How big was that huge volcanic eruption which happened 90 thousand years ago?

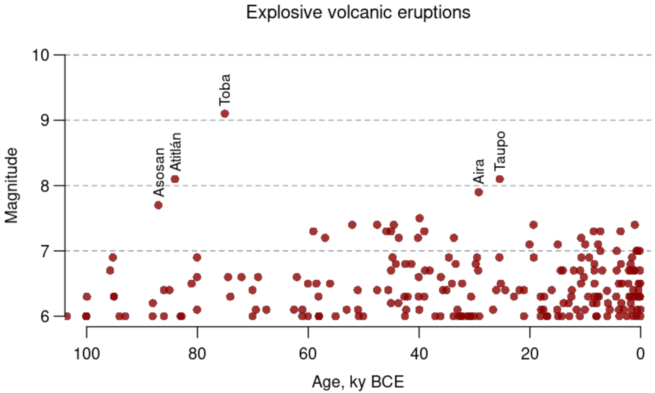
Jonathan Rougier

Rougier Consulting Ltd
& University of Bristol

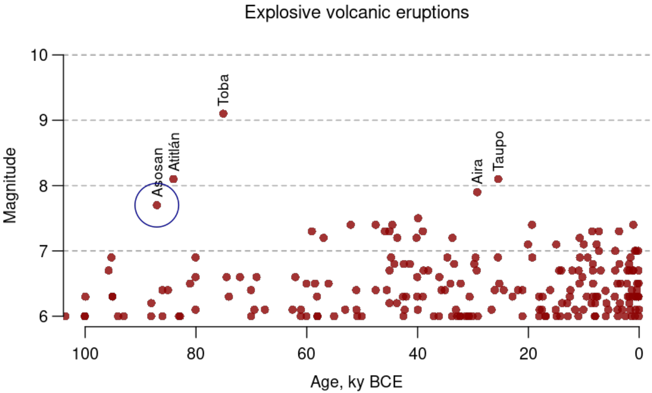
Jointly with Steve Sparks, Willy Aspinall, Sue Mahony

Representing and quantifying uncertainty in complex systems
RSS Annual Conference, Sep 2021

