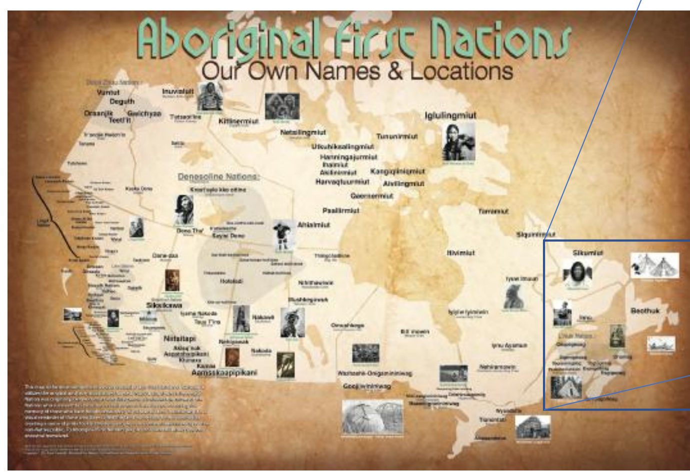
Uncertainties

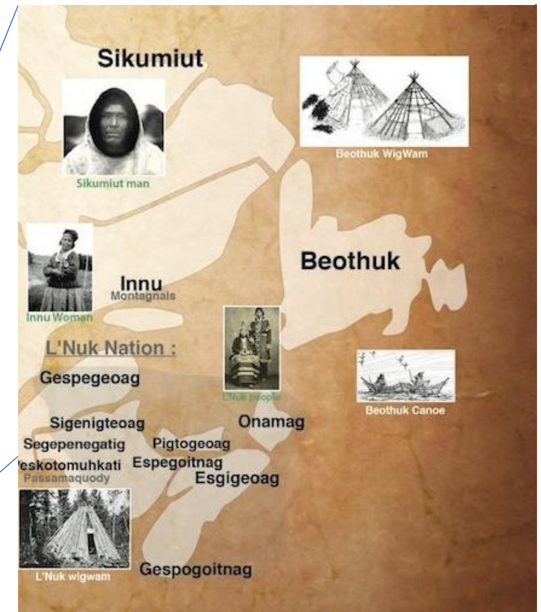
Alison Malcolm

Memorial University of Newfoundland and Labrador

St John's

https://www.britannica.com/place/Newfoundland-and-Labrador





https://www.ammsa.com/sites/default/files/articles/article-photos/11-sn_ws_can-givennames.jpg

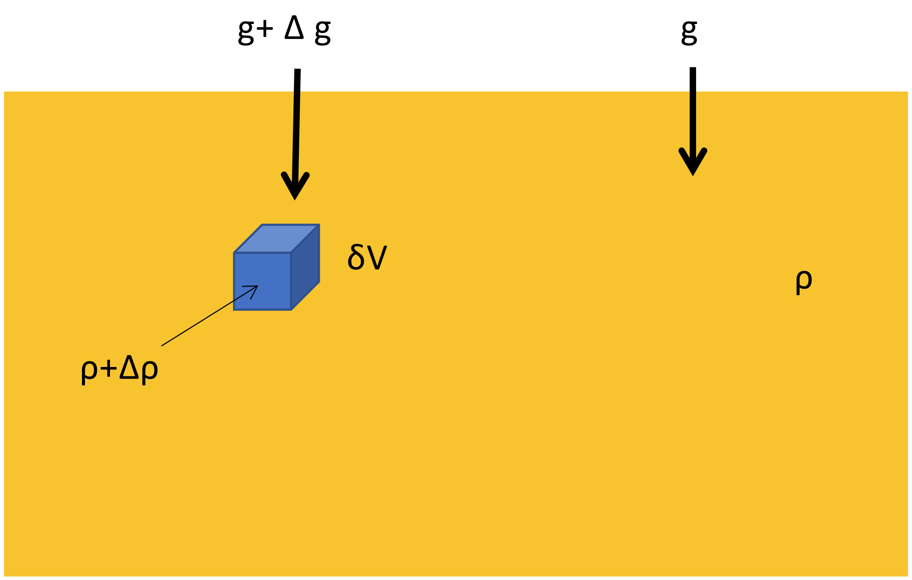
Example – Gravity

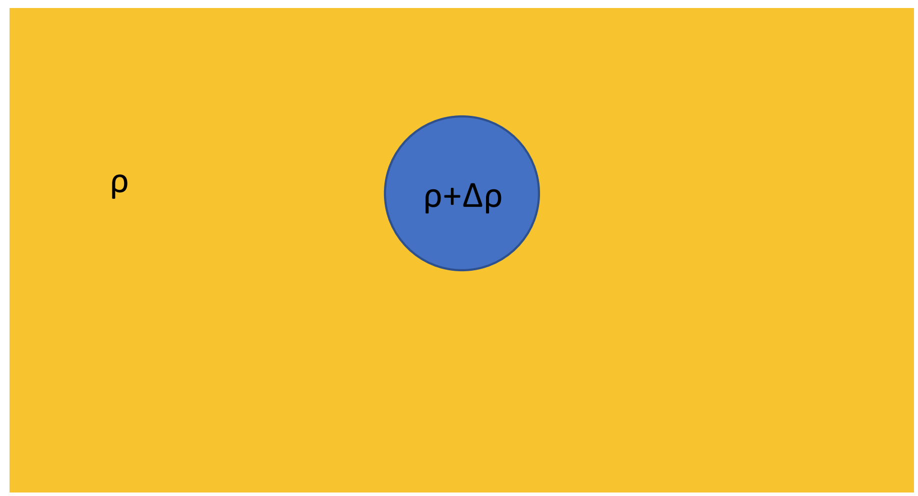
For a small volume δV, the excess mass is:
$\Delta M = \Delta\rho \; \delta V$

And the gravity anomaly is: $\Delta g = G \, \frac{\Delta M}{R^2}$

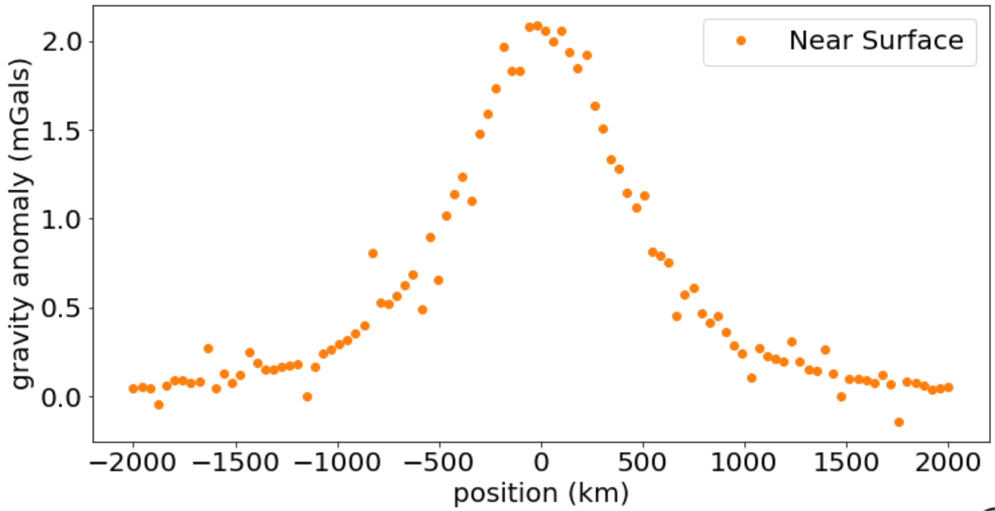
For a sphere, the gravitational anomaly is:

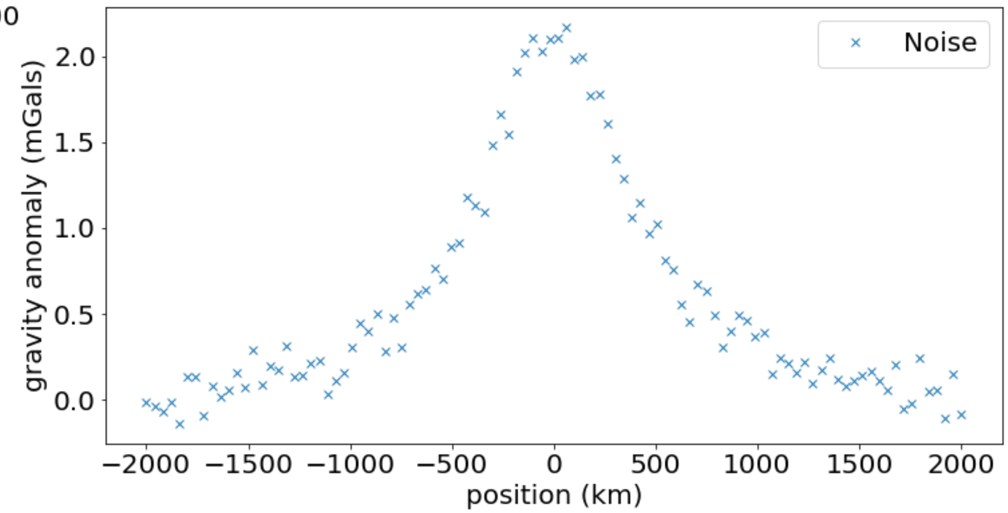
$$\Delta g(x, z) = G \, \frac{4\pi R^3 \Delta\rho}{3(x^2 + z^2)}$$

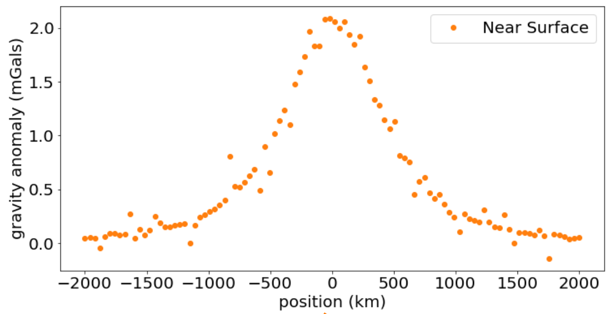




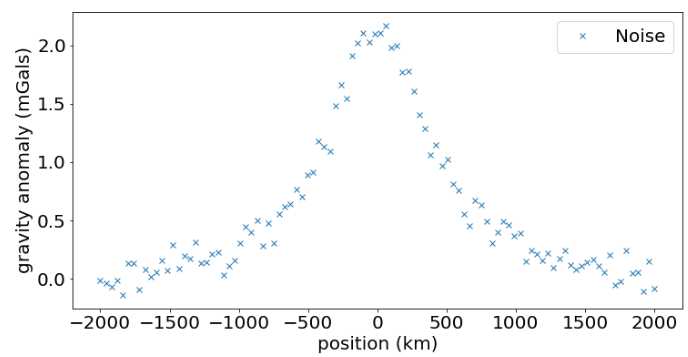
Uncertainties

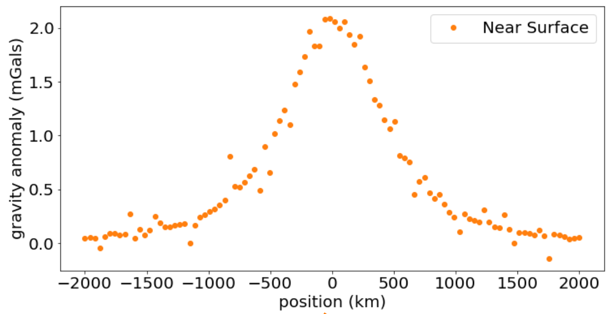




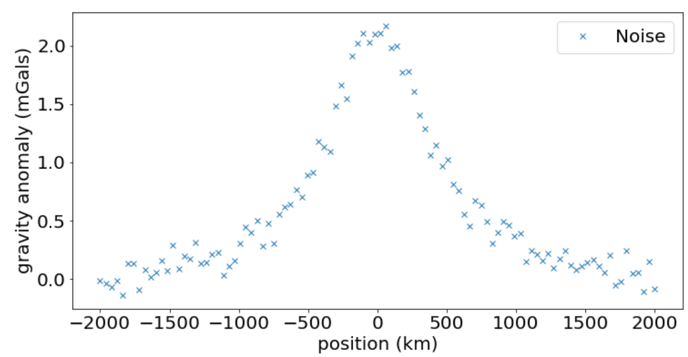


$$\chi^2(\mathbf{m}) = \frac{1}{2}\left(\mathbf{m} - \mathbf{m}^{prior}\right)^T C_{\mathbf{m}}^{-1}\left(\mathbf{m} - \mathbf{m}^{prior}\right) + \frac{1}{2}\left(\mathbf{G}\mathbf{m} - \mathbf{d}^{obs}\right)^T C_{\mathbf{d}}^{-1}\left(\mathbf{G}\mathbf{m} - \mathbf{d}^{obs}\right)$$





$$\chi^2(\mathbf{m}) = \frac{1}{2}\left(\mathbf{m} - \mathbf{m}^{prior}\right)^T C_{\mathbf{m}}^{-1}\left(\mathbf{m} - \mathbf{m}^{prior}\right) + \frac{1}{2}\left(\mathbf{Gm} - \mathbf{d}^{obs}\right)^T C_d^{-1}\left(\mathbf{Gm} - \mathbf{d}^{obs}\right)$$



Problem: we need to know **both** $C_{\mathbf{m}}^{-1}$ and $C_{\mathbf{d}}^{-1}$ before we can evaluate our objective function!

