Inference for Extreme Earthquake Magnitudes

How to use and learn from small seismic events

Zak Varty, Jonathan Tawn, Peter Atkinson & Stijn Bierman

M_{max} Workshop, June 2022.

Lancaster University, Shell

Outline

Aim is not only to learn about $M_{\max}$, but also how we approach it.

1. Primer on Extreme Value Theory

2. Learning from Small Magnitude Events

3. Outcomes for Groningen

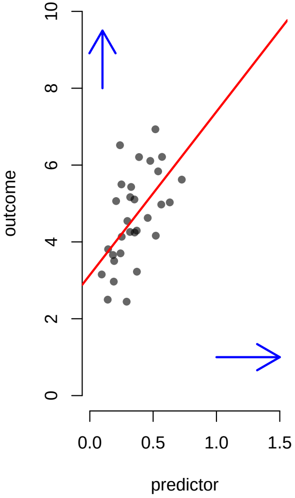
4. Further Work: Past and Present

A Primer on Extreme Value Theory

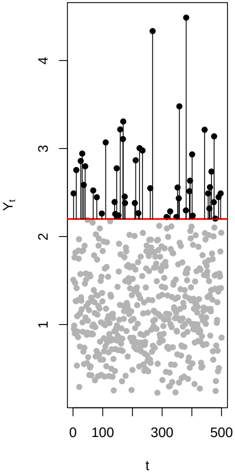
Standard Statistical Approaches

- Data $= g(\text{Signal, Noise})$
- **Aim of Inference:** identify the signal and describe the noise

- Standard methods describe typical values of a process:
  - Linear Regression, t-tests, ANOVA;
  - GLMs, GAMs, Random Forests...
- Fitting and evaluation driven by central values.
  Extrapolate at your own risk!

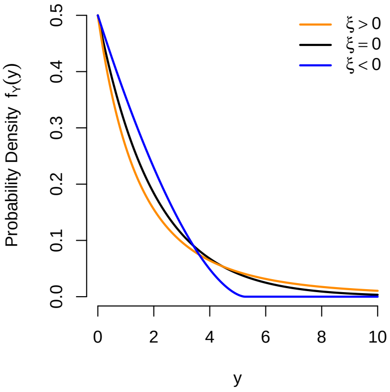


What to do instead: Asymptotically Motivated Model



- If we want to model big values then consider only big values.
- Define 'big' as exceedances of some high threshold u.
- In the limit as $u \to \infty$ then the distribution of the suitably rescaled threshold excesses converges to a single probability distribution, **regardless of the initial distribution**.
- Extreme Value Theory tells us that this is the **Generalised Pareto Distribution**.

The Generalised Pareto Distribution



- $\xi = 0$: Exponential distribution $\iff$ GR Law
- $\xi > 0$: Pareto distribution $\iff$ Power Law
- $\xi < 0$: Finite upper endpoint, similar to exponential taper model.

The Generalised Pareto Distribution

Distribution function:

For GPD(σ_u, ξ) exceedances of u

$$F(y) = 1 - \left[1 + \xi \frac{y - u}{\sigma_u}\right]_+^{-1/\xi} \text{ for } y > u$$

where $\sigma_u > 0$ and $x_+ = \max(x, 0)$.

Threshold stability property:

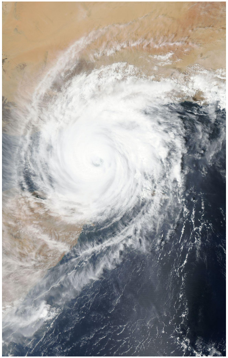
If GPD$(\sigma_u,\ \xi)$ above u then for any $v > u$

$$Y - v \mid Y > v \ \sim \ \text{GPD}(\sigma_u + \xi(v - u),\ \xi) \ = \ \text{GPD}(\sigma_v,\ \xi.)$$

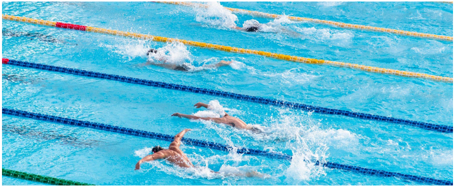
Applying EVT to finite samples

- Apply this tractable asymptotic model at finite levels.
- Using assumptions to buy certainty:
  - central limit theorem
  - elastic thin sheet of infinite extent
- Not a spherical chicken in a vacuum, requires only very light assumptions on the original distribution.
- Extends to non-i.i.d. data by having threshold and parameters as a function of time or covariates.