Large explosive volcanic eruptions



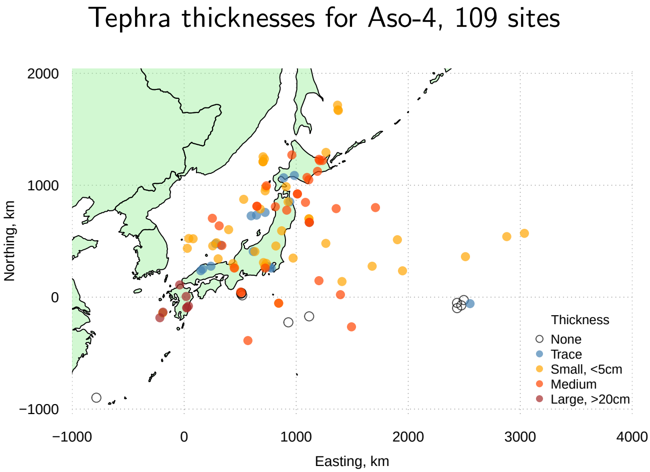
Large explosive volcanic eruptions



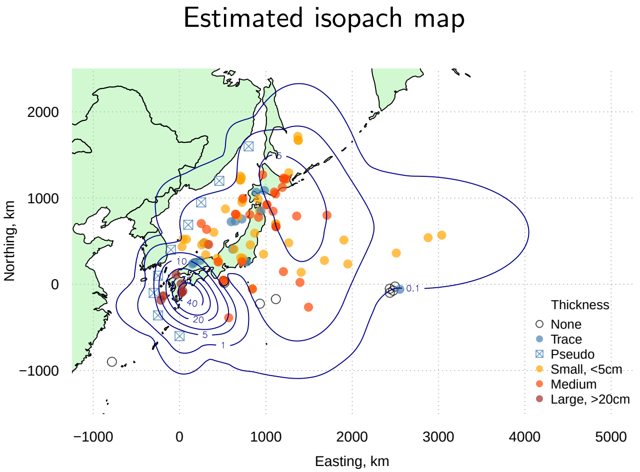
Large explosive volcanic eruptions

Aso volcano today

Large explosive volcanic eruptions



Large explosive volcanic eruptions



Outline

- The deposition of tephra fall on land and sea is complex, the archive is prone to disturbances, and the measurements are sparse and inaccurate.

Outline

- The deposition of tephra fall on land and sea is complex, the archive is prone to disturbances, and the measurements are sparse and inaccurate.
- Current practice for volume estimation involves a high degree of expert judgement, being based on a small number of hand-drawn isopachs (= contours of equal thickness).

The interplay of measurement and judgement is not transparent, and the approach is not amenable to a formal assessment of uncertainty.

Outline

- The deposition of tephra fall on land and sea is complex, the archive is prone to disturbances, and the measurements are sparse and inaccurate.
- Current practice for volume estimation involves a high degree of expert judgement, being based on a small number of hand-drawn isopachs (= contours of equal thickness).

  The interplay of measurement and judgement is not transparent, and the approach is not amenable to a formal assessment of uncertainty.
- We use modern flexible-fitting methods from Statistics and Machine Learning, we represent judgements using pseudo-measurements, and we quantify variability using a bootstrap 95% CI.

A spatial model for thickness

We insist from the outset that thickness z is non-negative:

$$\mathrm{BC}(z(s);\lambda) = \mathrm{BC}(0;\lambda) + \sum_{j=1}^{k} \beta_j \phi_j(s) + r(s), \quad s \in \mathcal{S} \subset \mathbb{R}^2,$$

where $\phi_j \geq 0$ with compact support, $\beta_j \geq 0$, and 'BC' is the Box-Cox transformation,

$$\mathrm{BC}(z;\lambda) = \begin{cases} \log(z+1) & \lambda = 0 \\ \frac{(z+1)^\lambda - 1}{\lambda} & \lambda \neq 0. \end{cases}$$

- This model is a bit bespoke, but we need $\hat{z}(s) \geq 0$ everywhere, *and* $\hat{z}(s) = 0$ for all locations outside the footprint of $\bigcup_j \phi_j$, regardless of how we transform z.

Modelling options

Reminder of the model:

$$\mathrm{BC}(z(s);\lambda) = \mathrm{BC}(0;\lambda) + \sum_{j=1}^{k} \beta_j \phi_j(s) + r(s), \quad s \in \mathcal{S} \subset \mathbb{R}^2.$$

Modelling options

Reminder of the model:

$$\mathrm{BC}(z(s);\lambda) = \mathrm{BC}(0;\lambda) + \sum_{j=1}^{k} \beta_j \phi_j(s) + r(s), \quad s \in \mathcal{S} \subset \mathbb{R}^2.$$

- Choice of basis functions, ϕ_j (multitudinous)

Modelling options

Reminder of the model:

$$\mathrm{BC}(z(s);\lambda) = \mathrm{BC}(0;\lambda) + \sum_{j=1}^{k} \beta_j \phi_j(s) + r(s), \quad s \in \mathcal{S} \subset \mathbb{R}^2.$$

- Choice of basis functions, ϕ_j (multitudinous)
- Choice of fitting method (spoiler alert, LASSO)

Modelling options

Reminder of the model:

$$\mathrm{BC}(z(s);\lambda) = \mathrm{BC}(0;\lambda) + \sum_{j=1}^{k} \beta_j \phi_j(s) + r(s), \quad s \in \mathcal{S} \subset \mathbb{R}^2.$$

- Choice of basis functions, ϕ_j (multitudinous)
- Choice of fitting method (spoiler alert, LASSO)
- Choice of transformation, λ

