Shaking Things Up: Statistical Modelling of Earthquakes

RSS Avon Local Group Meeting

Zak Varty

Talk Structure

1. Seismology 101 & Earthquake Data
2. Building Blocks of Statistical Seismology
3. Current Work
4. Future Directions

Aim: Know where we’re at, where we’re going, how to join in.

1: Seismology 101 & Earthquake Data

What is an Earthquake?

What is an Earthquake?

Vibrations in the ground:

- Jumping up and down
- Jack-hammer
- Traffic and Trains
- Sea Storms

Image by wirestock on Freepik

Natural Seismicity

- Tectonic motion
- Volcanic activity
- Landslides

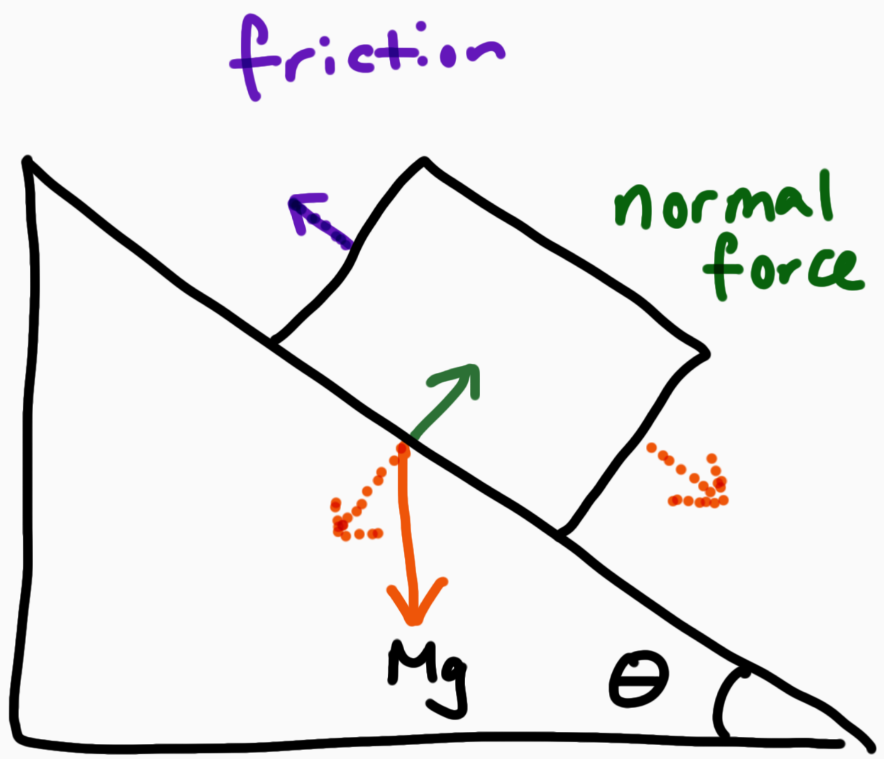
Volcanic Eruptions in Iceland. Image source - ABC.

Induced Seismicity

- Pop concerts
- Nuclear tests
- Mine blasts
- Fluid injection / extraction
- Fracking

Lumen Field, Seattle. Image source - NME

Physics Recap



Cats on a Roof

What is a Fault?

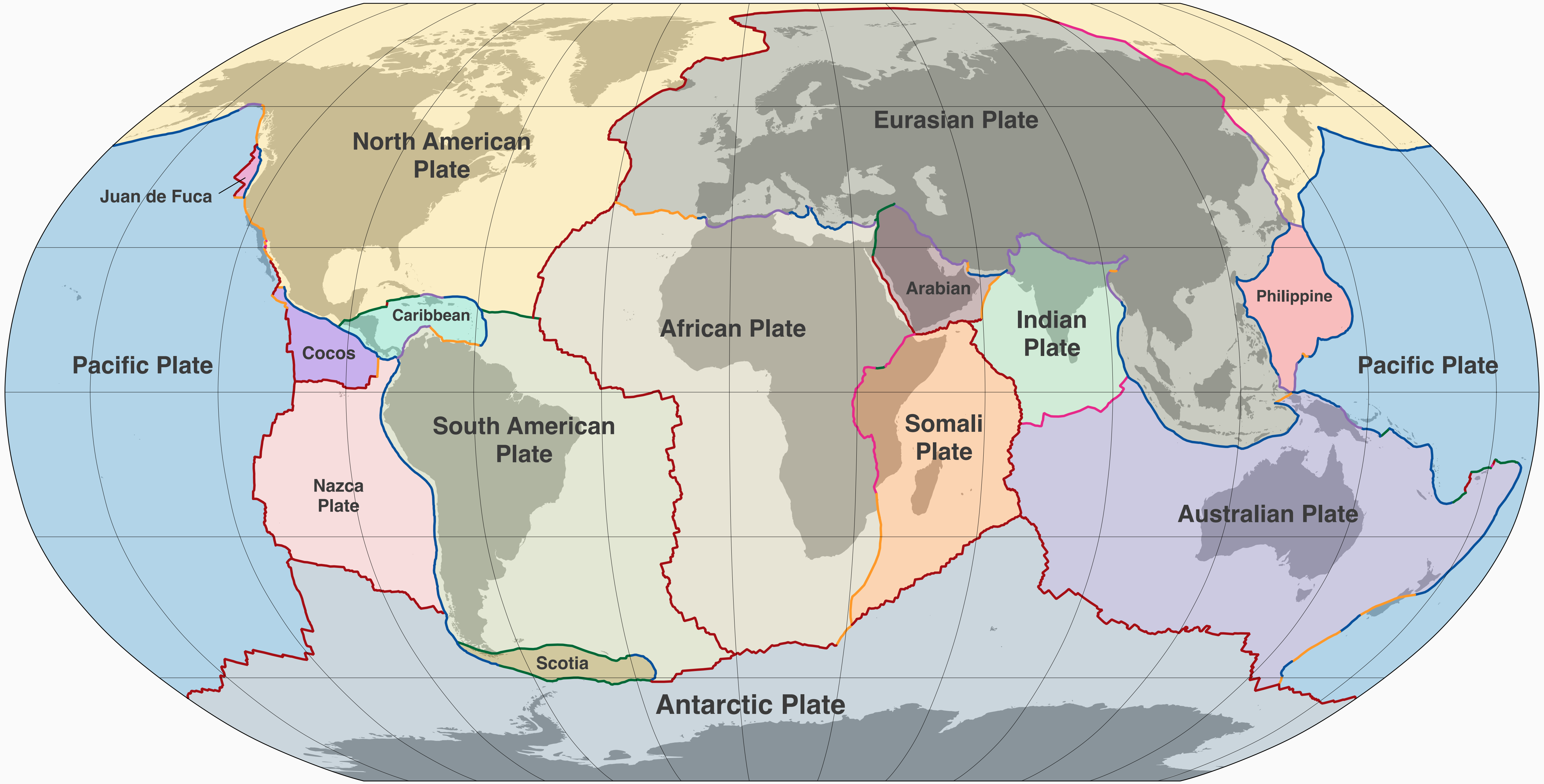
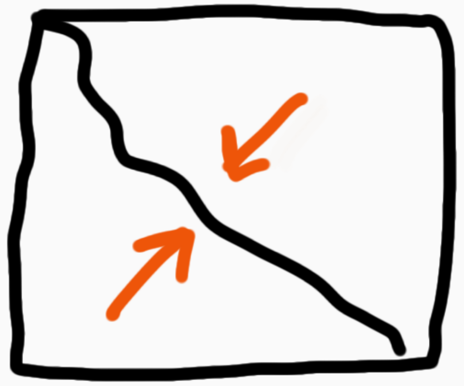
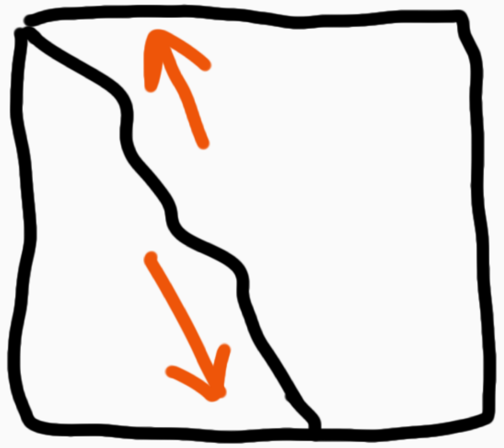