Applications of Extreme Value Theory: Natural Hazards



Applications of Extreme Value Theory: Elsewhere



Extreme Value Theory: Summary

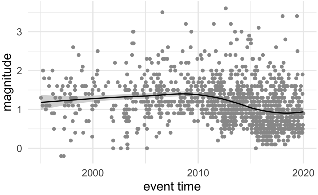
- EVT is a mathematically justified means of extrapolation.
- Extreme value models are used as standard across many other disciplines to reflect uncertainty in the tail shape as well as its scale.
- Many intrinsic parallels between EVT and seismicity models:
  - threshold selection $\iff$ magnitude of completion;
  - modelling conditional on being sufficiently large;
  - GPD $\iff$ power law, Gutenberg-Richter, tapered magnitudes.

What can small events tell us about large ones?

Groningen Earthquake Catalogue

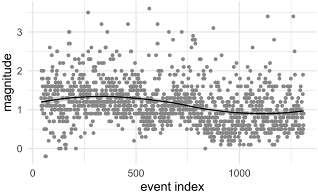
Recording:

- Time changing measurement process; more sensors and improved sensitivity.
- Earthquakes missing-not-at-random.
- Rounding to 0.1 M_L.



Magnitude of completion $m_c(t)$:

- Smallest magnitude at which earthquake in region is certain to be recorded



Work in 'event time' and back-transform

Aims of Inference

1. Automate estimation of a time-varying m_c.
2. Estimate and lower $m_c(t)$ from current standard
   - $m_c = 1.45\ M_L$, after 2014 reduced to $m_c = 0.95\ M_L$.
3. Use additional information to **estimate the upper tail of magnitude distribution** - high quantiles and $m_{\max}$.
4. Test consistency with Gutenberg Richter Law.
5. Develop method to handle:
   - trade off in quality of fit vs inference uncertainty
   - rounding
   - non-stationarity
   - informative missingness

Existing Methods & Limitations

Threshold Selection

- Parameter stability plots
- Linearity of magnitude frequency relationship
- PP and QQ plots

- Summaries such as Anderson-Darling
- Rolling quantile

Inference

- Seismic models underestimate epistemic uncertainty
- Roounding and Censoring

New Strategy: Model Structure

- Data $(t_i, x_i) : i = 1, \ldots, n$
- x_i rounded magnitudes, rounding to nearest 2δ
- y_i true magnitudes
- $y_i \in (x_i - \delta, x_i + \delta]$
- event index τ
- threshold (unknown) $v(\tau) \equiv (v_1, \ldots, v_n)$
- $Y_i - u \mid Y_i > u \sim \text{GPD}(\sigma_u, \xi)$ for $u < \min(v_1, \ldots, v_n)$

New Strategy: Model Structure

- Data $(t_i, x_i) : i = 1, \ldots, n$
- x_i rounded magnitudes, rounding to nearest 2δ
- y_i true magnitudes
- $y_i \in (x_i - \delta, x_i + \delta]$
- event index τ
- threshold (unknown) $v(\tau) \equiv (v_1, \ldots, v_n)$
- $Y_i - u \mid Y_i > u \sim \text{GPD}(\sigma_u, \xi)$ for $u < \min(v_1, \ldots, v_n)$

Three cases:

- $x_i > v_i + \delta \implies y_i > v_i$
- $x_i < v_i - \delta \implies y_i < v_i$
- $\mid x_i - v_i \mid < \delta \implies y_i > v_i$ with probability w_i

