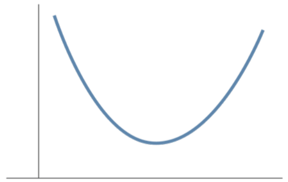
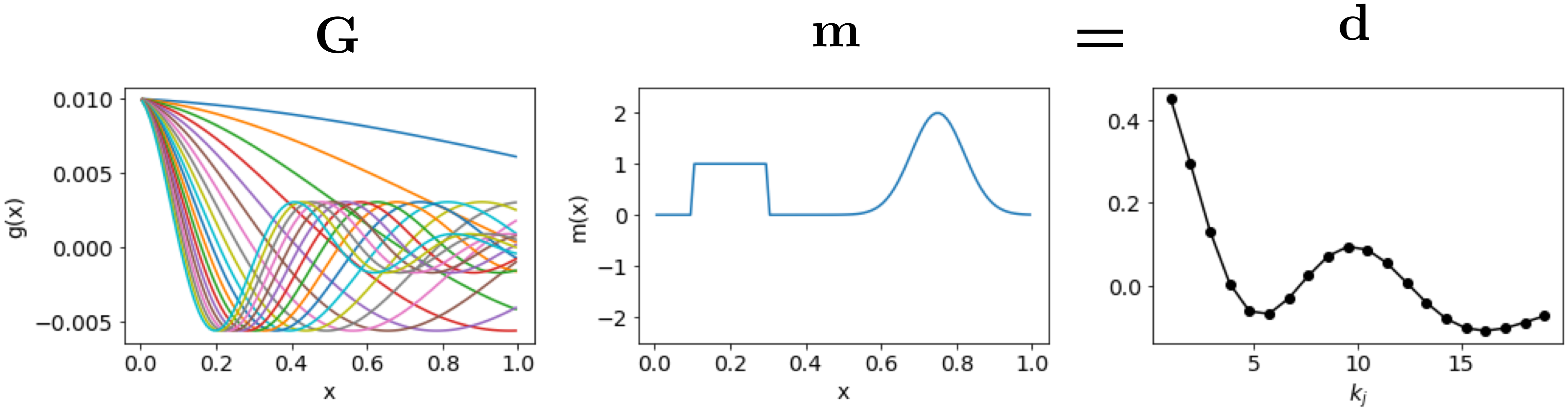
Linear L_2 inversion



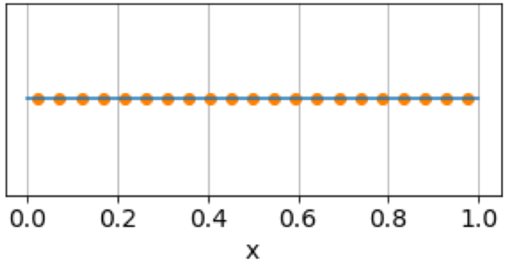
Outline

- Linear inversion using L2 norms
- 1D inversion app
- 2D cross-well tomography
- Incorporating prior information
  - Alpha parameters
  - Reference model
- 3D magnetics
  - Role for weightings