Modelling options

Reminder of the model:

$$\mathrm{BC}(z(s);\lambda) = \mathrm{BC}(0;\lambda) + \sum_{j=1}^{k} \beta_j \phi_j(s) + r(s), \quad s \in \mathcal{S} \subset \mathbb{R}^2.$$

- Choice of basis functions, ϕ_j (multitudinous)
- Choice of fitting method (spoiler alert, LASSO)
- Choice of transformation, λ
- Addition of pseudo-measurements 🤔

Modelling options

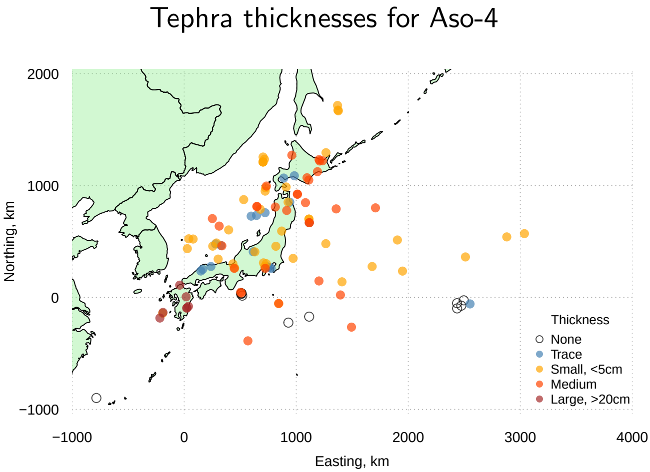
Reminder of the model:

$$\mathrm{BC}(z(s);\lambda) = \mathrm{BC}(0;\lambda) + \sum_{j=1}^{k} \beta_j \phi_j(s) + r(s), \quad s \in \mathcal{S} \subset \mathbb{R}^2.$$

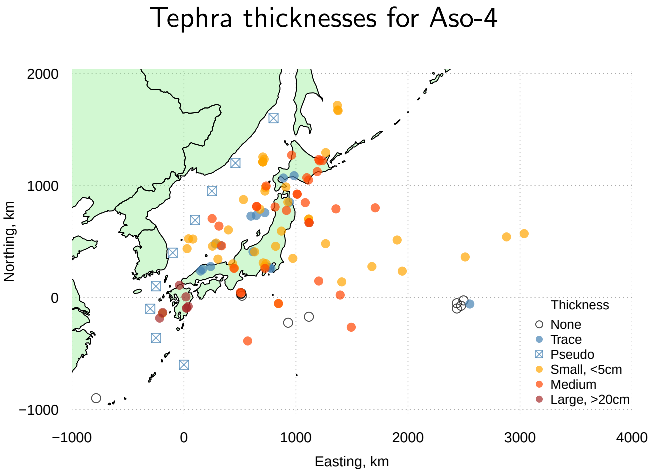
- Choice of basis functions, ϕ_j (multitudinous)
- Choice of fitting method (spoiler alert, LASSO)
- Choice of transformation, λ
- Addition of pseudo-measurements 🤔

Let's take these in reverse order ...

Modelling options, pseudo-measurements



Modelling options, pseudo-measurements



Modelling options, transformation

$$\mathsf{BC}(z(s);\lambda) = \mathsf{BC}(0;\lambda) + \sum_{j=1}^{k} \beta_j \phi_j(s) + r(s), \quad s \in \mathcal{S} \subset \mathbb{R}^2.$$

Remembering that $z \geq 0$,

- $\lambda = 1$ (linear) tends to over-weight the largest thicknesses, which is problematic because at least some of the measurement error is proportional.