Image Source - Wikimedia

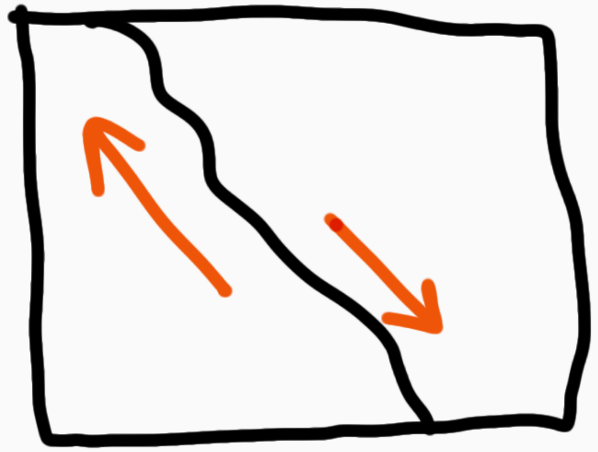
Earthquakes as Fault Slip



normal force



driving force



friction

$$\mu_s > \mu_d$$

Fault Slip

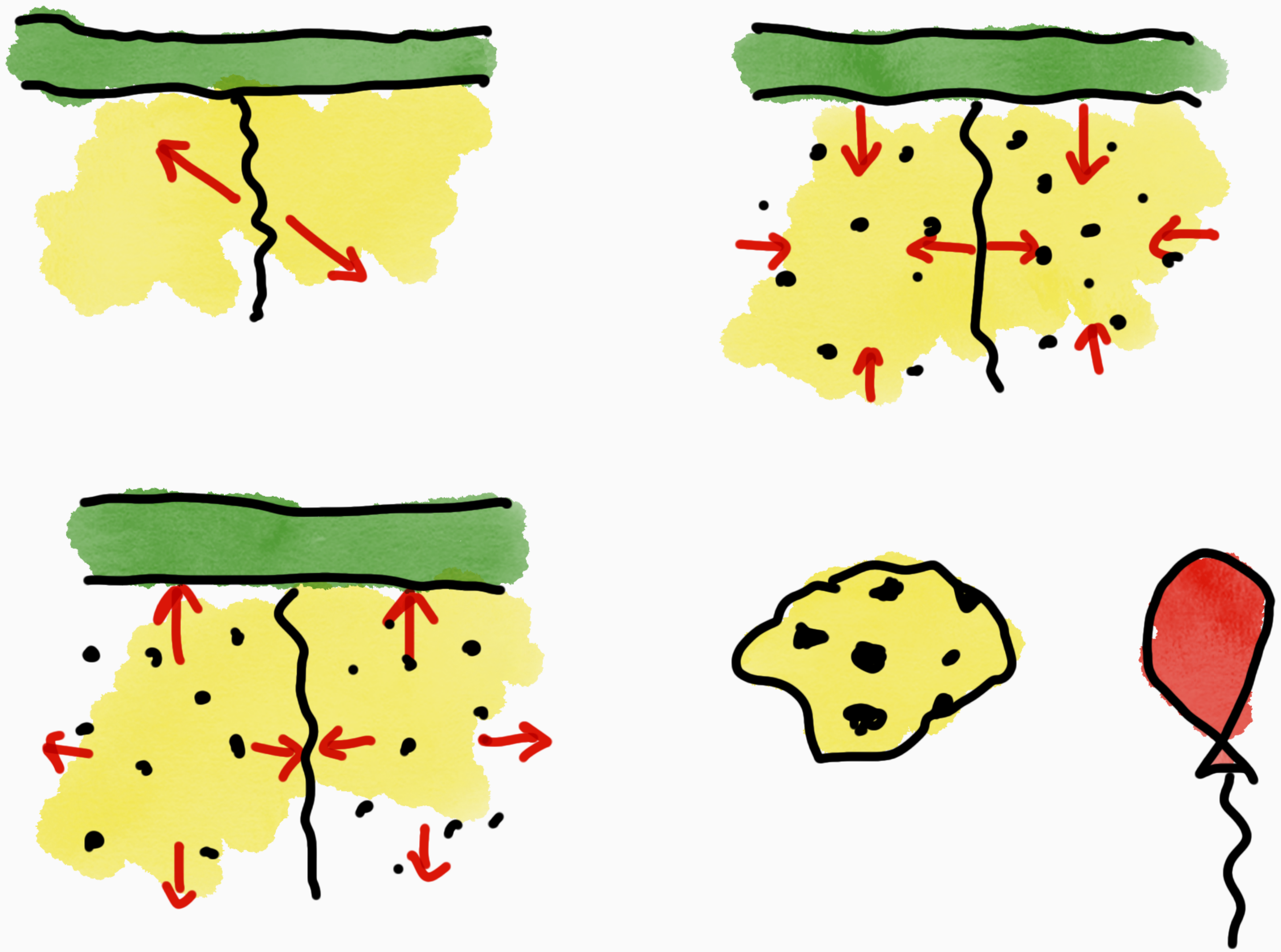
How does it happen?