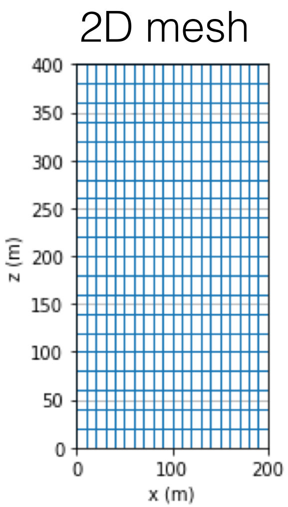
1D linear example

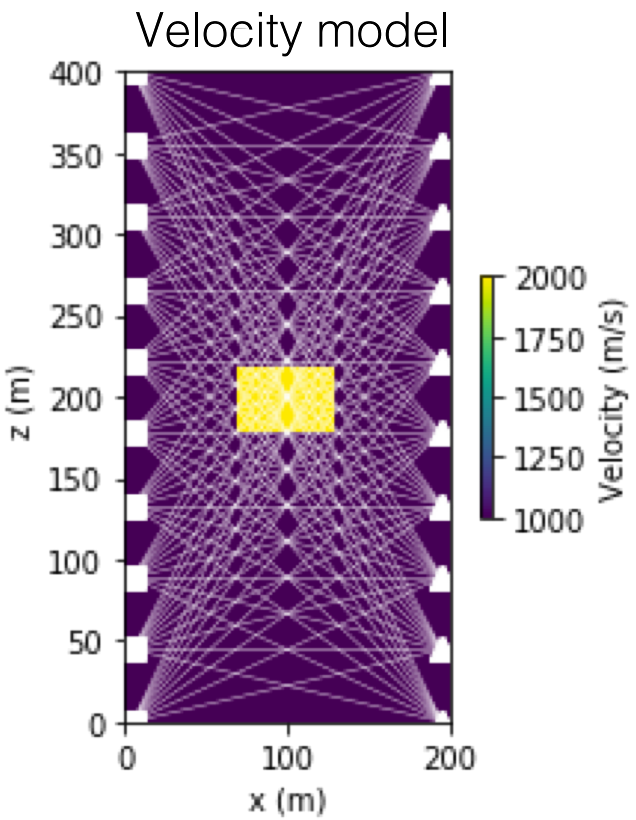


1D mesh

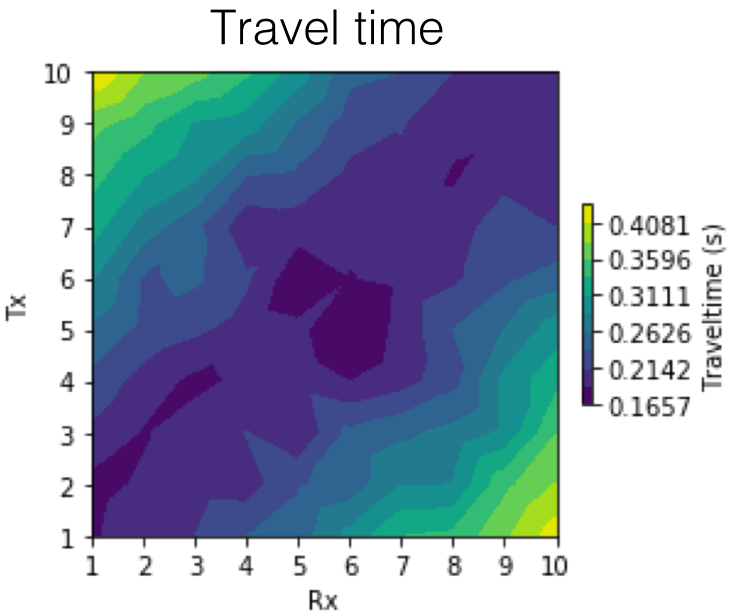


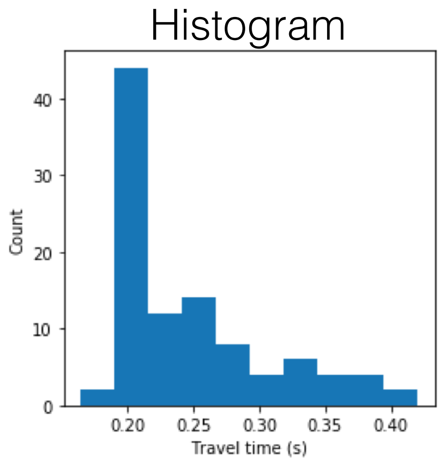
2D Crosswell Tomography





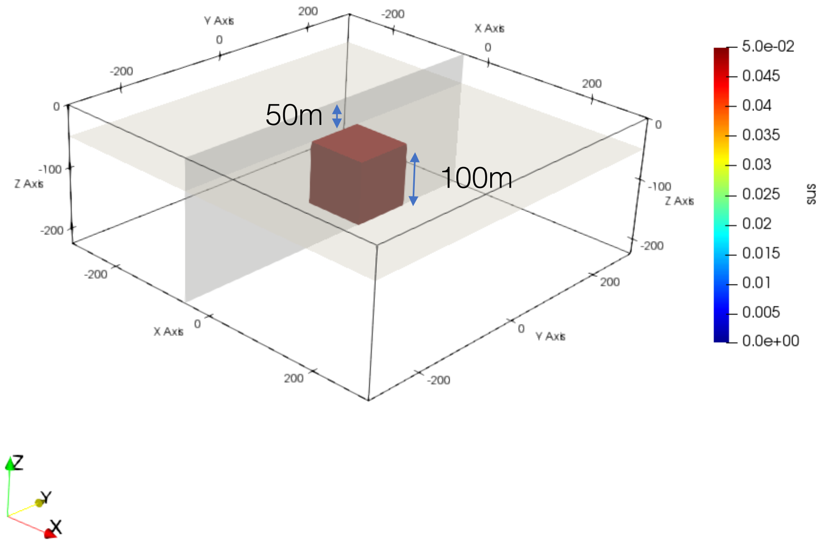
$$\mathbf{Gm} = \mathbf{d}$$





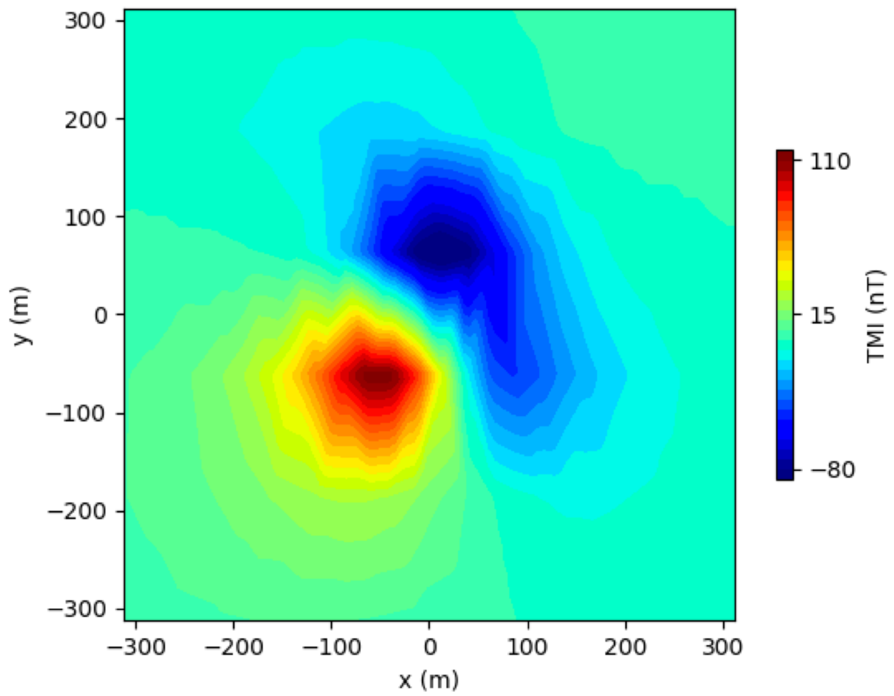
3D Magnetic

Susceptibility model

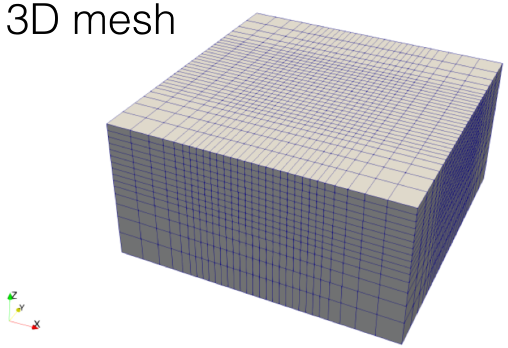


$$\mathbf{Gm} = \mathbf{d}$$

Magnetic data



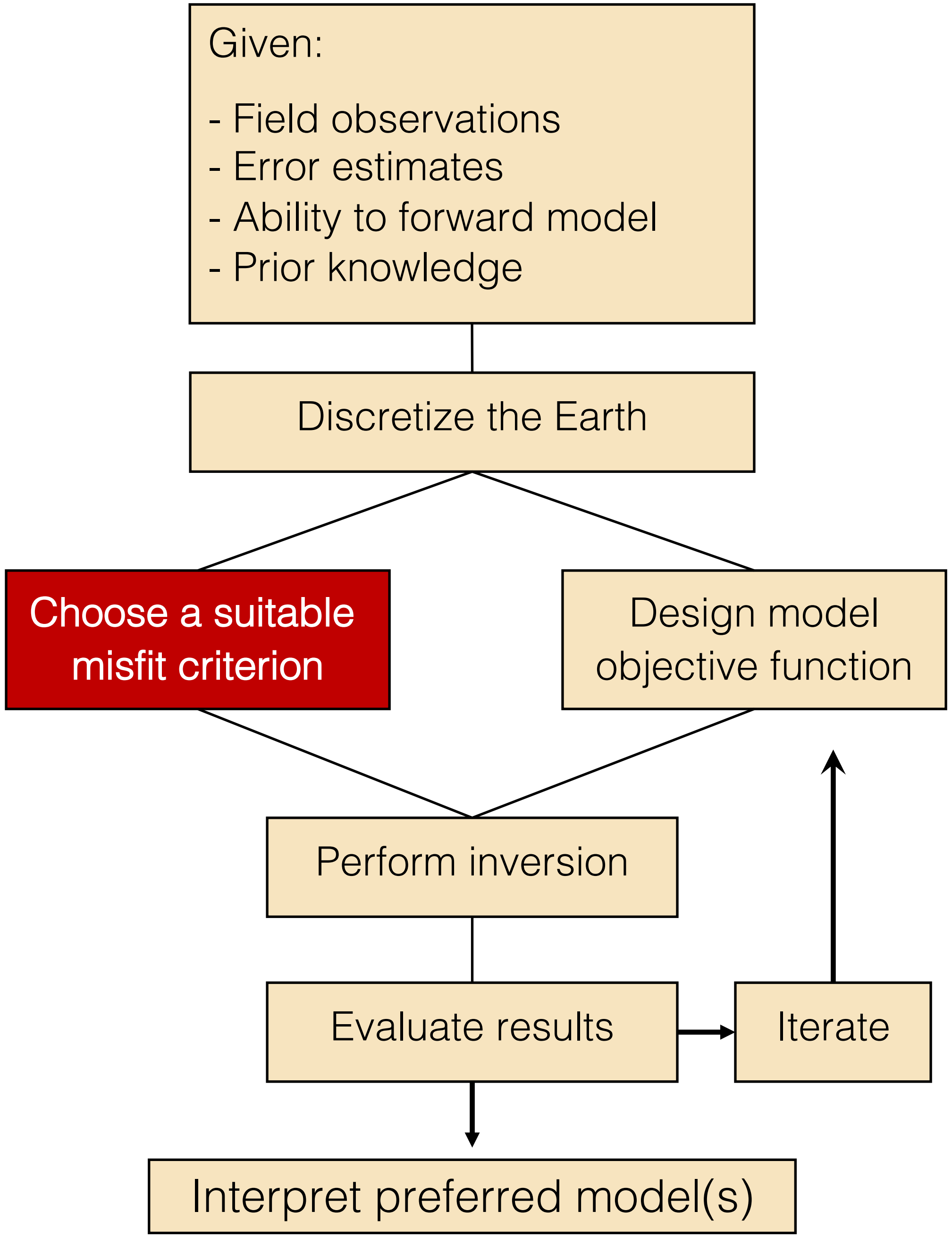
3D mesh



$$\triangle \vec{B} \simeq \vec{B}_A \cdot \hat{B}_0$$

$\vec{B}_0$: Earth magnetic field
$\vec{B}_A$: Anomalous magnetic field

Flow chart for inverse problem



Dealing with uncertainties

Observed datum: $$d_j^{obs} = \mathcal{F}_j[m] + \tilde{n}_j$$

$\mathcal{F}[m]$: ??
m: our sought function
$\tilde{n}$: additive noise

We want to find the model m that gave rise to the data, d_j^{obs}

Let $F_j[m]$ be our mathematical representation for forward operator

It is never exact: $$\mathcal{F}_j[m] = F_j(m) + \hat{n}_j$$

$\hat{n}_j$ are discrepancies

- wrong dimension (1D or 3D)
- incomplete physics
- discretization errors

Dealing with uncertainties

$$\mathcal{F}_j[m] = F_j(m) + \hat{n}_j$$

Statistics of $\hat{n}_j$

- not easily quantified
- biased, convoluted
- problem dependent

Final problem: $$d_j^{obs} = F_j(m) + \underbrace{\hat{n}_j + \tilde{n}_j}_{n_j}$$

$$d_j^{obs} = F_j(m) + n_j$$

n_j: statistical variable that accounts for all "noise"

Assume

Gaussian distribution with zero mean, standard deviation (ϵ_j)

Dealing with uncertainties

Consider $\quad \text{random variable, } x_j \in \mathcal{N}(0,1)$

Define $$\chi_N^2 = \sum_{j=1}^{N} x_j^2$$

Chi-squared statistic with *N* degrees of freedom

$$\begin{cases} \text{Expected value: } E(\chi_N^2) = N \\ \text{Variance: } \text{Var}(\chi_N^2) = 2N \\ \text{Standard deviation: } \text{std}(\chi_N^2) = \sqrt{2N} \end{cases}$$

Misfit function

Crucial steps for any misfit:

(1) Specify the metric used
(2) Determine target misfit

We use L_2 norm (least squares statistic)

Define data misfit: $\phi_d = \sum_{j=1}^{N} \left(\frac{F_j(m) - d_j^{obs}}{\epsilon_j} \right)$

Define $\mathbf{W}_d = \mathbf{diag}(1/\epsilon_1, \ldots, 1/\epsilon_N)$

$$\phi_d = \|\mathbf{W}_d(F[\mathbf{m}] - \mathbf{d}^{obs})\|_2^2$$

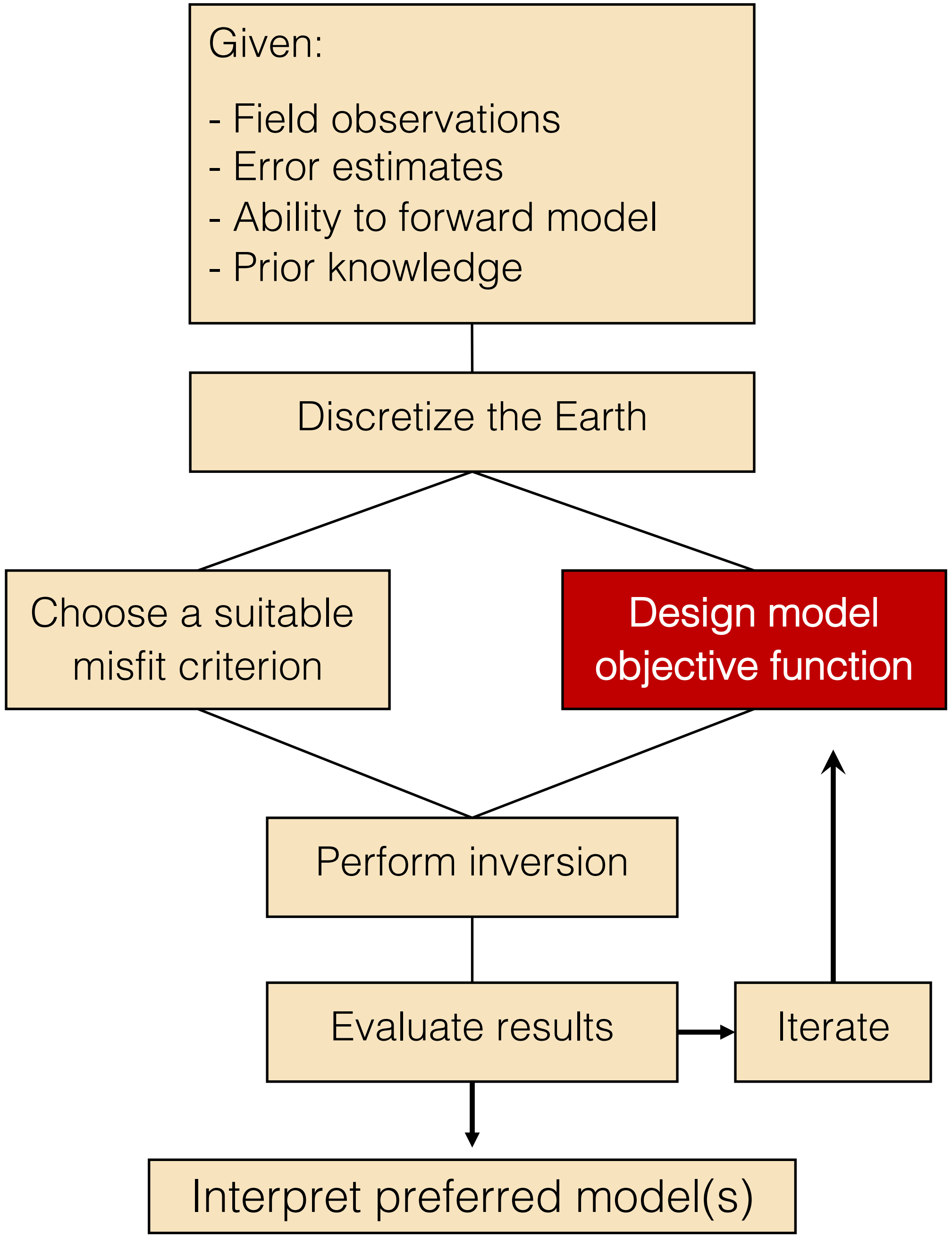
$$E[\phi_d] \simeq N$$

ϕ_d is now a χ_N^2 variable

<u>Reality</u>: we do not know the uncertainties

<u>Try</u>: $\epsilon_j = \%|d_j^{obs}| + \text{floor}$

Flow chart for inverse problem



Model objective function

- 1D problem
- Discretizing
- Roles of alphas

Model norms

Smallest model: $$\phi_m = \int m^2 dx$$

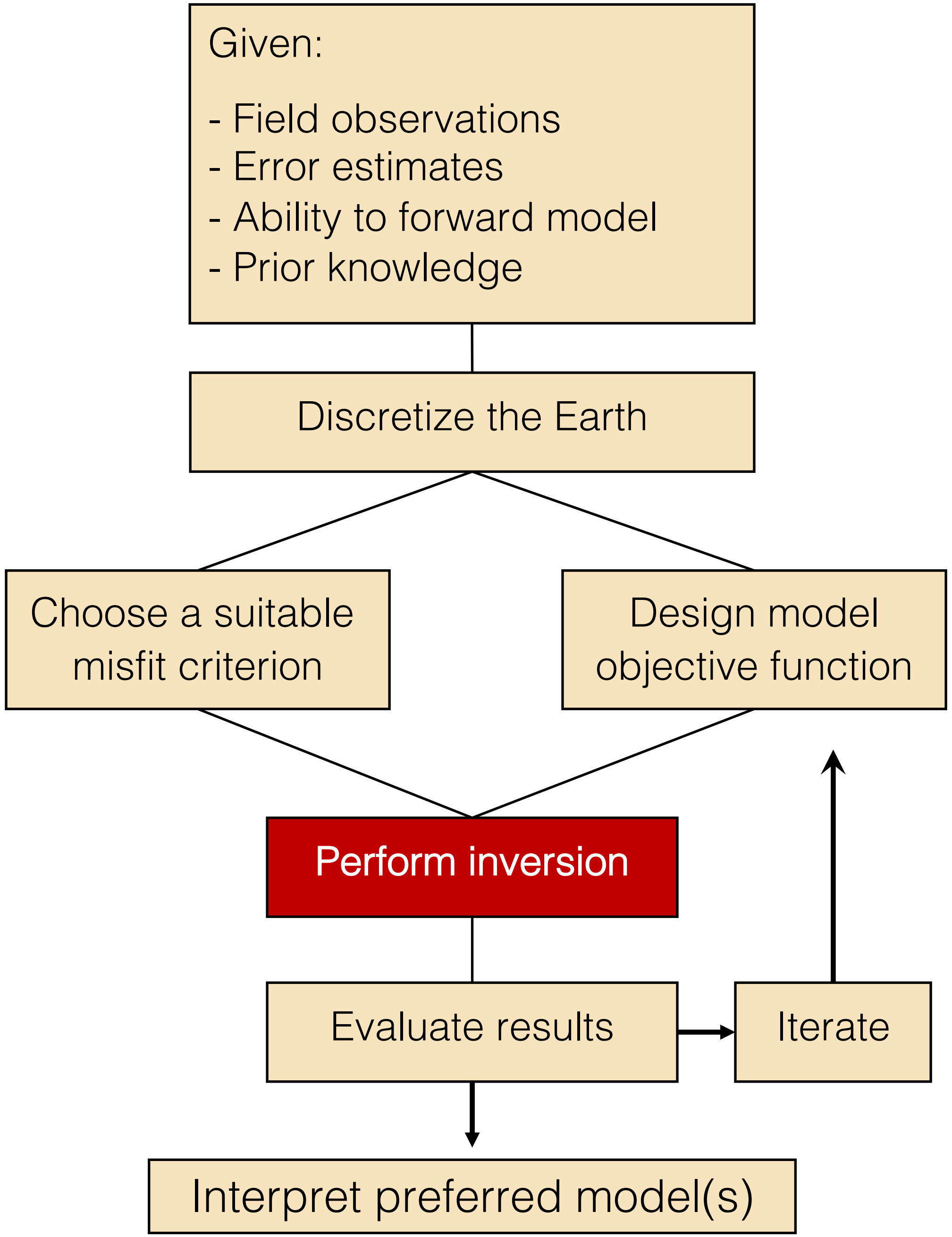
Smallest with reference: $$\phi_m = \int (m - m_{ref})^2 dx$$

Smoothest model: $$\phi_m = \int \left(\frac{dm}{dx}\right)^2 dx$$

Combination: $$\phi_m = \alpha_s \int (m - m_{ref})^2 dx + \alpha_x \int \left(\frac{dm}{dx}\right)^2 dx$$

