Volcanoes, Earthquakes, and Deposition as Intensity-Driven Processes

(with a focus on Volcanoes)

Christopher Harper and Leif Karlstrom

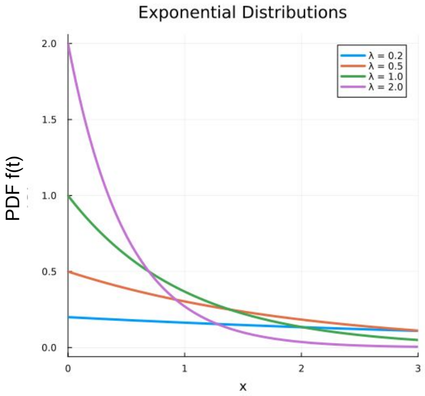
Motivation within an SZ4D context

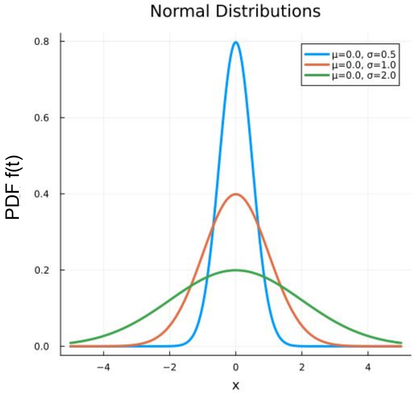
A Need for Common Statistical Framing in SZ4D

- SZ4D is concerned wide range of processes: **volcanism**, **seismology**, and **surface processes.**
- Each of these systems has been modeled with many valid, but different statistical approaches.
- This results in similar concepts sometimes used in incompatible ways, causing ambiguity and confusion.
- Record Completeness is a great example:
  - **sedimentologist:** Completeness = fraction of time preserved in the stratigraphic record; missing time due to erosion or hiatus
  - **seismologist**: Completeness = probability of detecting an event above a magnitude threshold; missing events, not time.
  - To a **volcanologist** it's both!
- Problems like these could be avoided if we had a shared framework, and similarities between the probabilistic (non-deterministic) elements of our system would be emphasized
- The framework we propose is based on the notion of intensity.
- **How does adding another "framework" to the pile accomplish this?**

The Intensity Framework Is Uniquely Well Situated

- Each system will have random variables that capture the probabilistic behavior
  - **Interarrival Times *T:*** The time between eruptions, earthquakes or depositional periods (hiatuses)
  - Distance between vents or epicenters, alignment of dikes, thickness of deposition
- Random variables are usually thought of in terms of Probability Density Functions (PDF)
  - Any Random Variable that has a PDF has an intensity
- **An intensity emphasizes how risk builds through space/time within a single event**; a PDF suppresses that very information to emphasize the final outcomes across many events
- This is particularly important for Eruptive processes and Earthquakes, where the physics driving the process is unfolding between events.
- An Intensity framework has many advantages for information poor or unreliable records





What is an Intensity?

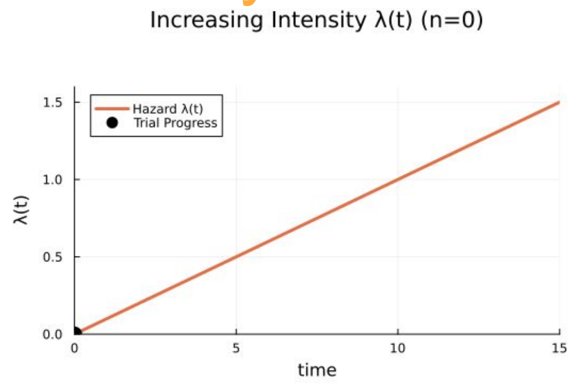
Three Examples

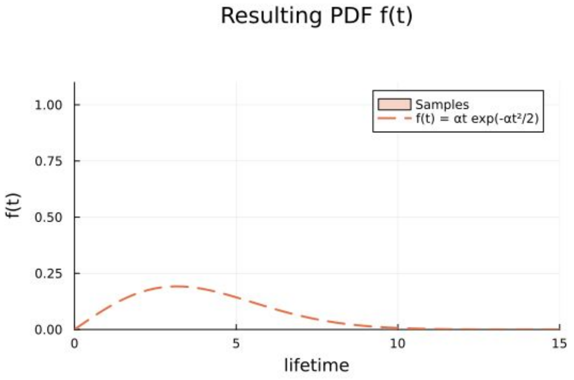
Great, but what **is** Intensity: Increasing Intensity