- Energy stored in rock under tension.
- Movement when driving force exceeds static friction.
- Depends on a lot of factors:
  - Material constants
  - Slip area and distance
  - Fault orientation
  - …

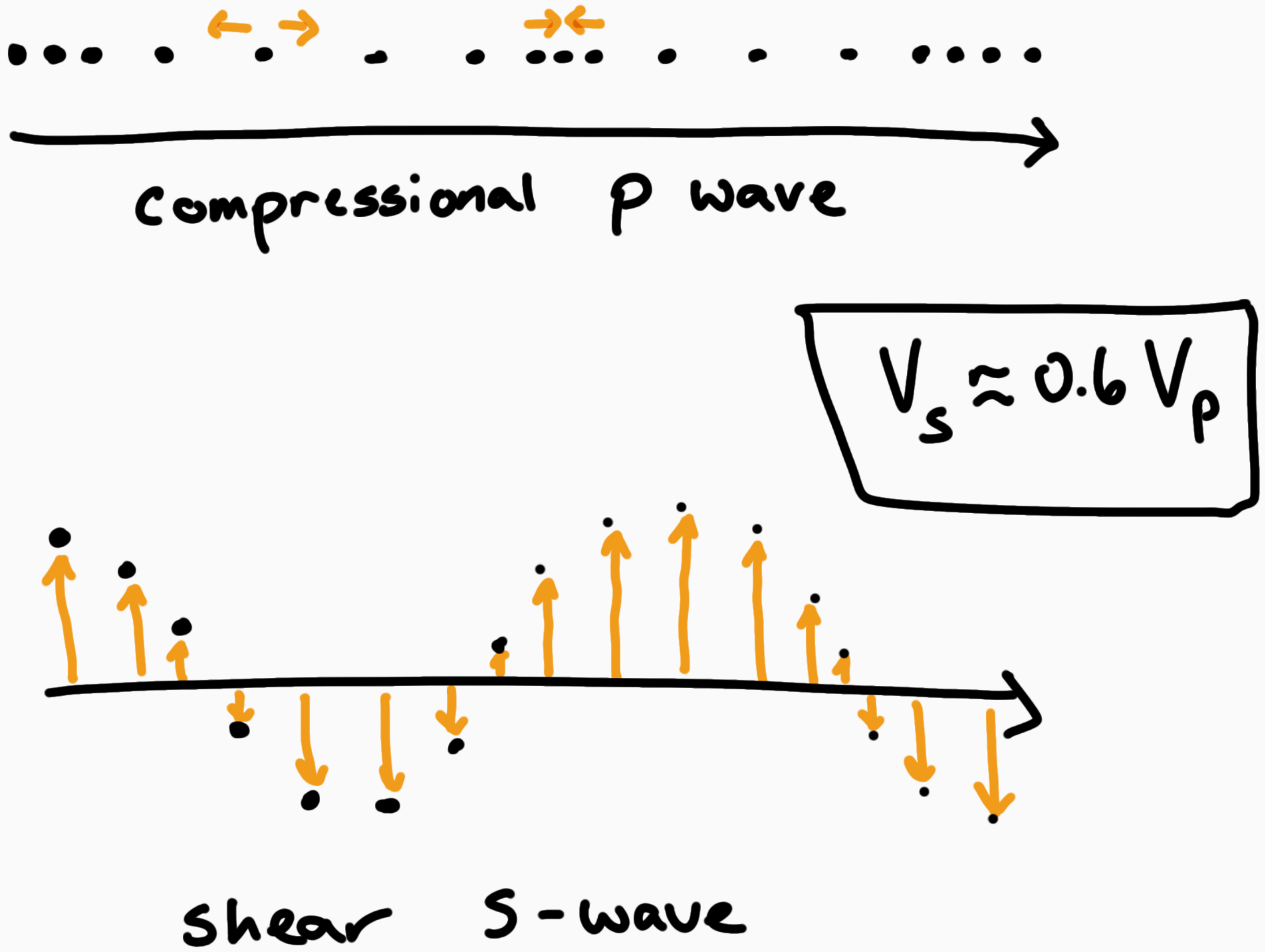
** What do we care about?**

- Total energy released
- Power (W = J/s)
- Aseismic creep

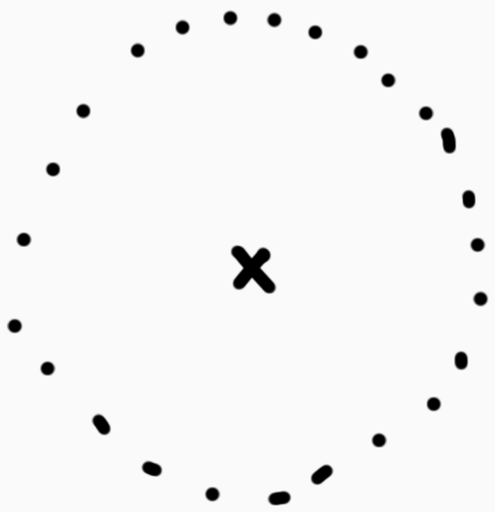
What provides the driving force?



How do we measure earthquakes?