Modelling options, transformation

$$\mathrm{BC}(z(s);\lambda) = \mathrm{BC}(0;\lambda) + \sum_{j=1}^{k} \beta_j \phi_j(s) + r(s), \quad s \in \mathcal{S} \subset \mathbb{R}^2.$$

Remembering that $z \geq 0$,

- $\lambda = 1$ (linear) tends to over-weight the largest thicknesses, which is problematic because at least some of the measurement error is proportional.
- But $\lambda = 0$ (logarithmic) tends to overweight the smallest measurements, which is problematic because they can also be quite inaccurate, and they contribute least to the estimate of volume.

Modelling options, transformation

$$\mathrm{BC}(z(s);\lambda) = \mathrm{BC}(0;\lambda) + \sum_{j=1}^{k} \beta_j \phi_j(s) + r(s), \quad s \in \mathcal{S} \subset \mathbb{R}^2.$$

Remembering that $z \geq 0$,

- $\lambda = 1$ (linear) tends to over-weight the largest thicknesses, which is problematic because at least some of the measurement error is proportional.
- But $\lambda = 0$ (logarithmic) tends to overweight the smallest measurements, which is problematic because they can also be quite inaccurate, and they contribute least to the estimate of volume.
- So after some experimentation, we have currently settled on $\lambda = \frac{1}{2}$ (square-root), which seems to give reasonable results, as judged by the volcanologists.

Modelling options, basis functions

We want bell-shaped basis functions with compact support.

Modelling options, basis functions

We want bell-shaped basis functions with compact support.

- I like the bisquare,

$$b(s) = \left(1 - \|s\|^2\right)^2, \quad \|s\| \le 1$$

and $b(s) = 0$ otherwise, which lives on the unit disk.

Modelling options, basis functions

We want bell-shaped basis functions with compact support.

- I like the bisquare,

$$b(s) = \left(1 - \|s\|^2\right)^2, \quad \|s\| \leq 1$$

  and $b(s) = 0$ otherwise, which lives on the unit disk.

- But then map the unit disk to ellipses,

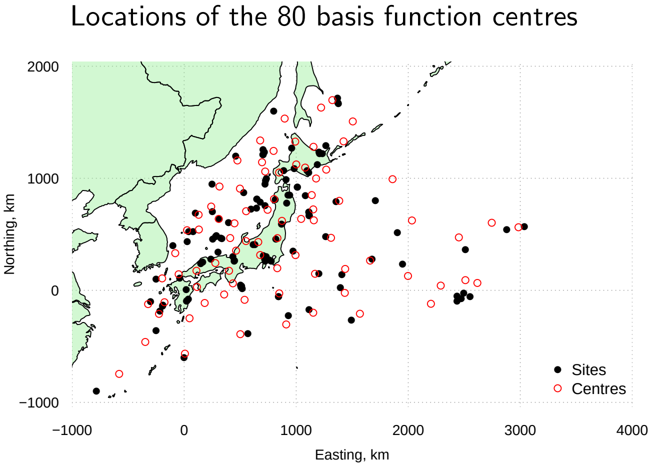
$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

  and use a lot of them.

Modelling options, basis functions

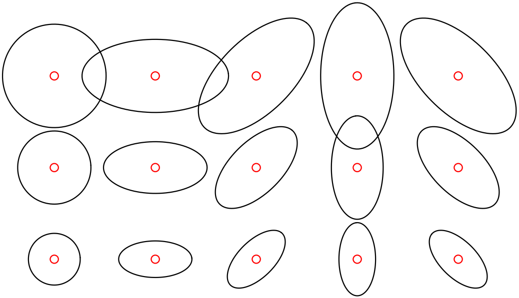
- Where to put these ellipses?
  - Putting them at the data sites leads to over-fitting,
  - And yet we want more where the density of sites is high,
  - So I ended up using the Delaunay triangulation of the sites.
  - Specifically, the centroids of the Delaunay tiles with minimum edge length of at least 100 km.

