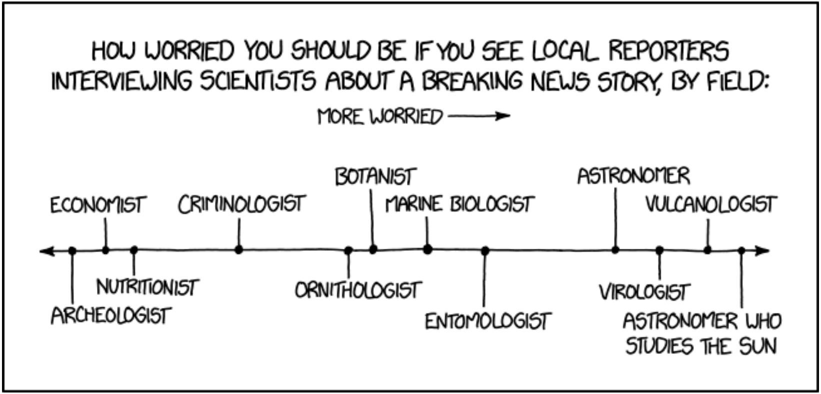
Uncertainty in Hazard Forecasting: Or where will you go when the volcano blows?

Elaine Spiller

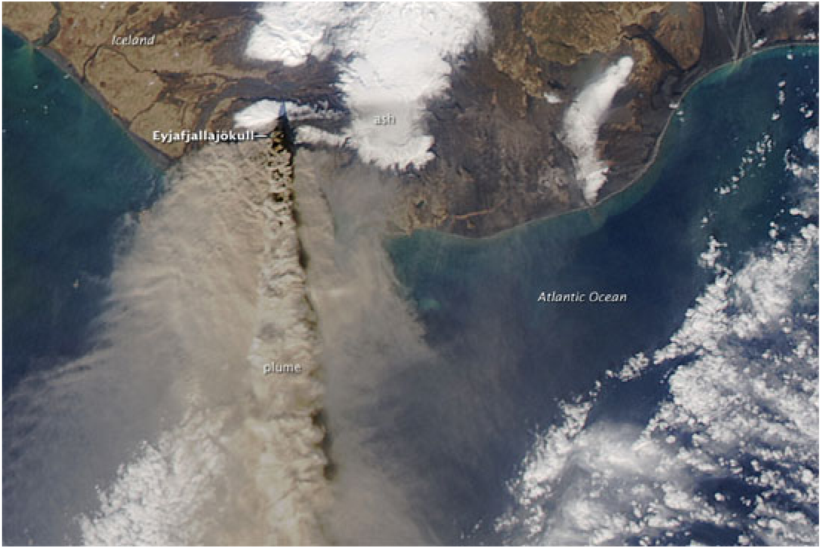
Marquette University

October 3, 2017

XKCD comic from September 28



Eyjafjallajökull – $2-5bn



Nevada del Ruiz – 23,000 fatalities

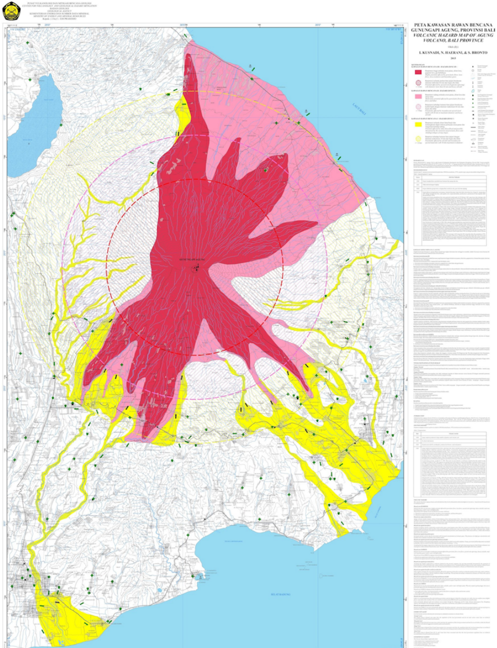
Mount Pelée, Martinique – 30,000 fatalities