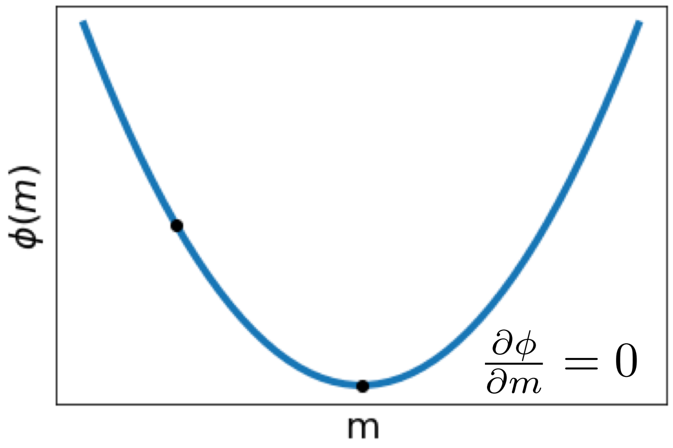
Discretize: $$\phi_m = \alpha_s \|\mathbf{W}_s(\mathbf{m} - \mathbf{m}_{ref})\|_2^2 + \alpha_x \|\mathbf{W}_x(\mathbf{m})\|_2^2$$

Flow chart for inverse problem



Perform inversion

$$\phi(\mathbf{m}) = \frac{1}{2}\|(\mathbf{G}\mathbf{m} - \mathbf{d}^{obs})\|_2^2 + \frac{1}{2}\|\mathbf{W}_m(\mathbf{m} - \mathbf{m}_{\text{ref}})\|_2^2$$



Quadratic objective function
(for a single variable)

$$\mathbf{g} = \nabla_m \phi \quad \mathbf{g} = \mathbf{G}^T\mathbf{W}_d^T\mathbf{W}_d(\mathbf{G}\mathbf{m} - \mathbf{d}^{obs}) + \beta\mathbf{W}_m^T\mathbf{W}_m(\mathbf{m} - \mathbf{m}_{\text{ref}})$$

$$\mathbf{g} = 0 \qquad \left(\mathbf{G}^T\mathbf{W}_d^T\mathbf{W}_d\mathbf{G} + \beta\mathbf{W}_m^T\mathbf{W}_m\right) = \mathbf{G}^T\mathbf{W}_d^T\mathbf{W}_d\mathbf{d}^{obs} + \beta\mathbf{W}_m^T\mathbf{W}_m\mathbf{m}_{\text{ref}}$$

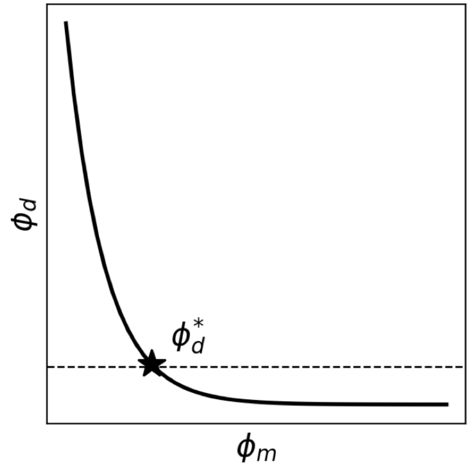
$$\mathbf{A}\mathbf{m} = \mathbf{b}$$

$$\begin{cases} \mathbf{A} \in \mathbb{R}^{M\times M} \text{ is full rank} \\ \mathbf{m}, \mathbf{b} \in \mathbb{R}^M \end{cases}$$

$$\mathbf{m} = \mathbf{A}^{-1}\mathbf{b}$$

The role of β: managing misfit

- Our inverse problem
  - Find the model (m): $\text{minimize} \quad \phi_d + \beta\phi_m$
  - Which beta to use?



- If standard deviations of data are known,

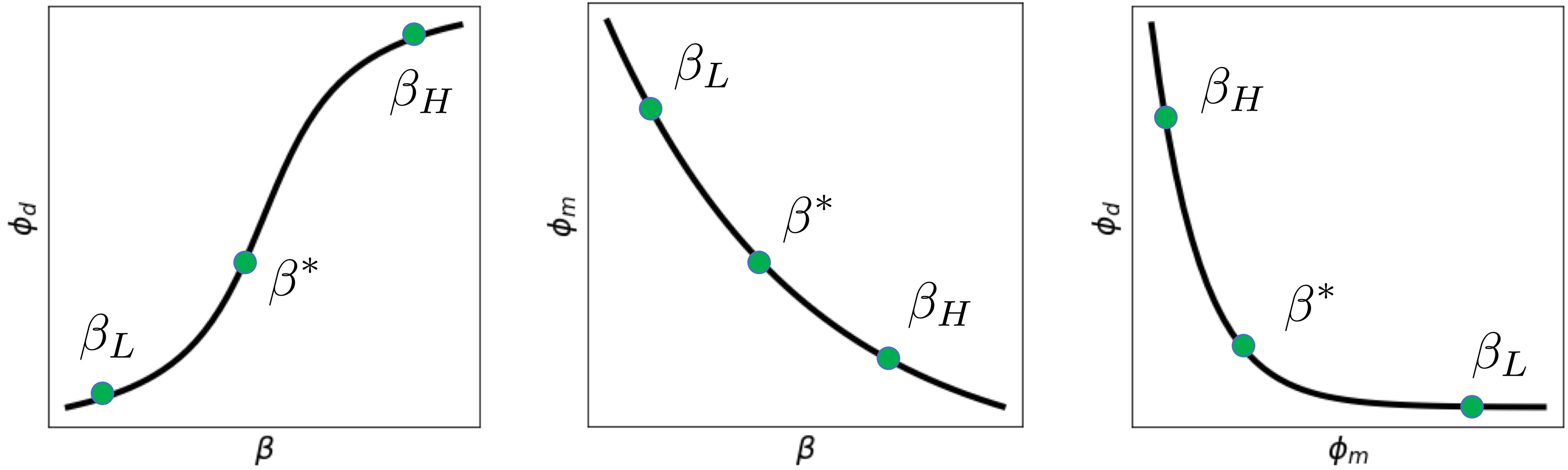
$$E[\phi_d] = N \qquad E[\phi_d]: \text{ expected value of } \phi_d$$

- Desired misfit is $\phi_d^* \simeq N$

- Choose β so that $\phi_d(m) = \phi_d^*$

Choosing beta

Minimize $$\phi(m) = \phi_d(m) + \beta\phi_m(m)$$



Options for choosing β

1) Fix β (Guess)

2) Solve $\mathbf{A}(\beta)\mathbf{m} = \mathbf{b}$

For a logarithmic number of β, choose $\text{(a)}\ \beta^* \quad \phi_d \simeq \phi_d^*$

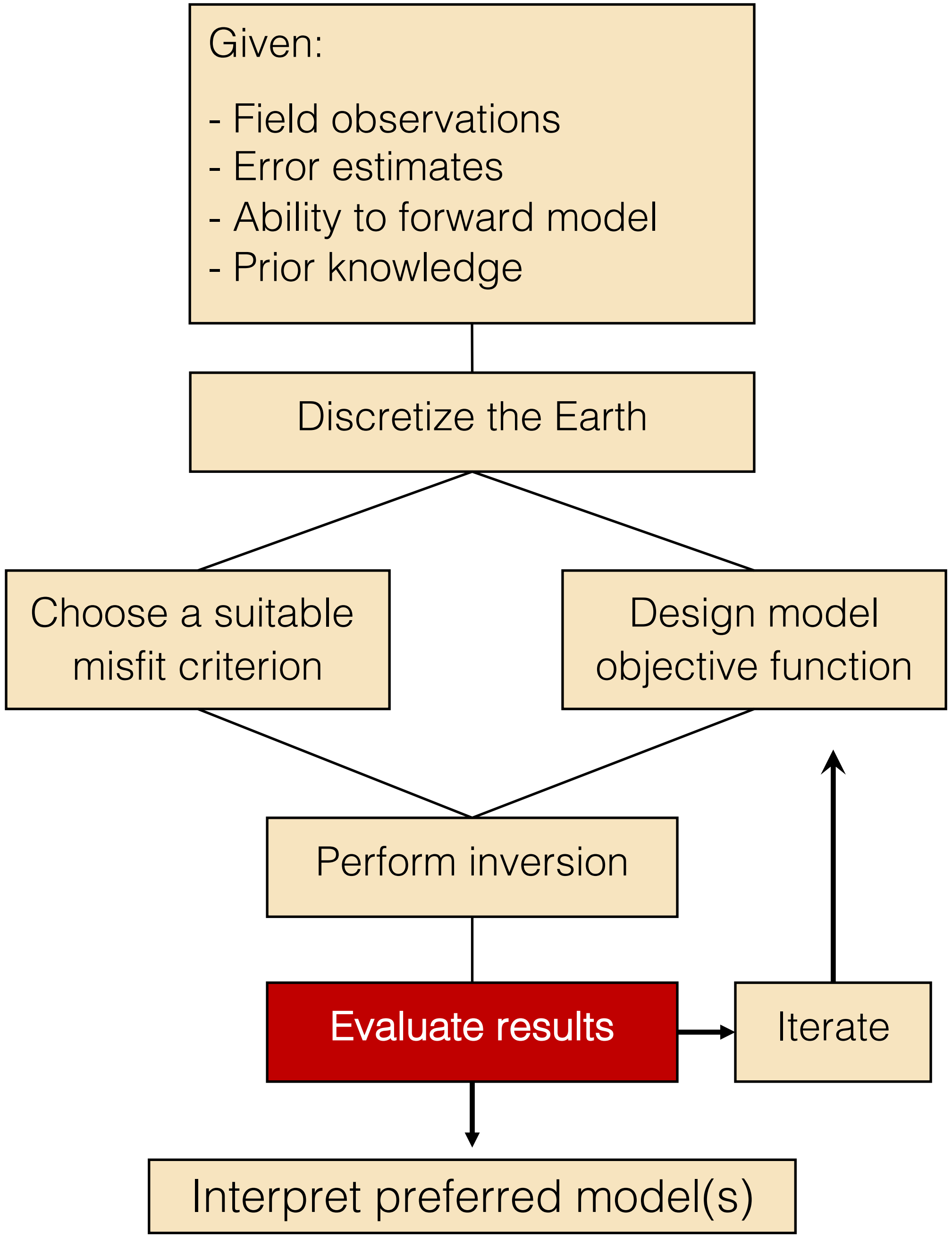
$\text{(b) ``Corner'' (well-known technique)}$

3) GCV (Generalized cross-validation)