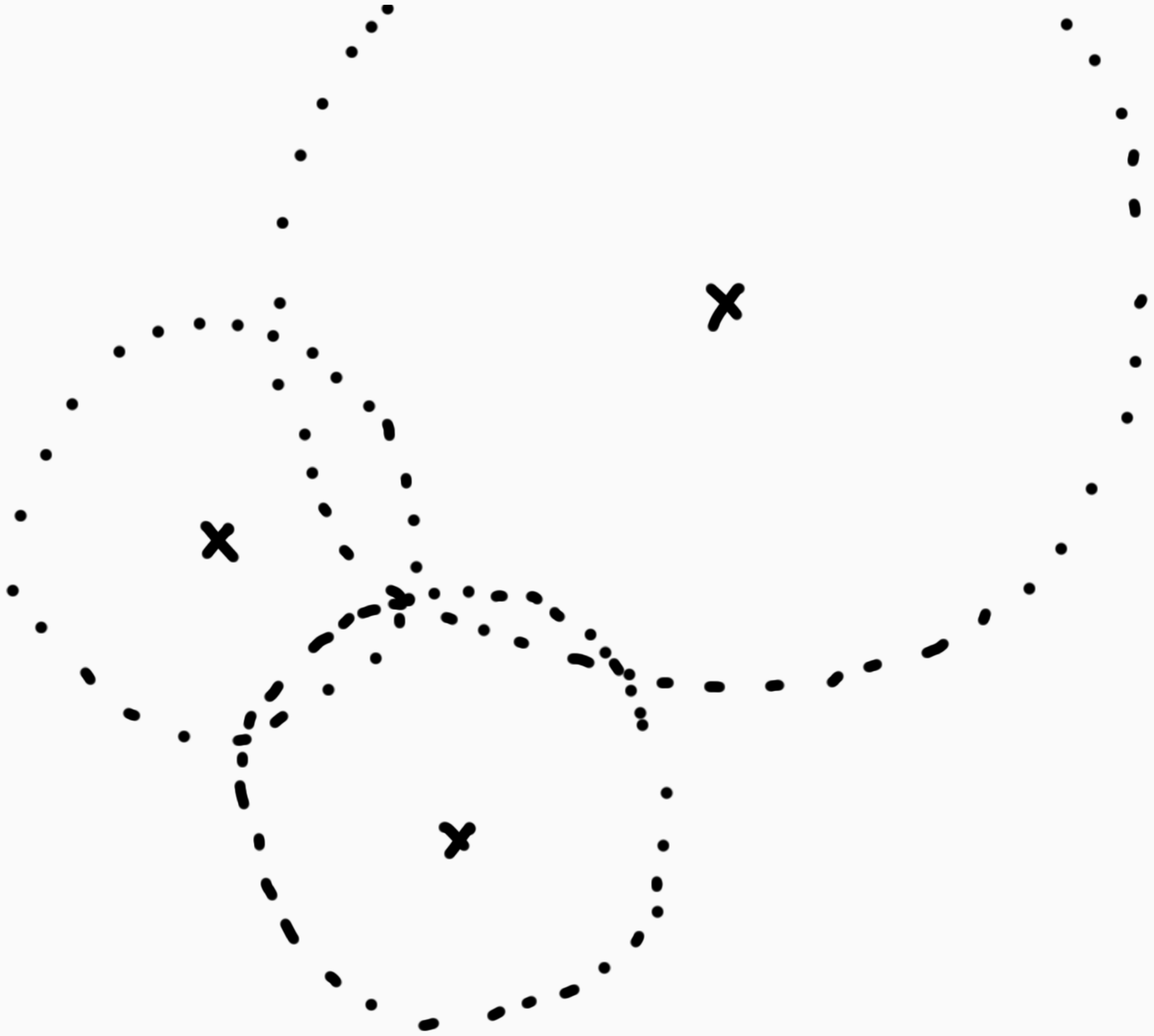


How do we locate earthquakes? (1)

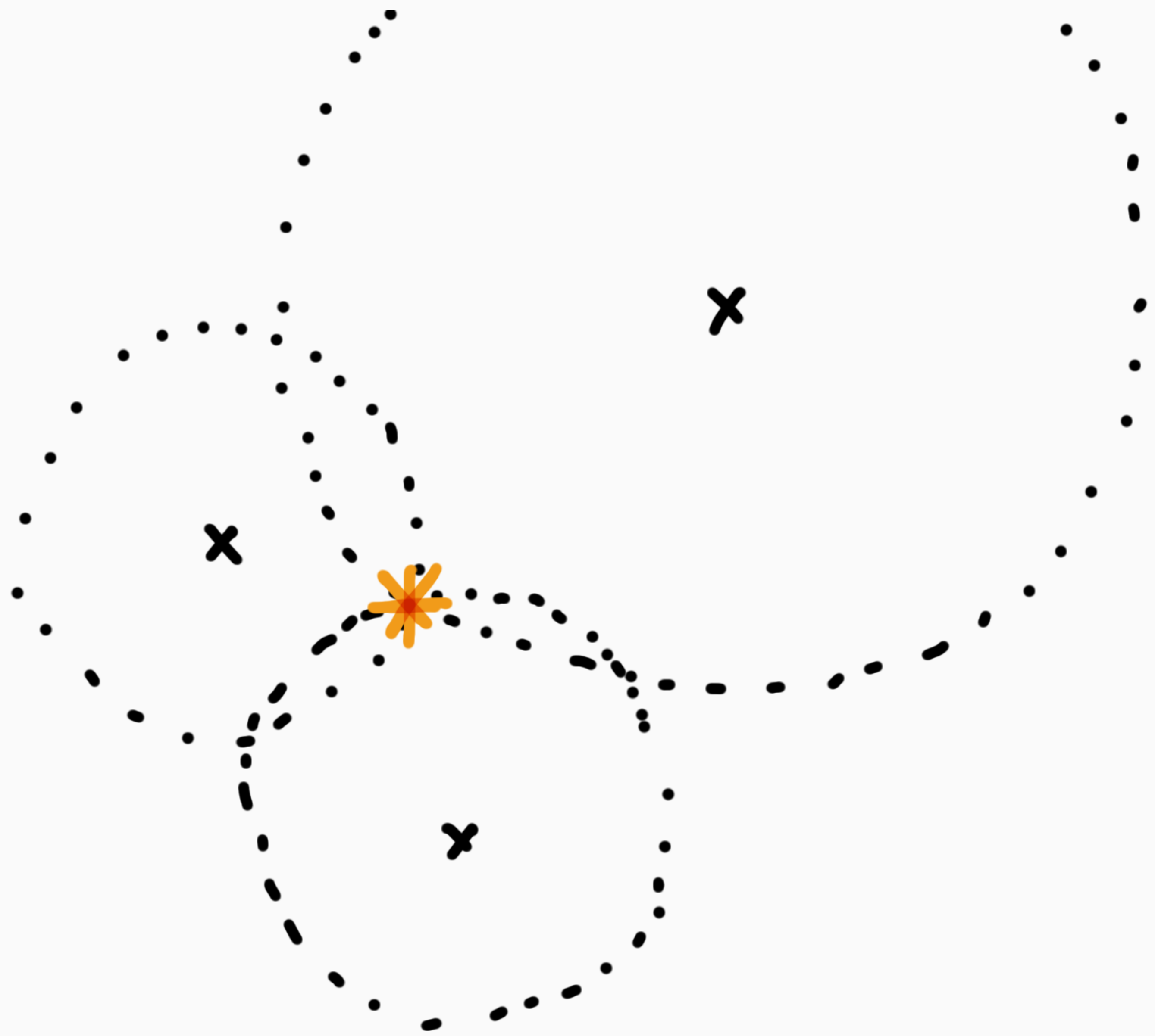


How do we locate earthquakes? (2)

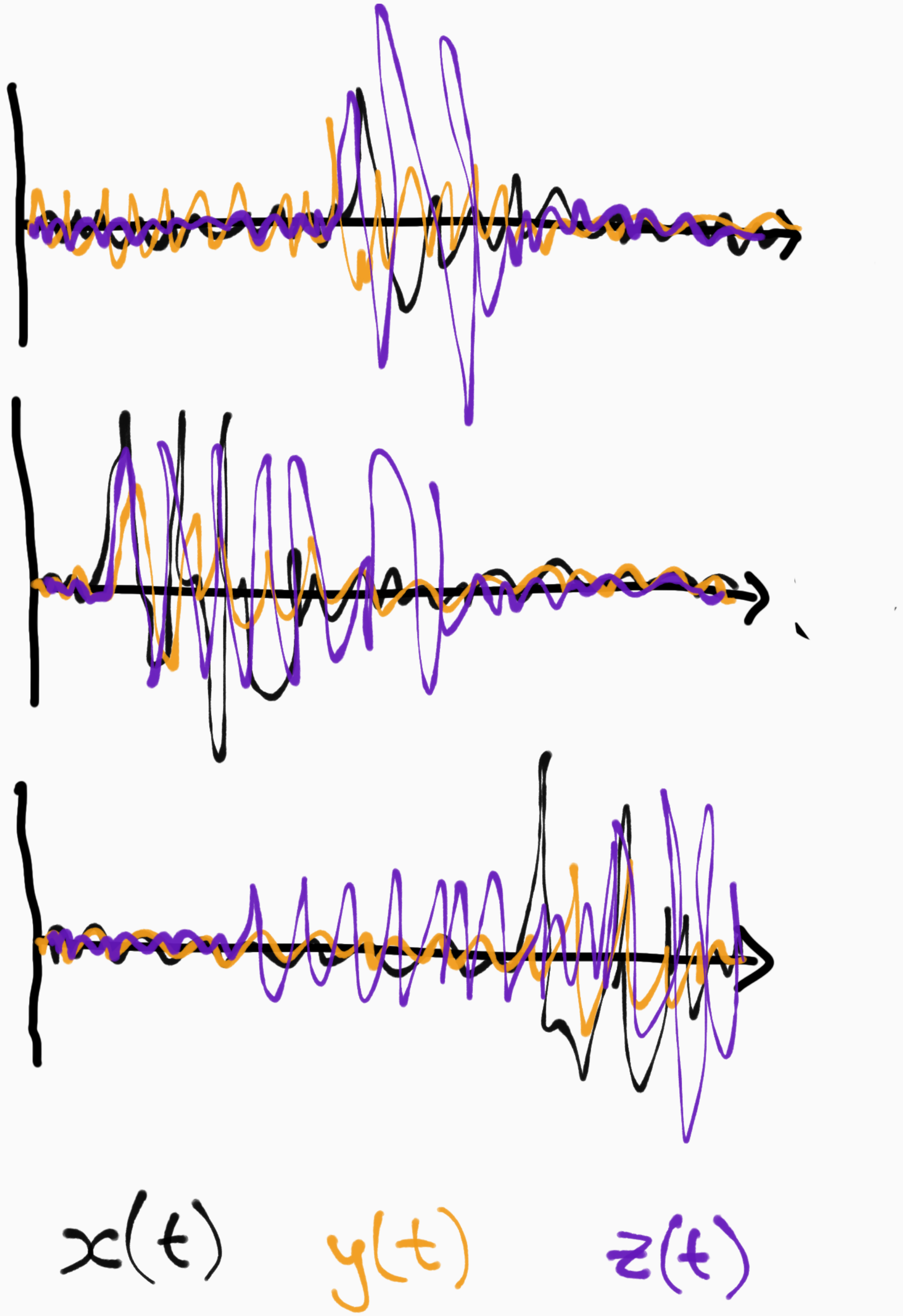
How do we locate earthquakes? (3)