Modelling options, basis functions



Modelling options, basis functions

Multi-resolution: I put 15 ellipses at every centre,



where the area of the largest subset (top) was treated as a bandwidth parameter, to be learnt and plugged-in; $15 \times 80 = 1200$ basis functions altogether. 😳

Modelling options, bandwidth

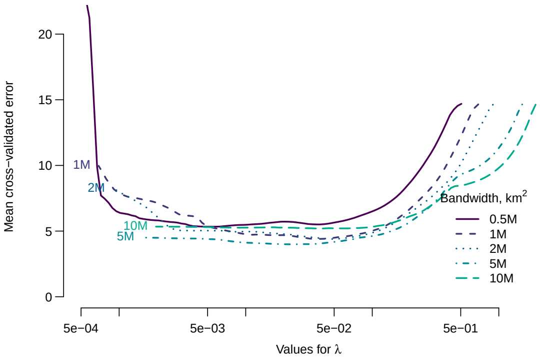
We have 1200 basis functions, an unknown bandwidth parameter, and only 118 (very imprecise) measurements.

- Downweight the 'trace' measurements using $w_i = 1/4$

Modelling options, bandwidth

We have 1200 basis functions, an unknown bandwidth parameter, and only 118 (very imprecise) measurements.

- Downweight the 'trace' measurements using $w_i = 1/4$
- Fit with an L_1 penalty, and use ten-fold cross-validation to choose both the sparsity coefficient λ and the bandwidth:

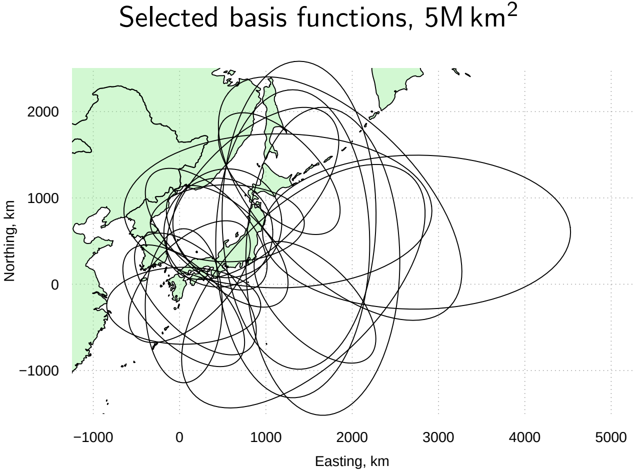


Modelling options, fitting

That gets us to our fitted model and our isopach map.

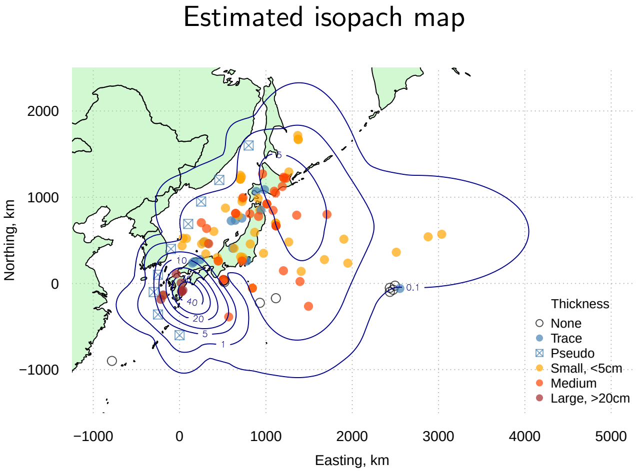
Modelling options, fitting

That gets us to our fitted model and our isopach map.