4) Cooling strategy: successive inversions

$$\beta_k = \frac{\beta_{k-1}}{\gamma}, \quad \gamma > 1$$

Flow chart for inverse problem

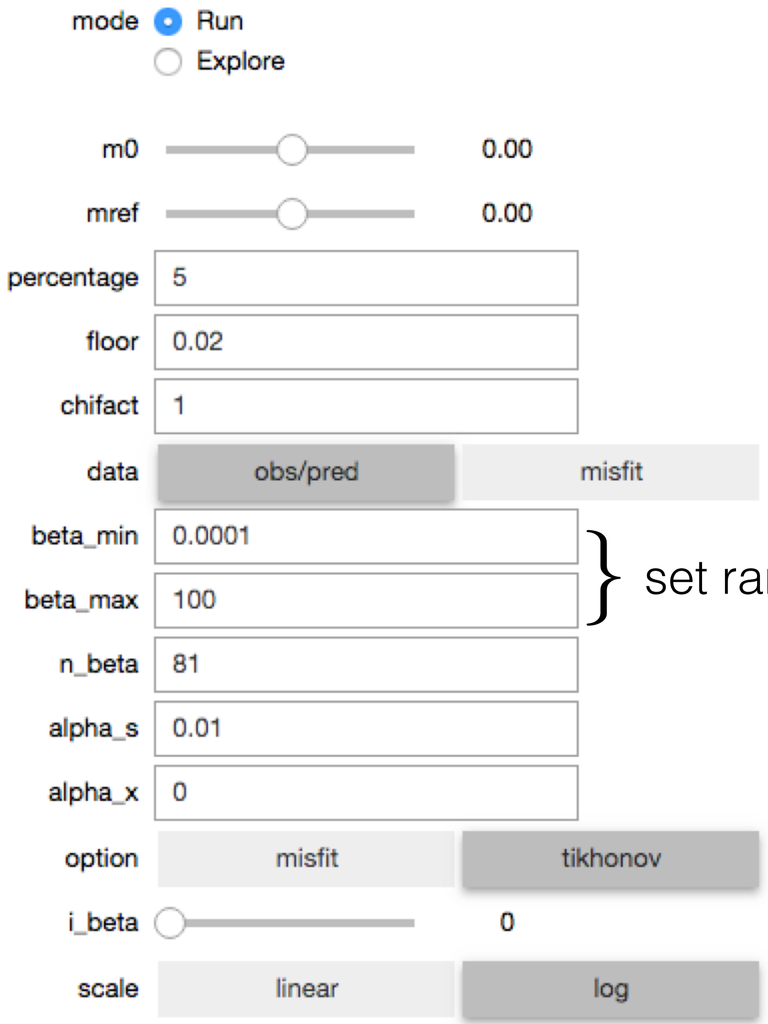


Evaluate results

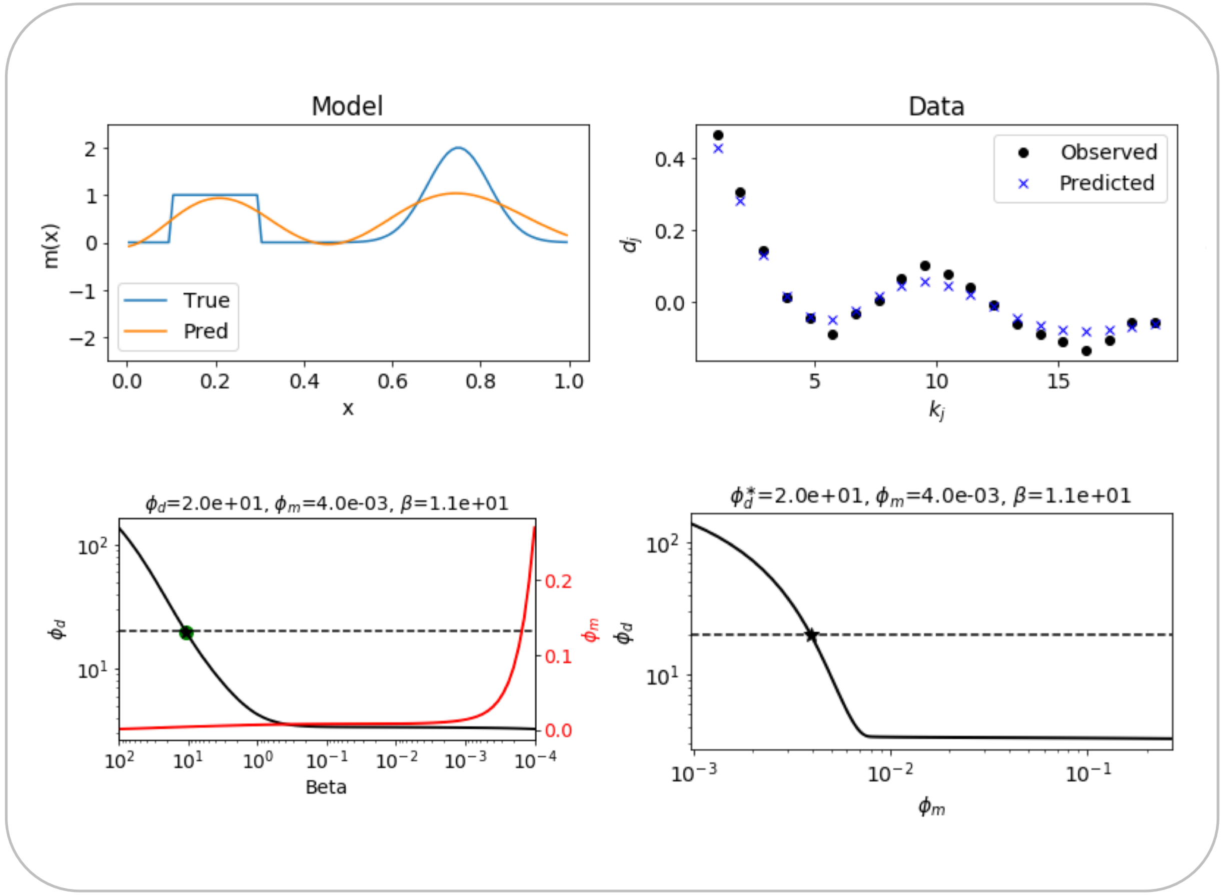
- Evaluate the global misfit
- Plot the Tikhonov curve
- Look at the model: too rough?, too smooth? …
- Look at normalized misfit
- Look at other models with different beta. (three panel plots of model, Tikhonov curve and misfit

- Do I want to change something? (uncertainties, something in the model norm?)

Inversion app (demo)



} set range of β



β is the trade-off parameter

- Solve

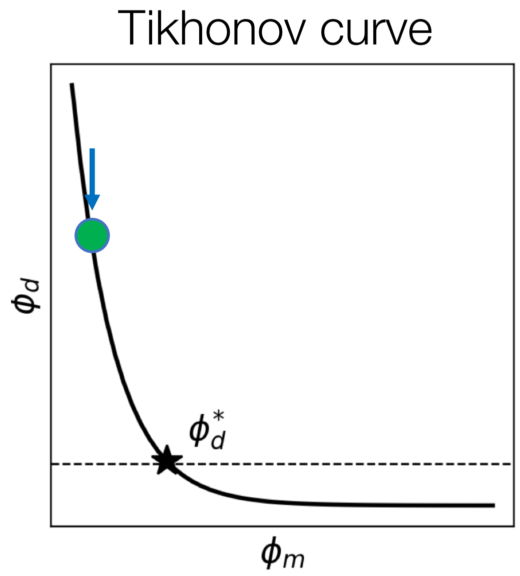
$$\text{minimize} \quad \phi_d + \beta\phi_m$$

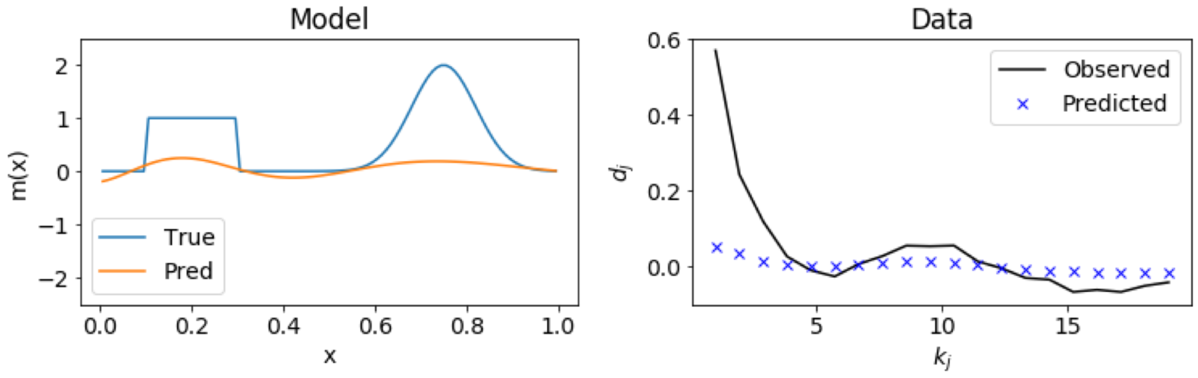
- β too large → **underfitting**
  - Structural information lost

- β too small → overfitting the data
  - Noise becomes imaged as structure

- β just right → optimal fit
  $\phi_d(m) \simeq N$

  - Best estimate of a model which adequately re-creates the observations





β is the trade-off parameter

- Solve

$$\text{minimize} \quad \phi_d + \beta\phi_m$$

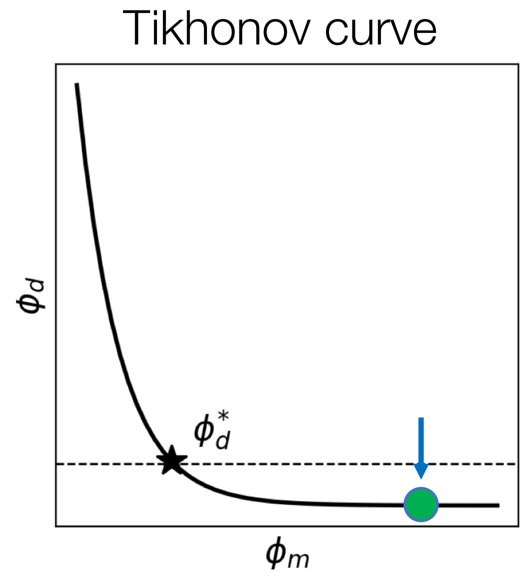
- β too large → underfitting
  - Structural information lost

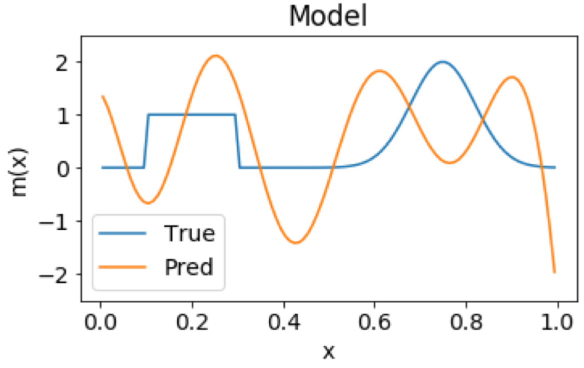
- β too small → **overfitting** the data
  - Noise becomes imaged as structure

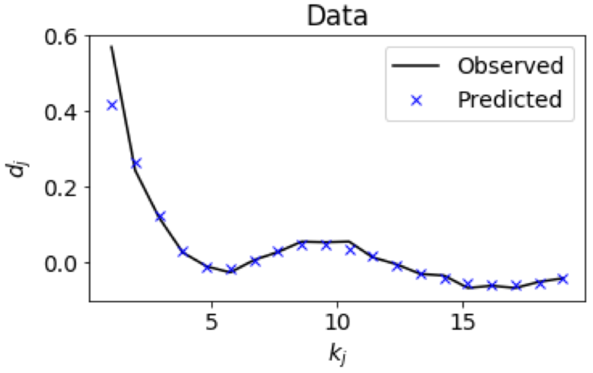
- β just right → optimal fit

$$\phi_d(m) \simeq N$$

  - Best estimate of a model which adequately re-creates the observations







β is the trade-off parameter

- Solve

$$\text{minimize} \quad \phi_d + \beta\phi_m$$

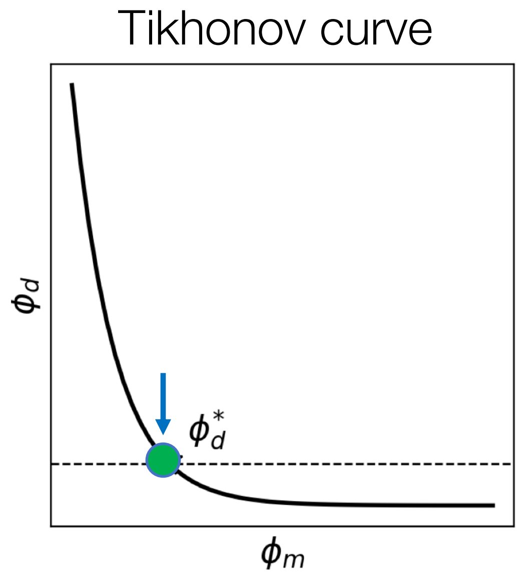
- β too large → underfitting
  - Structural information lost

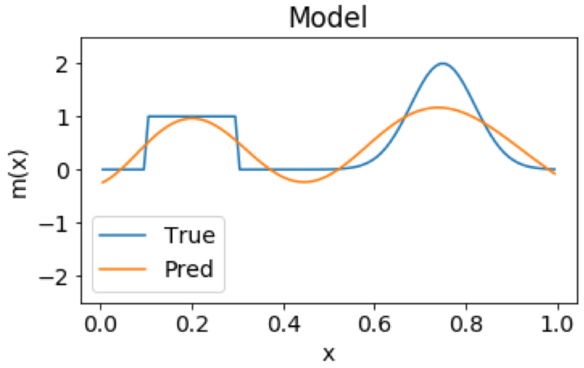
- β too small → overfitting the data
  - Noise becomes imaged as structure

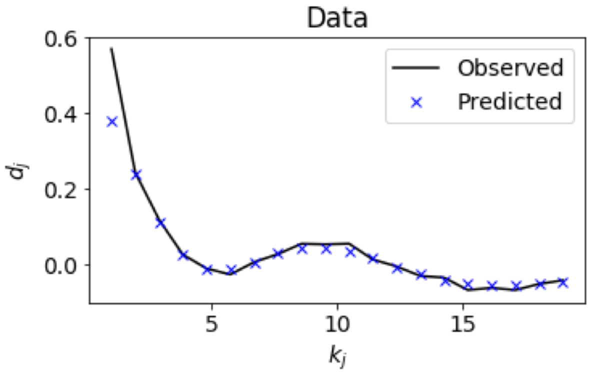
- β just right → **optimal fit**

  $\phi_d(m) \simeq N$

  - Best estimate of a model which adequately re-creates the observations







Summary

- Solved linear inverse problem when misfit and model norm are L_2
- Solution is obtained in one step
- Choose a value for the regularization parameter
- To address non-uniqueness we need a priori information

What information is available?

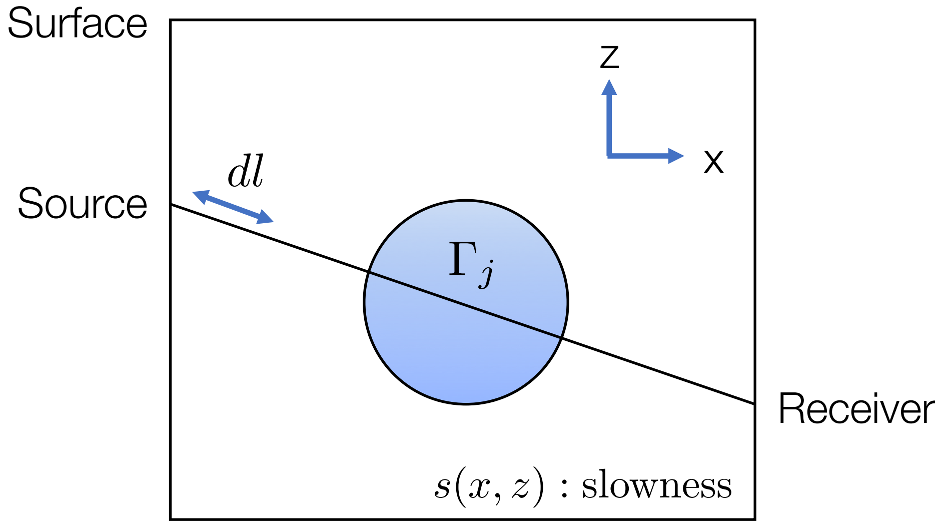
m_{ref}

- General geologic structure (smooth, discrete boundaries)
- Background model
- Geologic constraints
- Physical property constraints (bounds)
- Local well-log
- Multiple data sets (joint inversion)

- What information can be added while still keeping the system quadratic?