How do we locate earthquakes? (4)



It’s a bit more complicated than that



- 3 Dimensions
- Energy dissipation
- Background noise
- Material Inhomogeneity
- Reflection, refraction and interference
- Expert interpretation
- DL: Yoon et al (2023)

Earthquake Catalogues

- **x** Easting / Longitude
- **y** Northing / Latitude
- **z** Depth (often nominal)
- **t** Time (YYYY-MM-DD HH:MM:SS)
- **m** Magnitude

A note on Magnitudes

Magnitudes are measured on logarithmic scales and typically reported to 1 decimal place.

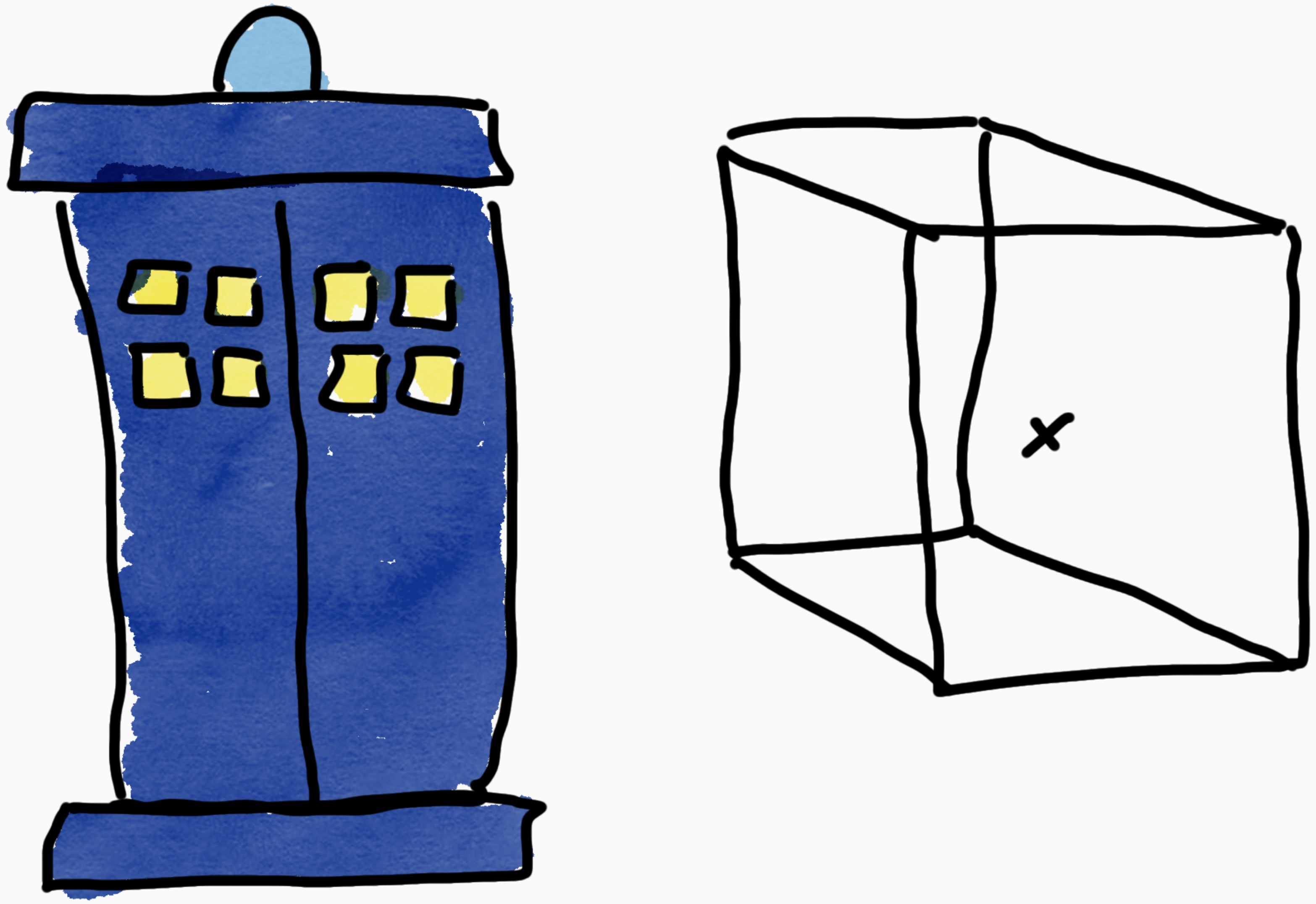
$$\text{Richter Scale} \rightarrow \text{Local Magnitude} \sim \text{Amplitude}.$$

$$\text{Moment Magnitude}$$

$$E \propto A^{1.5} \text{ so } 10\times \text{ or } 32\times \text{ increase.}$$

What is an Aftershock?

What is an Aftershock?



2: Earthquake Modelling

Locations: Point Processes

- Stochastic process $\square = \{X_1, \dots, X_N\}$
  - Values represent locations in time / space, number of values N is also random.

