Uncertainty in Hazard Forecasting: Or where will you go when the volcano blows?

Elaine Spiller

Marquette University

October 3, 2017
How worried you should be if you see local reporters interviewing scientists about a breaking news story, by field:

More worried →

Economist  Criminologist  Botanist  Marine Biologist  Astronomer
Nutritionist  Archeologist  Ornithologist  Entomologist  Vulcanologist  Virologist  Astronomer who studies the Sun
Eyjafjallajökull – $2-5bn
Nevada del Ruiz – 23,000 fatalities
“One hundred years ago, government officials in Martinique made the mistake of assuming that, despite signs to the contrary, Mount Pelée would behave in 1902 as it had in 1851 – when a rain of ash from what they considered a benign volcano surprised, but did not harm those living under its shadow.”

(Cristina Reed, *Geotimes* 2002)
Pyroclastic – “broken fire” – flows
Montserrat – A volcanologist playground
Data at Montserrat – valleys traversed by PFs
Data at Montserrat – PF frequency and volume

![Graph showing the relationship between PF Volume (m³) and Annual Rate. The x-axis represents PF Volume (m³) with values ranging from 5e+04 to 5e+08, and the y-axis represents Annual Rate with values ranging from 0.02 to 50.00.]
Data at Montserrat – (negative) slope $\alpha$

![Graph showing a negative slope relationship between PF Volume (m$^3$) and Annual Rate.](image-url)
Learning about $\alpha$ from data

Bayes Theorem

$$p(\alpha \mid \text{data}) \propto p(\text{data} \mid \alpha)p(\alpha)$$
Learning about $\alpha$ from data

Bayes Theorem

\[ p(\alpha \mid \text{data}) \propto p(\text{data} \mid \alpha)p(\alpha) \]

Typically can’t compute $p(\alpha \mid \text{data})$, but can sample
Data at Montserrat – $p(\alpha | \text{data})$
μMULOGO
Data and data models at Montserrat

α < 1 indicates so-called heavy tails
Data and data models at Montserrat

\( \alpha < 1 \) indicates so-called heavy tails
Rare Events

Pareto model is much more likely to observe future volumes that far exceed those in the recent history...
Pareto model is much more likely to observe future volumes that far exceed those in the recent history...

Consider a record of 10 volumes \((V_1, \ldots, V_{10})\)
Pareto model is much more likely to observe future volumes that far exceed those in the recent history...

Consider a record of 10 volumes \((V_1, \ldots, V_{10})\)

non-heavy tailed:

\[
P(V_{11} > 10 \max(V_1, \ldots, V_{10})) = 1/200,000
\]
Pareto model is much more likely to observe future volumes that far exceed those in the recent history...

Consider a record of 10 volumes \((V_1, \ldots, V_{10})\)

**non-heavy tailed:**
\[
P(V_{11} > 10 \max(V_1, ..., V_{10})) = 1/200,000
\]

**heavy tailed:**
\[
P(V_{11} > 10 \max(V_1, ..., V_{10})) = 1/100
\]
What happens at larger-than-recorded volumes?

![Graph showing the relationship between PF Volume (m³) and Annual Rate. The graph includes a trend line and data points.](image-url)
We would like records from many volcanic eruptions
Best we can do: simulate replicate volcanic eruptions
Physics based models as a “lab”

Assume: flow layer thin relative to lateral extension

continuity

\[
\frac{\partial h}{\partial t} + \frac{\partial h u_x}{\partial x} + \frac{\partial h u_y}{\partial y} = e_s
\]

\(x\) momentum

\[
\frac{\partial h u_x}{\partial t} + \frac{\partial (h u^2_x + k_{ap} g_z h^2 / 2)}{\partial x} + \frac{\partial h u_y u_x}{\partial y} =
\]

\[
g_x h + u_x e_s - \frac{u_x}{\sqrt{u_x^2 + u_y^2}} (g_z + \frac{u^2_x}{\kappa_x}) h \tan(\phi_{bed}) - \text{sgn}(\partial u_x y) h k_{ap} \frac{\partial h g_z}{\partial y} \sin(\phi_{int})
\]

1. Gravitational driving force
2. Coulomb friction at the base – \(\phi_{bed}\)
3. Intergranular Coulomb force – \(\phi_{int}\)

due to velocity gradients normal to flow direction

(see Savage; Bursik; Pitman)
$\text{log}V = 6.3751$, Orientation = $60^\circ$
Simulated pyroclastic flows at four different inputs ($V, \theta$)