Outline

- Reminder/Expansion about covariance matrices
- What are uncertainties?
  - Data Uncertainties, $C_{\boldsymbol{d}}^{-1}$, and the $\chi^2 -$ test
  - Model Uncertainties, $C_{\mathbf{m}}^{-1}$
- Identifying what you want to learn and setting up your inverse problem accordingly
  - The null space of your problem
  - Choosing a good parameterization
  - Checking your results

What is a covariance matrix anyway?

$$C_x^{ij} = E[(x_i - E[x_i])(x_j - E[x_j])]$$

Generally we can assume that $E[x_i] = \bar{x}$ (technically it is only if we take 'enough' samples we expect to get the mean, but we can safely ignore this)

$$C_x^{ii} = E[(x_i - \bar{x})^2] = \sigma_i^2$$
$$C_x^{ij} = E\left[(x_i - \bar{x})\left(x_j - \bar{x}\right)\right] = \sigma_i \sigma_j$$

Some notes:

- We usually use a diagonal approximation $C_x^{ij} = \sigma_j^2 \delta^{ij}$
- We often use a single value $C_x^{ij} = \sigma^2 \delta^{ij}$

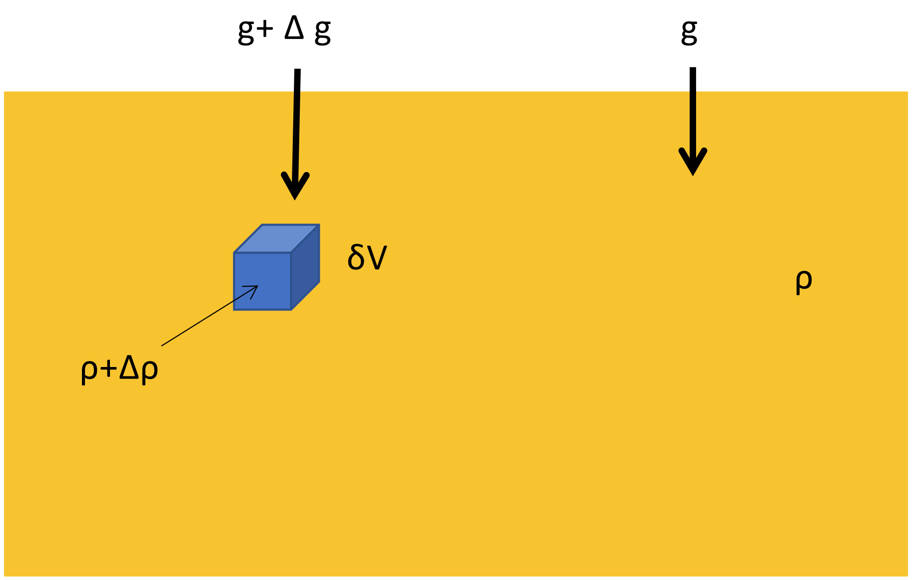
$C_{\mathbf{d}}$ – Data Uncertainties

- Measurement uncertainties
- Processing-induced uncertainties
- Propagating uncertainties
- Errors vs uncertainties

Example – Gravity

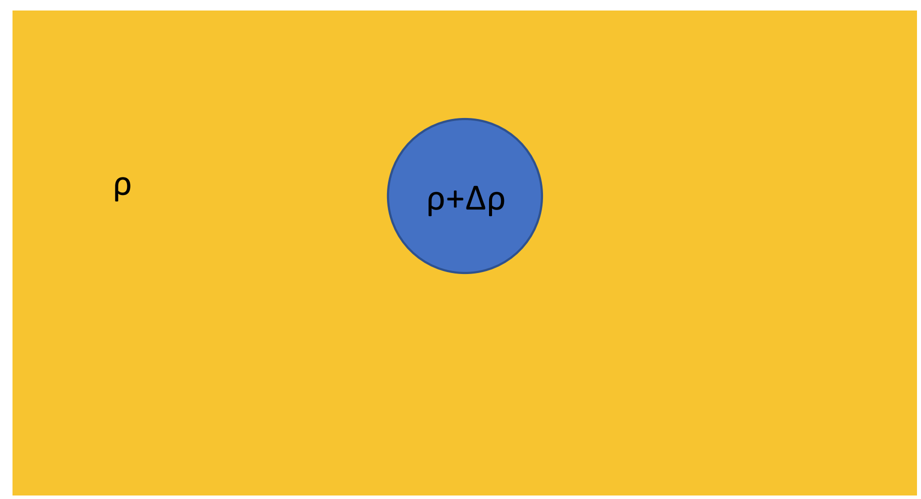
For a small volume δV, the excess mass is:
$\Delta M = \Delta\rho\ \delta V$

And the gravity anomaly is: $\Delta g = G\ \frac{\Delta M}{R^2}$

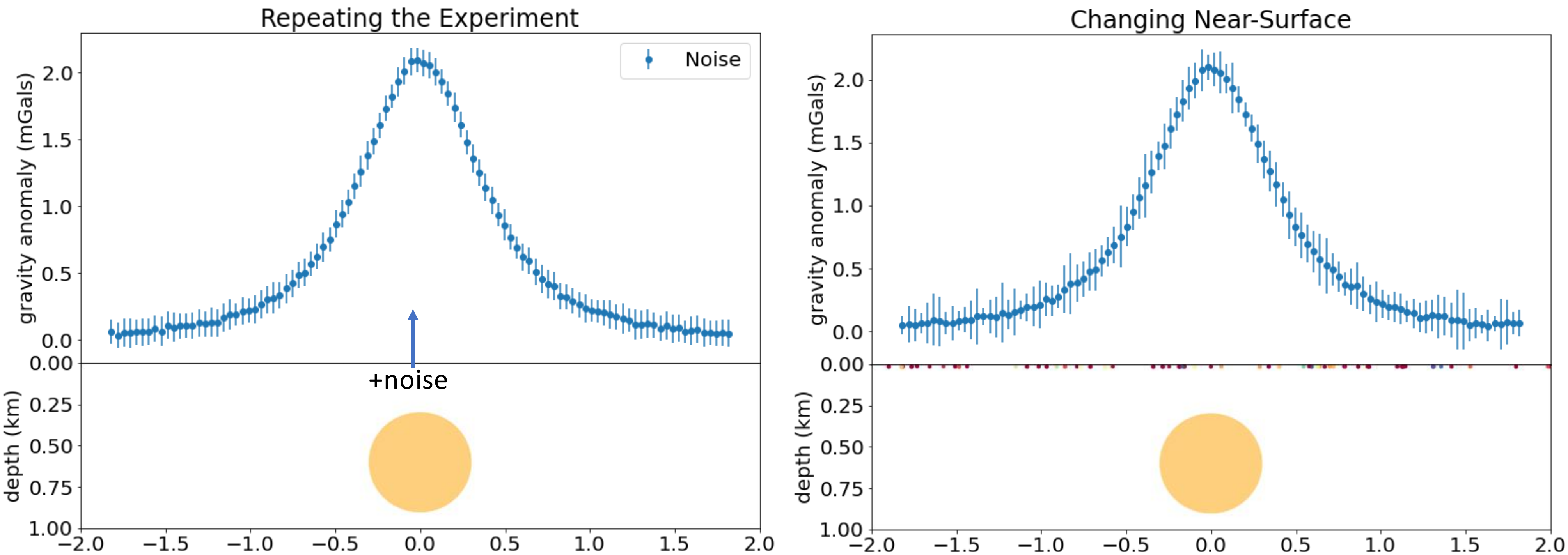


For a sphere, the gravitational anomaly is:

$$\Delta g(x,z) = G\,\frac{4\pi R^3 \Delta\rho}{3(x^2 + z^2)}$$

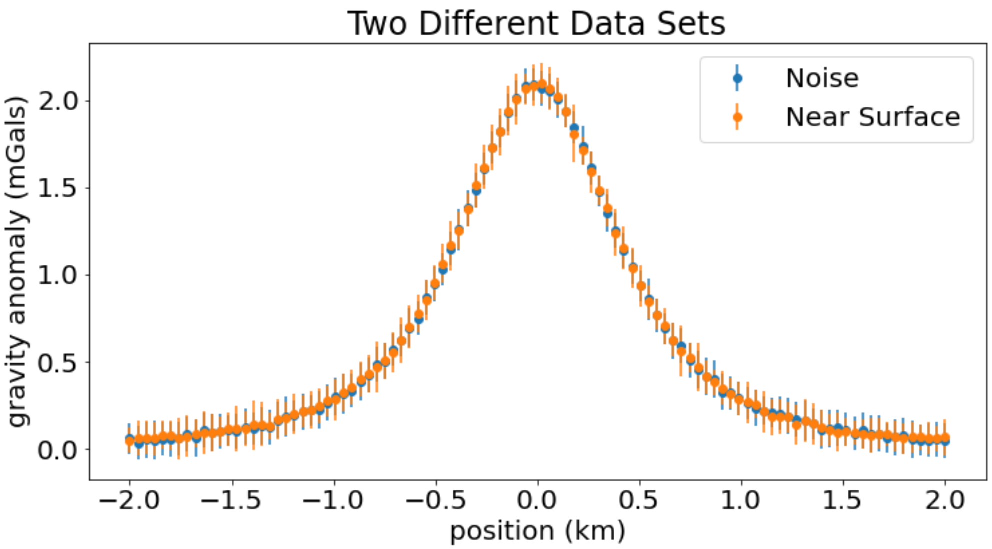


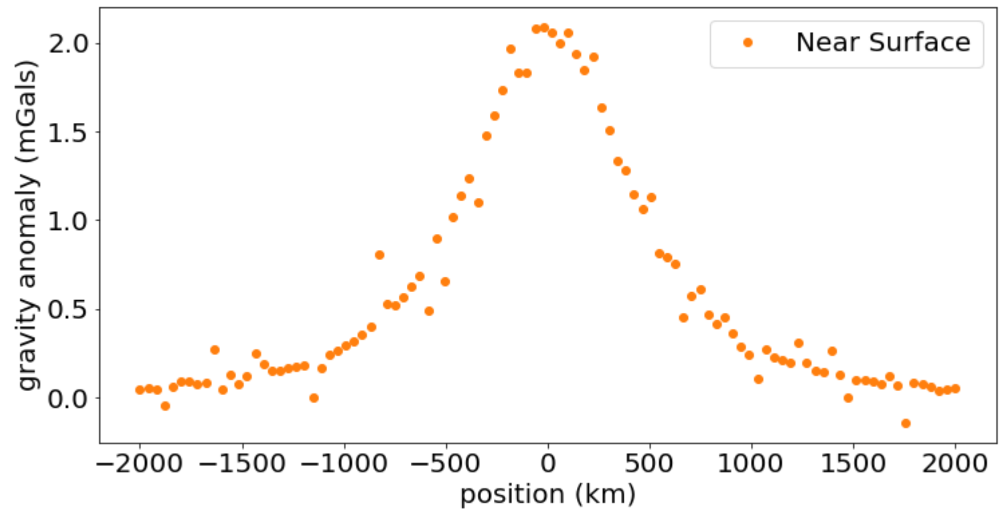
Data Uncertainties – How do we estimate them?

