

Statistical modelling of volcanoes, with some reflections on expert judgement and uncertainty

Jonathan Rougier

School of Mathematics, University of Bristol

Food security, Warwick, Apr 2015

Obligatory mention of food security

The screenshot shows a web browser displaying a Guardian article. The browser's address bar shows the URL: www.theguardian.com/world/2015/mar/28/are-we-ready-for-the-next-big-volcanic-eruption. The Guardian logo is prominently displayed at the top right, with the tagline "Winner of the Pulitzer prize". Below the logo is a navigation menu with categories like "UK", "election", "world", "sport", "football", "opinion", "culture", "economy", "lifestyle", "fashion", "environment", "tech", "travel", and "all sections".

The article's main heading is "Volcanoes" by "The Observer", with a sub-heading "Are we ready for the next volcanic catastrophe?". The lead text reads: "The largest eruption ever recorded, in Indonesia 200 years ago, wreaked havoc across the world, causing hunger, disease and death for years afterwards. When a volcanic event on that scale happens again - and it will - we should be prepared for serious disruption to our climate and food production".

The author's name, "Bill McGuire", is listed on the left, along with the date "Saturday 28 March 2015 19:30 GMT". Below this are social media sharing icons for Facebook, Twitter, Email, and LinkedIn, and a statistics box showing "Shares 3109" and "Comments 489".

The main image is a photograph of a large volcanic eruption, showing a massive plume of white ash and smoke rising from a mountain range under a clear blue sky. In the bottom right corner, there is a "Most popular" section with two items: "What happened next to the Premier League's 10 youngest players? | Nick Ames" and "Can you solve the maths question for Singapore schoolkids that went".

Obligatory mention of food security

1. Laki (Iceland, 1783, M6.2?). 50% of livestock died within one year, and the total population decreased by 25%. Winter temperatures in the Eastern USA about 4.8°C lower, and about 1°C for the entire Northern hemisphere.

Obligatory mention of food security

1. Laki (Iceland, 1783, M6.2?). 50% of livestock died within one year, and the total population decreased by 25%. Winter temperatures in the Eastern USA about 4.8°C lower, and about 1°C for the entire Northern hemisphere.
2. Pinatubo (Philippines, 1991, M6.1). 800 km^2 of agricultural crops destroyed, and 800,000 livestock and poultry killed. Global temperatures about 0.5°C lower 1991–1993. Large decrease in ozone levels in the atmosphere.

Obligatory mention of food security

1. Laki (Iceland, 1783, M6.2?). 50% of livestock died within one year, and the total population decreased by 25%. Winter temperatures in the Eastern USA about 4.8°C lower, and about 1°C for the entire Northern hemisphere.
2. Pinatubo (Philippines, 1991, M6.1). 800 km^2 of agricultural crops destroyed, and 800,000 livestock and poultry killed. Global temperatures about 0.5°C lower 1991–1993. Large decrease in ozone levels in the atmosphere.
3. Tambora (Indonesia, 1815, M7.0). An eruption column 45 km tall. 1816 was 'the year without summer', with frost and snow in June in the Eastern USA, and three month cold spell that ruined most agricultural crops in North America. More than a century before the volcano was resettled.

My collaborators

Prof. Steve Sparks



Prof. Kathy Cashman



both in the Department of Earth Sciences at the University of Bristol.

A potted history of our progress so far

We want to produce a first-stage hazard assessment for active volcanoes, expressed as a return period curve for each volcano.

A potted history of our progress so far

We want to produce a first-stage hazard assessment for active volcanoes, expressed as a return period curve for each volcano. The cool new statistical idea is to treat subsets of volcanoes as similar but not identical (i.e. exchangeable).

A potted history of our progress so far

We want to produce a first-stage hazard assessment for active volcanoes, expressed as a return period curve for each volcano. The cool new statistical idea is to treat subsets of volcanoes as similar but not identical (i.e. exchangeable).

1. The proof of concept paper on Japanese stratovolcanoes 1600CE–present was almost written but we needed ...

A potted history of our progress so far

We want to produce a first-stage hazard assessment for active volcanoes, expressed as a return period curve for each volcano. The cool new statistical idea is to treat subsets of volcanoes as similar but not identical (i.e. exchangeable).

1. The proof of concept paper on Japanese stratovolcanoes 1600CE–present was almost written but we needed ...
2. an informed prior distribution, of the large-eruption rate of active stratovolcanoes. So we wrote another paper about this, using all volcanoes (active and inactive) back to 1600CE. This paper has been completely written except for ...