Likelihood Based Inference

For rounded GPD data and a given threshold $v(\tau)$:

$$\begin{aligned}
\ell(\boldsymbol{\theta}|\boldsymbol{x},\boldsymbol{v}) &= \sum_{i=1}^{n} w_i \log \Pr(X_i = x_i | Y_i > v_i, \boldsymbol{\theta}) \\
&= \sum_{i=1}^{n} w_i \log \Pr(\max(v_i, x_i - \delta) < Y_i < x_i + \delta | \boldsymbol{\theta})
\end{aligned}$$

where

$$w_i = \frac{\Pr(\max(v_i, x_i - \delta) < Y_i < x_i + \delta | \boldsymbol{\theta})}{\Pr(x_i - \delta < Y_i < x_i + \delta | \boldsymbol{\theta})}.$$

Likelihood Based Inference

For rounded GPD data and a given threshold $v(\tau)$:

$$\begin{aligned}\ell(\boldsymbol{\theta}|\boldsymbol{x},\boldsymbol{v}) &= \sum_{i=1}^{n} w_i \log \Pr(X_i = x_i | Y_i > v_i, \boldsymbol{\theta}) \\ &= \sum_{i=1}^{n} w_i \log \Pr(\max(v_i, x_i - \delta) < Y_i < x_i + \delta | \boldsymbol{\theta})\end{aligned}$$

where

$$w_i = \frac{\Pr(\max(v_i, x_i - \delta) < Y_i < x_i + \delta | \boldsymbol{\theta})}{\Pr(x_i - \delta < Y_i < x_i + \delta | \boldsymbol{\theta})}.$$

Issue: Changing sample size rules out standard model comparison

Measuring Fit: Metric definitions

Calculated on standard exponential scale (threshold invariance and PIT).

$$d_0 = d(q,1) = \frac{1}{m}\sum_{j=1}^{m} |-\log(1-p_j) - Q(p_j)|$$

$$d_0 = d(q,2) = \frac{1}{m}\sum_{j=1}^{m} (-\log(1-p_j) - Q(p_j))^2.$$

- Sequence of probabilities $p_i = i/(m+1)$
- Q is empirical quantile function
- PP methods- much less effective

New Threshold Selection Strategy

For given threshold choice for $v(\tau)$:

- Assess fit using QQ or PP plot
- Use metric to summarise difference between model and empirical - d_0
- Parametric bootstrapped replicates of $X \implies \hat{\theta}_{GPD} \implies Y$
  - Sample size varies (due to rounding)
  - Take expectation over latent Y and $\hat{\theta}_{GPD}$ variables

$\implies$ Output $d = E_{Y,\hat{\theta}_{GPD}|x,v(\tau)}(d_0)$ (subject to Monte Carlo noise)

New Threshold Selection Strategy

For given threshold choice for $v(\tau)$:

- Assess fit using QQ or PP plot
- Use metric to summarise difference between model and empirical - d_0
- Parametric bootstrapped replicates of $X \implies \hat{\theta}_{GPD} \implies Y$
  - Sample size varies (due to rounding)
  - Take expectation over latent Y and $\hat{\theta}_{GPD}$ variables

$\implies$ Output $d = E_{Y,\hat{\theta}_{GPD}|x,v(\tau)}(d_0)$ (subject to Monte Carlo noise)

Automatic selection of $v(\tau)$:

- Parametric $v(\tau)$ with parameters θ_v
- Minimise d over θ_v
- Minimise using grid-search for low dimensional θ_v
- Minimise using Bayesian optimisation methods for higher dimensional θ_v

Parametric Threshold Models

Flat: $\theta_v = (v)$

$$v(\tau) = v.$$

Stepped: $\theta_v = (v_L, v_R, \tau^*)$

$$v(\tau) = \begin{cases} v_L & \text{if } \tau \leq \tau^* \\ v_R & \text{if } \tau > \tau^* \end{cases}$$