- Intensity is the instantaneous rate of “occurence” given that the event has not yet occurred.

$$\lambda(t) = \lim_{dt \to 0} \frac{\mathbb{P}\{T \in [t, t+dt) \mid T > t, \mathcal{H}_t\}}{dt}$$

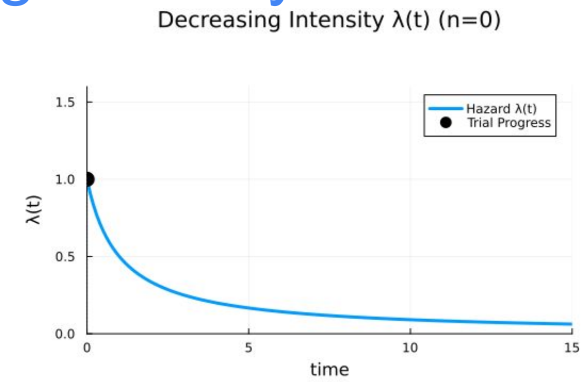
- We categorize an increasing intensity process as “*aging*” .
- This creates a central mode .
- The shape of failure in one event drives the overall shape of the distribution.

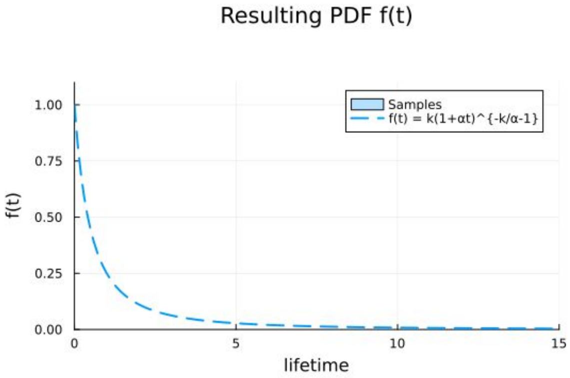




Great, but what **is** Intensity: decreasing intensity ?

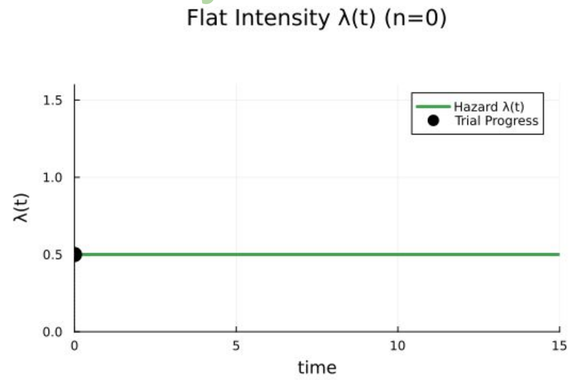
- Different shaped intensity means different pdf
- We categorize a decreasing intensity process as "*clumpy*".
- This creates a threshold before which you are more likely than random to fail, and after your are less likely.
- Both examples are independent in interarrival times (T):
    - The intensity curve does not change between failures.

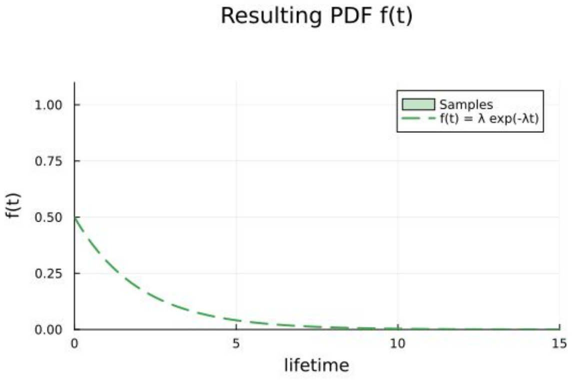




Great, but what **is** Intensity: Constant Intensity

- Flat intensity through time produces exponential distribution for failure
- A flat intensity is the only one independent in number of events N(t):
  - Point processes built from the interarrival times of the previous two slides would be dependent.
- This question of what **is** and **is not** independent, motivates the need for a clear and uniform framework.