Modelling options, fitting

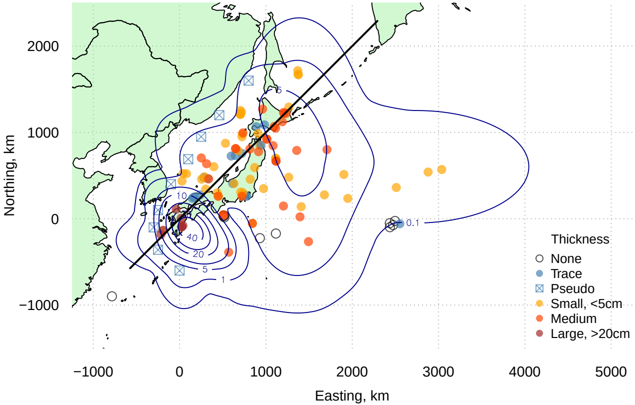
That gets us to our fitted model and our isopach map.



Modelling options, fitting

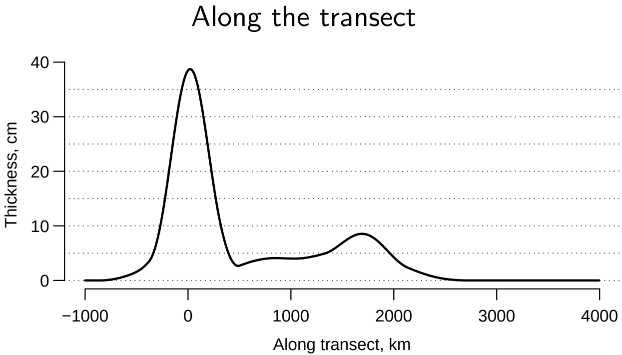
That gets us to our fitted model and our isopach map.

Along the transect

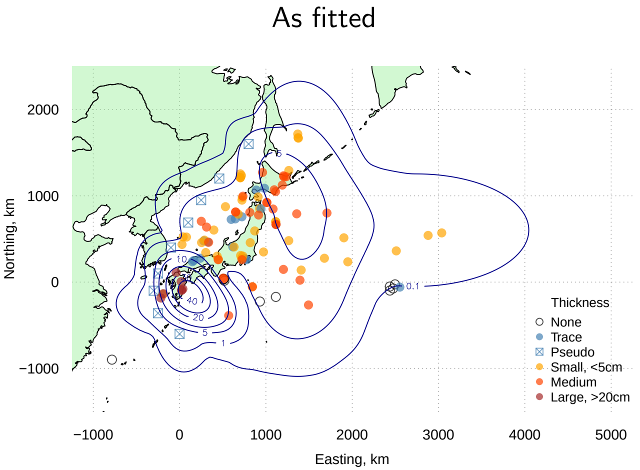


Modelling options, fitting

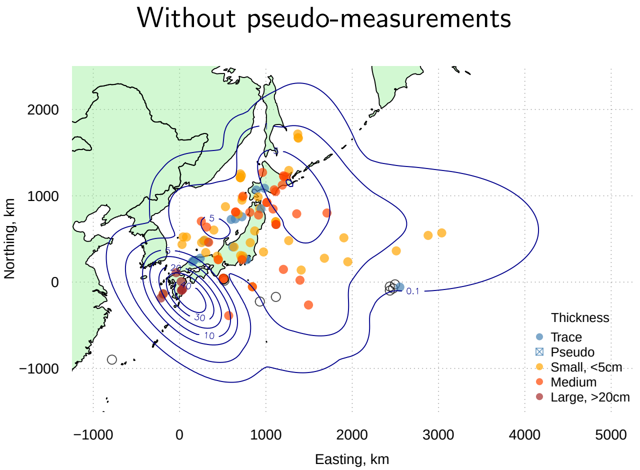
That gets us to our fitted model and our isopach map.