"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

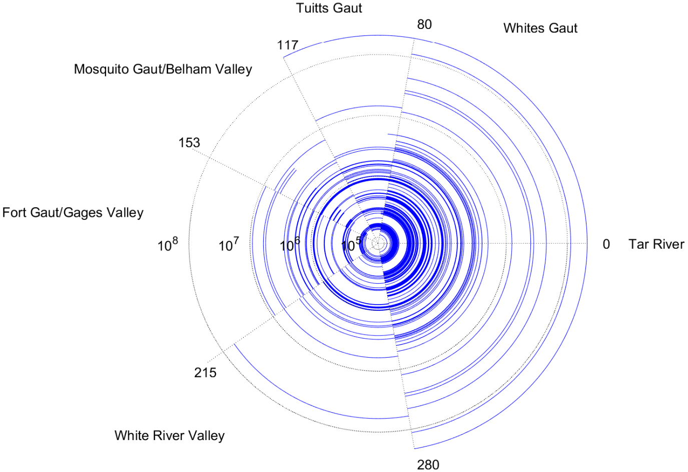
Mount Agung

Mount Agung

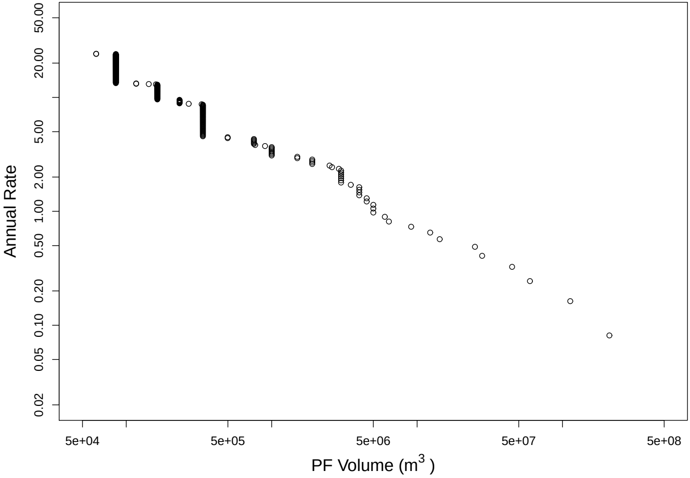

PETA KAWASAN RAWAN BENCANA GUNUNGAPI AGUNG, PROVINSI BALI
VOLCANIC HAZARD MAP OF AGUNG VOLCANO, BALI PROVINCE
I. KUSNADI, N. HAERANI, & S. BRONTO
2015

Montserrat – A volcanologist playground

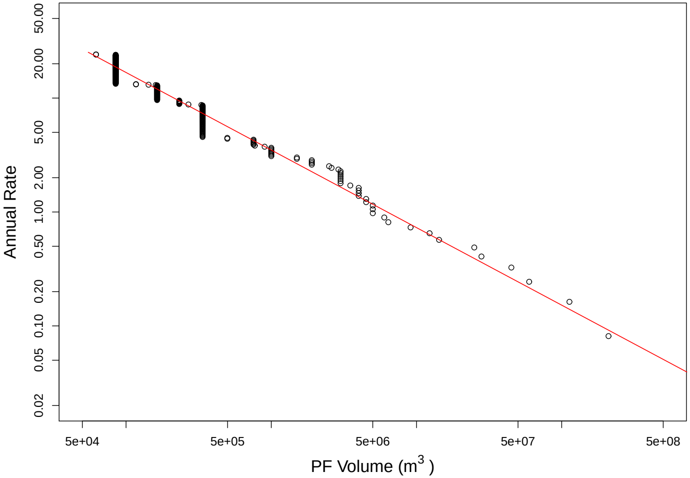
Data at Montserrat – valleys traversed by PFs



Data at Montserrat – PF frequency and volume



Data at Montserrat – (negative) slope α



Bayes Theorem

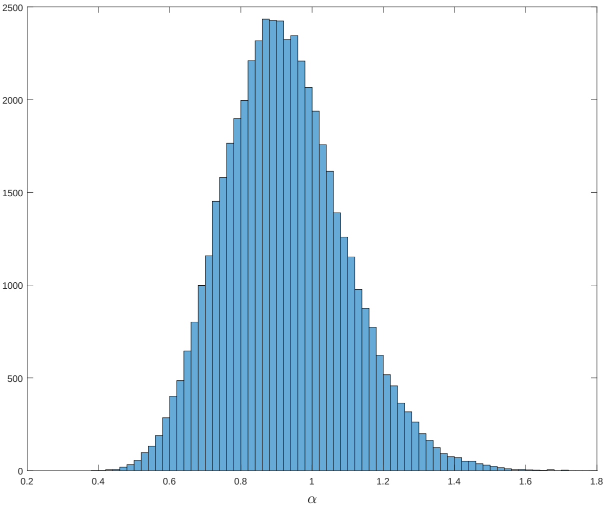
$$p(\alpha \mid \mathrm{data}) \propto \mathrm{p}(\mathrm{data} \mid \alpha)\mathrm{p}(\alpha)$$

Bayes Theorem

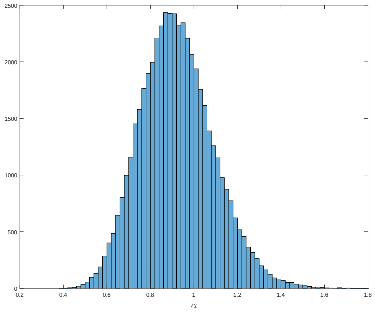
$$p(\alpha \mid \mathrm{data}) \propto \mathrm{p}(\mathrm{data} \mid \alpha)\mathrm{p}(\alpha)$$

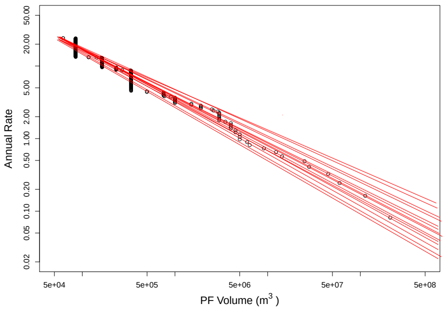
Typically can't compute $p(\alpha \mid \mathrm{data})$, but can sample

Data at Montserrat – $p(\alpha \mid data)$

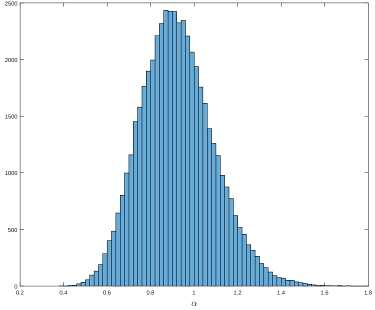


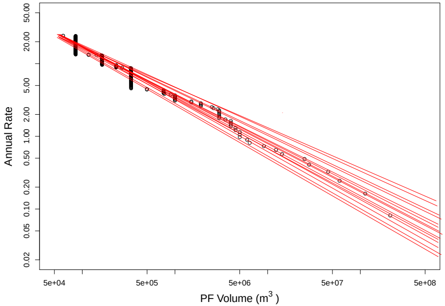
Data and data models at Montserrat





Data and data models at Montserrat





$\alpha < 1$ indicates so-called *heavy tails*

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:

$P(V_{11} > 10 \max(V_1, ..., V_{10})) = 1/200,000$

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

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

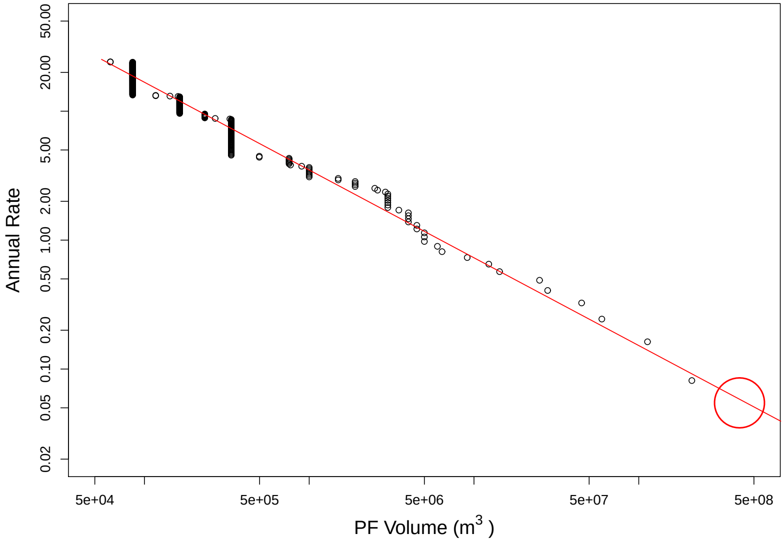
non-heavy tailed:

$P(V_{11} > 10\,\text{max}(V_1, ..., V_{10})) = 1/200,000$

heavy tailed:

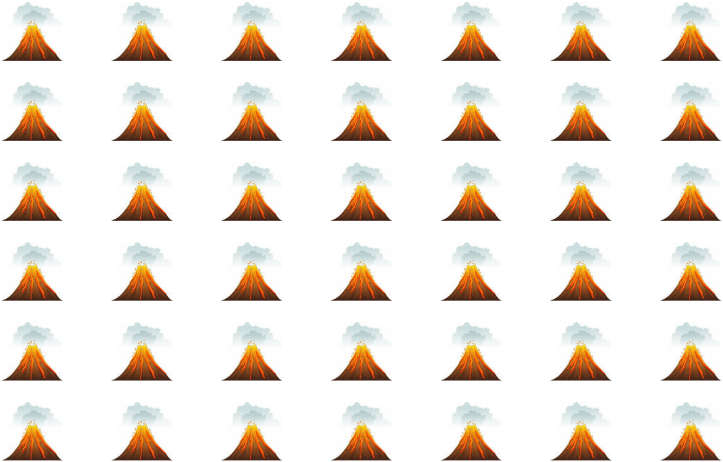
$P(V_{11} > 10\,\text{max}(V_1, ..., V_{10})) = 1/100$

What happens at larger-than-recorded volumes?



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

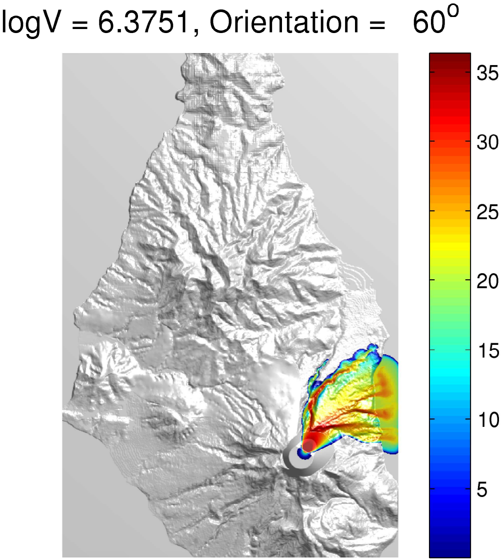
continuity
$$\frac{\partial h}{\partial t} + \frac{\partial h u_x}{\partial x} + \frac{\partial h u_y}{\partial y} = e_s$$

x momentum
$$\frac{\partial h u_x}{\partial t} + \frac{\partial (h u_x^2 + k_{ap} g_z h^2/2)}{\partial x} + \frac{\partial h u_y u_x}{\partial y} =$$

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

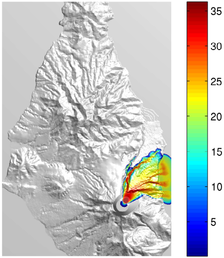
1. Gravitational driving force
2. Coulomb friction at the base – ϕ_{bed}
3. Intergranular Coulomb force – ϕ_{int}
   due to velocity gradients normal to flow direction

(see Savage; Bursik; Pitman)

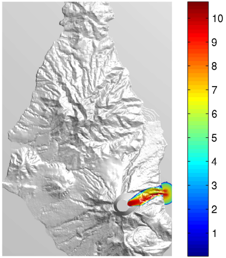

logV = 6.3751, Orientation = 60°
35
30
25
20
15
10
5

Simulated pyroclastic flows at four different inputs (V, θ)

