lambda \Delta t$ and so the probability of

$$P[ \text{At least one PF during time } \Delta t ] = 1 - \exp(-\lambda \Delta t)$$

Or about $1 - e^{-22/365} \approx 5.8\%$ for one day with the SHV data from 1995–2010.
A Summary of our Stochastic Model:

The data suggest a (provisional) model in which:

1. PF Volumes are iid from a Pareto \( V \sim Pa(\alpha, \epsilon) \) distribution for some shape parameter \( \alpha (\approx 0.63) \) and minimum flow \( \epsilon \approx 5 \cdot 10^4 \text{m}^3 \); and
2. PF Initiation Angles have a von Mises \( \theta \sim vM(\mu, \kappa) \) distribution on \([0, 2\pi)\) with \( \mu \approx 0 \) and \( \kappa \approx 0.4 \); and
3. PF Arrival Times follow a stationary Poisson process at some rate \( \lambda \approx 22/\text{yr} \).

These have the beautiful and simplifying consequence that the number of PFs in any region of three-dimensional \((V \times \theta \times \tau)\) space has a Poisson probability distribution—so, we can evaluate:

Posterior distribution of \((\alpha, \lambda)\)

For a given minimum volume \( \epsilon > 0 \) and period \((0, t]\), the sufficient statistics and Likelihood Function are

\[
J = \text{Number of PF’s in } (0, t], \quad \text{and} \\
S = \sum \log(V_j), \text{ the log-product of their volumes} \\
L(\alpha, \lambda) \propto (\lambda \alpha)^j \exp \{- \lambda t \epsilon^{-\alpha} - \alpha S\}
\]

Objective Priors:

- Jeffreys prior is \( \pi_j(\alpha, \lambda) \propto l(\alpha, \lambda)\sqrt{2/\alpha} \propto \alpha^{-4} \epsilon^{-\alpha} \);
- Reference priors:
  - \( \alpha \) of interest: \( \pi_{R1}(\alpha, \lambda) \propto \lambda^{-1/2} \alpha^{-1} \epsilon^{-\alpha/2} \)
  - \( \lambda \) of interest: \( \pi_{R2}(\alpha, \lambda) \propto \lambda^{-1/2} (\alpha^{-2} + (\log \epsilon)^2)^{1/2} \epsilon^{-\alpha/2} \)
   which is also Jeffreys’ independent prior.

Posterior: \( \pi(\alpha, \lambda \mid \text{data}) \propto L(\alpha, \lambda) \pi(\alpha, \lambda) \), quite tractable.

Hazard

P[ Catastrophy within \( T \) Years ]

\[
= 1 - P[Y_T = 0] \quad (\text{where } Y_T \text{ is the number of catastrophes}) \\
= 1 - \exp \left(-\lambda T e^\alpha \int_0^{2\pi} \Psi(\theta)^{-\alpha} \pi_\kappa(d\theta)\right)
\]

Which we can compute pretty easily on a computer.

Accommodating uncertainty in \( \lambda, \kappa, \alpha, \text{ etc.} \) is easy in Objective Bayesian statistics— we use Reference Prior distributions, and simulation-based methods (MCMC) to evaluate the necessary integrals.

Computing the probability of catastrophe

To compute \( \Pr(\text{at least one catastrophic event in } T \text{ years } \mid \text{ data}) \) for a range of \( T \), an importance sampling estimate is

\[
P(T) \approx 1 - \frac{\sum_i \exp \left[-\lambda_i T e^\alpha \hat{\Psi}(\alpha_i)\right]}{\sum_i \pi^*(\alpha_i, \lambda_i) g(\alpha_i, \lambda_i)},
\]

where

- \( \hat{\Psi}(\alpha) \) is an MC estimate of \( \int_0^{2\pi} \Psi(\theta)^{-\alpha} \pi_\kappa(d\theta) \) based on draws \( \theta_i \sim vM(\mu, \kappa) \) (or use quadrature);
- \( \pi^*(\alpha, \lambda) \) is the un-normalized posterior density;
- \( \{(\alpha_i, \lambda_i)\} \) are iid draws from the importance sampling density \( g(\alpha, \lambda) = t_2(\alpha, \lambda \mid \hat{\mu}, \hat{\Sigma}, 3) \), with d.f. 3, mean \( \hat{\mu}^t = (\hat{\alpha}, \hat{\lambda}) \), and scale \( \hat{\Sigma} = \text{inverse of observed Information Matrix.} \)
We have argued that:

- Hazard assessment of catastrophic events (in the absence of lots of extreme data) requires
  - Mathematical computer modeling to support extrapolation beyond the range of the data;
  - Statistical modeling of available (possibly not extreme) data to determine input distributions;
  - Statistical development of emulators of the computer model to determine critical event contours.

- Major sources of uncertainty can be combined and incorporated with Objective Bayesian analysis.

- Extend the methodology to create entire Hazard Maps (i.e., find hazard for all locations $x$ simultaneously).

- Go beyond stationarity with
  - Change-point model for intensity $\lambda_t$;
  - Model (heavy-tailed) duration of activity;
  - Model caldera evolution ($\mu, \kappa$ for vM).

- Reflect uncertainty and change in DEMs.
A Collaborative Effort...

Thanks— to Organizers and Collaborators!
For more, see: www.stat.duke.edu/~rlw/ or www.mvo.ms