The effect of the pseudo-measurements

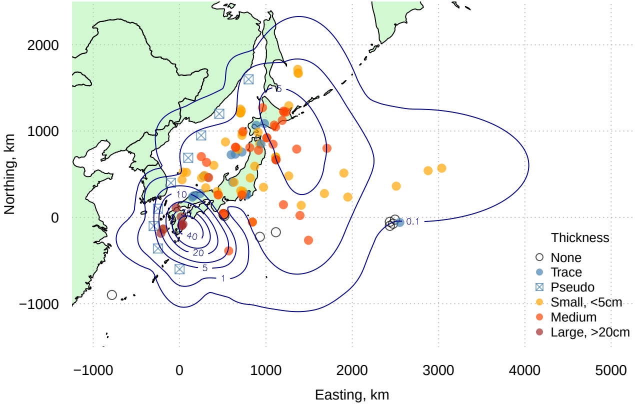


The effect of the pseudo-measurements



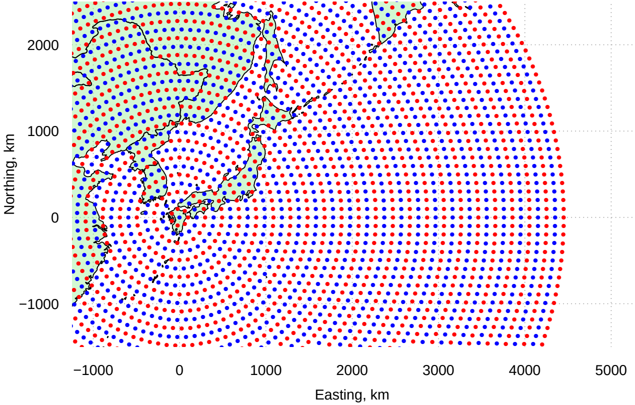
Estimating volume

I didn't overthink this . . .



Estimating volume

I didn't overthink this . . .



What about variability?

Want to report a 95% CI for volume. We have developed an algorithm: $\mathcal{X} \mapsto$ volume; we're going to bootstrap our way to a 95% confidence interval.

What about variability?

Want to report a 95% CI for volume. We have developed an algorithm: $\mathcal{X} \mapsto$ volume; we're going to bootstrap our way to a 95% confidence interval.

1. 'Wild' resampling scheme to allow for heteroskedasticity.

   Basically, fit the model and then generate replicates $\mathcal{X}^*$ by flipping the residual; see Davidson and Flachaire (2008).

What about variability?

Want to report a 95% CI for volume. We have developed an algorithm: $\mathcal{X} \mapsto$ volume; we're going to bootstrap our way to a 95% confidence interval.

1. 'Wild' resampling scheme to allow for heteroskedasticity.

   Basically, fit the model and then generate replicates $\mathcal{X}^*$ by flipping the residual; see Davidson and Flachaire (2008).

2. Bootstrap to estimate the standard error of log-volume. Then convert the Wald 95% confidence interval for log-volume back to volume.

What about variability?

Want to report a 95% CI for volume. We have developed an algorithm: $\mathcal{X} \mapsto$ volume; we're going to bootstrap our way to a 95% confidence interval.

1. 'Wild' resampling scheme to allow for heteroskedasticity.

   Basically, fit the model and then generate replicates $\mathcal{X}^*$ by flipping the residual; see Davidson and Flachaire (2008).

2. Bootstrap to estimate the standard error of log-volume. Then convert the Wald 95% confidence interval for log-volume back to volume.

3. Fix the bandwidth at 5M km^2 throughout, but let the basis function selection be driven by the resampled dataset.

What about variability?

Want to report a 95% CI for volume. We have developed an algorithm: $\mathcal{X} \mapsto$ volume; we're going to bootstrap our way to a 95% confidence interval.

1. 'Wild' resampling scheme to allow for heteroskedasticity.