Renewal Process: A special type of intensity (I.I.D)

- If the intensity does not change between events we have a **renewal process**, and our interarrival times are Independent Identically Distributed (IID) random variables

- Renewal assumption is important for this work, and while many distributions can be formed from IID interaarrivals, many important ones cannot.
  - **Some Distributions formed from IID T:** All families of Weibull , Gamma, Lognormal, and their mixtures
  - **Some Distributions Which do NOT satisfy renewal process:** non stationary Poisson (Hawkes) and Markov modulated processes (hidden states)

Renewal Assumption allows us to link Infensity and PDFs

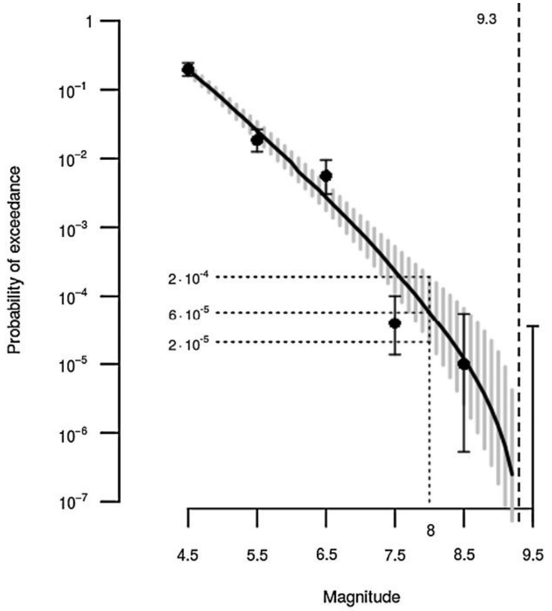
CDF (F):	$F(t) = P(T \leq t)$
Survival Function (S):	$S(t) = P(T > t) = 1 - F(t)$
Accumulated Intensity (Λ):	$\Lambda(t) = \int_0^t \lambda(u)\, du$
Map from S to Λ:	$S(t) = e^{-\Lambda(t)}$
Map from f to λ:	$f(t) = \lambda(t) S(t) = \lambda(t) e^{-\Lambda(t)}$

Next we show this frameworks utility through an application with the volcanic record.