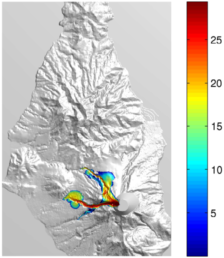
logV = 6.3751, Orientation = 60^o



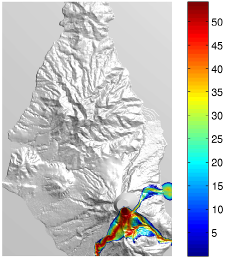
logV = 5.5779, Orientation = 16^o



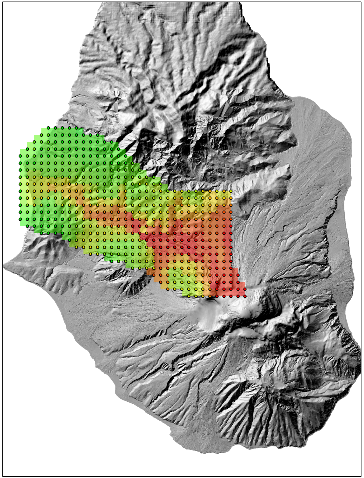
logV = 6.0987, Orientation = 166^o



logV = 6.6456, Orientation = 286^o

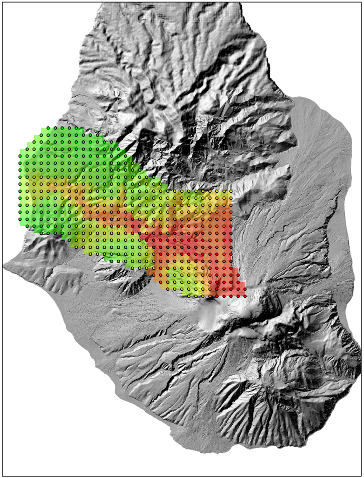


Belham Valley Probabilistic Hazard Map (t=2.5 yrs)



- 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)



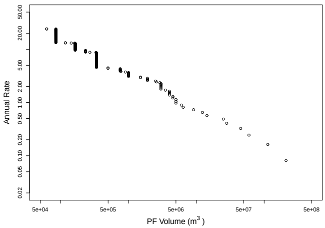
- incorporate any/all sources of knowledge
  - physics of granular flow
  - data on frequency/size of flows
- avoid one-off simulations

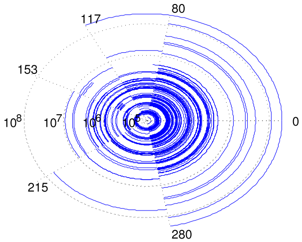
- Methodology developed for hazard mapping works for UQ

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^{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

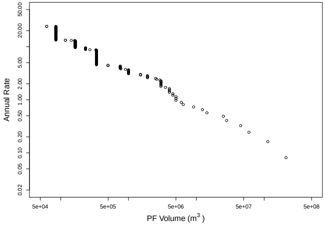
Physical scenarios: data and models of $p(V, \theta)$

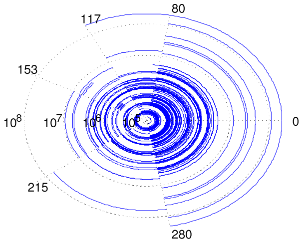




Possible statistical models for physical scenarios, $p(V, \theta)$

Physical scenarios: data and models of $p(V, \theta)$





Possible statistical models for physical scenarios, $p(V, \theta)$

- linear volume model

  (2 parameters)
- frequency model, rate$= \lambda$

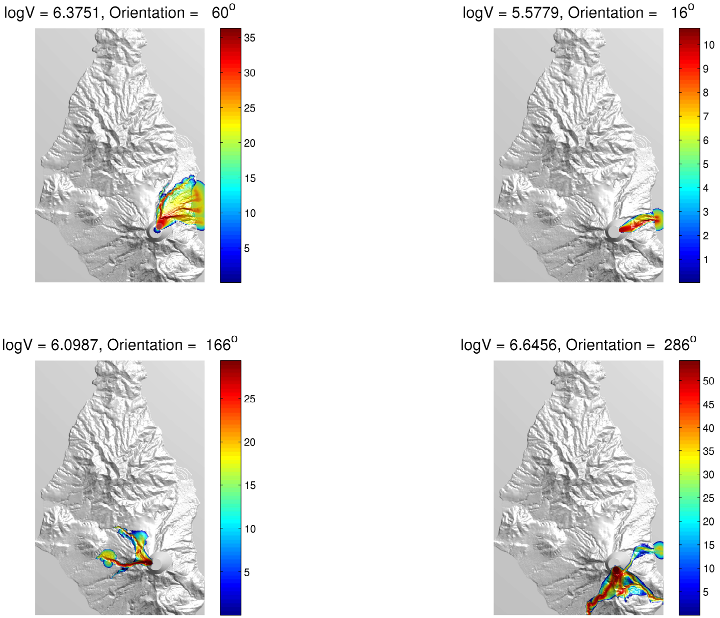
  (1 parameter)

- uniform
- Von Mises (2 pars) (Gaussian on a circle)

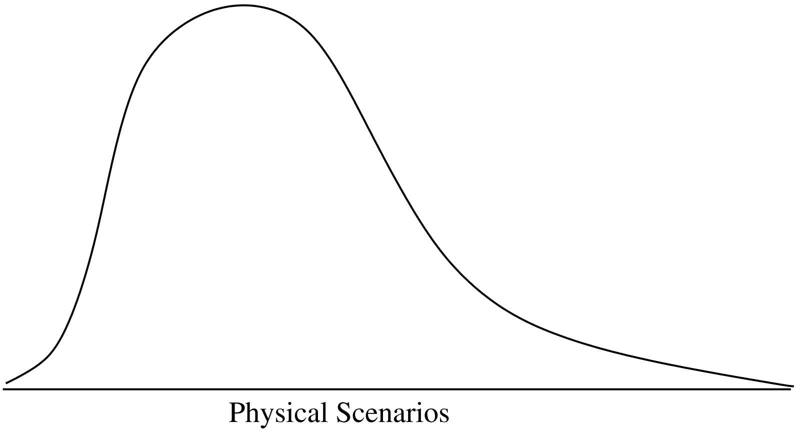
Idea literally named for gambling

- “roll” the “die” N times
- “die” is probabilistic scenario model
- “roll” is the flow model exercised at a sampled scenario
- P(hazard)= (# of catastrophes)$/N$