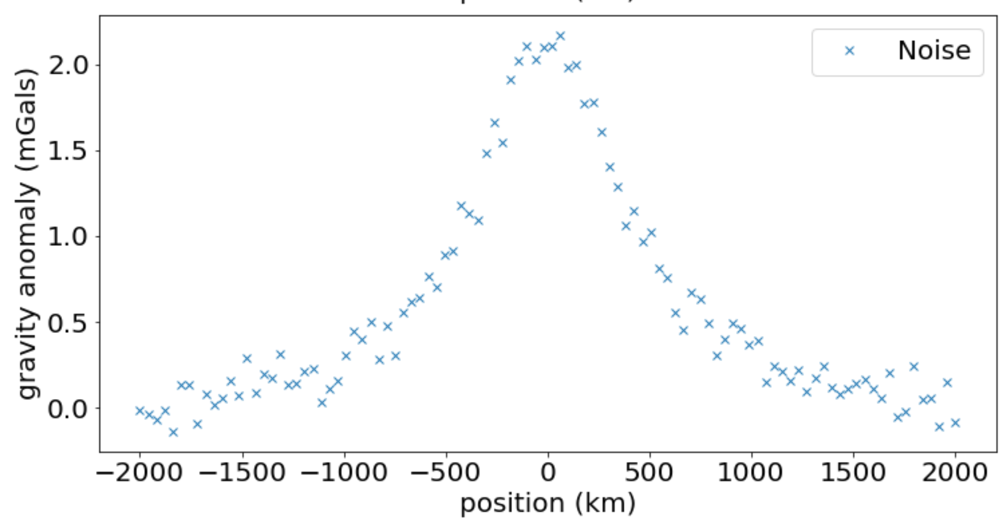


Data Uncertainties – How do we estimate them?

Two different sources of error give essentially the same distribution

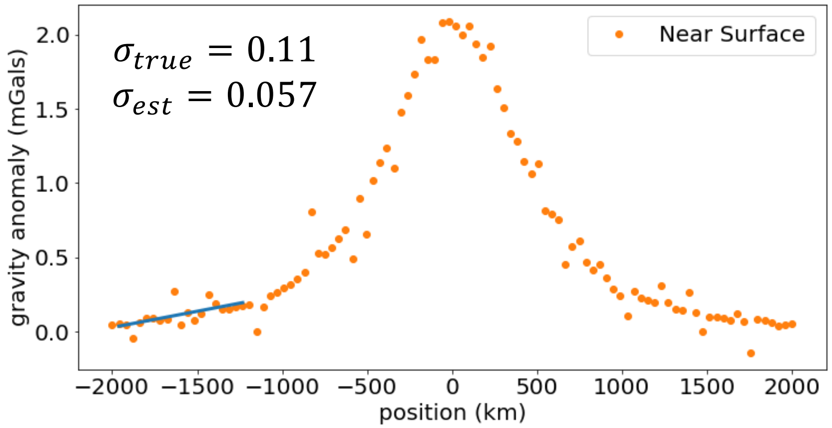






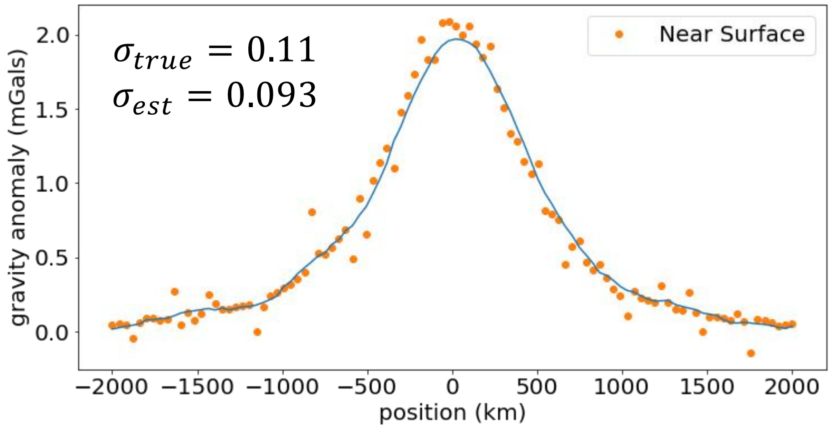
Data Uncertainties – How do we estimate them?

- It's difficult to estimate uncertainties accurately without repeated experiments
- Most techniques to do so are somewhat ad-hoc
  - Look before the first-arrival for fluctuations in the system
  - Smooth the data, use this as the mean, calculate a σ of sorts
  - Repeat a few measurements (if you can), or search for similar parameters (using e.g. reciprocity as in Cai & Zelt, Geophysics, 2022)
  - Guess 🙃



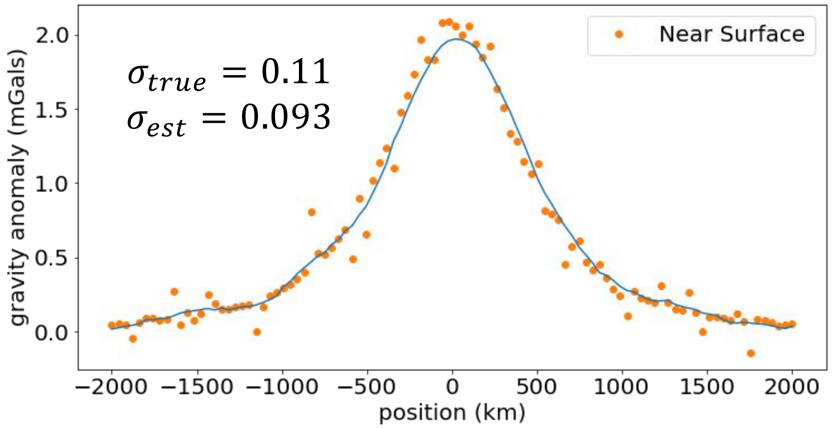
Data Uncertainties – How do we estimate them?

- It's difficult to estimate uncertainties accurately without repeated experiments
- Most techniques to do so are somewhat ad-hoc
  - Look before the first-arrival for fluctuations in the system
  - Smooth the data, use this as the mean, calculate a σ of sorts
  - Repeat a few measurements (if you can), or search for similar parameters (using e.g. reciprocity as in Cai & Zelt, Geophysics, 2022)
  - Guess 🙃



Data Uncertainties – How do we estimate them?

- It’s difficult to estimate uncertainties accurately without repeated experiments
- Most techniques to do so are somewhat ad-hoc
  - Look before the first-arrival for fluctuations in the system
  - Smooth the data, use this as the mean, calculate a σ of sorts
  - Repeat a few measurements (if you can), or search for similar parameters (using e.g. reciprocity as in Cai & Zelt, Geophysics, 2022)
  - Guess 🙃
- Think about correlations between data points



Model Uncertainties – Things to think about

- How certain is our model? (E.g. anisotropy, attenuation etc)
- What do you know about it beforehand? (E.g. velocity ranges, density is positive)
- What is the resolution, or parameterisation that you are interested in?

What goes into your prior vs your model covariance?

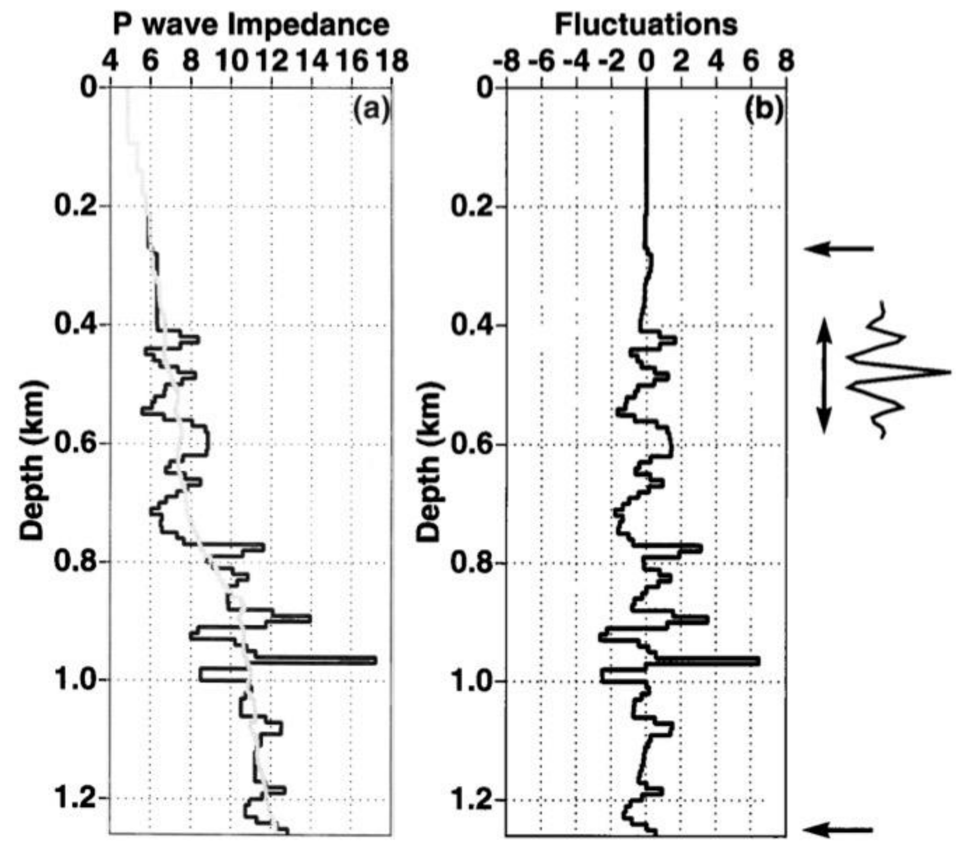
How to estimate $C_{\mathbf{m}}$

- Often we just assume $C_{\mathbf{m}}$ is just the identity, because we have nothing else to put in there, but it can be estimated
- Gouveia & Scales stimate $C_{\mathbf{m}}$ by first getting m_{prior} by smoothing the log, then getting the std from the fluctuations about that mean
- This is capturing sub-seismic resolution changes in the model

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 103, NO. B2, PAGES 2759–2779, FEBRUARY 10, 1998