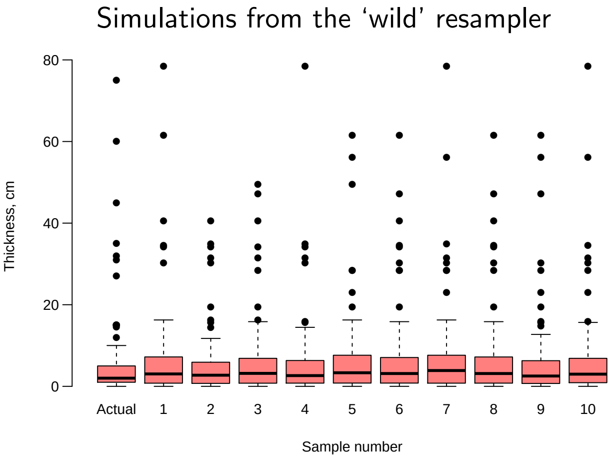
   Basically, fit the model and then generate replicates $\mathcal{X}^*$ by flipping the residual; see Davidson and Flachaire (2008).

2. Bootstrap to estimate the standard error of log-volume. Then convert the Wald 95% confidence interval for log-volume back to volume.

3. Fix the bandwidth at $5\mathrm{M}\,\mathrm{km}^2$ throughout, but let the basis function selection be driven by the resampled dataset.

4. The resulting 95% CI is $[220\,\mathrm{km}^3, 370\,\mathrm{km}^3]$.

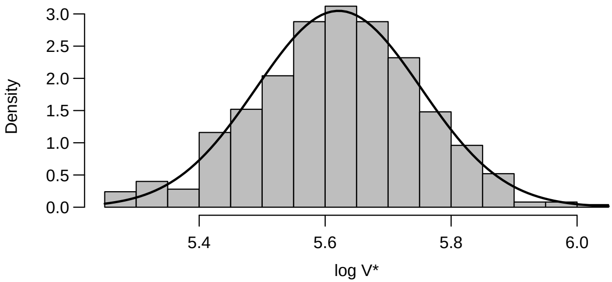
What about variability? (cont)



The slightly lower IQR for the 'Actual' is due to treatment of the 'trace' measurements; has negligible effect on the volume estimate.

What about variability? (cont)

Histogram of replicates of log-volume



This looks OK – thank goodness. Originally I used the variance-stabilized Studentized Pivotal Bootstrap, but it gave roughly the same 95% CI.

Take-home messages

Aimed more at non-statisticians who need to quantify uncertainty from limited measurements of a complex system.

1. If you have an algorithm for transforming a dataset into an estimate, then one simple and transparent way to include expert judgement is through adding pseudo-measurements.

Take-home messages

Aimed more at non-statisticians who need to quantify uncertainty from limited measurements of a complex system.

1. If you have an algorithm for transforming a dataset into an estimate, then one simple and transparent way to include expert judgement is through adding pseudo-measurements.
2. If you have an algorithm for transforming a dataset into an estimate, then you can quantify uncertainty as a 95% CI using a resampling method, such as a 'bootstrap'.

Take-home messages

Aimed more at non-statisticians who need to quantify uncertainty from limited measurements of a complex system.

1. If you have an algorithm for transforming a dataset into an estimate, then one simple and transparent way to include expert judgement is through adding pseudo-measurements.
2. If you have an algorithm for transforming a dataset into an estimate, then you can quantify uncertainty as a 95% CI using a resampling method, such as a 'bootstrap'.
3. Modern methods from Statistics and Machine Learning can provide flexible and data-driven methods for constructing an algorithm for transforming a dataset into an estimate.

Take-home messages

Aimed more at non-statisticians who need to quantify uncertainty from limited measurements of a complex system.

1. If you have an algorithm for transforming a dataset into an estimate, then one simple and transparent way to include expert judgement is through adding pseudo-measurements.
2. If you have an algorithm for transforming a dataset into an estimate, then you can quantify uncertainty as a 95% CI using a resampling method, such as a 'bootstrap'.
3. Modern methods from Statistics and Machine Learning can provide flexible and data-driven methods for constructing an algorithm for transforming a dataset into an estimate.

And don't be surprised if the resulting 95% CI is quite large!

References

Davidson, R. and Flachaire, E. (2008). The wild bootstrap, tamed at last. *Journal of Econometrics*, 146:162–169.