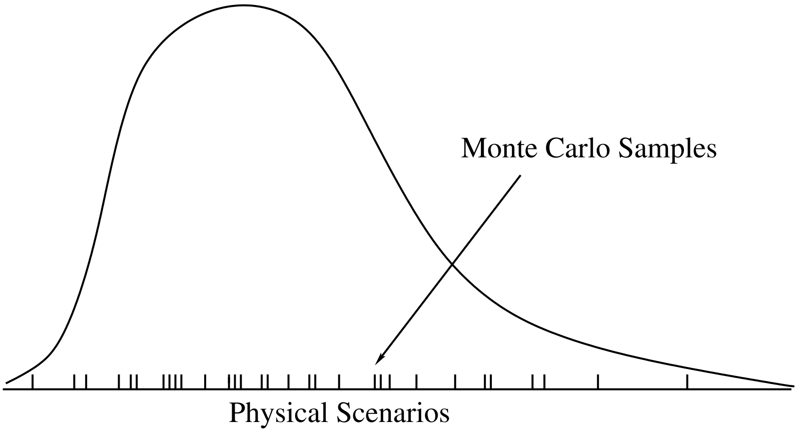
four draws from $p(V, \theta)$



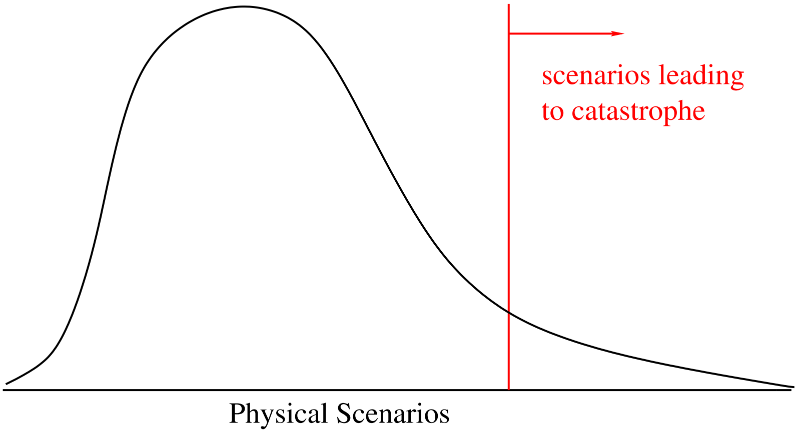
Cartoon p(physical scenario)



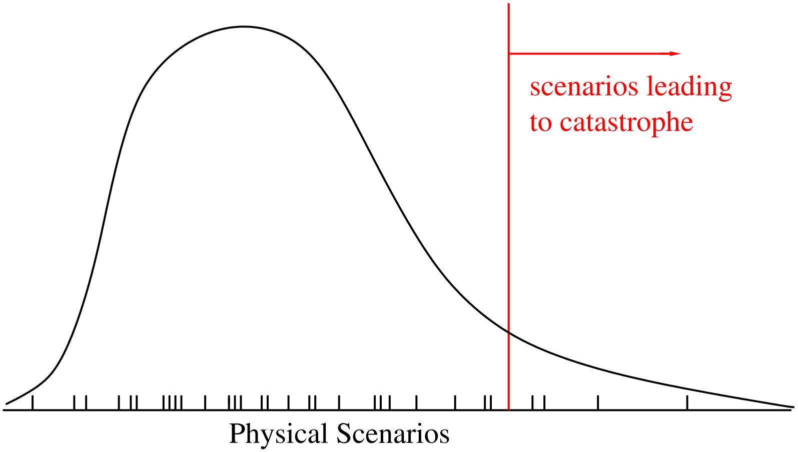
Cartoon p(physical scenario)



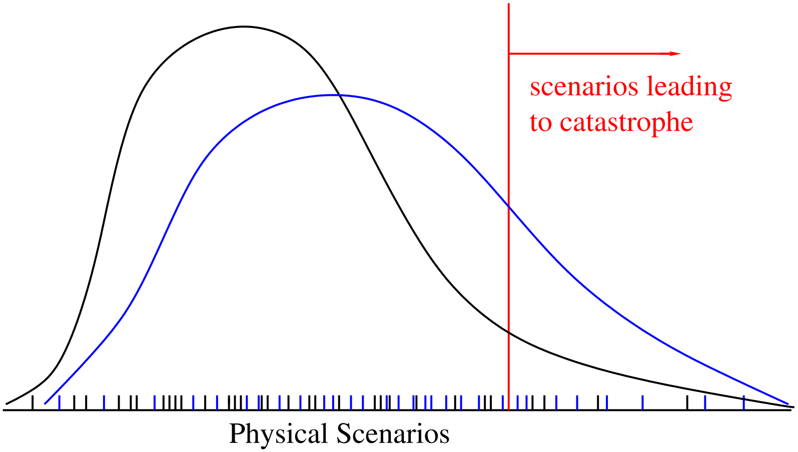
Cartoon p(physical scenario)



Cartoon p(physical scenario)



Cartoon p(physical scenario)



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

Strategy: separate physical model from probabilistic models

Idea: A given $V-\theta$ pair will either result in inundation or not *independent* of how probable that event is