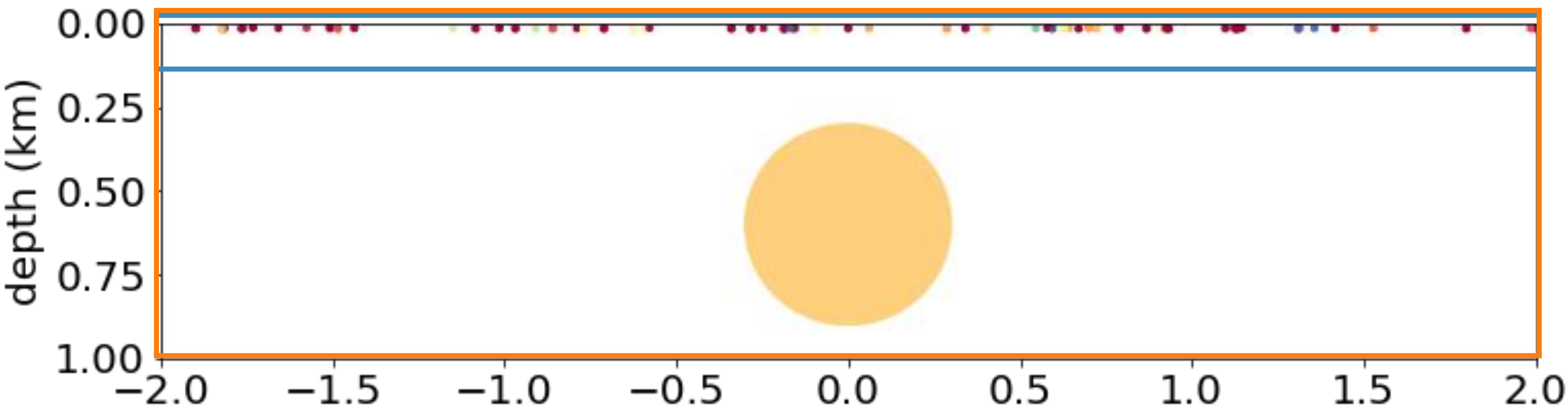
Bayesian seismic waveform inversion: Parameter estimation and uncertainty analysis

Wences P. Gouveia[1] and John A. Scales
Department of Geophysics, Center for Wave Phenomena, Colorado School of Mines, Golden

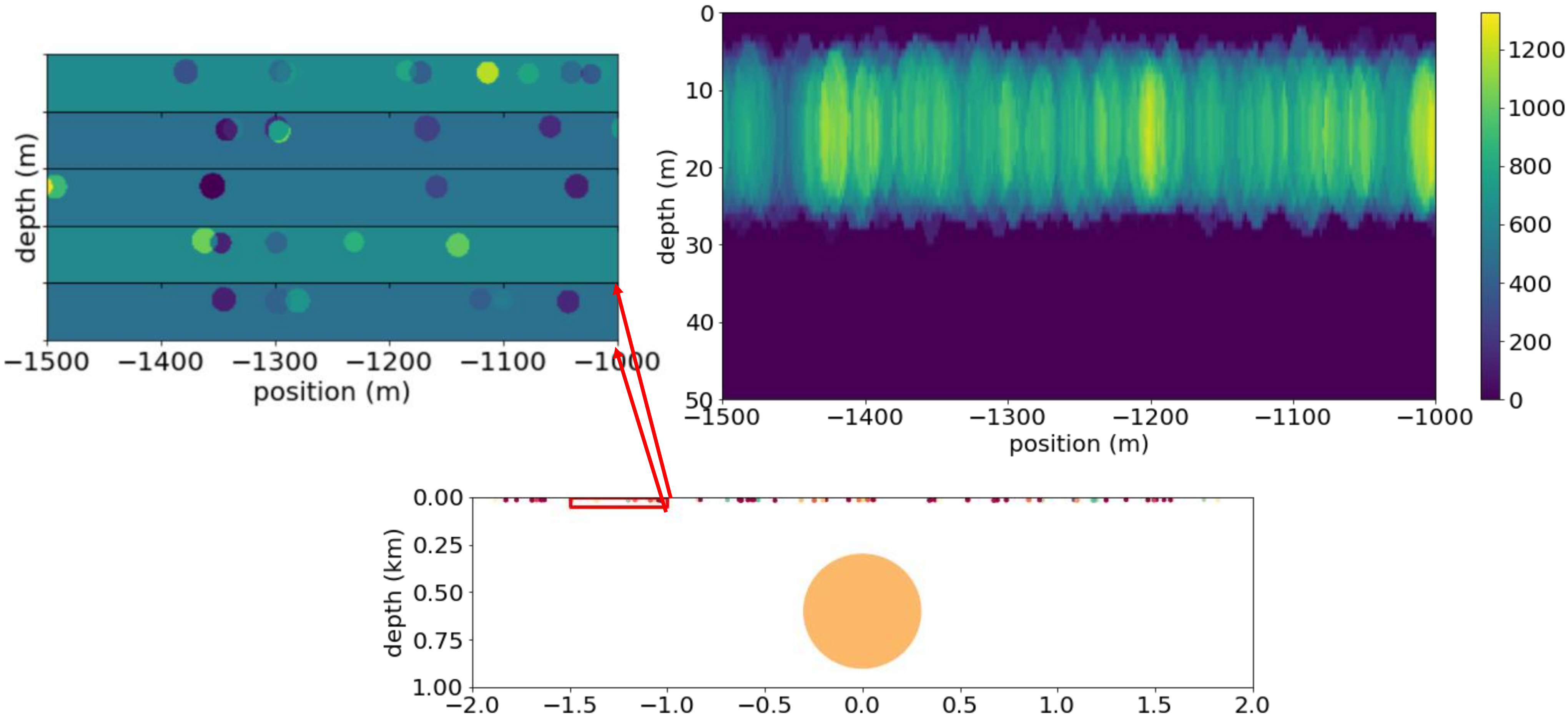


Model Covariances – Gravity example

- Over the volume of the spheres: $C_m = 1500 \text{ kg/m}^3$
- Within the blue box: $C_m = 6.9 \text{ kg/m}^3$
- Within the orange box: $C_m = 0.69 \text{ kg/m}^3$

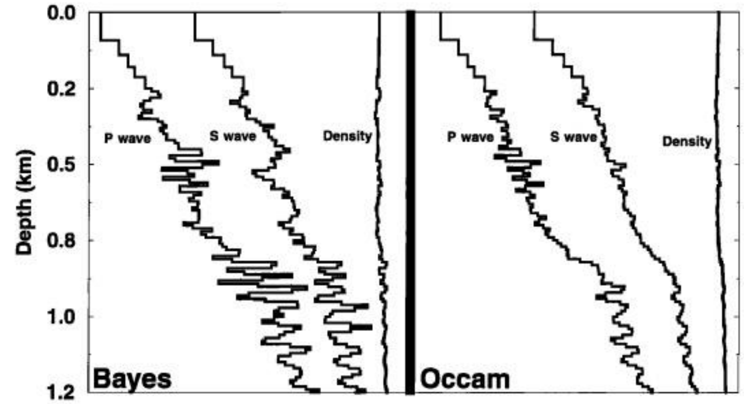


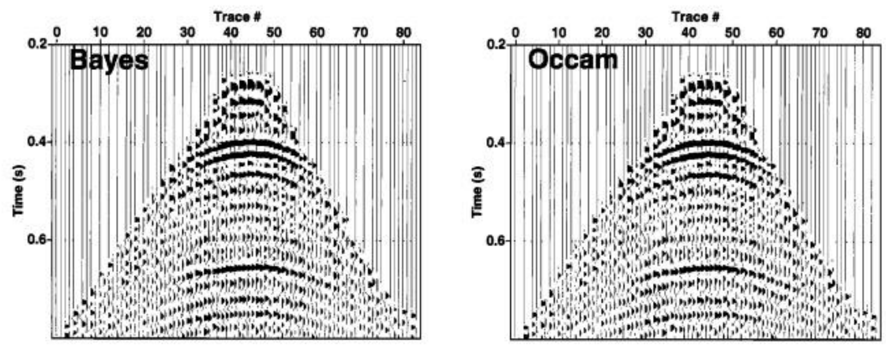
Model Covariances – Gravity example



Uncertainties – A Comprehensive Example

- Gouveia & Scales method:
  - Combining well-log and seismic data
  - Split the errors into 4 parts:
    - Random noise – use data from before the first arrival
    - Near-surface heterogeneities – model many different scenarios
    - Modelling errors – model with many different discretisations
    - Scaling factor (to match field and synthetic data) – try many, compute mean and σ
    - Add these together, which assumes each component is Gaussian





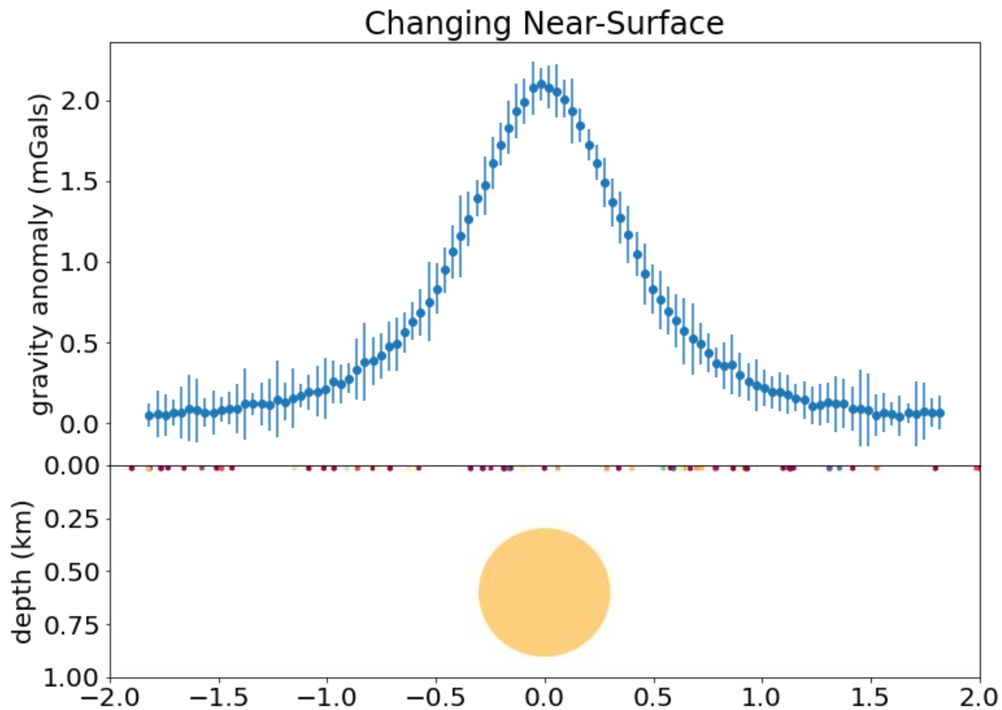
JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 103, NO. B2, PAGES 2759–2779, FEBRUARY 10, 1998

Bayesian seismic waveform inversion: Parameter estimation and uncertainty analysis

Wences P. Gouveia[1] and John A. Scales
Department of Geophysics, Center for Wave Phenomena, Colorado School of Mines, Golden

Prior Uncertainties and χ^2

- Suppose we now have an estimate of our data uncertainties.
- We can often fit this noise with model details
- This is obviously not realistic



Prior Uncertainty and χ^2

Start from our 'usual' objective function:

$$\chi^2(\mathbf{m}) = \frac{1}{2}\left(\mathbf{m} - \mathbf{m}^{prior}\right)^T C_{\mathbf{m}}^{-1}\left(\mathbf{m} - \mathbf{m}^{prior}\right) + \frac{1}{2}\left(\mathbf{Gm} - \mathbf{d}^{obs}\right)C_{\mathbf{d}}^{-1}\left(\mathbf{Gm} - \mathbf{d}^{obs}\right)$$

We want to minimize this, but not 'all the way' because we don't want to fit the noise. We use the χ^2 test.

$$\chi^2(\mathbf{m}) \approx N$$

Here $N = N_{data} - N_{constraints} = N_{degrees\ of\ freedom}$

A good explanation of the details is here: P 79 of http://experimentationlab.berkeley.edu/sites/default/files/pdfs/Bevington.pdf