Sigmoid: $\theta_v = (v_L, v_R, \tau^*, \gamma)$

$$v(\tau) = v_R + (v_L - v_R)\Phi\left(\frac{\tau^* - \tau}{\gamma}\right)$$

Threshold Selection on Simulated Catalogues

Simulated Catalogue Structure

Censoring methods: (i) hard and (ii) phased. In phased censoring the detection probability of each event is

$$\alpha(y_i, v_i) = \exp(-\lambda[v_i - y_i]_+), \text{ where } \lambda > 0.$$

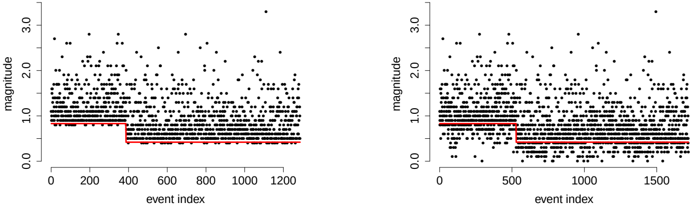


Figure 1: Example simulated catalogues with hard censoring [left] and phased censoring [right] for stepped thresholds of $(v_L, v_R) = (0.83, 0.42)$, shown as a red line, and phasing parameter $\lambda = 7$.

Flat Threshold, Hard Censoring: Single Catalogue

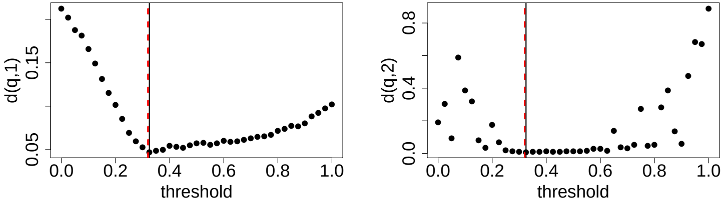


Figure 2: Flat threshold selection on a simulated catalogue. Top row: expected mean absolute [left] and expected mean squared [right] QQ-distances against threshold value. Selected and true thresholds are indicated by solid black and dashed red lines.

Flat Threshold, Hard Censoring: Replicate Results

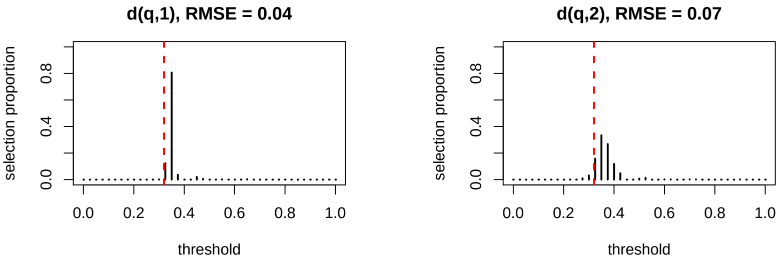


Figure 3: Sampling distribution of threshold selection methods for quantile-based metrics over 500 simulated catalogues with constant threshold and hard censoring. The true threshold is shown by a dashed red line and the root mean squared error (RMSE) for each method is given in plot titles.

Focus now only the absolute error QQ metric

Comparison of Hard and Phased Censoring

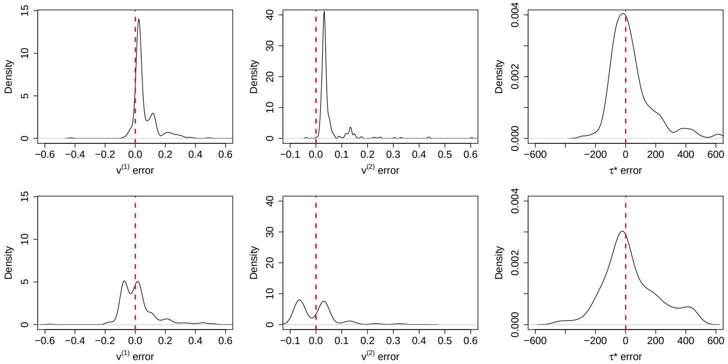


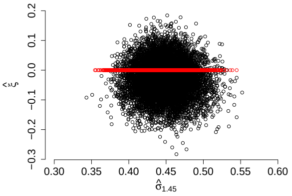
Figure 4: Marginal sampling distributions of errors in the selected values of $v^{(1)}$ (left), $v^{(2)}$ (center) and τ^* (right) for 500 simulated catalogues with change-point type thresholds and hard (top row) or phased (bottom row) censoring.