An Application to Volcanic Systems

Intensity Framework Allows Us to Do Regional Statistics

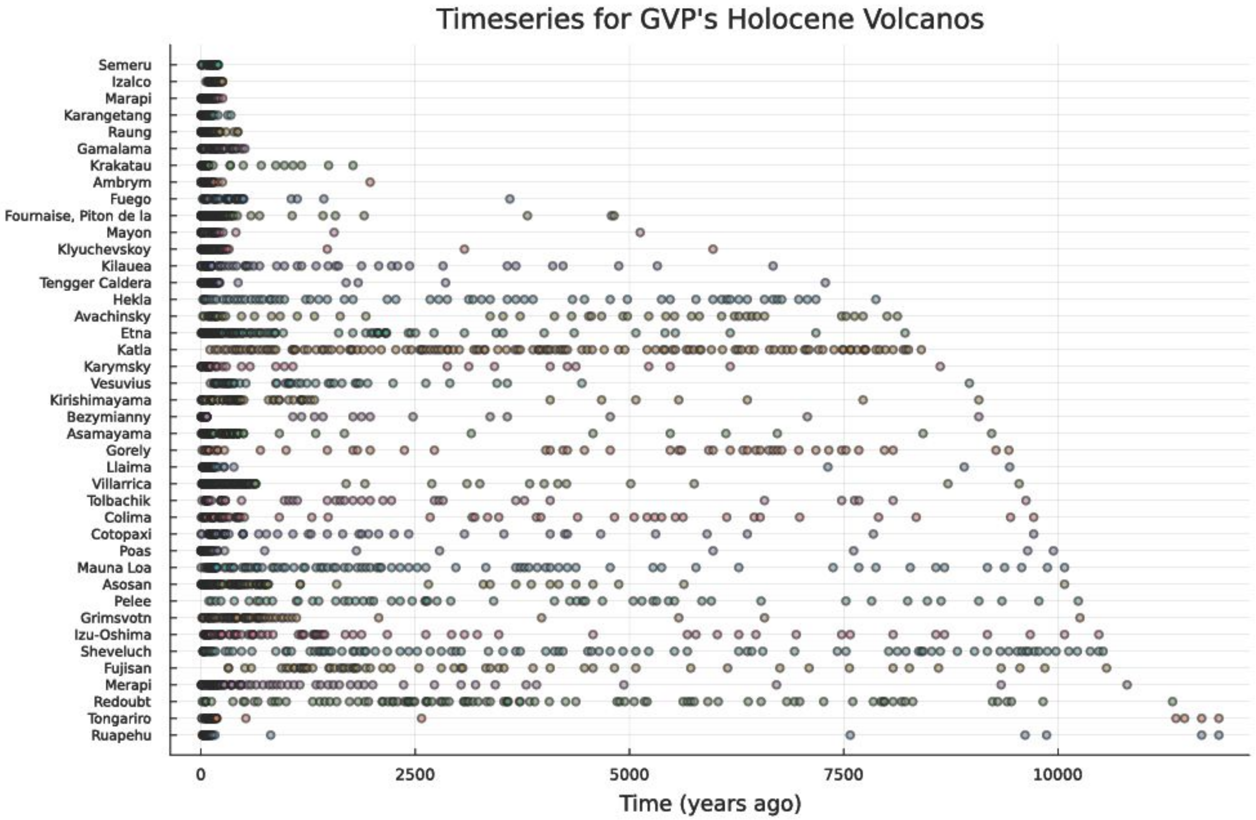
- **Sparse Data:** Aggregating globally, Poisonian behavior is recovered (Papale 2018).
- We can even recover a magnitude frequency relationship, but ~40 million entries are needed for completeness (Papale and Marzocchi, 2021).
- Many interesting questions are only resolved on the sub "the entire earth" scale.
- Non-Poissonian nature of individual volcanic records is well-documented (e.g., Cole and Sparks, 2006).
- Even testing the null hypothesis of Poisson is beyond many records capabilities.
- **We provide a regional characteriszation of volcanic records** which rejects H_o = Poisson **AND** provides an alterative H_1.
- This Characterization does NOT address magnitude.



(Rougier et al., 2017)

Smithsonian Gloval Volcano Project (GVP)

- Time Series of GVP’s most active holocene volcanos
- highlights the catalog's extreme heterogeneity
- remarkably varied eruptive behavior across systems
- the record exposes a strong preservation bias
- **What portion of the record do we use?**