Prior Uncertainty and χ^2

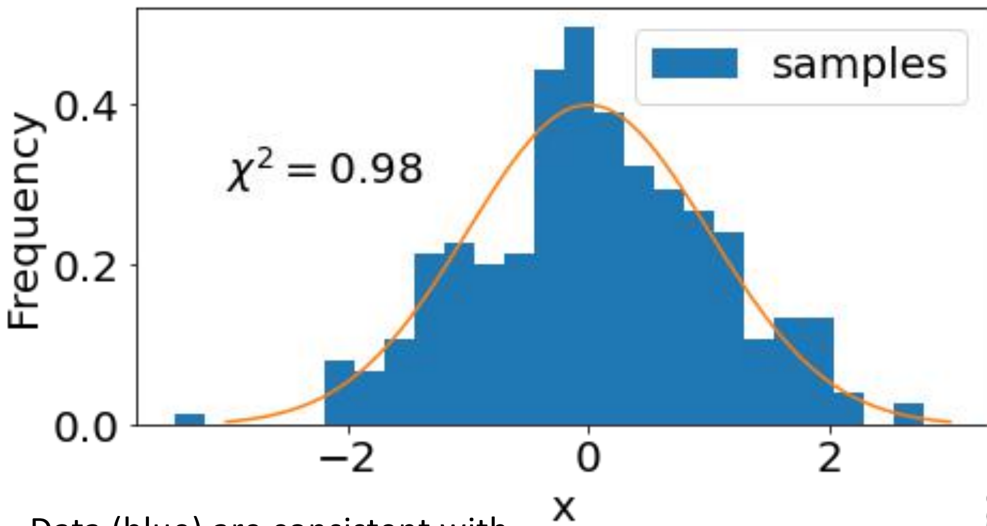
- Try a simpler problem to see where this comes from:

  - Suppose we have N samples of a distribution $P(x)$. Suppose we've discretized x into k possible outcomes. We'd expect to observe x_j a number of times determined by $P(x)$, more precisely $\mathrm{N}P(x)$ times. But this is of course not exactly what we observe instead we observe $h_j(x)$. The χ^2 test checks how close our estimate $h_j(x)$ is to $\mathrm{N}P(x)$, more specifically we calculate:

$$\chi^2(x) = \sum_{j=0}^{k} \frac{\left(h_j(x) - NP(x)\right)^2}{N\sigma^2} = \frac{measured\ data\ spread}{predicted\ data\ spread} \approx 1$$

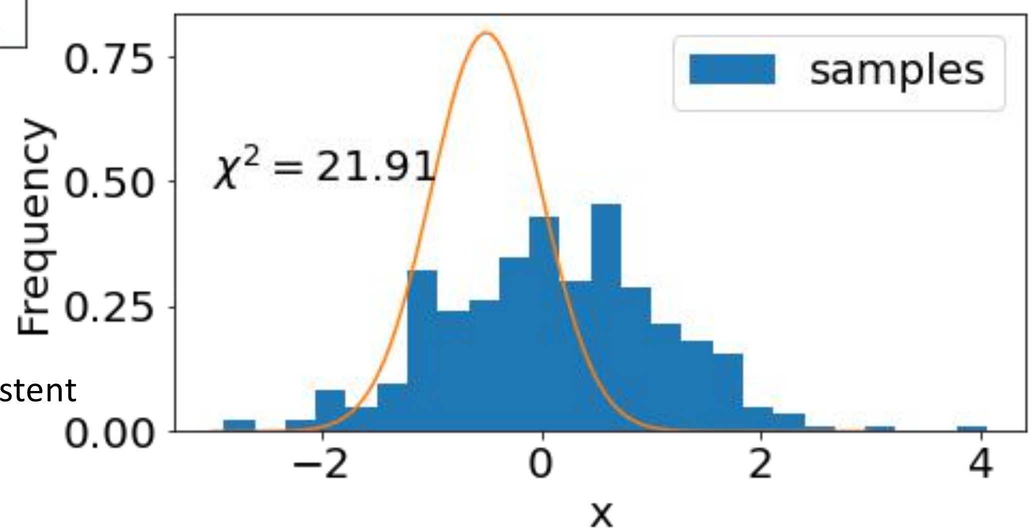
A good explanation of the details is here: P 79 of http://experimentationlab.berkeley.edu/sites/default/files/pdfs/Bevington.pdf

Prior Uncertainty and χ^2



Data (blue) are consistent with the underlying distribution (orange)

What we are testing, is are our modelled data 'statistically consistent' with our field data.



Data (blue) are NOT consistent with the underlying distribution (orange)

A good explanation of the details is here: P 79 of http://experimentationlab.berkeley.edu/sites/default/files/pdfs/Bevington.pdf

Summary so far

- We split our **prior** uncertainties into $C_{\mathbf{m}}$ and $C_{\mathbf{d}}$
- Neither is easy to estimate
- There are ad hoc ways to do so
- We don't want to fit our data perfectly – use χ^2 test

Onwards to incorporate these errors into our final models!

Prior uncertainties -> Posterior uncertainties

Likelihood – where all our carefully estimated covariances come in

Prior – this is what we know before we experiment

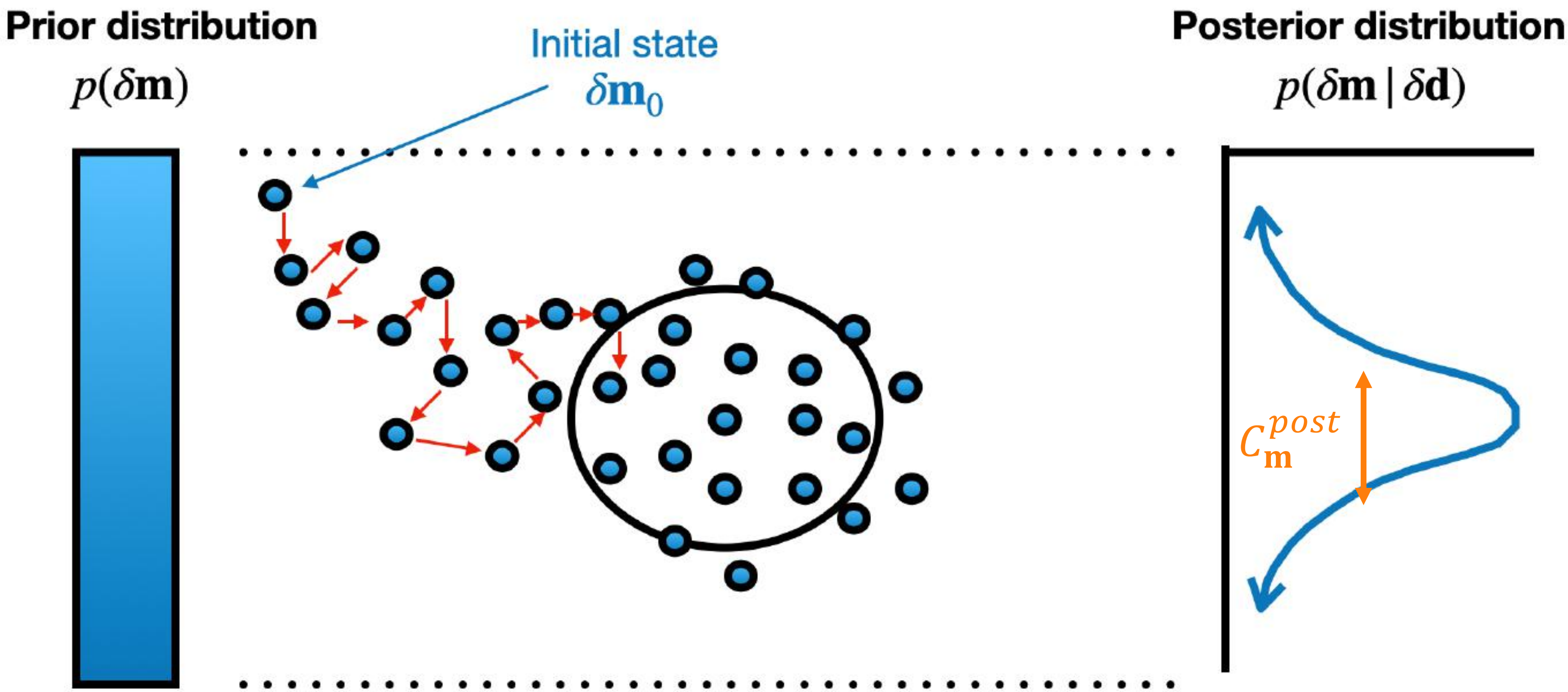
$$p(m|d) = \frac{p(d|m)p(m)}{p(d)}$$

Posterior – its covariance is often what we're after

This is often ignored (I will follow Andreas and leave this to Thomas)

Prior uncertainties → Posterior uncertainties

the sampling edition



Prior uncertainties $\rightarrow$ Posterior uncertainties
the linear optimization edition

- We can show that:

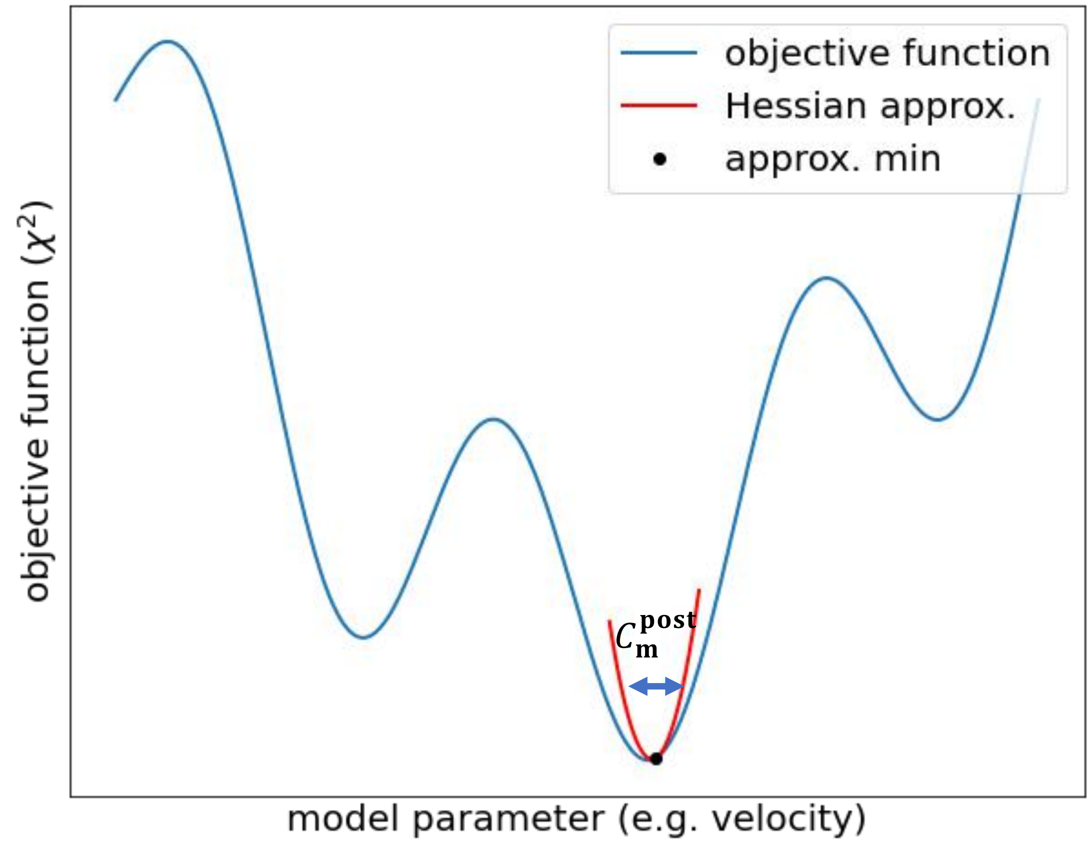
$$C_{\mathbf{m}}^{post} = (H + (C_{\mathbf{m}}^{prior})^{-1})^{-1}$$

- Just differentiate the misfit function and you find this relationship (or look at section 3.4 of Tarantola's 2005 book)
- Intuitively:
  - The Hessian measures the (local) curvature of the misfit function
  - The (inverse) covariance measures the curvature of a distribution
  - OR The Hessian measures how two points in our forward model are related to one another and the covariance measures how two points in our model space are correlated

Prior uncertainties $\rightarrow$ Posterior uncertainties

the linear optimization edition

- If we find $C_{\mathbf{m}}^{\mathbf{post}}$ this way, we are approximating our objective function locally by a parabola
- This gives us an estimate of how well resolved our model parameters are



We have an a posteriori covariance … are we done?