2D Crosswell Tomography

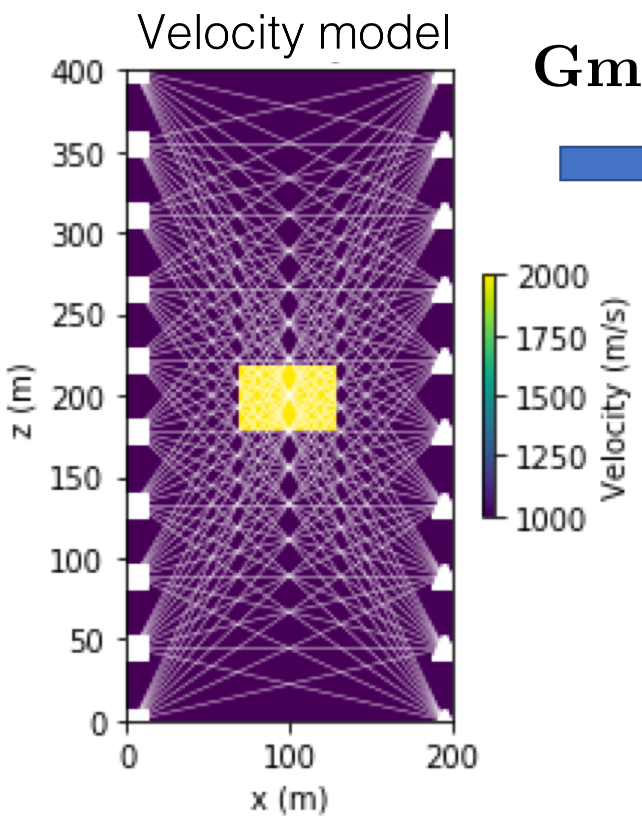
Background physics + forward modelling



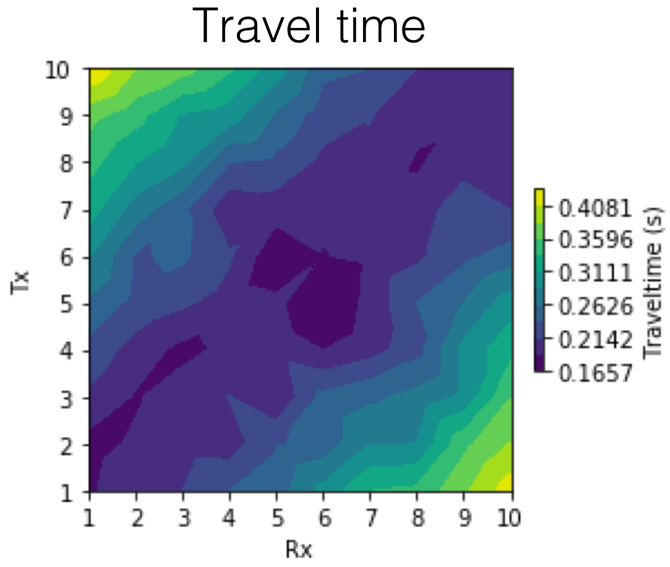
Travel time: $t_j = \int_{\Gamma_j} s(x, z)dl$

s: slowness $(=v^{-1})$

Linear equation: $\mathbf{Gm} = \mathbf{d}$



$\mathbf{Gm} = \mathbf{d}$



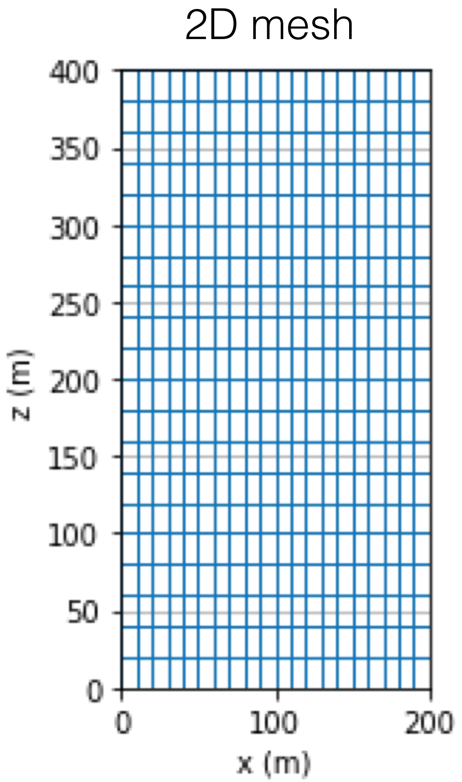
2D model objective function (regularization)

- Discretization

$$\phi_s = \int_V (m - m_{\text{ref}})^2 dv$$

$$\phi_x = \int_V \left(\frac{dm}{dx}\right)^2 dv$$

$$\phi_z = \int_V \left(\frac{dm}{dz}\right)^2 dv$$



- So the discrete form is:

$$\phi_m = \alpha_s \|\mathbf{W}_s(\mathbf{m} - \mathbf{m}_{\text{ref}})\|^2 + \alpha_x \|\mathbf{W}_x \mathbf{m}\|^2 + \alpha_z \|\mathbf{W}_z \mathbf{m}\|^2$$

Choosing α -parameters

Model objective function:

$$\phi_m = \alpha_s \int_v (m - m_{\text{ref}})^2 dv + \alpha_x \int_v \left(\frac{dm}{dx}\right)^2 dx + \alpha_z \int_v \left(\frac{dm}{dz}\right)^2 dz$$

Dimensionally, $\alpha_s(\triangle m)^2 A \quad + \quad \alpha_x(\frac{\triangle m}{\mathrm{L}_x})^2 A \quad + \quad \alpha_z(\frac{\triangle m}{\mathrm{L}_z})^2 A$

L_x: scale length for x direction $[m]$
L_z: scale length for z direction $[m]$
A: area $[m^2]$

If smallest norm $\simeq$ flattest norm

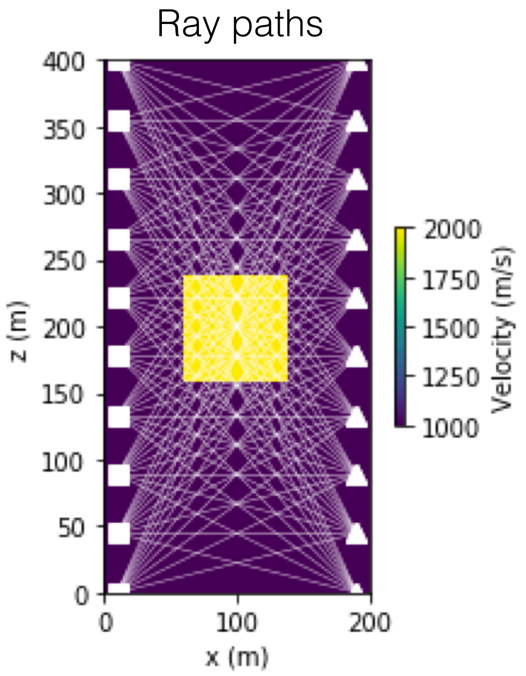
$$\alpha_s = \frac{\alpha_x}{(\mathrm{L}_x)^2}$$

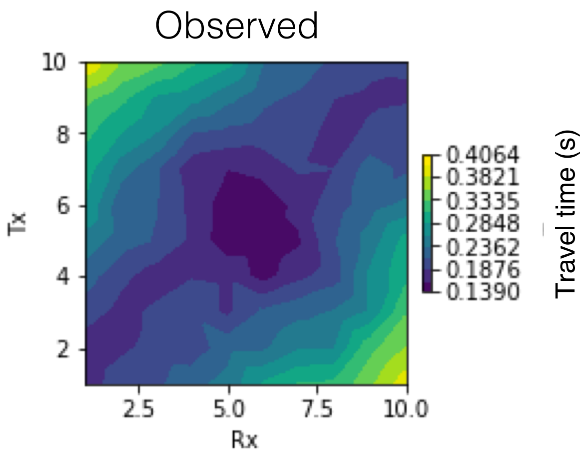
Generic inversion

$$\phi_m = \alpha_s \|\mathbf{W}_s(\mathbf{m} - \mathbf{m}_{\mathrm{ref}})\|^2 + \alpha_x \|\mathbf{W}_x \mathbf{m}\|^2 + \alpha_z \|\mathbf{W}_z \mathbf{m}\|^2$$

Choose

$$\mathbf{m}_{\mathrm{ref}} = 1/1000$$
$$\alpha_s = 1/1000$$
$$\alpha_x = \alpha_z = 1$$



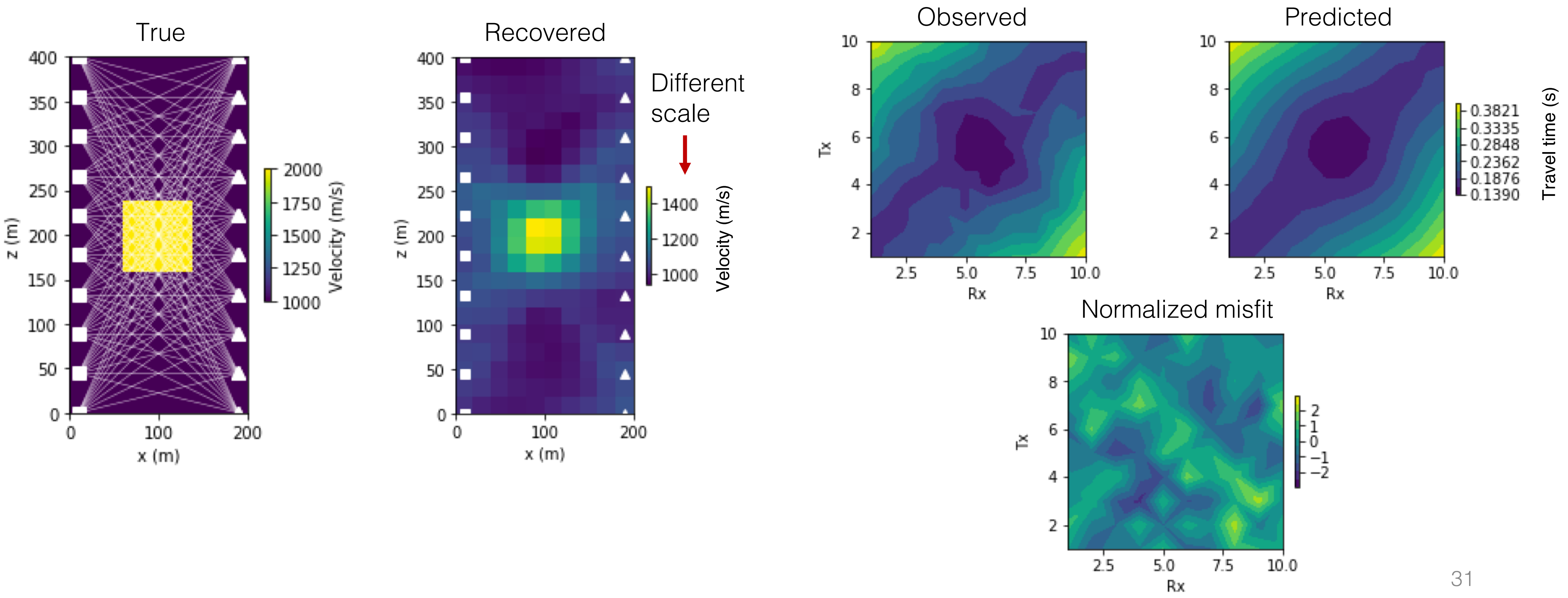


Generic inversion

$$\phi_m = \alpha_s \|\mathbf{W}_s(\mathbf{m} - \mathbf{m}_{\text{ref}})\|^2 + \alpha_x \|\mathbf{W}_x \mathbf{m}\|^2 + \alpha_z \|\mathbf{W}_z \mathbf{m}\|^2$$