A potted history of our progress so far

We want to produce a first-stage hazard assessment for active volcanoes, expressed as a return period curve for each volcano. The cool new statistical idea is to treat subsets of volcanoes as similar but not identical (i.e. exchangeable).

1. The proof of concept paper on Japanese stratovolcanoes 1600CE–present was almost written but we needed . . .
2. an informed prior distribution, of the large-eruption rate of active stratovolcanoes. So we wrote another paper about this, using all volcanoes (active and inactive) back to 1600CE. This paper has been completely written except for . . .
3. an informed under-recording model for the period 1600–1950CE. So we wrote another paper to study the global large-eruption recording rates back to 1050CE. This paper is currently at late-draft stage.

A potted history of our progress so far

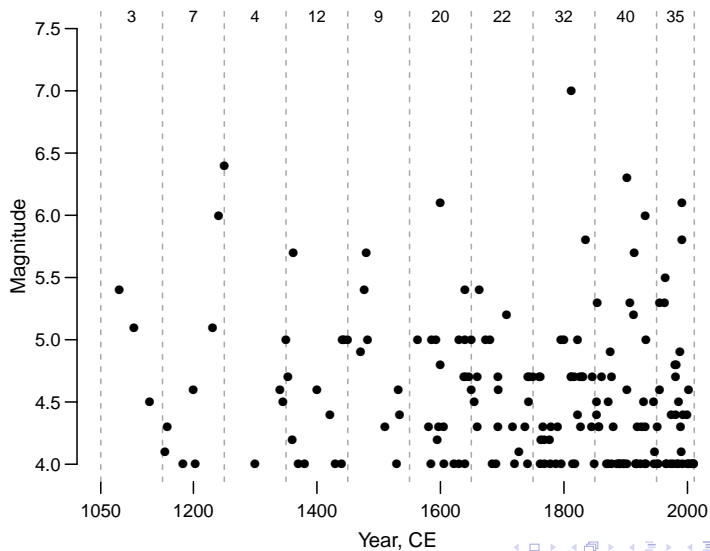
We want to produce a first-stage hazard assessment for active volcanoes, expressed as a return period curve for each volcano. The cool new statistical idea is to treat subsets of volcanoes as similar but not identical (i.e. exchangeable).

1. The proof of concept paper on Japanese stratovolcanoes 1600CE–present was almost written but we needed ...
2. an informed prior distribution, of the large-eruption rate of active stratovolcanoes. So we wrote another paper about this, using all volcanoes (active and inactive) back to 1600CE. This paper has been completely written except for ...
3. an informed under-recording model for the period 1600–1950CE. So we wrote another paper to study the global large-eruption recording rates back to 1050CE. This paper is currently at late-draft stage.

Duration of the project so far: several years, although this strand has taken about six months.

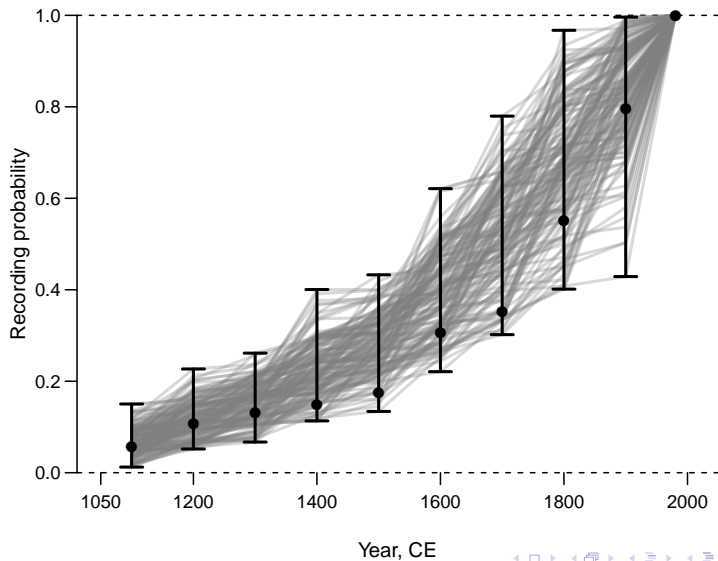
Statistical alchemy

Raw dataset (from the LaMEVE database)