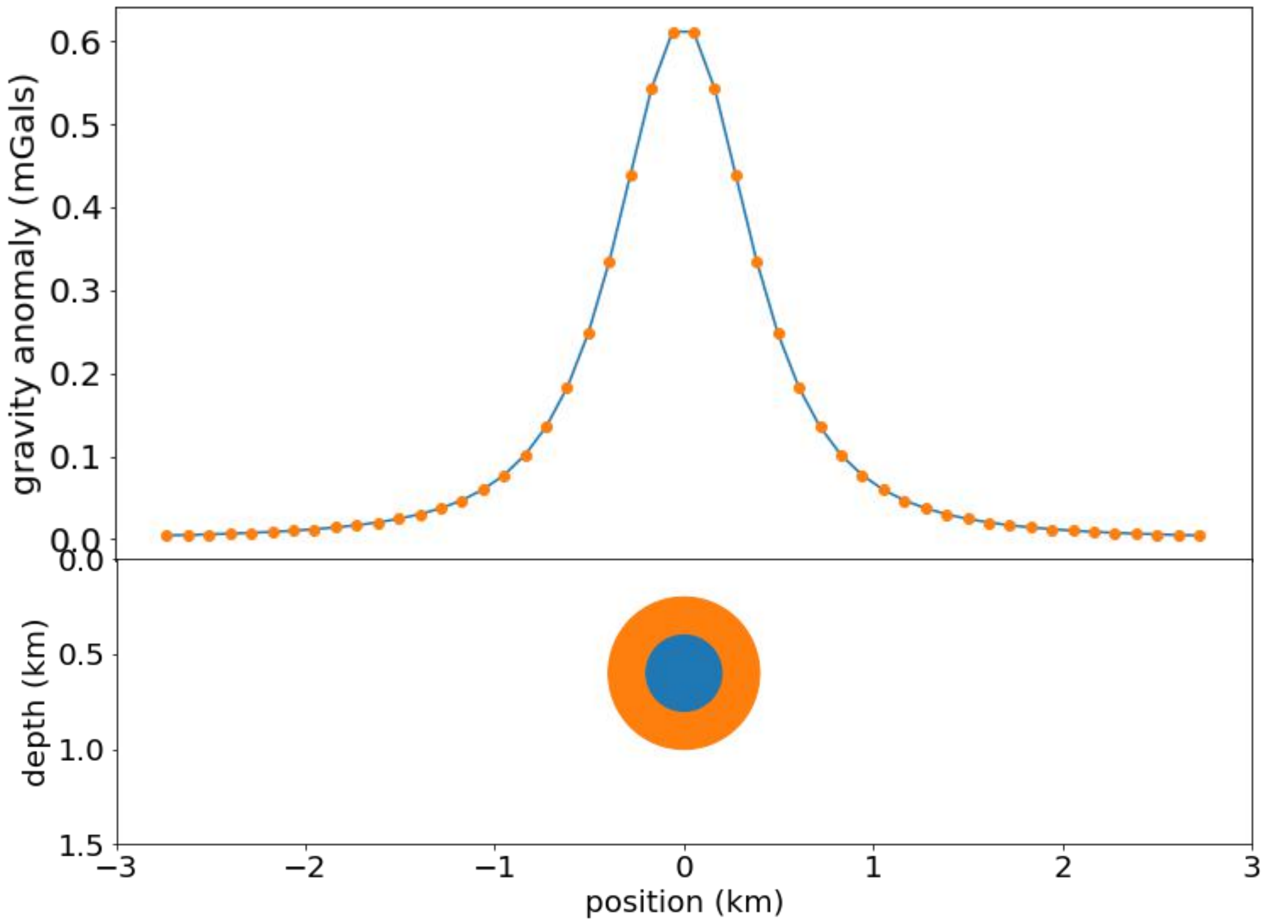
What about:

- The nullspace?
- Did I estimate the full covariance?
- How do I take this covariance and use it to answer my original questions?

The Nullspace – Definition

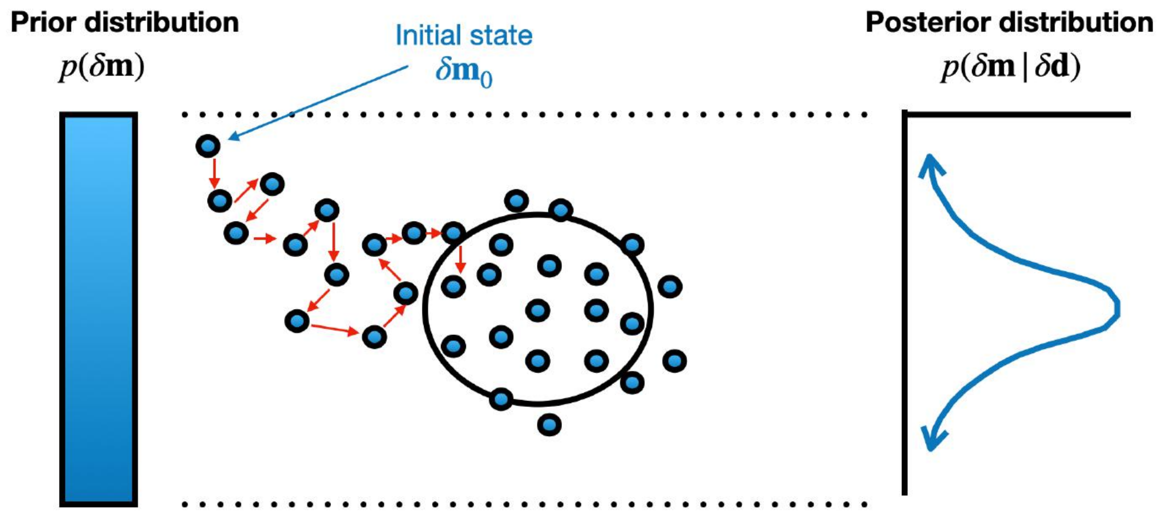
- Unresolvable model differences
- Multiple models give exactly the same likelihood
- For gravity of a sphere: $\{(R,\rho)|R^3\rho = k\}$

How would you find and characterize the nullspace of your problem?



The Nullspace
the MCMC edition

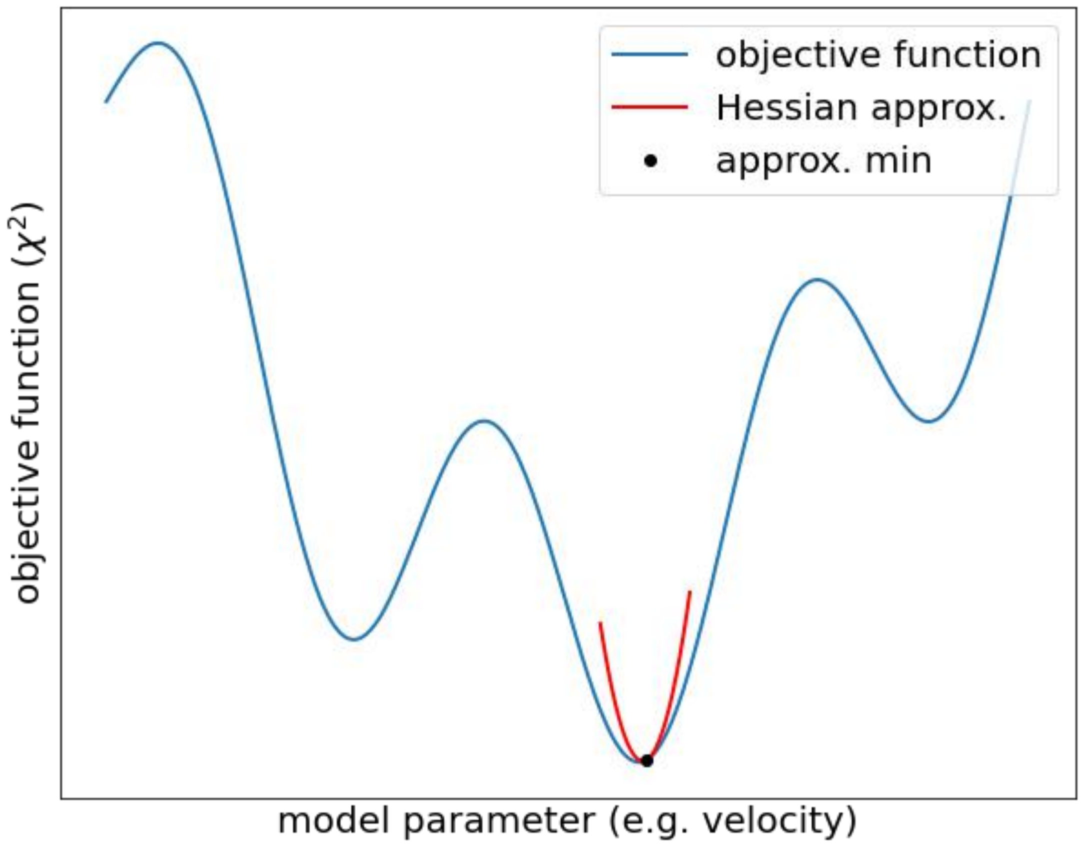
- Here we will sample the nullspace, but what parts and how completely may depend on many things.
- This will also be true as we think about variational inference etc
- Everything with MCMC is in the 'infinite samples' limit



The Nullspace

the linear optimization edition

- If we just find $C_{\mathbf{m}}^{\mathbf{post}}$ we are approximating our objective function locally by a parabola
- **This tells us nothing about the nullspace**
- It does tell us how well resolved our model parameters are



We have an a posteriori covariance … are we done?

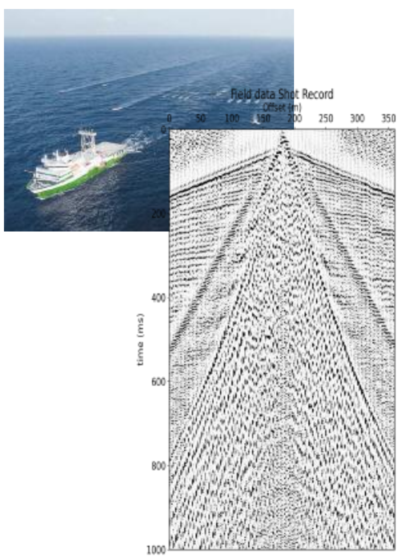
What about:

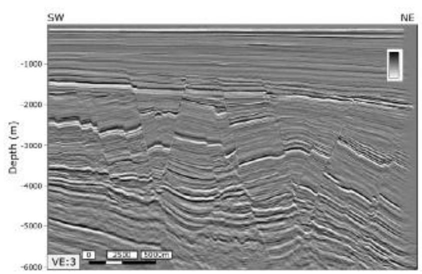
- The nullspace?
- **Did I estimate the full covariance?**
- How do I take this covariance and use it to answer my original questions?