Choose

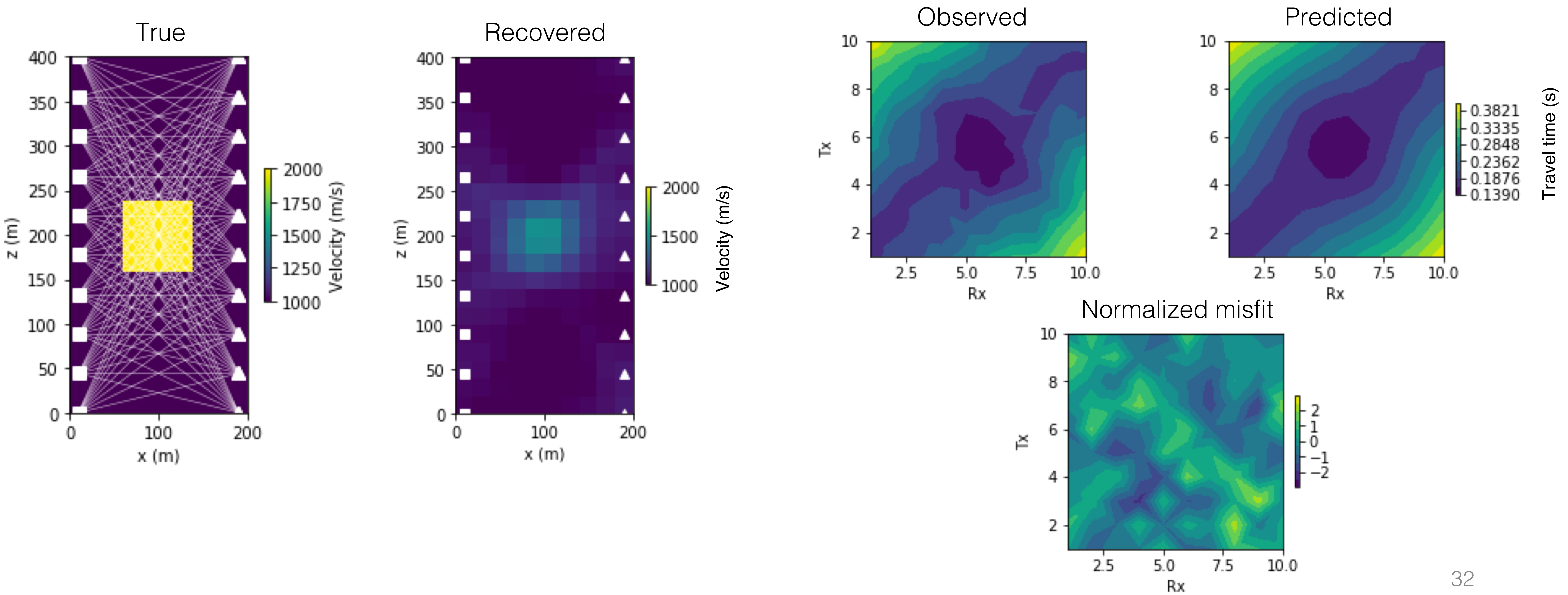
$$\mathbf{m}_{\text{ref}} = 1/1000$$
$$\alpha_s = 1/1000$$
$$\alpha_x = \alpha_z = 1$$



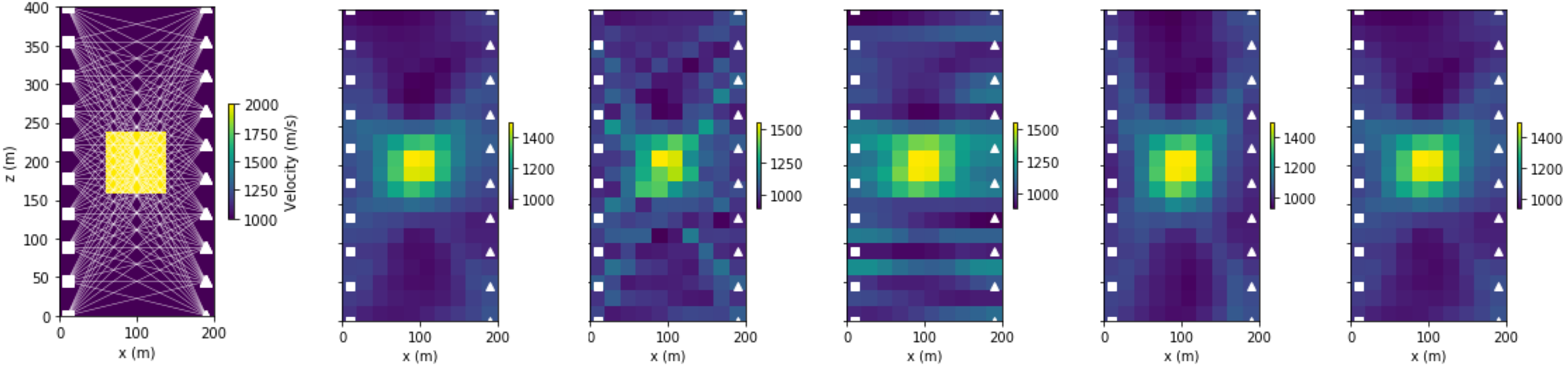
Generic inversion

$$\phi_m = \alpha_s \|\mathbf{W}_s(\mathbf{m} - \mathbf{m}_{\text{ref}})\|^2 + \alpha_x \|\mathbf{W}_x \mathbf{m}\|^2 + \alpha_z \|\mathbf{W}_z \mathbf{m}\|^2$$

Choose
$$\mathbf{m}_{\text{ref}} = 1/1000$$
$$\alpha_s = 1/1000$$
$$\alpha_x = \alpha_z = 1$$



Exploring α-parameters



True

$\alpha_s = 2.5\times10^{-3}$
$\alpha_x = \alpha_z = 1$

$\alpha_s = 1$
$\alpha_x = \alpha_z = 0$

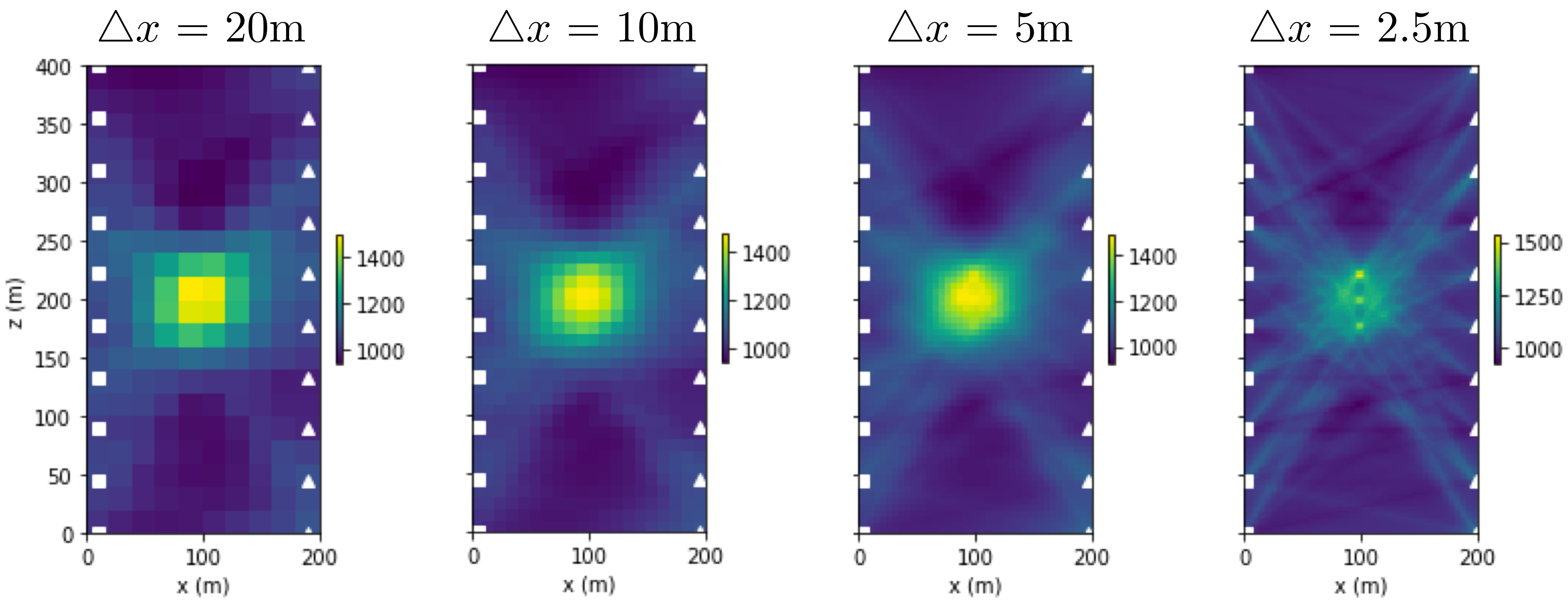
$\alpha_s = 10^{-8}$
$\alpha_x = 1$
$\alpha_z = 0$

$\alpha_s = 10^{-8}$
$\alpha_x = 0$
$\alpha_z = 1$

$\alpha_s = 10^{-8}$
$\alpha_x = \alpha_z = 1$

Cell size: effects

- Inverse problem can be regularized by cell size $\triangle x = \triangle z$



Role of the reference model

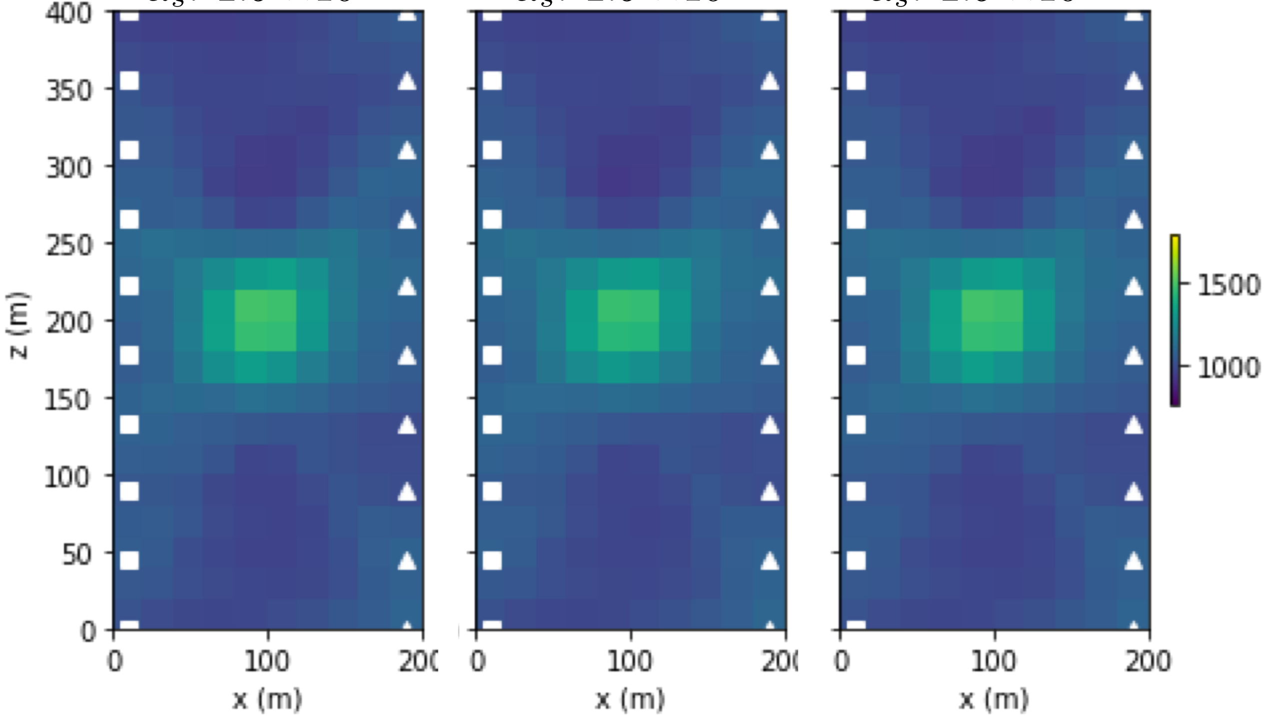
$$\phi_m = \alpha_s \|\mathbf{W}_s(\mathbf{m} - \mathbf{m}_{\text{ref}})\|^2 + \alpha_x \|\mathbf{W}_x \mathbf{m}\|^2 + \alpha_z \|\mathbf{W}_z \mathbf{m}\|^2$$

α_s: must be large enough to be influential

$\mathbf{m}_{\text{ref}} = 1/1000$
α_s: 2.5×10^{-3}

$\mathbf{m}_{\text{ref}} = 1/750$
α_s: 2.5×10^{-3}

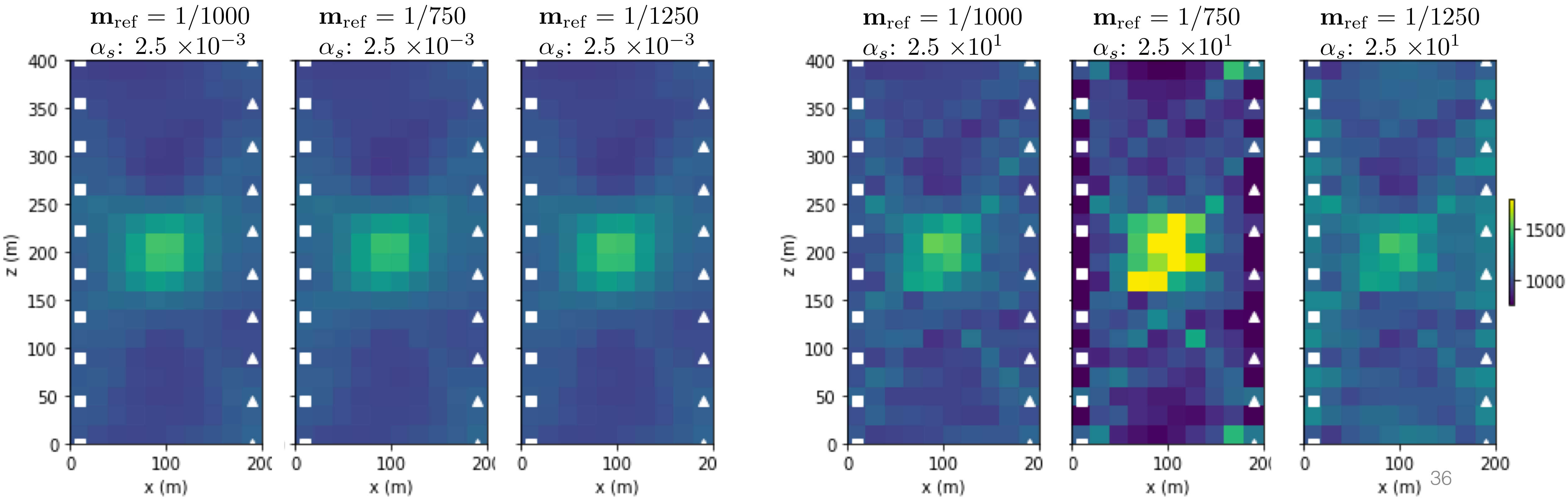
$\mathbf{m}_{\text{ref}} = 1/1250$
α_s: 2.5×10^{-3}