Freedman-Diaconis as a Measure of record quality

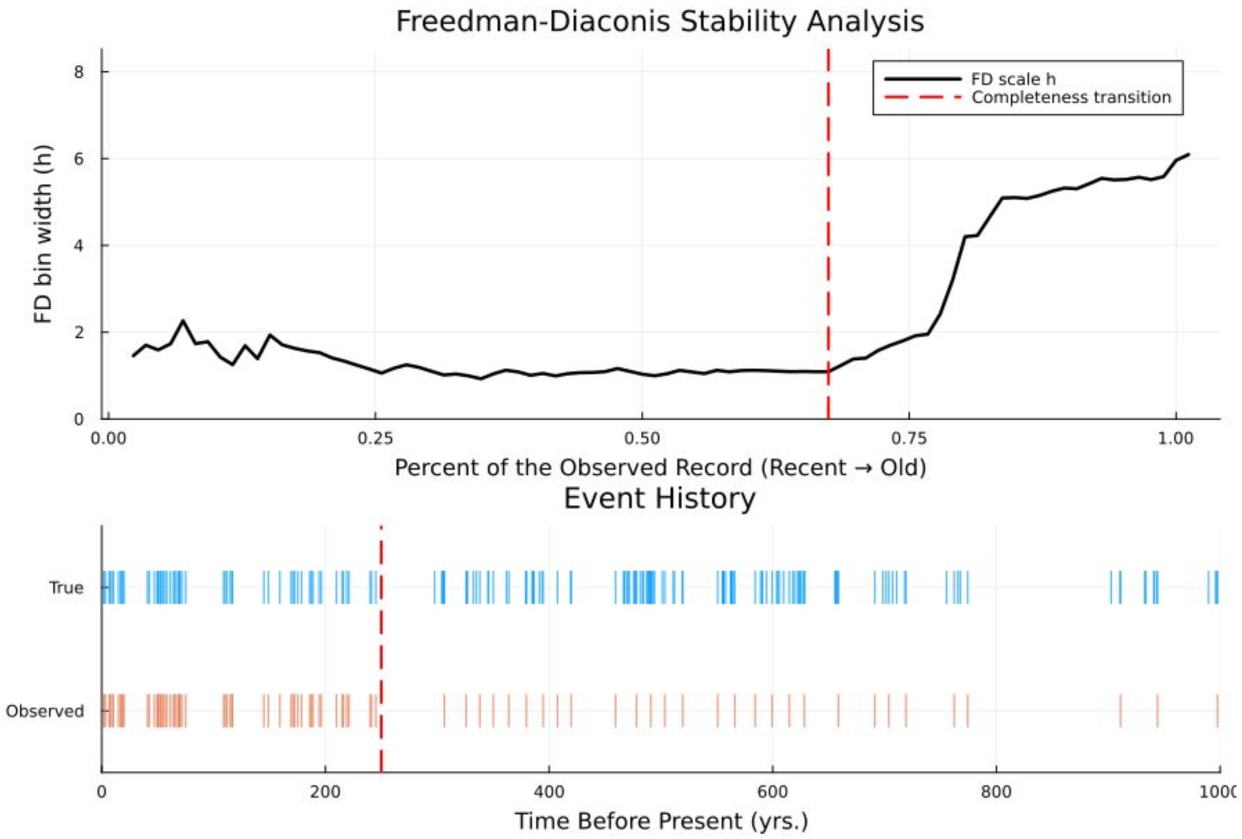
- In looking for portions of the record which satisfy the **renewal assumption** we manage to address many of the underlying issues with the record:
  - Stationarity
  - Reliability of Recording
  - Preservation Bias (more likely to be missing events as you go further back in time)
- To test what portion of a record (if any) exhibits the renewal assumption we use a sability measure produced by the freedman-diaconis bin-width *h*

$$\underbrace{\frac{1}{kh}}_{bias} + \underbrace{\int_I \left(f_h(x) - f(x)\right)^2 dx}_{variance}$$

- More than just a histogram binwidhth picker it finds the scale at which signal can be separated from noise!

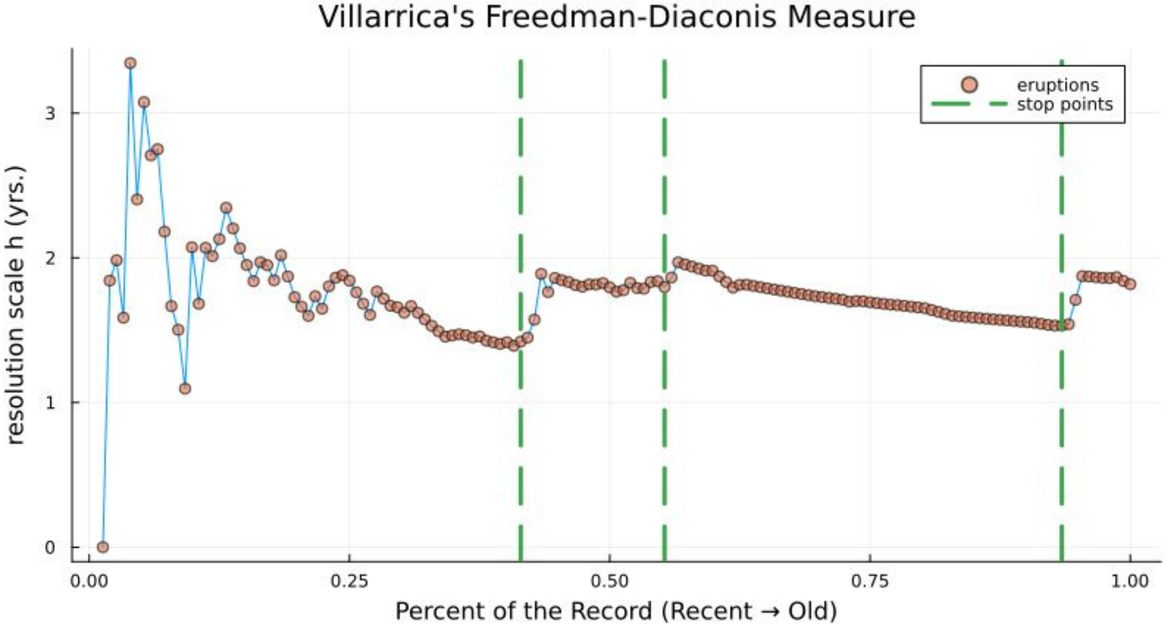
Freedman Diaconis: A synthetic Demonstration