Example 1: 4D seismic

1. Collect data, and form an image
2. Change something (e.g. CO_2)
3. Re-collect the same data, matching everything you can, and form another image
4. Subtract the two resulting images

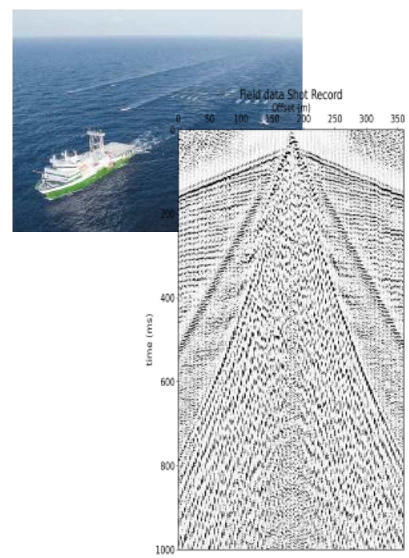
Before:

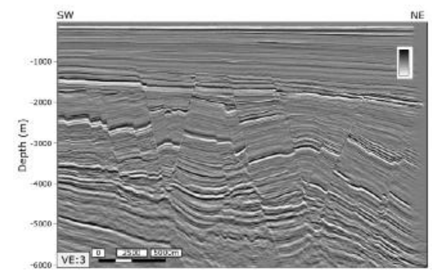




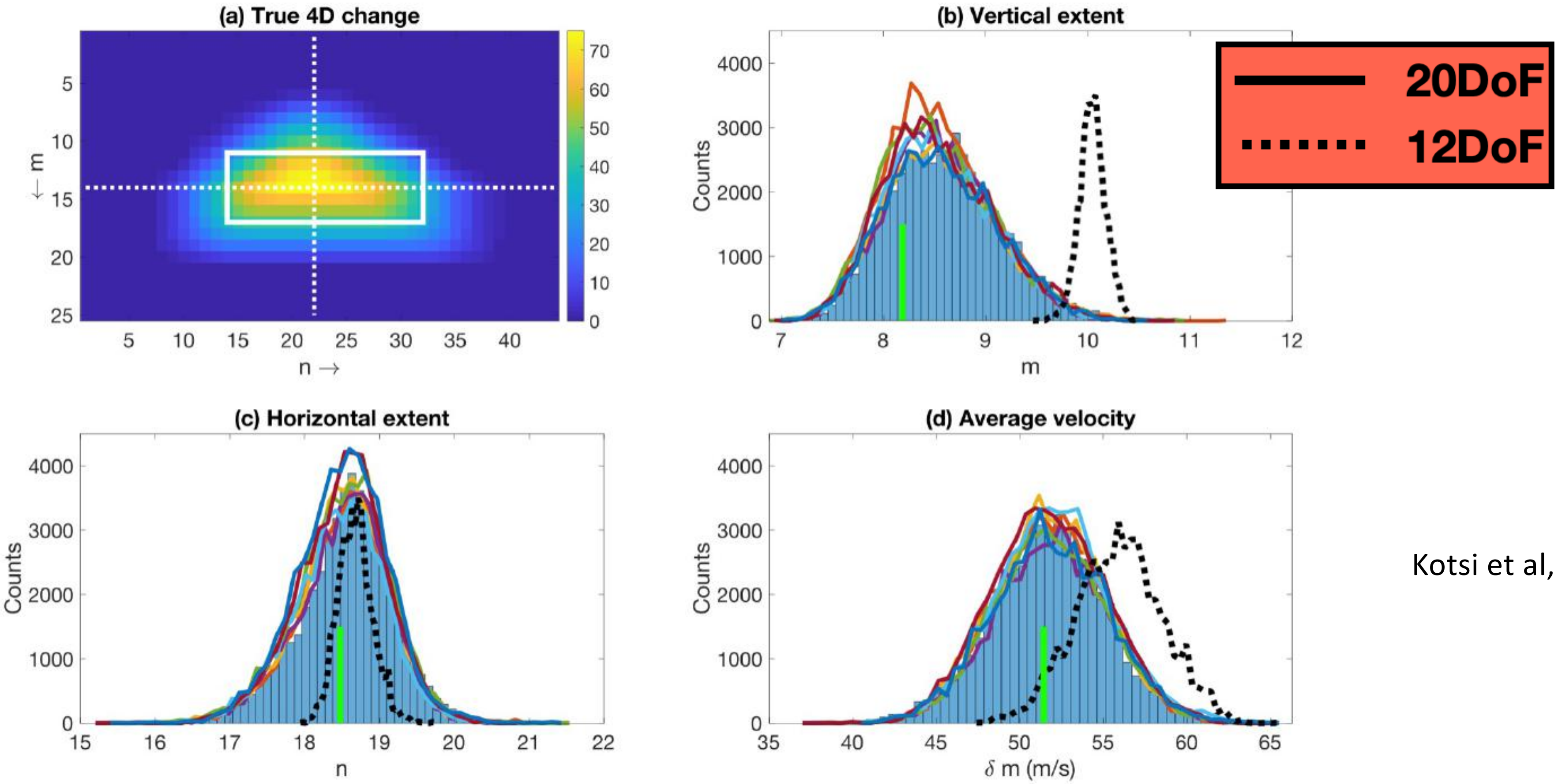
Injection:

After:





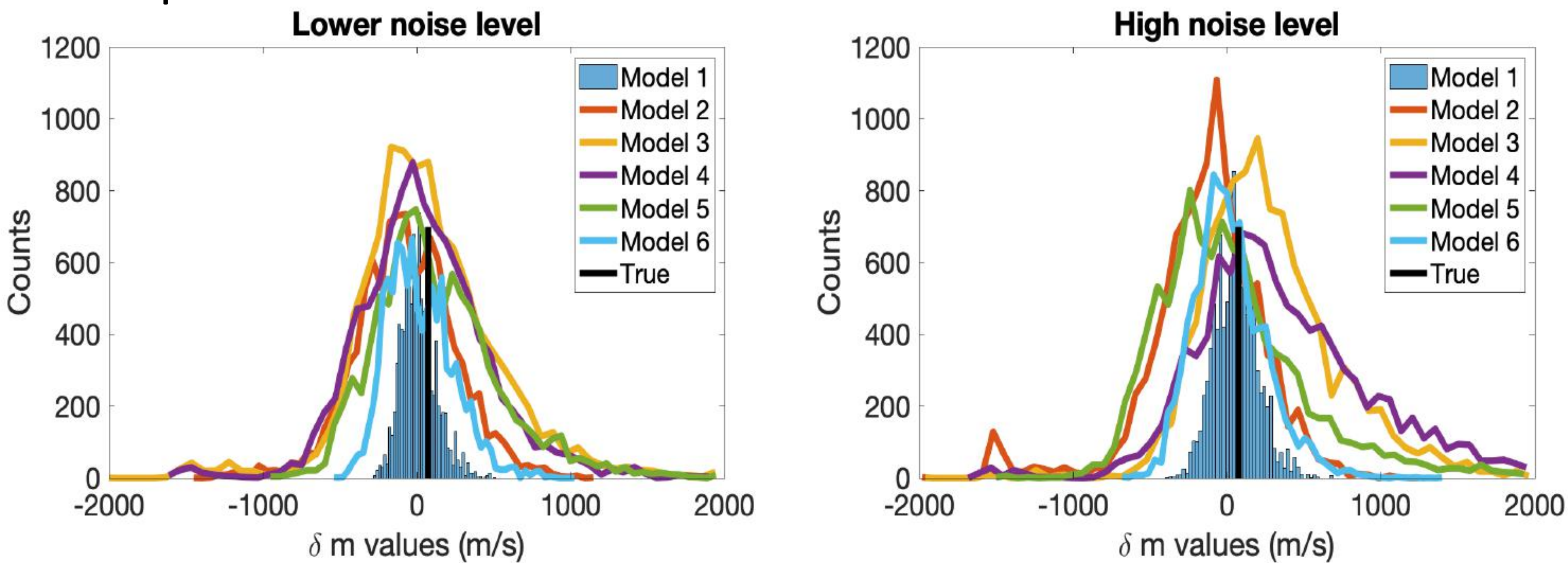
Example 1: 4D Seismic



Kotsi et al, GJI, 2020

What I want to know controls how I choose my model parameters

Example 1: 4D Seismic

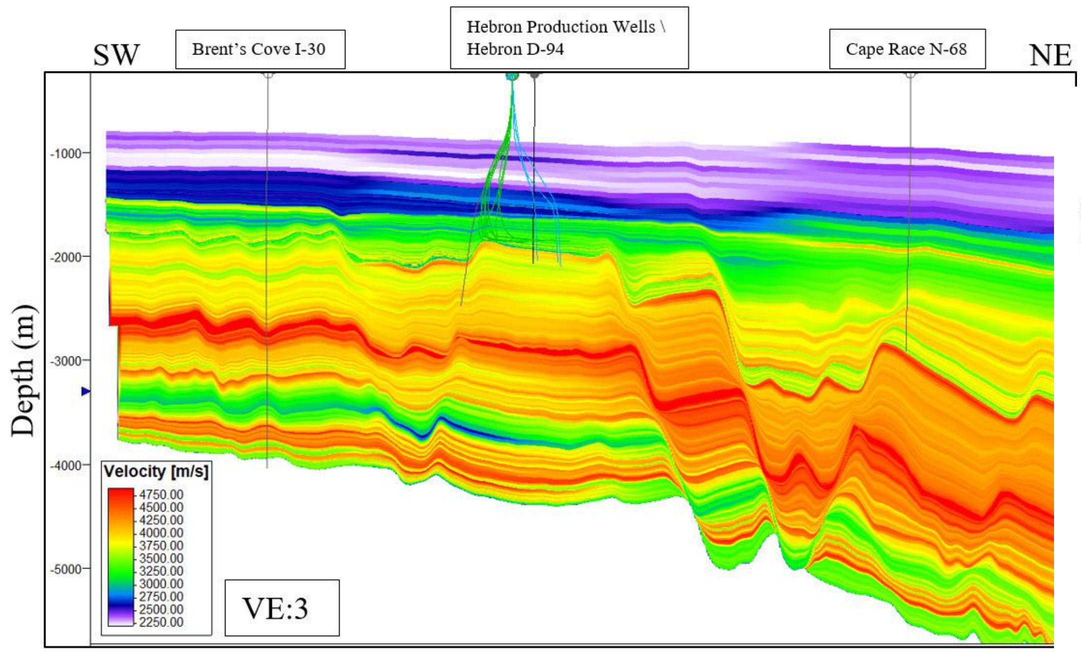


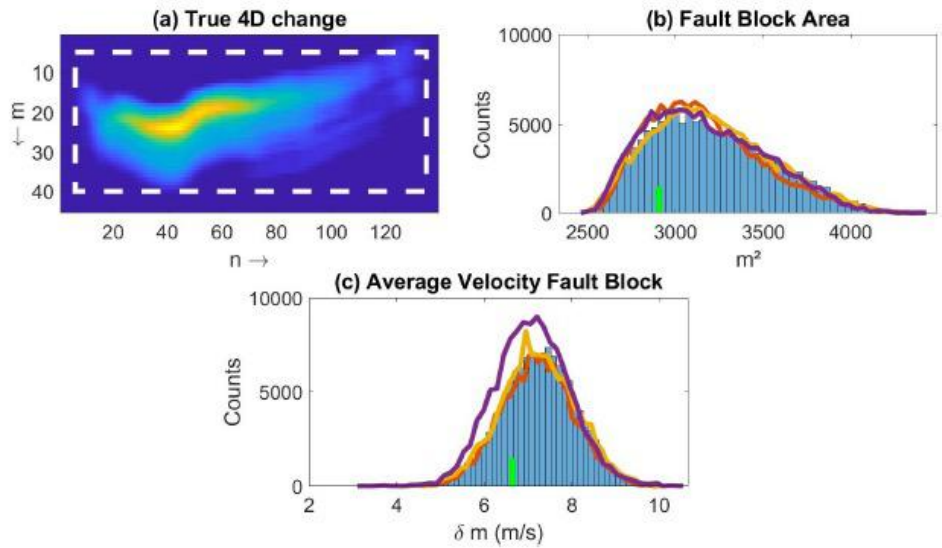
Here we started from many different baseline models, to attempt to propagate uncertainties from one part of the process to the next.

Some uncertainties are unimportant for our final answer!