Strategy: separate physical model from probabilistic models

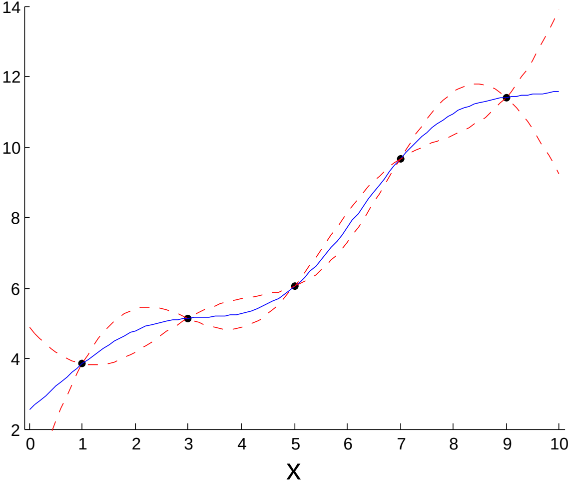
Idea: A given $V - \theta$ pair will either result in inundation or not *independent* of how probable that event is

- Run TITAN2D at (V, θ) 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

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

- 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

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

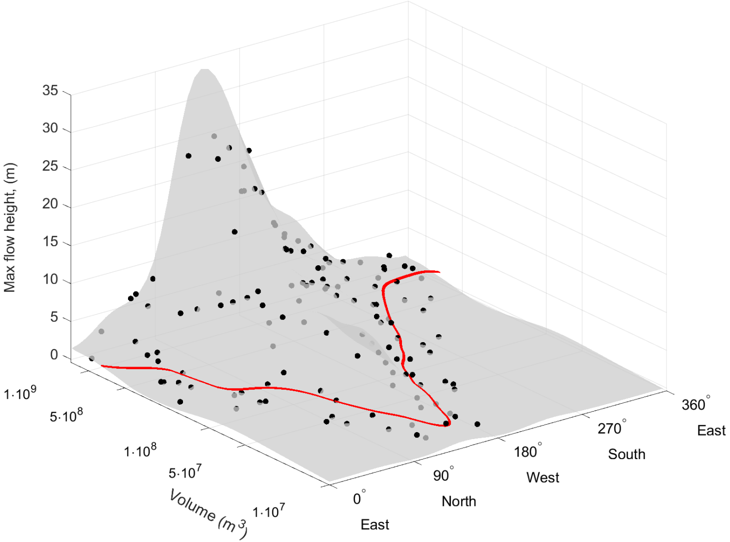
- Many scenarios lead to no flow at many locations

  -run physical model a N “spread out” scenarios

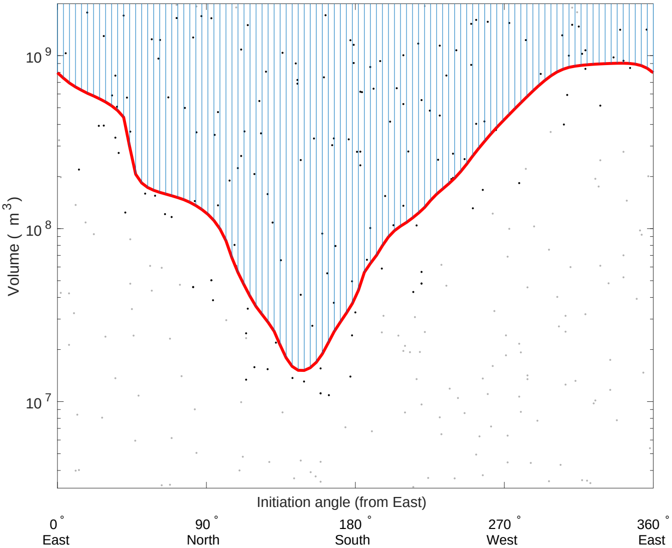
  -choose site-specific subdesigns from N model runs

  -include “important” runs resulting in no-flow

Emulator at one map site



Emulator at one map site



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

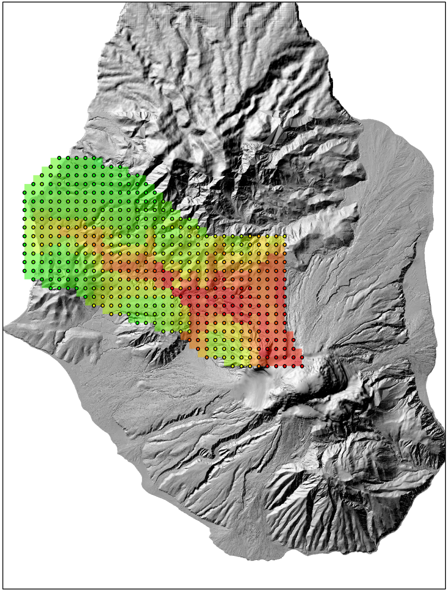
1. Run $N = 2048$ TITAN2D, store data for each location
2. Repeat following process, in parallel, for each site
   1. choose subdesign
   2. fit emulator
   3. draw catastrophic contours, $\psi(\theta)$'s
3. Choose model for aleatory variability of scenarios
4. Run probability calculations, in parallel,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

Note

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

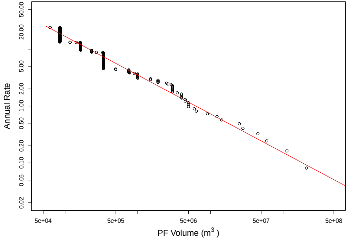
details in *SIAM/ASA JUQ* (Spiller 2014), overview in *IJUQ* (Bayarri 2015)

P(catastrophe in 2.5 years)



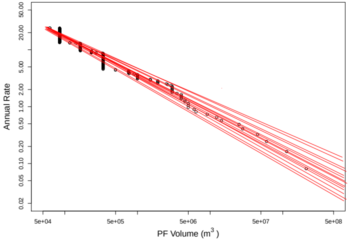
Epistemic uncertainty: uncertainty in probability model

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



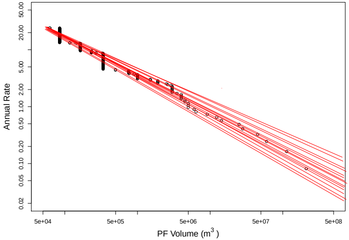
Epistemic uncertainty: uncertainty in probability model