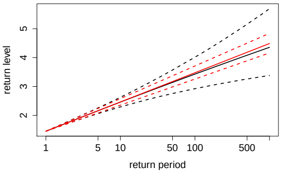
Application to Groningen Catalogue

Application to Groningen catalogue

Comparing GPD vs Gutenberg-Richter above conservative threshold:

- Conservative threshold: $v_C = 1.45 M_L$
- 311 exceedances
- $\hat{\xi} = -0.018$
- (if rounding ignored $\hat{\xi} = -0.027$)
- Bootstrap 95% CI: $(-0.147; 0.086)$
- Can't rule out Gutenberg-Richter law (Exponential tail)





GPD fits well above v_C

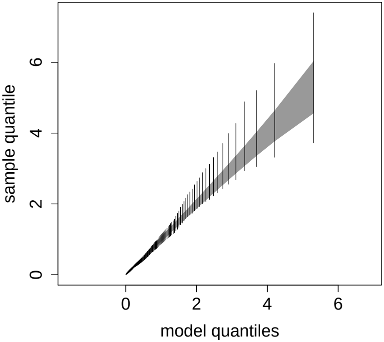
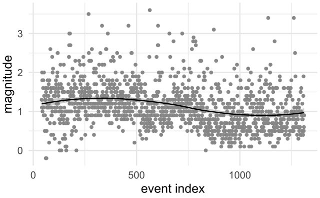


Figure 5: QQ plot for Groningen magnitudes exceeding $1.45\mathrm{M_L}$ under the GPD model. Grey regions show 95% tolerance intervals while vertical lines show 95% confidence intervals on sample probabilities / quantiles. All confidence intervals overlap with the associated tolerance intervals.

Flat Threshold Selection



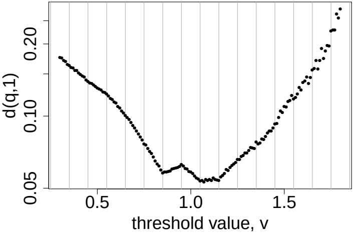


Figure 6: [Left] Data, [Right] Grid search to minimise $d(q, 1)$ over threshold values flat threshold. Metric values are shown on log-scale and vertical lines mark the edges of magnitude rounding intervals

- $v_C = 1.45$ is a poor choice
- Two minima of interest

Sigmoid Threshold Selection

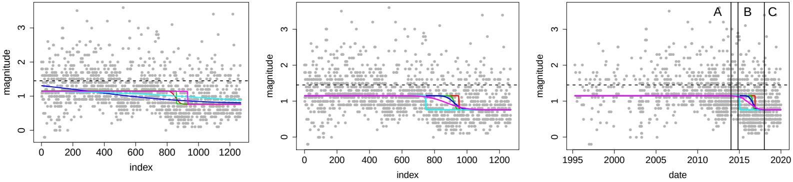


Figure 7: Selected sigmoid thresholds using Bayesian optimisation. [left] Optimising over all thresholds parameters. [centre, right] Optimising over (τ^*, γ) and fixing $(v_L, v_R) = (1.15, 0.76)$ on index- (centre) and natural- (right) timescales. Dates: (A) network development begins, (B) first additional sensors activated, (C) upgrade complete.