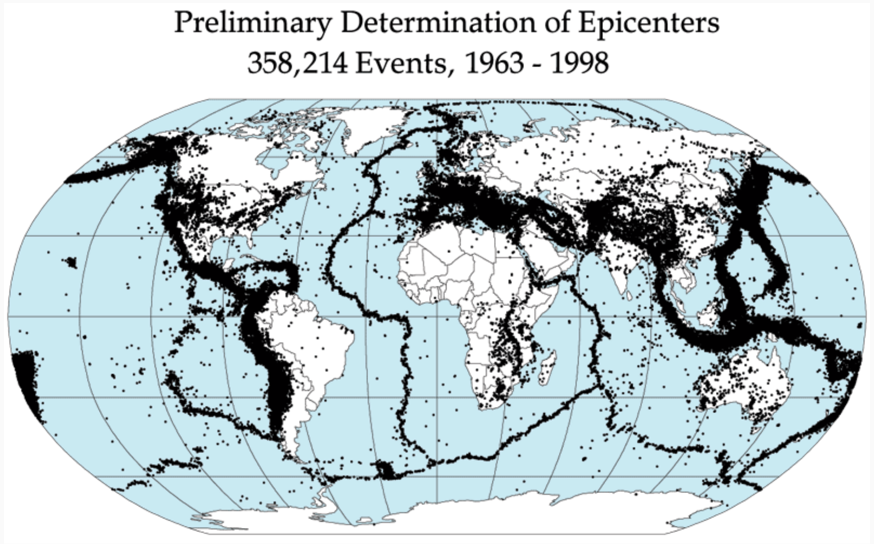


Image source - Wikimedia

- Homogeneous Poisson Process on A:
  - $N(A) \sim \text{Pois}(\lambda|A|)$
  - $(X_i | N(A) = n) \overset{\text{i.i.d}}{\sim} \text{Unif}(A).$

Locations: Inhomogeneous Poisson Processes

Introduce an intensity function $\lambda(a) : A \to \mathbb{R}_0^+$.

$$N(A) \sim \text{Pois}\left(\Lambda(A)\right) \quad \text{where} \quad \Lambda(A) = \int_A \lambda(a)da.$$

Then

$$X_i \overset{i.i.d}{\sim} f_X(a) = \frac{\lambda(a)}{\Lambda(a)}.$$

Physically motivated forms for $\lambda(a)$, e.g. Bourne and Oates (2018):

$$\Lambda \propto \exp(b_0 + b_1 z(a))$$

Locations: Adding in Aftershocks

Hawkes Processes add self excitation.

$$\lambda(t; \square_t) = \mu + \sum_{i:t_i<t} \alpha \exp\{-\beta(t - t_i)\}.$$

Locations: Adding in Aftershocks - a better way

$$
\begin{aligned}
\lambda(t) &= \mu + \sum_{i:t_i<t} \alpha \exp\{-\beta(t - t_i)\} \\
&= \mu + \sum_{i:t_i<t} \frac{\alpha}{\beta} \beta \exp\{-\beta(t - t_i)\} \\
&= \mu + \sum_{i:t_i<t} \alpha' \beta \exp\{-\beta(t - t_i)\} \\
&= \mu + \sum_{i:t_i<t} \varkappa(\cdot; \alpha') h(t - t_i; \beta)\}.
\end{aligned}
$$

What about magnitudes?

Gutenberg-Richter Law:

$$\log_{10} N = a - bM.$$

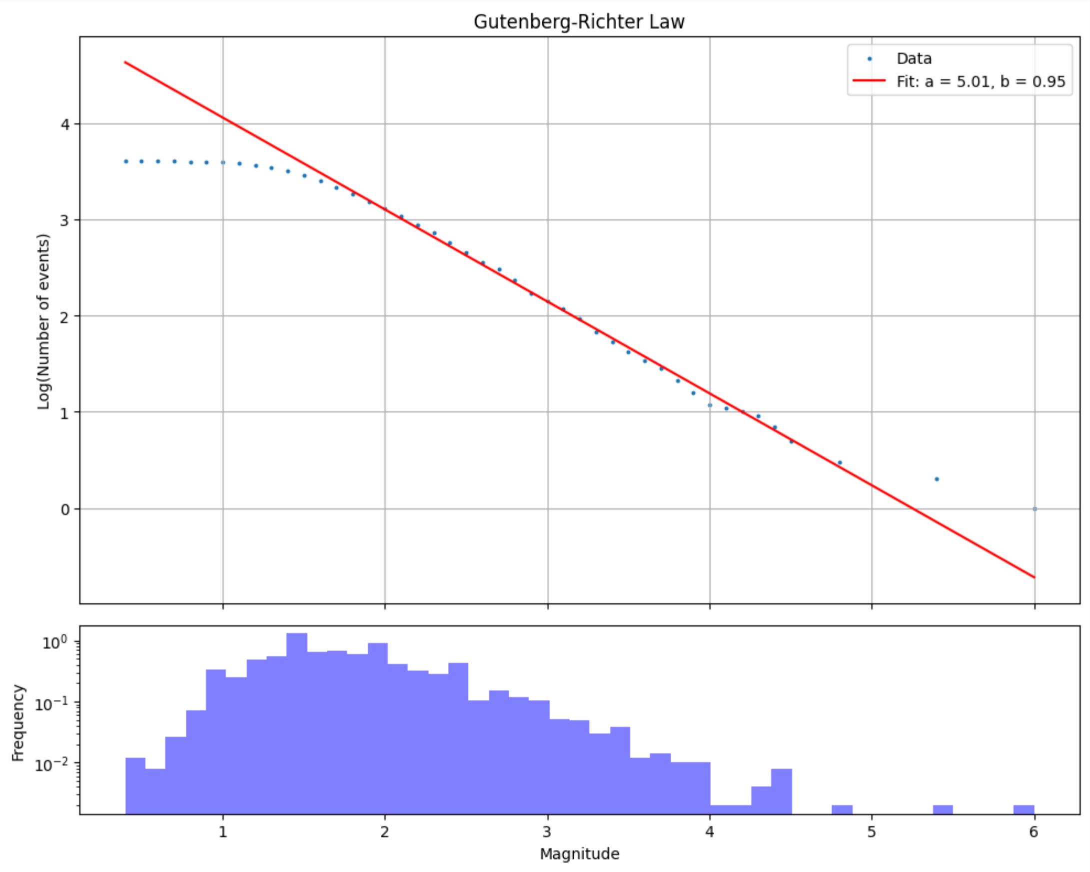


Image source - Wikimedia

For Statisticians:

$$M_i - m_c \,|\, M_i > m_c \sim \mathrm{Exp}(\beta).$$

Biased estimation if rounding ignored.