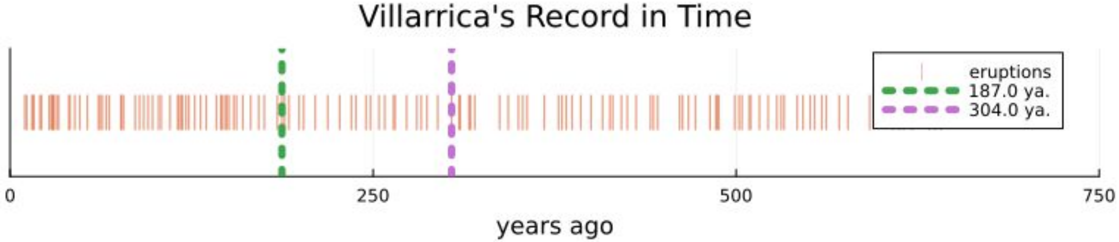
- True clustered process (blue)
- Sampled geologic record (orange)
- FD selects the appropriate observational scale
- Detects shifts in record completeness / quality
- FD lets you know when and at what scale you can incorporate a less reliable portion of the record
- Indicates when and where the record remains usable



Freedman Diaconis: Villarrica’s Record

- 3 “stable” portions detected
- Zooming in reveals differences:
  - Dense Bands
  - Regular Occurences
  - Bigger gaps
- We choose the first portion to be commensurate with current recording practices

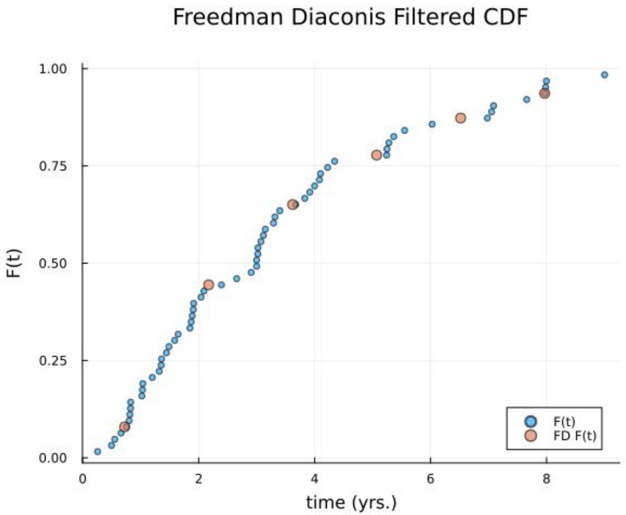


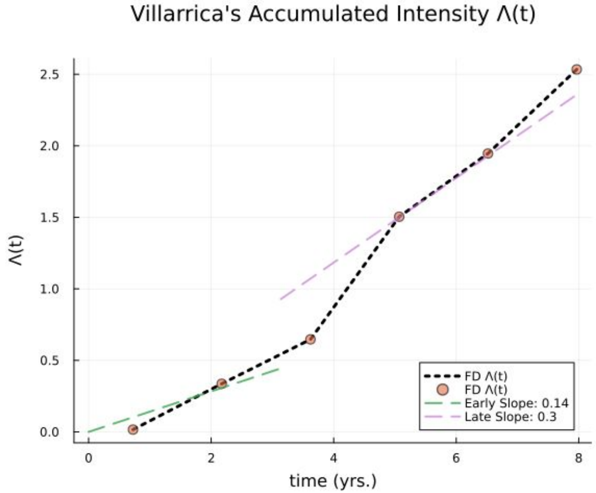


From Record to Λ(t): Villarrica

- CDF is pointsampled at *h*
- S = 1 - CDF, and Λ(t) = -log(S)
- Characterization is based off of a global description of Λ's concavity
- This is accomplished through a normalized relative difference of an early (⅓) and late (⅔) slope
- We call this the Concavity Index CI