- $\log V = 6.3751$, Orientation = $60^\circ$
- $\log V = 5.5779$, Orientation = $16^\circ$
- $\log V = 6.0987$, Orientation = $166^\circ$
- $\log V = 6.6456$, Orientation = $286^\circ$
Belham Valley Probabilistic Hazard Map (t=2.5 yrs)

- incorporate any/all sources of knowledge
  - physics of granular flow
  - data on frequency/size of flows
- avoid one-off simulations
Belham Valley Probabilistic Hazard Map (t=2.5 yrs)

- incorporate any/all sources of knowledge
  - physics of granular flow
  - data on frequency/size of flows
- avoid one-off simulations

Methodology developed for hazard mapping works for UQ
Simulation details: TITAN2D (Patra)

- Large scale computations to produce realistic simulations of mass flows — depth average hyperbolic balance laws
  - like shallow water with dissipative friction terms
  - finite-volume 2\textsuperscript{nd} order Godunov solver
  - integrated with GIS to obtain terrain data
  - local, adaptive mesh refinement
- High performance techniques for efficiency
  - parallel
  - dynamic load balancing
- \( \sim 1 \) hr run time
- each initialized with volume and initial direction
Possible statistical models for physical scenarios, $p(V, \theta)$
Possible statistical models for physical scenarios, $p(V, \theta)$

- linear volume model
  (2 parameters)
- frequency model, rate $= \lambda$
  (1 parameter)
- uniform
- Von Mises (2 pars)
  (Gaussian on a circle)
Monte Carlo Simulation

Idea literally named for gambling

- “roll” the “die” $N$ times
- “die” is probabilistic scenario model
- “roll” is the flow model exercised at a sampled scenario
- $P(\text{hazard}) = \frac{\text{(# of catastrophes)}}{N}$
four draws from $p(V, \theta)$
Cartoon $p(\text{physical scenario})$
Cartoon $p(\text{physical scenario})$

Monte Carlo Samples

Physical Scenarios
Cartoon $p(\text{physical scenario})$

Physical Scenarios

scenarios leading to catastrophe
Cartoon $p(\text{physical scenario})$

Physical Scenarios

scenarios leading to catastrophe
Cartoon $p(\text{physical scenario})$

Physical Scenarios

scenarios leading to catastrophe
Uncertainty

- **Aleatory variability** — random scenarios
  - volume
  - initiation angle
  - frequency
Uncertainty

- **Aleatory variability** — random scenarios
  - volume
  - initiation angle
  - frequency

- **Epistemic uncertainty** — imperfect descriptions
  - probabilistic models (of random scenarios)
  - numerical resolution
  - physical parameters
Idea: A given $V - \theta$ pair will either result in inundation or not independent of how probable that event is.
Idea: A given $V - \theta$ pair will either result in inundation or not independent of how probable that event is.

- Run TITAN2D at $(V, \theta)$ pairs spread over “physical scenario” space, collect max height of resulting flow around volcano.

- Interpolate between these runs to predict which locations would be inundated for any $V - \theta$ flow.

- **Statistical emulator** – interpolation & uncertainty estimates
Emulator – statistical model of physical model
Challenges for hazard mapping

- Emulate whole map at once?
Challenges for hazard mapping

- Emulate whole map at once? Huge matrix inversion. Complicated, topography dependent, spatial footprints

- "Spread out" scenarios
- Choose site-specific subdesigns from model runs
- Include "important" runs resulting in no-flow
Challenges for hazard mapping

- Emulate whole map at once? Huge matrix inversion. Complicated, topography dependent, spatial footprints
  - treat each site individually, build $M$ GaSP in parallel
Challenges for hazard mapping

  - treat each site individually, build $M$ GaSP in parallel

- Many scenarios lead to no flow at many locations
Challenges for hazard mapping

- Emulate whole map at once? Huge matrix inversion. Complicated, topography dependent, spatial footprints
  - treat each site individually, build $M$ GaSP in parallel

- Many scenarios lead to no flow at many locations
  - run physical model a $N$ “spread out” scenarios
  - choose site-specific subdesigns from $N$ model runs
  - include “important” runs resulting in no-flow
Emulator at one map site
Emulator at one map site

![Graph showing volume (m^3) vs. initiation angle (from East)](image-url)

- Volume (m^3) is plotted on the y-axis.
- Initiation angle (from East) is plotted on the x-axis.
- The graph shows a peak volume near 0° and 360° with minor fluctuations.
Making a hazard map