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



Epistemic uncertainty: uncertainty in probability model

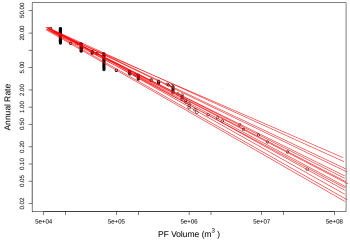
Recall volume data used to characterize $p(V, \theta)$ (aleatory variability)



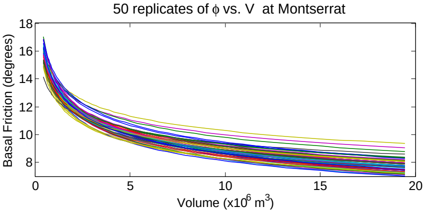
- each red curve corresponds to a different slope $p(V, \theta|\alpha)$
- now probability calculation is cheap —
  we can find P(hazard) for each α!

Epistemic uncertainty: in probability & physical models

uncertainty in prob model



uncertainty in phys model

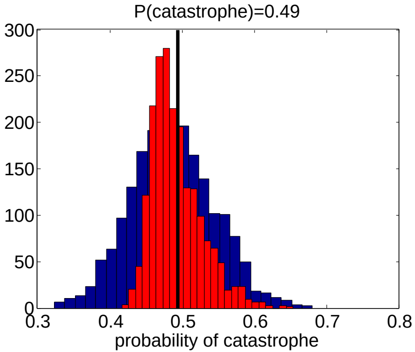


Repeat probability calculations many times

- vary α – probability model
- vary friction uncertainty – physical model

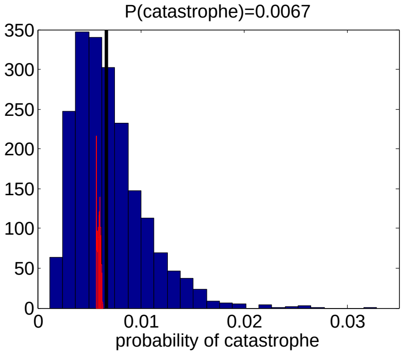
Histograms of catastrophic probabilities – close

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



Histograms of catastrophic probabilities – far

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

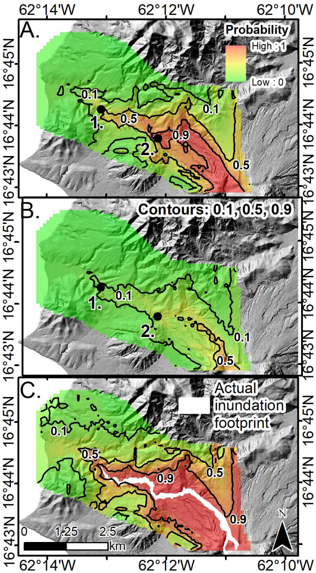
A retrospective "validation"

- use data from 1995-2003 to estimate Poisson frequencies for

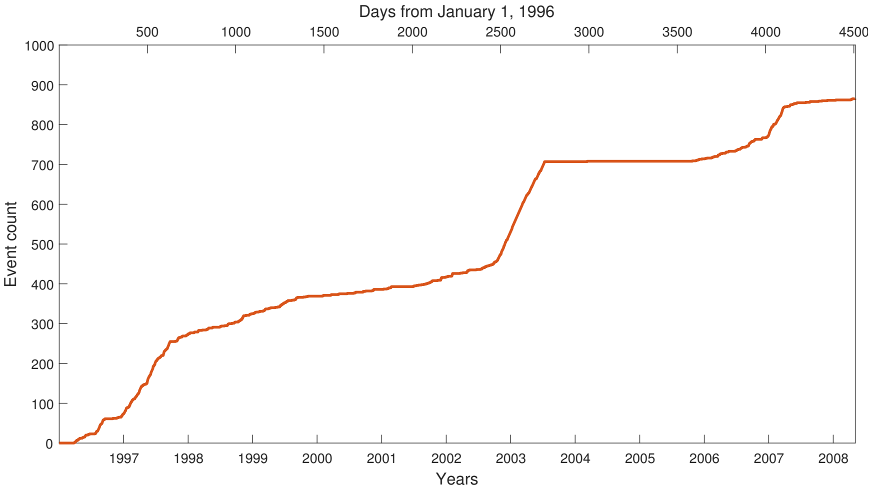
  (top, stationary)
  (mid, low activity)
  (bottom, high activity)

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

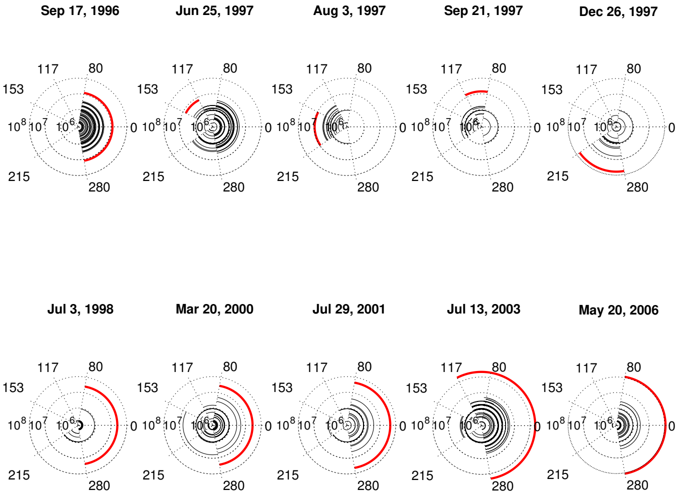
- white overlay, extent of deposits for 2004-2010