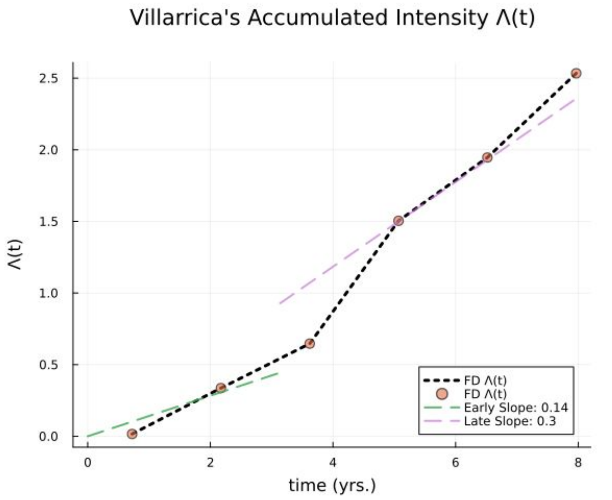
$$CI = \frac{(\text{late slope - early slope})}{(\text{late slope + early slope})}$$

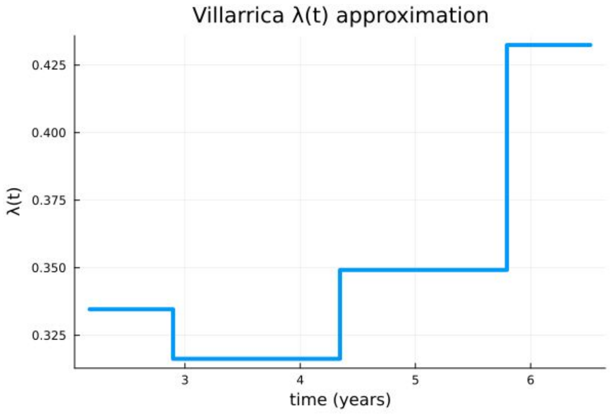


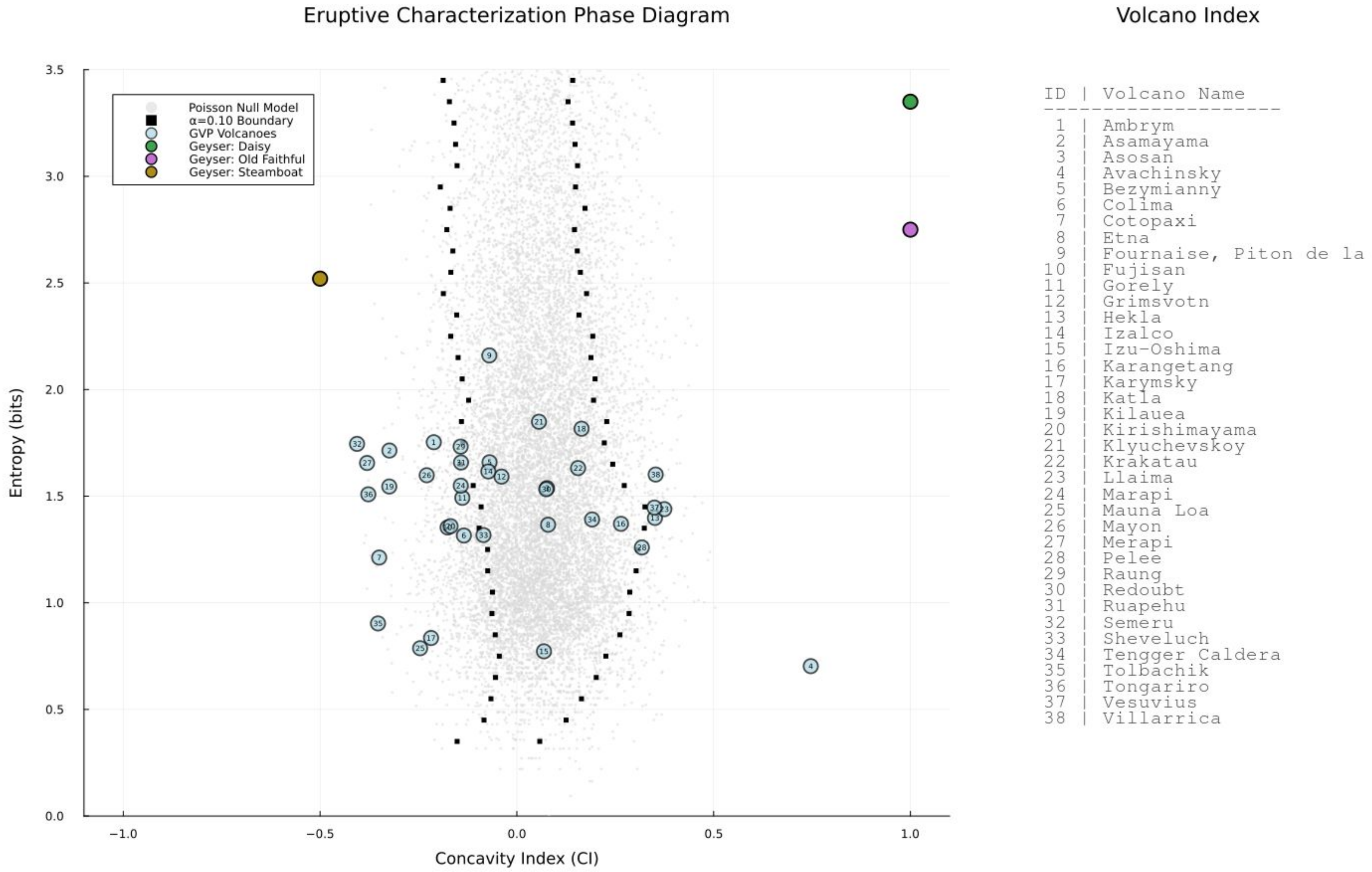


Concavity of Λ(t) as a Characterization of behavior

NR-Slope Difference	Concavity	Character	Forecasting Implications
Values from -1 to 0	Concave Up (+)	Clumpy	Clustering timescale
Values near 0	Linear (0)	critical state/random	No structure. Simple average is your best predictor
Values from 0 to 1	Concave Down (-)	Aging/regular(ish)	Quasiperiodic timescale






Eruptive Characterization Phase Diagram
Poisson Null Model
α=0.10 Boundary
GVP Volcanoes
Geyser: Daisy
Geyser: Old Faithful
Geyser: Steamboat
Entropy (bits)
Concavity Index (CI)
Volcano Index
ID | Volcano Name
1 | Ambrym
2 | Asamayama
3 | Asosan
4 | Avachinsky
5 | Bezymianny
6 | Colima
7 | Cotopaxi
8 | Etna
9 | Fournaise, Piton de la
10 | Fujisan
11 | Gorely
12 | Grimsvotn
13 | Hekla
14 | Izalco
15 | Izu-Oshima
16 | Karangetang
17 | Karymsky
18 | Katla
19 | Kilauea
20 | Kirishimayama
21 | Klyuchevskoy
22 | Krakatau
23 | Llaima
24 | Marapi
25 | Mauna Loa
26 | Mayon
27 | Merapi
28 | Pelee
29 | Raung
30 | Redoubt
31 | Ruapehu
32 | Semeru
33 | Sheveluch
34 | Tengger Caldera
35 | Tolbachik
36 | Tongariro
37 | Vesuvius
38 | Villarrica

Deposition as an Intensity Process

Deposition as an Intensity Process (Schumer & Jerolmack 2008)

- Let T be a hiatus duration with intensity (t)
- Intensity describes how deposition likelihood evolves through time:

$$\lambda(t) = \frac{\alpha\gamma}{1+\alpha t}$$

- Long Hiatus implies “clumpy” process, tail exponent emerges as

$$\gamma = \lim_{t\to\infty} t\lambda(t)$$

- $$C(t_{obs}) = V_{obs}/V$$

$$C(t_{obs}) \propto t_{obs}^{\gamma-1}$$

Conclusion

- Characterizing the intensity emphasizes the accumulation of risk between events as opposed to a PDFs emphasis on how that risk affects outcomes across many events.
- If you can't curve fit a standard distribution curve, consider characterizing a more general description of intensity.
  - Despite Making Conservative choices at each step in the volcanic analysis, we were still left with meaningful information
- First order behavior of the intensity curve affects the pertinent "hazard question"

Thank You!