1. Run $N = 2048$ TITAN2D, store data for each location
2. Repeat following process, in parallel, for each site
   1. choose subdesign
   2. fit emulator
   3. draw catastrophic contours, $\psi(\theta)$’s
3. Choose model for aleatory variability of scenarios
4. Run probability calculations, in parallel, for each site

Note

step 1 is expensive,
Making a hazard map

1. Run $N = 2048$ TITAN2D, store data for each location
2. Repeat following process, in parallel, for each site
   1. choose subdesign
   2. fit emulator
   3. draw catastrophic contours, $\psi(\theta)$’s
3. Choose model for aleatory variability of scenarios
4. Run probability calculations, in parallel, for each site

Note

step 1 is expensive, 2 is parallelizable,
Making a hazard map

1. Run $N = 2048$ TITAN2D, store data for each location
2. Repeat following process, in parallel, for each site
   1. choose subdesign
   2. fit emulator
   3. draw catastrophic contours, $\psi(\theta)$’s
3. Choose model for aleatory variability of scenarios
4. Run probability calculations, in parallel, for each site

Note

step 1 is **expensive**, 2 is **parallelizable**, 4 is **post processing**!

details in *SIAM/ASA JUQ* (Spiller 2014), overview in *IJUQ* (Bayarri 2015)
$P(\text{catastrophe in 2.5 years})$
Recall volume data used to characterize $p(V, \theta)$ (aleatory variability)
Recall volume data used to characterize $p(V, \theta)$ (aleatory variability)
Epistemic uncertainty: uncertainty in probability model

Recall volume data used to characterize $p(V, \theta)$ (aleatory variability)

- each red curve corresponds to a different slope $p(V, \theta | \alpha)$
- now probability calculation is cheap — we can find $P($hazard$)$ for each $\alpha$!
Epistemic uncertainty: in probability & physical models

uncertainty in prob model

uncertainty in phys model

Repeat probability calculations many times

- vary $\alpha$ – probability model
- vary friction uncertainty – physical model
red – fix friction, vary $\alpha'$s

blue – fix $\alpha = \hat{\alpha}$, vary friction
red – fix friction, vary $\alpha'$s  

blue – fix $\alpha = \hat{\alpha}$, vary friction
A retrospective “validation”

- use data from 1995-2003 to estimate Poisson frequencies for
  (top, stationary)
  (mid, low activity)
  (bottom, high activity)

- forecast probabilities of inundation for 2004-2010 under these three scenarios
A retrospective “validation”

- Use data from 1995-2003 to estimate Poisson frequencies for
  (top, stationary)
  (mid, low activity)
  (bottom, high activity)

- Forecast probabilities of inundation for 2004-2010 under these three scenarios

- White overlay, extent of deposits for 2004-2010
Aleatoric variability – short term modeling

Sep 17, 1996
Jun 25, 1997
Aug 3, 1997
Sep 21, 1997
Dec 26, 1997

Jul 3, 1998
Mar 20, 2000
Jul 29, 2001
Jul 13, 2003
May 20, 2006

[Graph showing time series data with dates and values]
$\mu_o = 135^\circ$

$\mu_o = 0^\circ$
To wrap up

Take home message

- Emulator of physical model identifies important regions of state space independent of probabilistic model
- Enables fast, flexible direct or MC probability calculations w/o more physical simulations
- Framework for exploring multiple sources of epistemic uncertainty and aleatory variability
- Not a replacement, but a tool for civil protection and scientists to forecast dynamic hazards and quantify uncertainty
interdisciplinary research team

Along with many current and former students...
The End

https://sites.google.com/view/elainespiller

Thanks to NSF: DMS-0757549-0757367-0757527, DMS-1228317-1228265-1228217, EAR-1331353, DMS-1622403-1621853-1622467, SES-1521855
Hierarchical Linear Model Example: basal friction vs. volume

Treat slopes as draws from a common distribution

\[ \beta_j \sim N(\mu, \tau^2) \]

\[ \tau^2 \rightarrow 0 \quad \text{single regression} \]

\[ \tau^2 \rightarrow \infty \quad \text{separate regressions} \]
Montserrat

Volume (m$^3$) $2 \times 10^4$ $2 \times 10^5$ $2 \times 10^6$ $2 \times 10^7$ $2 \times 10^8$

Coefficient of Friction $1 \times 10^{-2}$ $1 \times 10^{-1}$ $1 \times 10^0$ $1 \times 10^1$ $1 \times 10^2$

Confidence Interval - HM
Confidence Interval - LR
Linear Regression
Median Regression Line - HM

Soufriere Hills
Montserrat