Kotsi et al, GJI, 2020

Example 1.5: Field Data

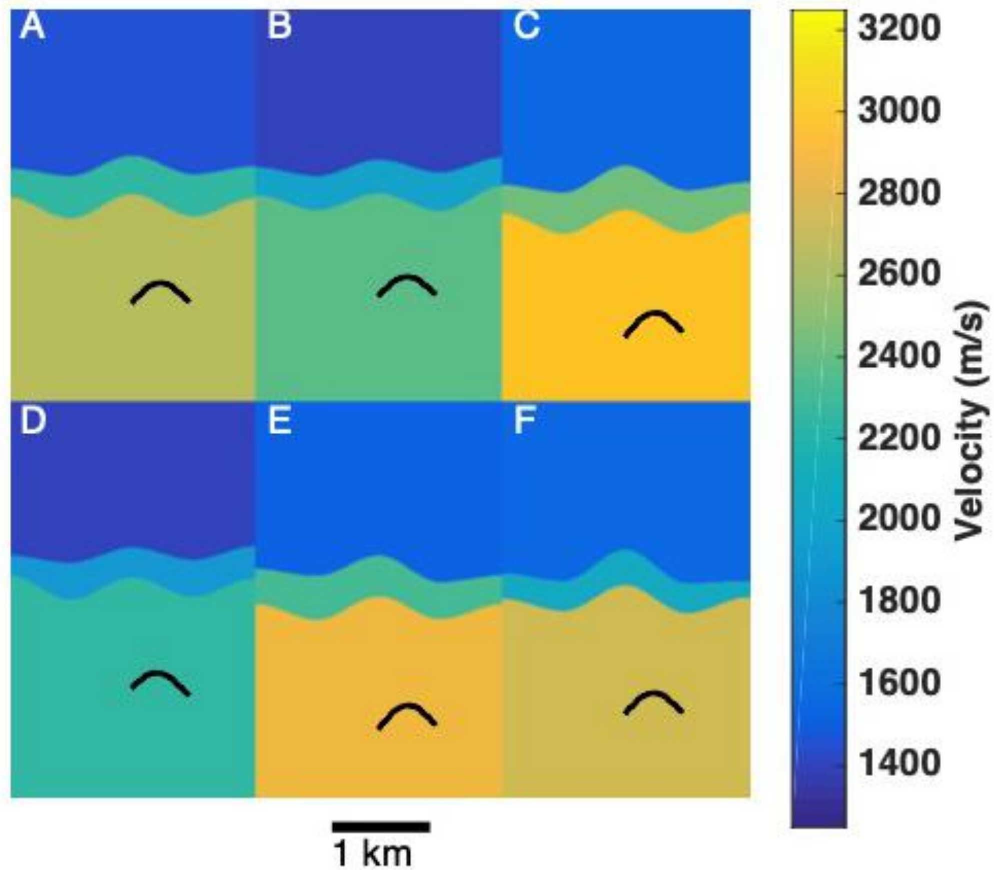




Lethbridge, MSc thesis, 2021.

Example 2: Seismic imaging/migration

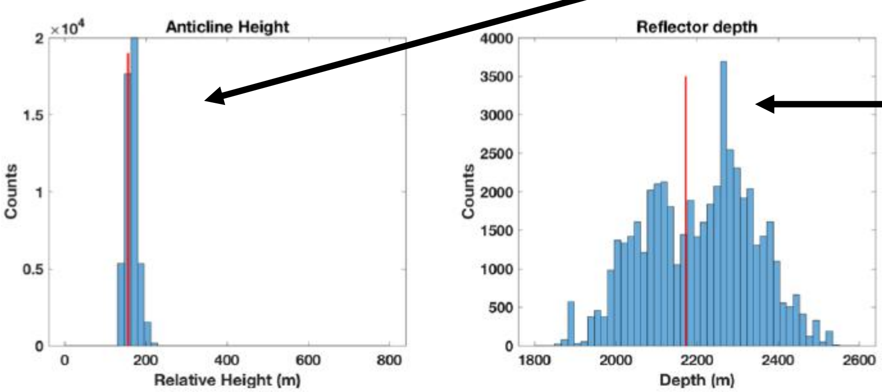
- These are all models, generated during a stochastic optimization, that fit the data to within our acceptable criteria
- How can we effectively show this?

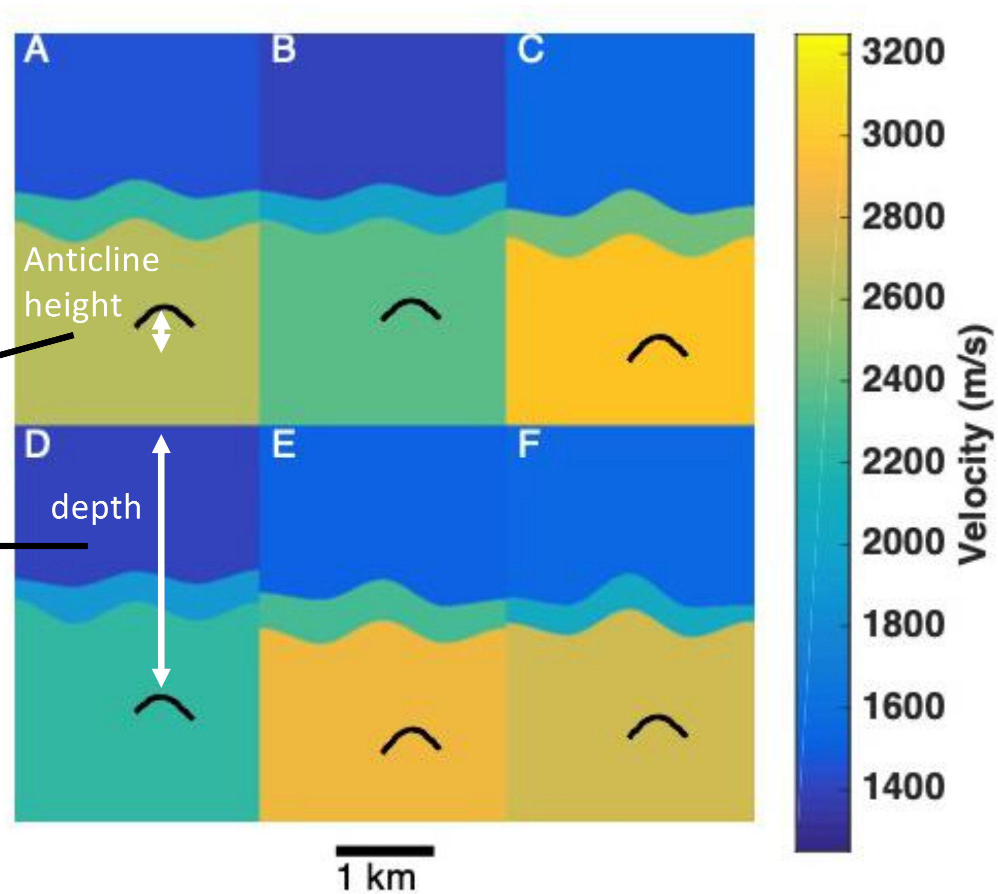


Ely et al, Geophysics, 2018

Example 2: Seismic imaging/migration

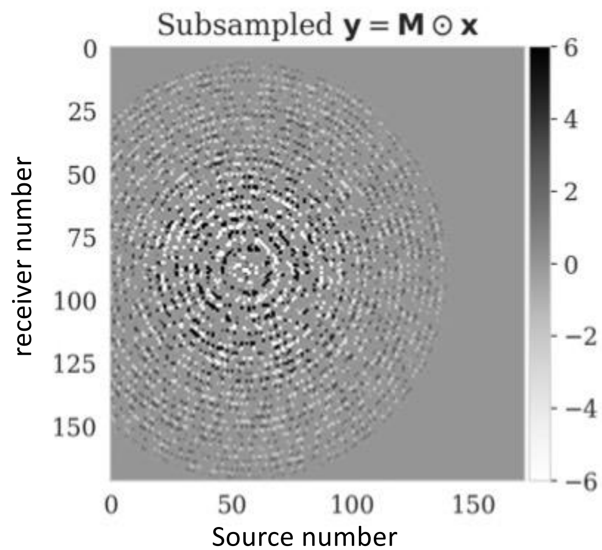
- These are all models, generated during a stochastic optimization, that fit the data to within our acceptable criteria
- How can we effectively show this?



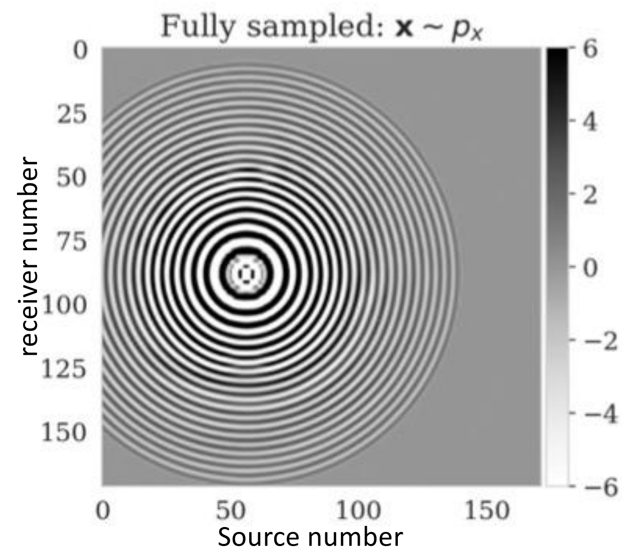


Ely et al, Geophysics, 2018

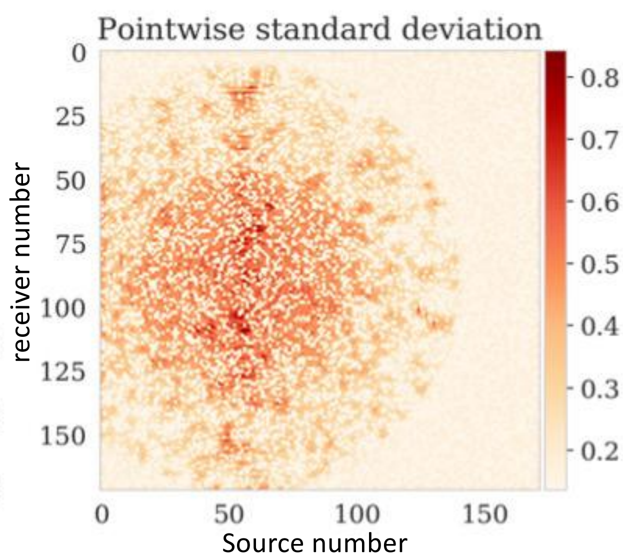
Example 3: Seismic interpolation



We use normalizing flows (a variational method) to interpolate data



Here we want the pointwise standard deviation because we will eventually use it as a weight in our velocity inversion



Kumar et al, 2022

Summary

- "There are known unknowns and unknown unknowns"

  (Maybe Donald Rumsfeld, 2002)
- It is important to characterize your uncertainties, but also to understand and convey what your method leaves out
- Before you start to estimate and quantify your uncertainties, think carefully about what you want to learn/understand/convey and make sure that you are estimating the right parameters, and the most important sources of uncertainty

Summary

- "There are known unknowns and ~~unknown unknowns~~"

  (Maybe Donald Rumsfeld, 2002)

- It is important to characterize your uncertainties, but also to understand and convey what your method leaves out
- Before you start to estimate and quantify your uncertainties, think carefully about what you want to learn/understand/convey and make sure that you are estimating the right parameters, and the most important sources of uncertainty

A very incomplete list of interesting references

- Gouveia & Scales, 1997, 1998
- Scales & Tenorio, 2001
- Zheng & Curtis, 2021
- Zhu et al, 2016
- Nawaz & Curtis, 2018
- Bui-Tanth et al, 2012
- Martin et al, 2012
- Mosegaard & Sambridge, Sambridge & Mosegaard, 2002