Role of the reference model

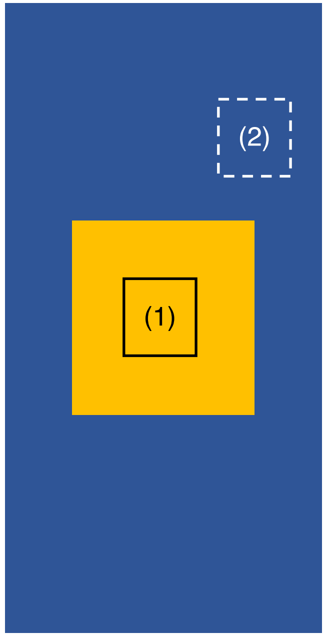
$$\phi_m = \alpha_s \|\mathbf{W}_s(\mathbf{m} - \mathbf{m}_{\text{ref}})\|^2 + \alpha_x \|\mathbf{W}_x \mathbf{m}\|^2 + \alpha_z \|\mathbf{W}_z \mathbf{m}\|^2$$

α_s: must be large enough to be influential



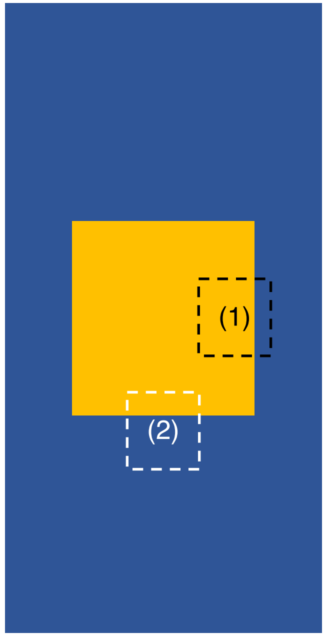
Enhanced functionality using ϕ_m

$$\phi_m = \alpha_s \int_v w_s(\mathbf{m} - \mathbf{m}_{\text{ref}})^2 dv + \alpha_x \int_v w_x \left(\frac{dm}{dx}\right)^2 dx + \alpha_z \int_v w_z \left(\frac{dm}{dz}\right)^2 dz$$



$w_s(x, z)$: local confidence in m_{ref}

(1) w_s: small
(2) w_s: large



(1) w_x: small, w_z: large $\rightarrow$ allows jump in x-direction

(2) w_x: large, w_z: small $\rightarrow$ allows jump in z-direction

Summary

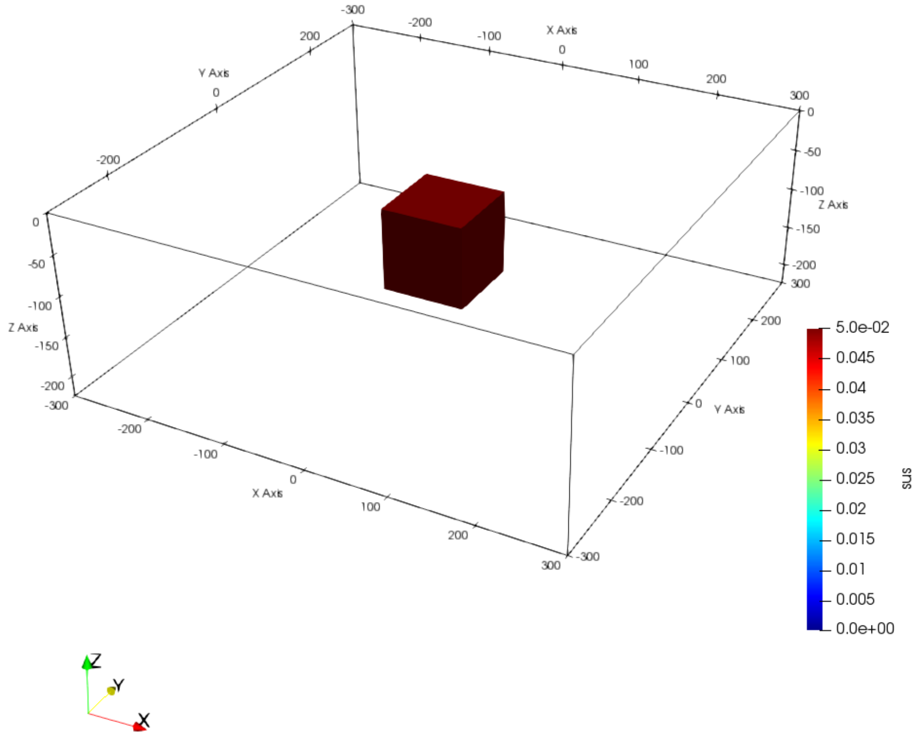
Character and details of recovered model is affected by

- Hyper-parameters (alpha)
- Reference model
- Local weightings

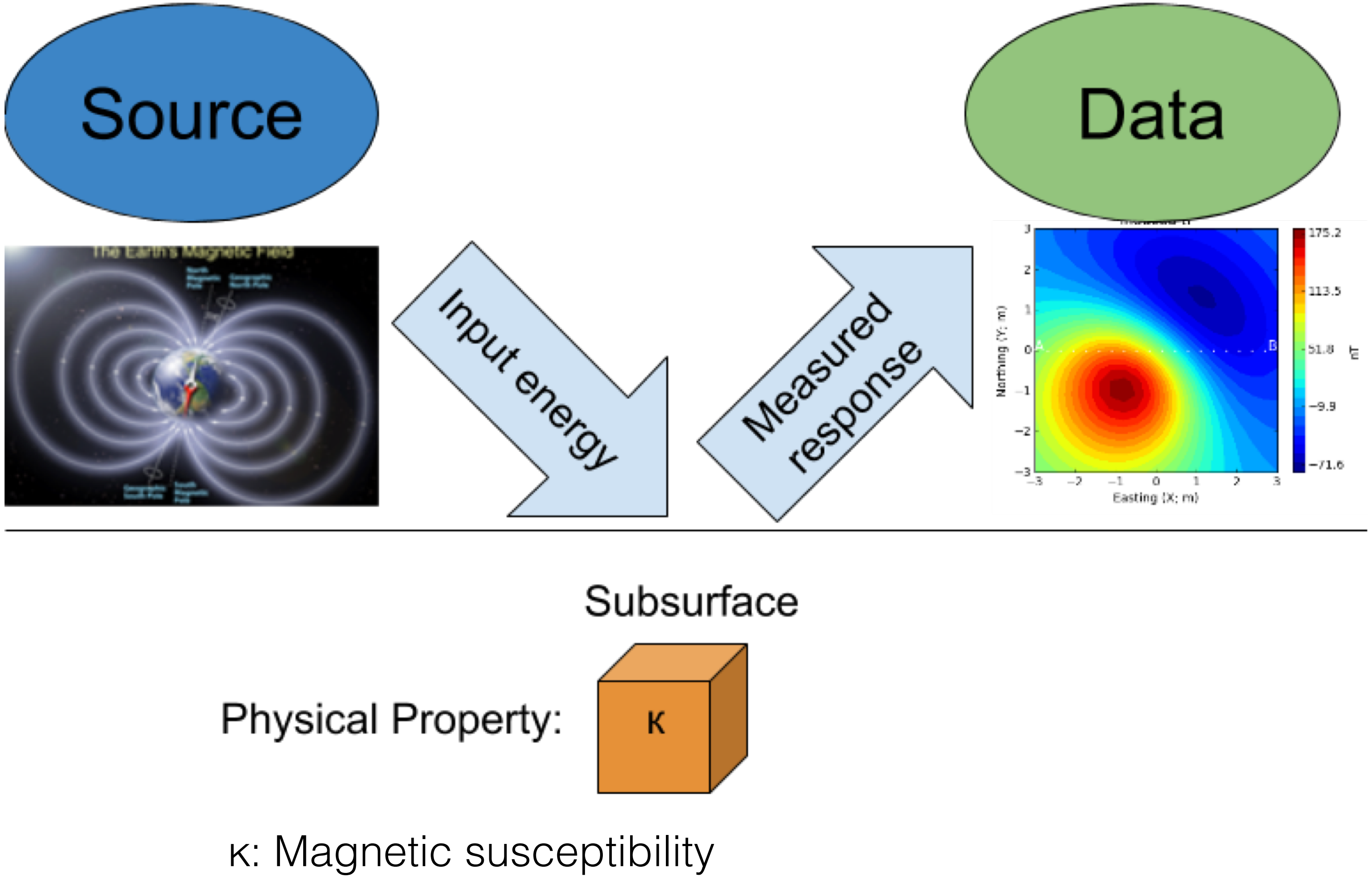
Changing parameters allows some exploration of model space

Not sufficient for many problems: next up is magnetics

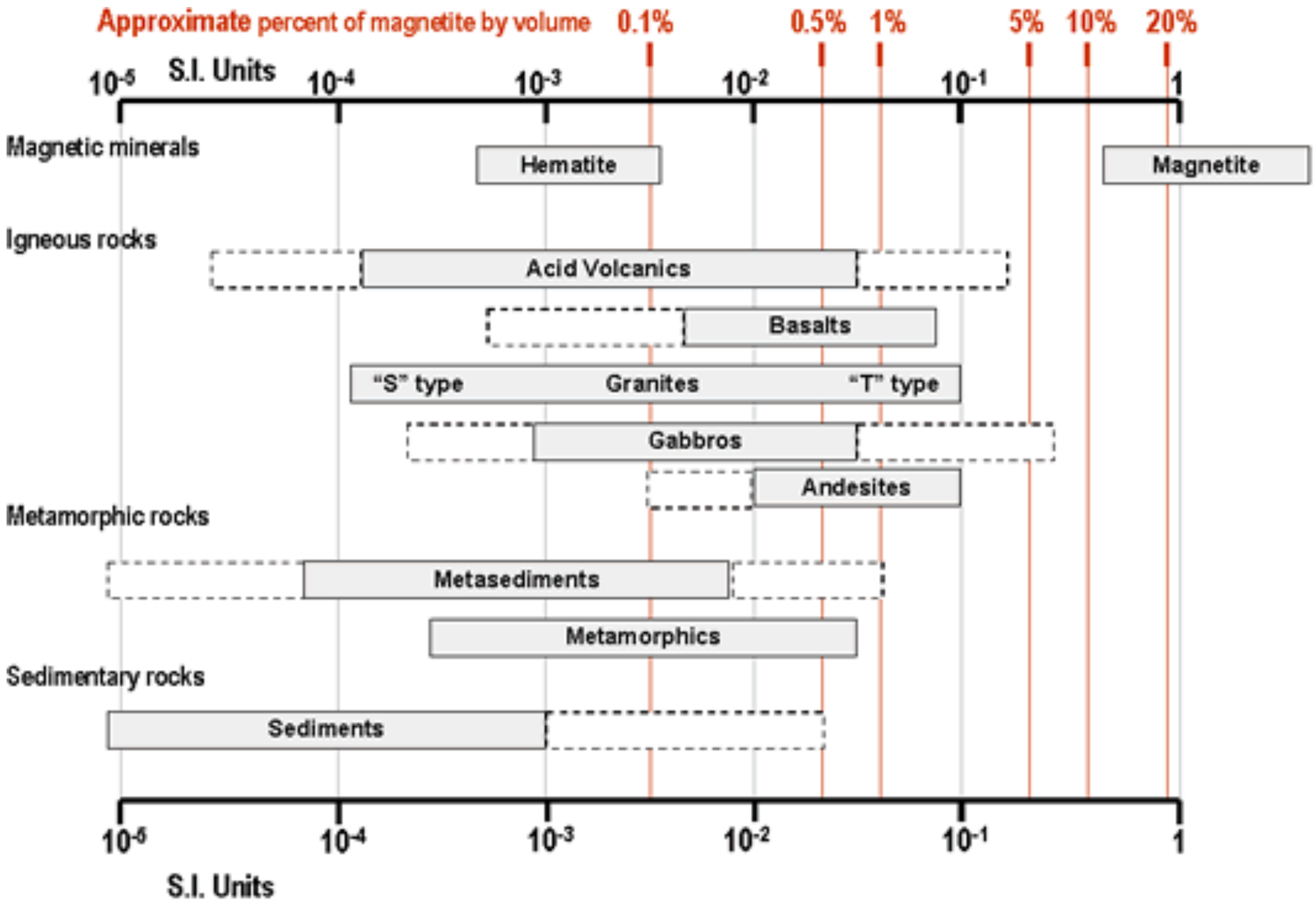
3D magnetics



Magnetic survey



Magnetic susceptibility



Adopted from Clark and Emerson, Exploration Geophysics (1991)