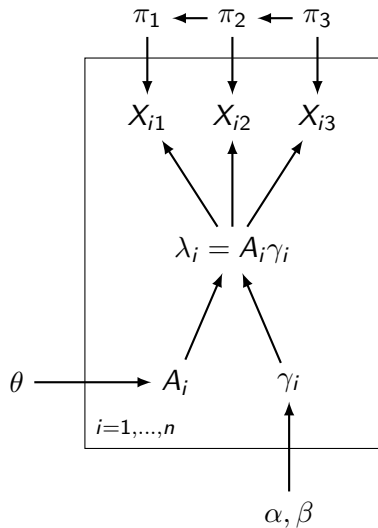


Statistical alchemy

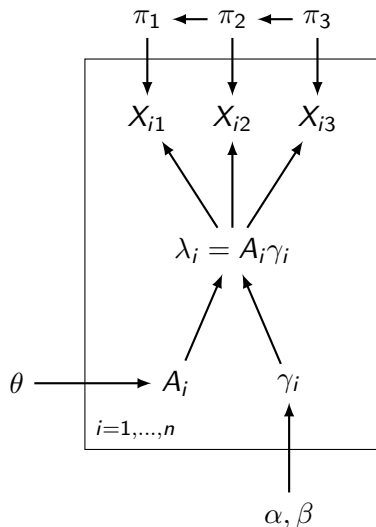
95% confidence set for the recording probability, M4+



Informed prior model



Informed prior model



$$\mathbf{X}_i | \boldsymbol{\pi}, \lambda_i \stackrel{\text{iid}}{\sim} \text{Pois}(\pi_i \lambda_i)$$

$$A_i | \theta \stackrel{\text{iid}}{\sim} \text{Ber}(\theta)$$

$$\gamma_i | \alpha, \beta \stackrel{\text{iid}}{\sim} \text{Gam}(\alpha, \beta)$$

$$\theta \sim \text{U}(0, 1)$$

$$(\pi_1, \pi_2, \pi_3) \sim \text{pr}_1$$

$$(\alpha, \beta) \sim \text{pr}_2$$

Our target is to compute

$$\text{Pr}^*(\gamma_{n+1} \leq v)$$

for values of $v \geq 0$.

Degrees of 'model'

Model An artefact designed to organise beliefs.

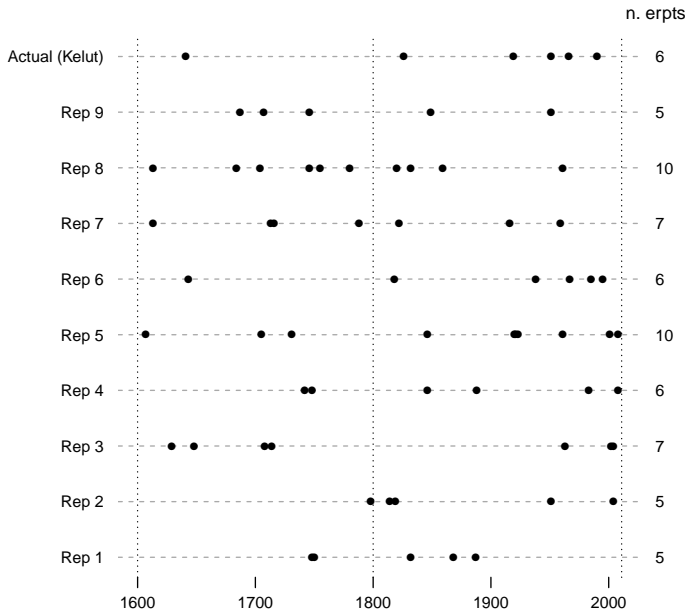
Degrees of 'model'

Model An artefact designed to organise beliefs.

Statistical model The vehicle through which we quantify beliefs, via the process of probabilistic conditioning.

'Turing test' for model adequacy

'Turing test' for model adequacy



Another pressing issue

- ▶ Statistical inference can only address simply-posed questions:

What is the return period curve for Mt Fuji, based on data in the LaMEVE database?

Another pressing issue

- ▶ Statistical inference can only address simply-posed questions:
 - What is the return period curve for Mt Fuji, based on data in the LaMEVE database?
- ▶ But the Mayor of Tokyo wants to know:
 - What is the return period curve for Mt Fuji, based on all available evidence?
- ▶ This involves a large amount of Mt-Fuji-specific information, some of it rather nebulous.

Another pressing issue

- ▶ Statistical inference can only address simply-posed questions:
 - What is the return period curve for Mt Fuji, based on data in the LaMEVE database?
- ▶ But the Mayor of Tokyo wants to know:
 - What is the return period curve for Mt Fuji, based on all available evidence?
- ▶ This involves a large amount of Mt-Fuji-specific information, some of it rather nebulous.
- ▶ **The only way to access this information is through structured expert elicitation.**

My collaboration with volcanologists exposes them to statistical concepts, and shapes their beliefs, but only indirectly does it address risk management questions.