All Together Now: ETAS model

$$\lambda(t; \square_t, \theta) = \mu + \sum_{i:t_i<t} \varkappa(m_i; K, a) h(t - t_i; c, p).$$

$$\lambda(t; \square_t, \theta) = \mu + \sum_{i:t_i<t} K e^{a(m_i - m_c)} (p-1) c^{p-1} \left(1 + \frac{t - t_i}{c}\right)^{-p}.$$

$$M_i \overset{\text{i.i.d}}{\sim} \mathrm{Exp}(\beta).$$

ETAS Extensions

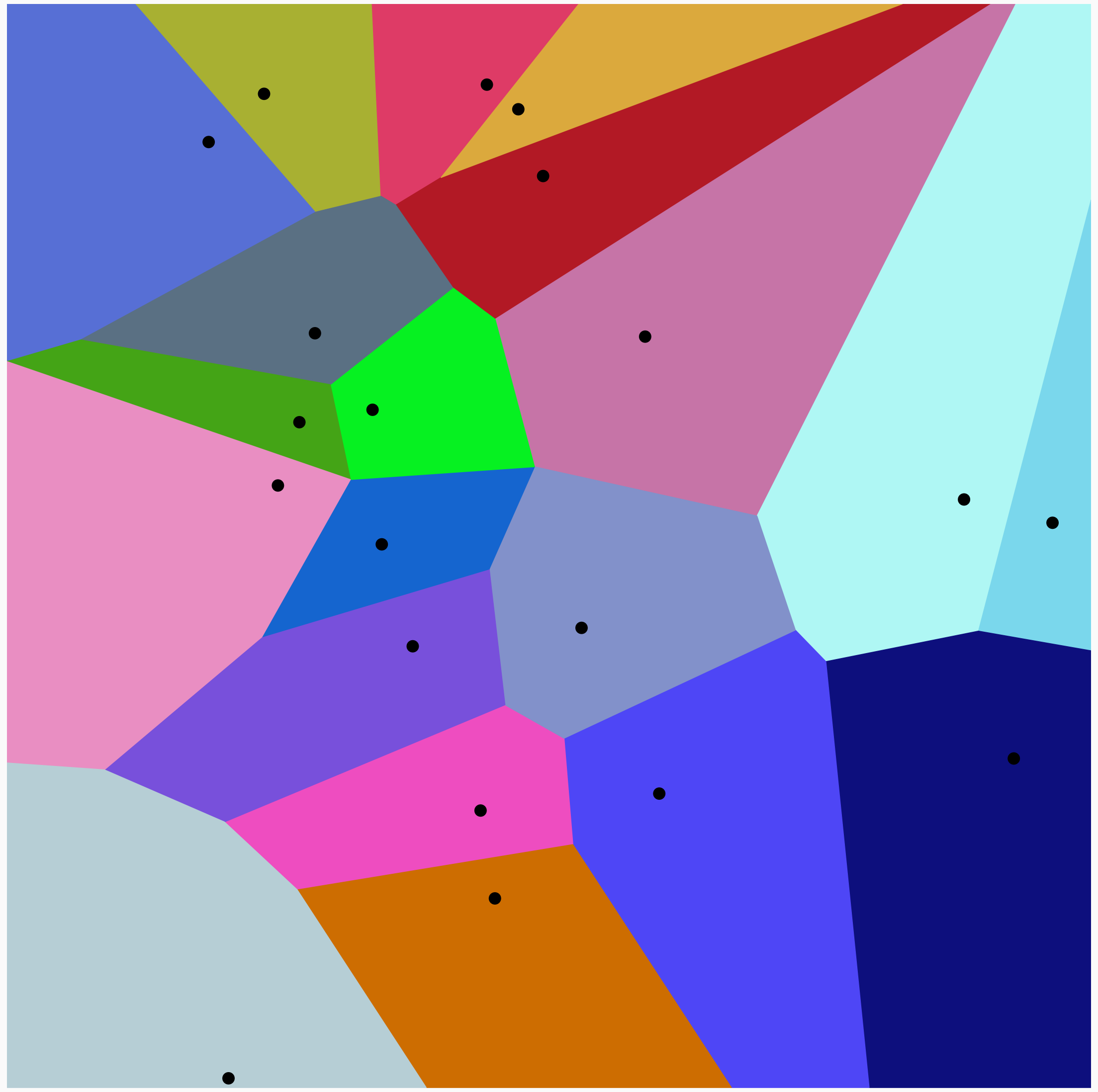


Image source - Wikimedia

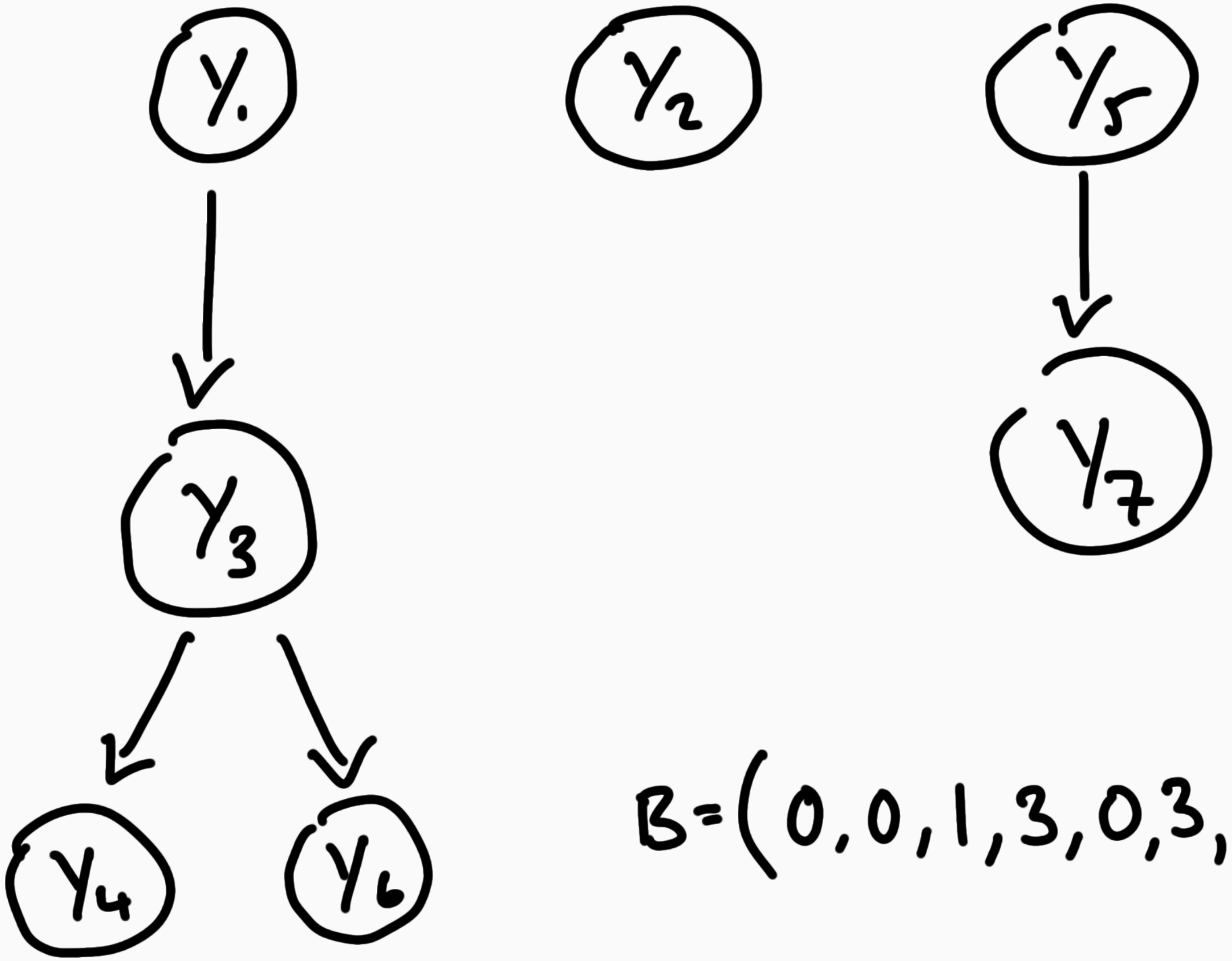
- Voroni Tesselation for μ.
- Covariate-driven background events.
- Flexible regression background (LGP).