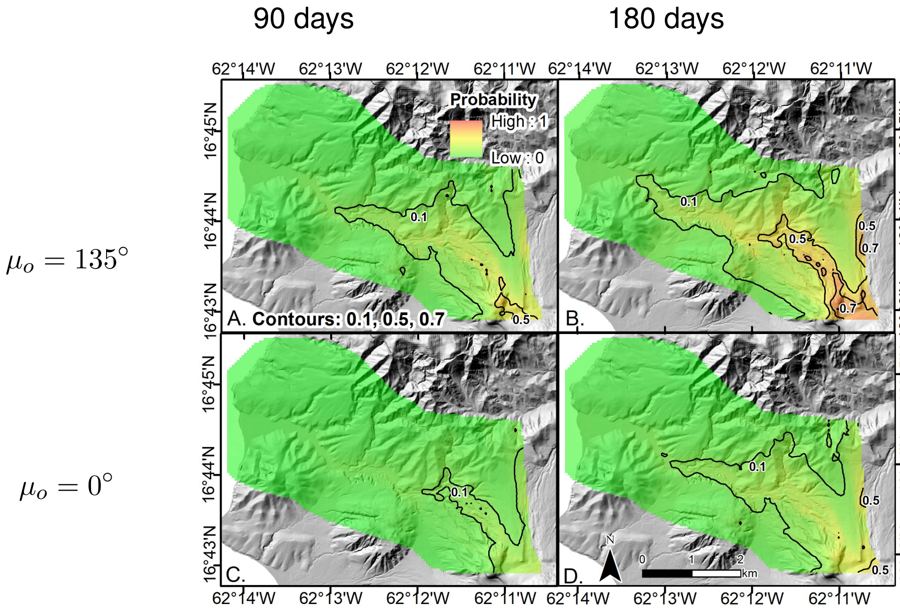
Aleatoric variability – short term modeling



Aleatoric variability – short term modeling



Short-term probabilistic hazard maps



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

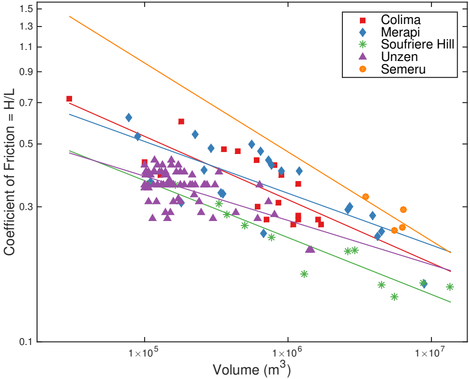
Along with many current and former students...

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

Beyond Montserrat: sparse data to inform scenarios

Hierarchical Linear Model Example: basal friction vs. volume



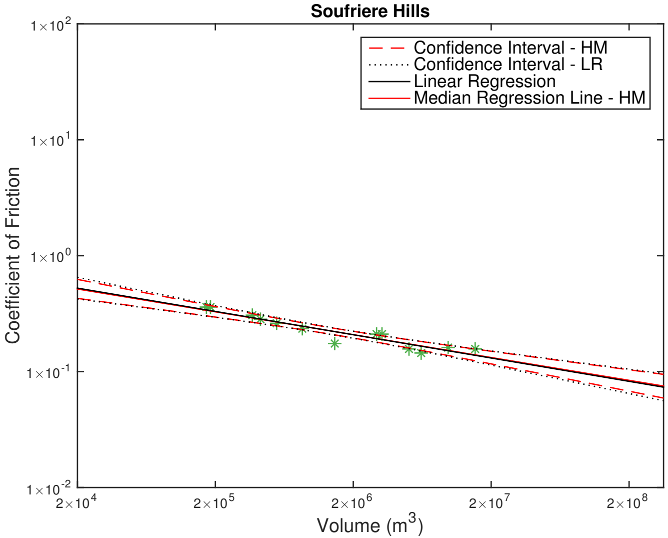
Treat slopes as draws from a common distribution

$\beta_j \sim N(\mu, \tau^2)$

$\tau^2 \to 0$ single regression

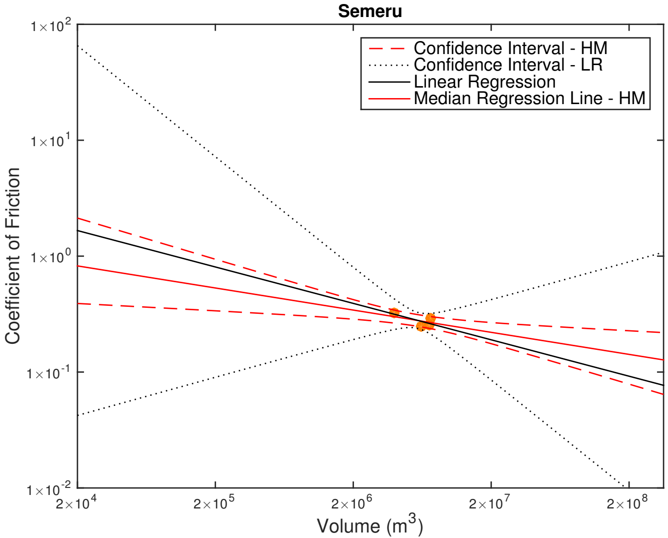
$\tau^2 \to \infty$ separate regressions

Montserrat



Soufriere Hills
Confidence Interval - HM
Confidence Interval - LR
Linear Regression
Median Regression Line - HM
Coefficient of Friction
1×10^2
1×10^1
1×10^0
1×10^-1
1×10^-2
2×10^4
2×10^5
2×10^6
2×10^7
2×10^8
Volume (m^3)

Montserrat



Ogburn *et. al. Journal of Statistics in Volcanology* 2016