Data used above thresholds and value of $d \times 1000$:

- Conservative choice: $n_v = 311$, $d = 91$
- Best flat: $n_v = 627$, $d = 54$
- Best sigmoid: $n_v = 702$, $d = 41$

Benefits of lower threshold

- Lowered $m_c(\tau) \implies$ more data to use
- Better parameter inference
- $\hat{\xi} = -0.069$ 95% CI $(-0.144, -0.008)$
- Excludes zero: weak evidence against Gutenberg-Richter law

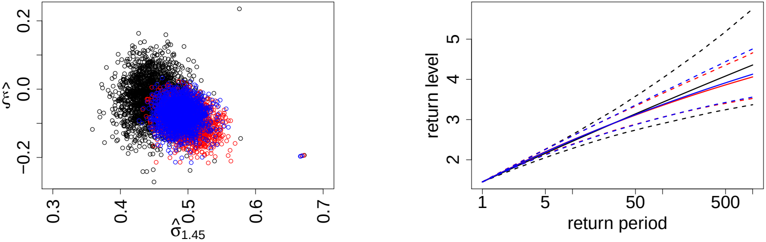


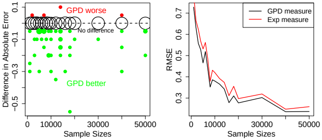
Figure 8: Bootstrap GPD parameter estimates based on exceedances of the conservative (black), flat and sigmoid thresholds [left]. Estimated return levels in M_L and 95% confidence intervals for magnitudes exceeding $1.45M_L$ [right].

Summary

1. The Generalised Pareto Distribution unifies magnitude-frequency relationships, better representing epistemic uncertainty.
2. Including small magnitude events is cost effective and informative.
3. Using a time varying threshold gives evidence of sub-exponential decay.
4. Magnitude-Frequency relationship is described as we approach $M_{\max}$, while properly accounting for uncertainties.

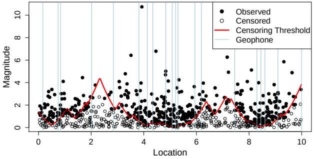
Further Work

- **Does choice of measurement scale matter? ✓**
  Hartog and Bierman investigated robustness to choice of measurement scale.



- **Are Exponential margins optimal?**
  Murphy simulation study shows smaller changes can be better.

- **Can we do this without aggregating over space? ✓**
  Murphy extending to spatial threshold selection.



Further, Further Work

- **Can we include stress-dependent magnitudes?**
  More challenging because requires separation of effects.

- **How does this impact prediction?**
  Combining models and propagating uncertainty about earthquake number, rate and size.

- **Can we demonstate effectiveness in other settings?**
  Also useful in more general EVT settings, improved detection common. Suggested applications welcome.

Thank you. Any Questions?

Varty, Z., Tawn, J. A., Atkinson, P. M., & Bierman, S. (2021). Inference for extreme earthquake magnitudes accounting for a time-varying measurement process. arXiv preprint arXiv:2102.00884.

Univariate Extremes

- Optimising the height of all coastal flood protection schemes in the UK, saving £200-300M over 13 years. [link to paper]
- Identified the likely cause of the sinking in 1980 of the MV Derbyshire for the £11M High Court Formal Investigation. [link to paper]
- *Calculated worldwide design standards for bulk carrier hold strength, impacted on the design of 6000 carriers - resulting in many saved ships/lives* [link to paper]

Multivariate/Spatial Extremes

- *Optimise the structural integrity of over 8% of worldwide offshore oil and gas facilities, saving £80M* [link to paper]
- Developing the widespread flooding scenarios for the UK Government's National Risk Assessment for river + coastal [link to paper]
- Developing spatial flood risk methods for the UK Government's 2016 National Flood Resilience Review:
  e.g., What is probability of a 1 in 100 year event at a site occurring anywhere in UK in a year? [link to paper]

International EVT Impact

- Netherlands: Coastal flooding [paper link]
- US: Hurricanes [paper link]
- France: Nuclear Safety [paper link]
- Japan: Earthquake (Annual M_{max} estimation) [paper link]

Non-Environmental EVT Impact

- Finance: Value-at-risk, portfolio selection [link to paper]
- Sport: rankings across events/records, ultimate performances [link to paper]
- Mortality: Upper bound on Oldest Human Ages [link to paper]