ETAS Model Concerns

Veen and Schoenberg (2012) highlight several issues with the ETAS likelihood:

$$\log \pi_{Y|\theta}(y) = -\mu t_{\max} - \sum_{i=1}^{n} \kappa(m_i|\theta_\kappa) H(t_{\max} - t_i|\theta_h) + \sum_{i=1}^{n} \log \left[\mu + \sum_{j:t_j<t_i} \kappa(m_j|\theta_\kappa) h(t_i - t_j|\theta_h) \right].$$

ETAS as a Branching Process



$$B = (0, 0, 1, 3, 0, 3, 5)$$

Conditional Inference

Conditioning on B provides very simple conditional distributions:

$$
\begin{aligned}
\mu | K, a, c, p, B &- \text{Homogeneous PP} \\
K, a | \mu, c, p, B &- \text{Poisson Regression} \\
c, p | \mu, K, a, B &- \text{Power-law} \\
B | \mu, K, a, c, p &- B_i \text{ multinomial}.
\end{aligned}
$$

Recent Developments

Unified Model and Dependence

Following from branching process representation Varty (2021) considers:

- GPD as unified parametrisation.

- Dependence between magnitudes.

Selecting M_c

Varty et al (2021) and Murphy et al (2023)

arXiv:2310.17999v1 [stat.ME] 27 Oct 2023

Automated threshold selection and associated inference uncertainty for univariate extremes

Conor Murphy*, Jonathan A. Tawn
Department of Mathematics and Statistics, Lancaster University
and
Zak Varty
Department of Mathematics, Imperial College London

October 30, 2023

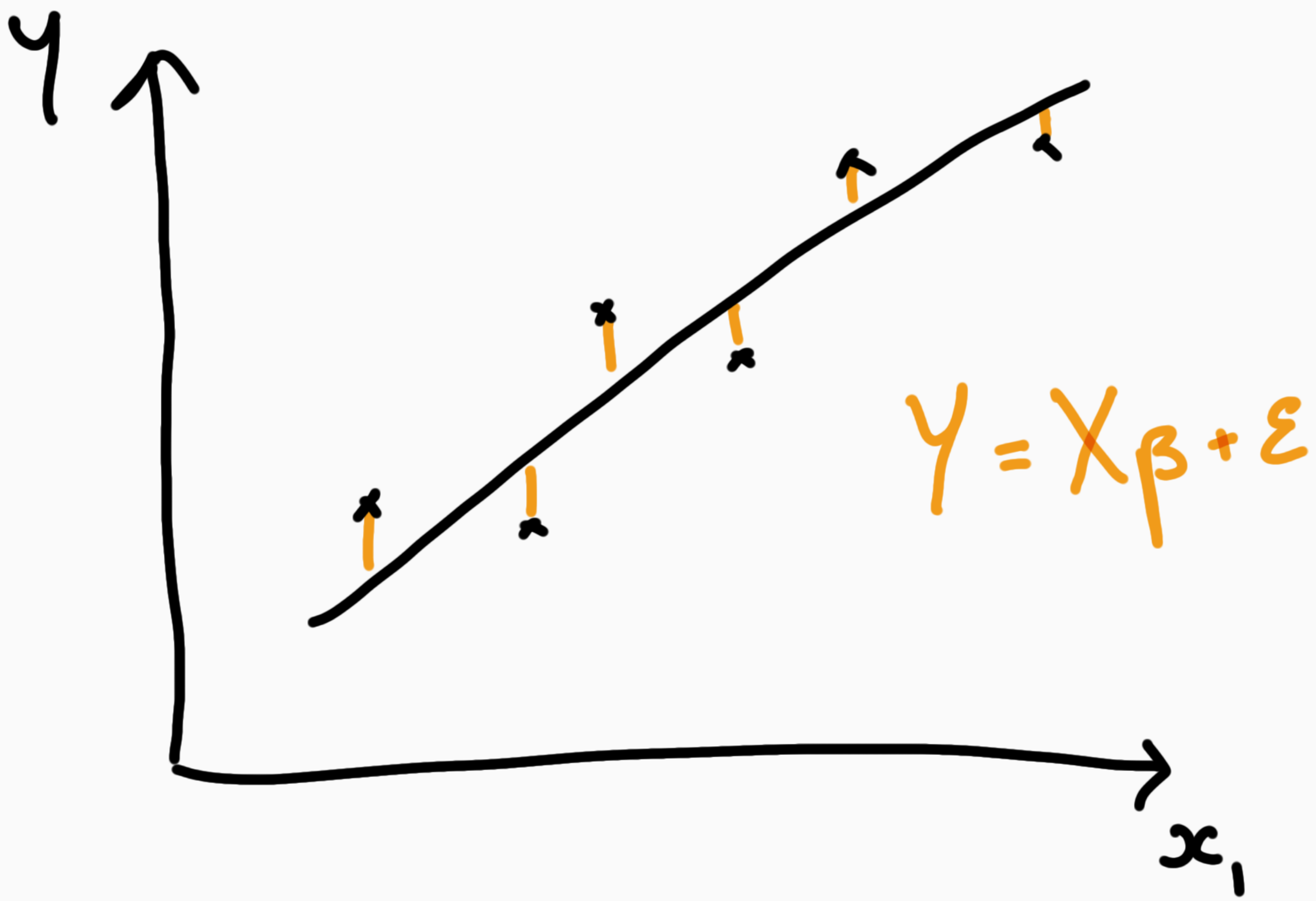
Abstract

Threshold selection is a fundamental problem in any threshold-based extreme value analysis. While models are asymptotically motivated, selecting an appropriate threshold for finite samples can be difficult through standard methods. Inference can also be highly sensitive to the choice of threshold. Too low a threshold choice leads to bias in the fit of the extreme value model, while too high a choice leads to unnecessary additional uncertainty in the estimation of model parameters. In this paper, we develop a novel methodology for automated threshold selection that directly tackles this bias-variance trade-off. We also develop a method to account for the uncertainty in this threshold choice and propagate this uncertainty through to high quantile inference. Through a simulation study, we demonstrate the effectiveness of our method for threshold selection and subsequent extreme quantile estimation. We apply our method to the well-known, troublesome example of the River Nidd dataset.

Keywords: extreme values, generalised Pareto distribution, river flows, return level, threshold selection, uncertainty quantification.

Errors in EV and PP models

There are an awful lot of measurement errors being ignored here… (Yue 2023-2025)



Future Directions

Wrapping Up

1. Earthquakes are significant natural hazards but have received relatively little attention.

2. State of the art models are reasonably accessible.

3. Lots of contributions that could be made to extend these models by statistically minded folks.