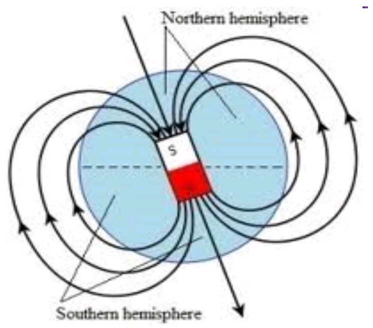
Magnetic surveying

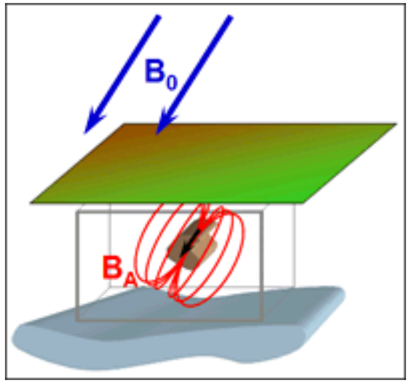
- Earth's magnetic field $\vec{B}_0$ is the source:
- Materials become magnetized

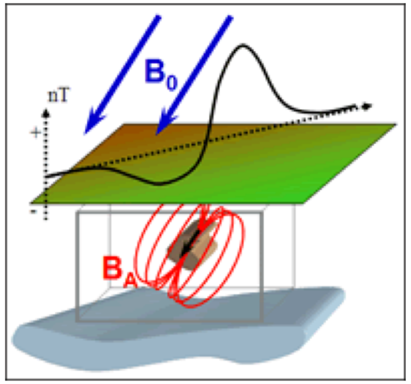
$$\vec{M} = \kappa \vec{H}_0 \text{ (magnetization)}$$
$$\vec{H}_0 = \vec{B}_0/\mu_0$$

- Create anomalous magnetic field
- Measure total magnetic field: $|\vec{B}| = |\vec{B}_0 + \vec{B}_A|$
- Total field anomaly: $\triangle \vec{B} = |\vec{B}_0 + \vec{B}_A| - |\vec{B}_0|$

$$\triangle \vec{B} \simeq \vec{B}_A \cdot \hat{B}_0 \qquad \text{where} \quad \hat{B}_0 = \frac{\vec{B}_0}{|\vec{B}_0|}$$



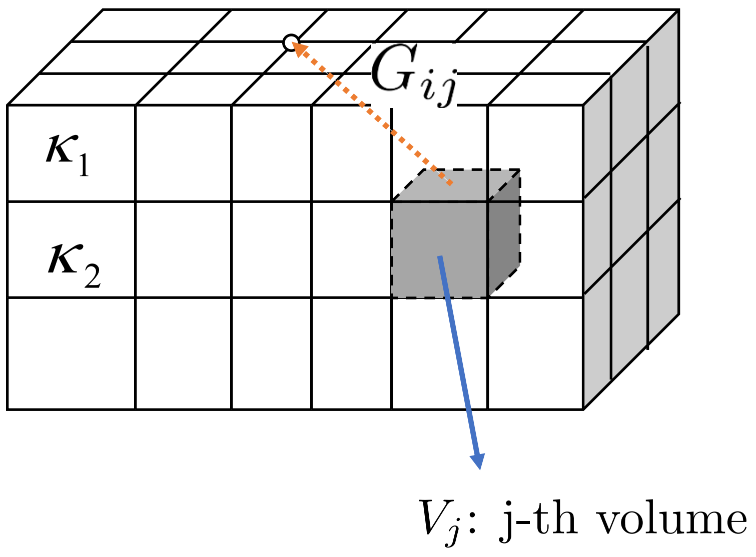




Forward modelling

- Discretize earth

$$\kappa_j \ (j = 1, \ldots, M)$$ susceptibility



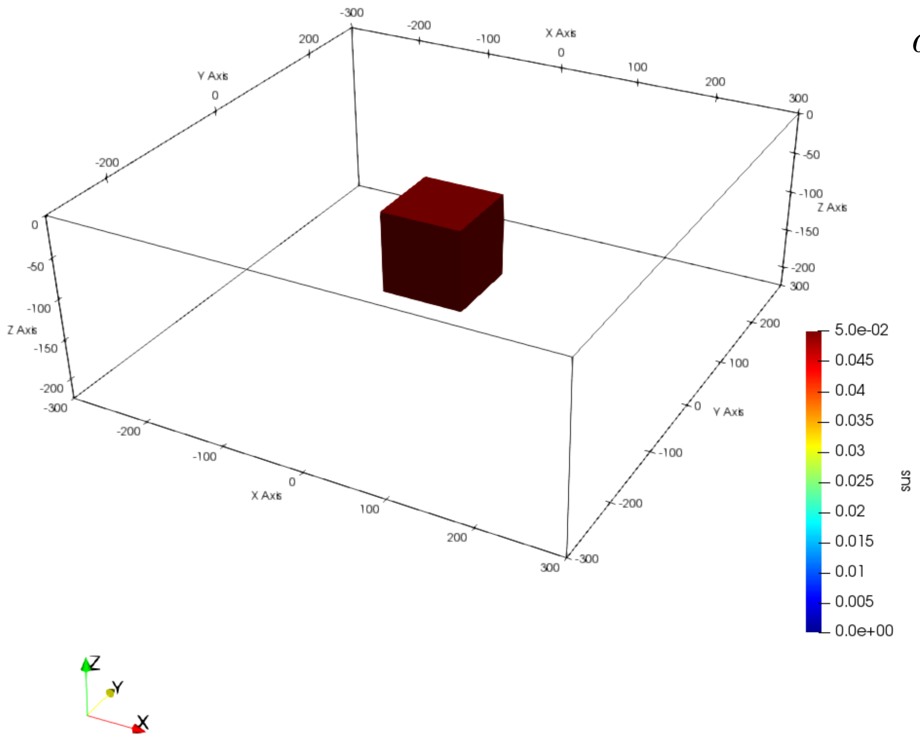
- Magnetic anomaly data are

$$d_i = \sum_{j=1}^{M} G_{ij} \kappa_j$$

$$\begin{cases} G_{ij} = \hat{B}_0 \cdot \left\{ \frac{\mu_0}{4\pi} \int_v \kappa \nabla\nabla \left(\frac{1}{r_i - r_j} \right) \ dV_j \right\} \cdot \hat{B}_0 \\ \hat{B}_0 = \frac{\vec{B}_0}{|\vec{B}_0|} \end{cases}$$

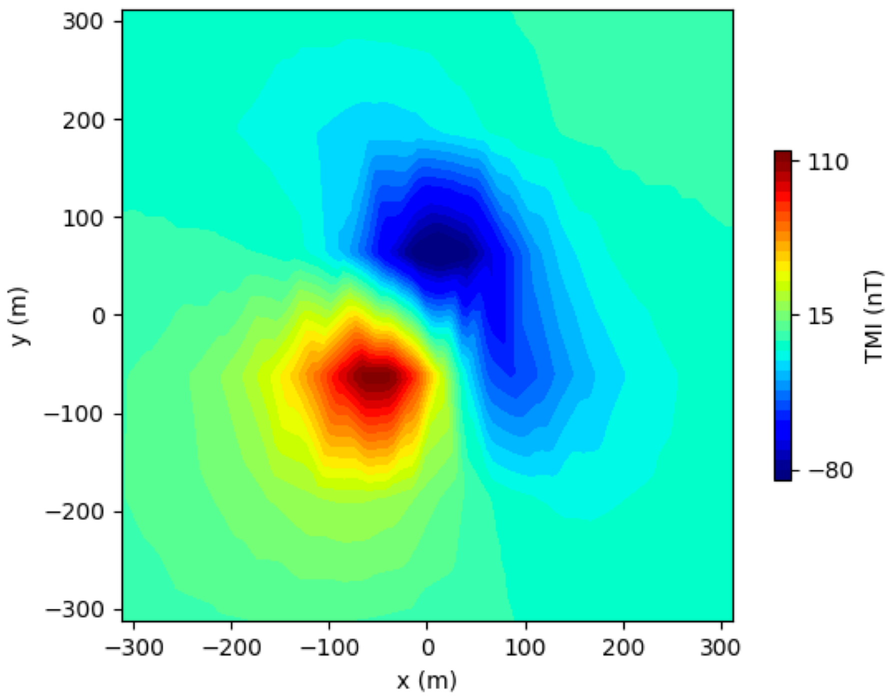
Forward modelling

Susceptibility model

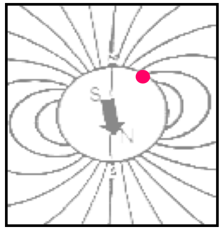


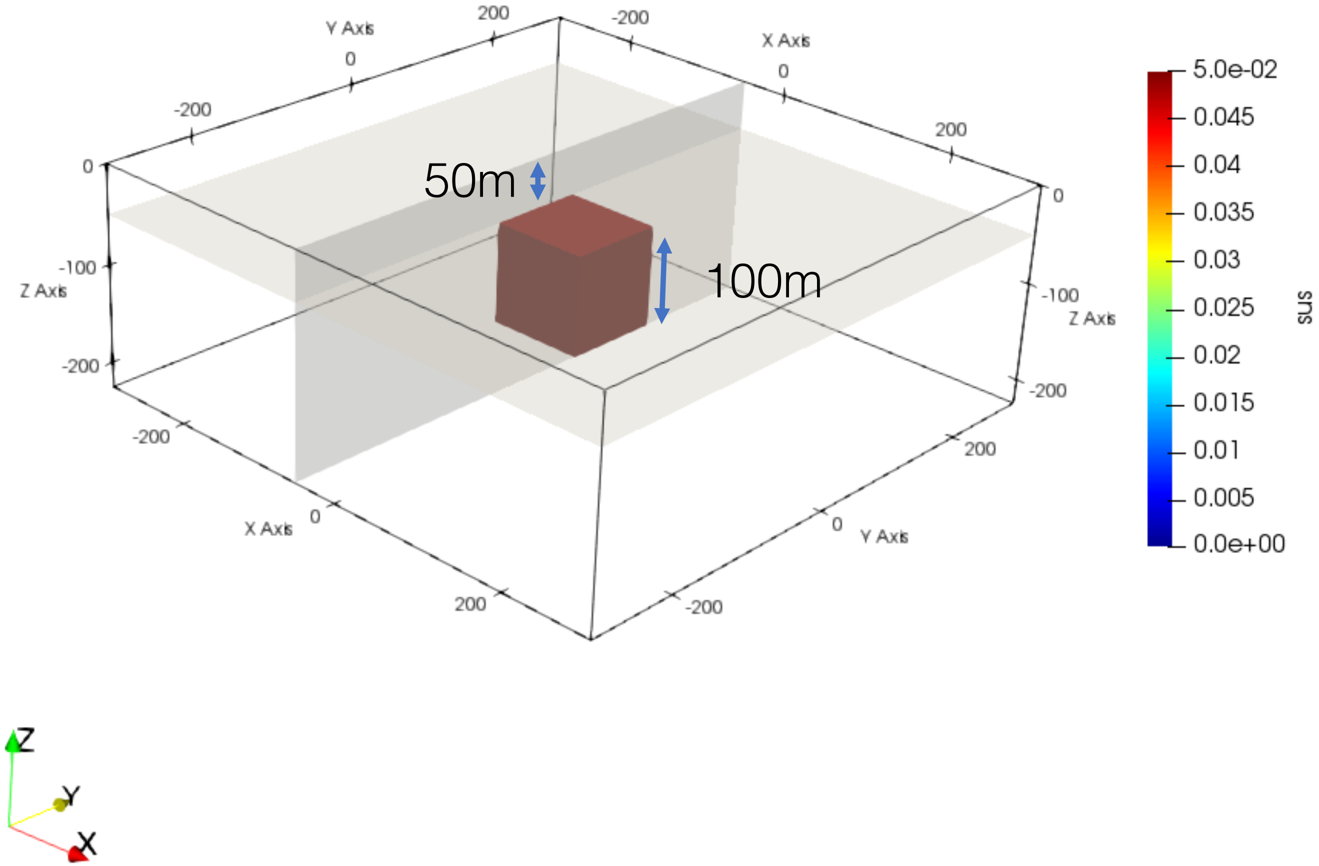
$$d_i = \sum_{j=1}^{M} G_{ij}\kappa_j$$

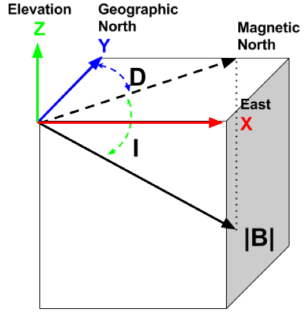
Magnetic data



Synthetic susceptibility model







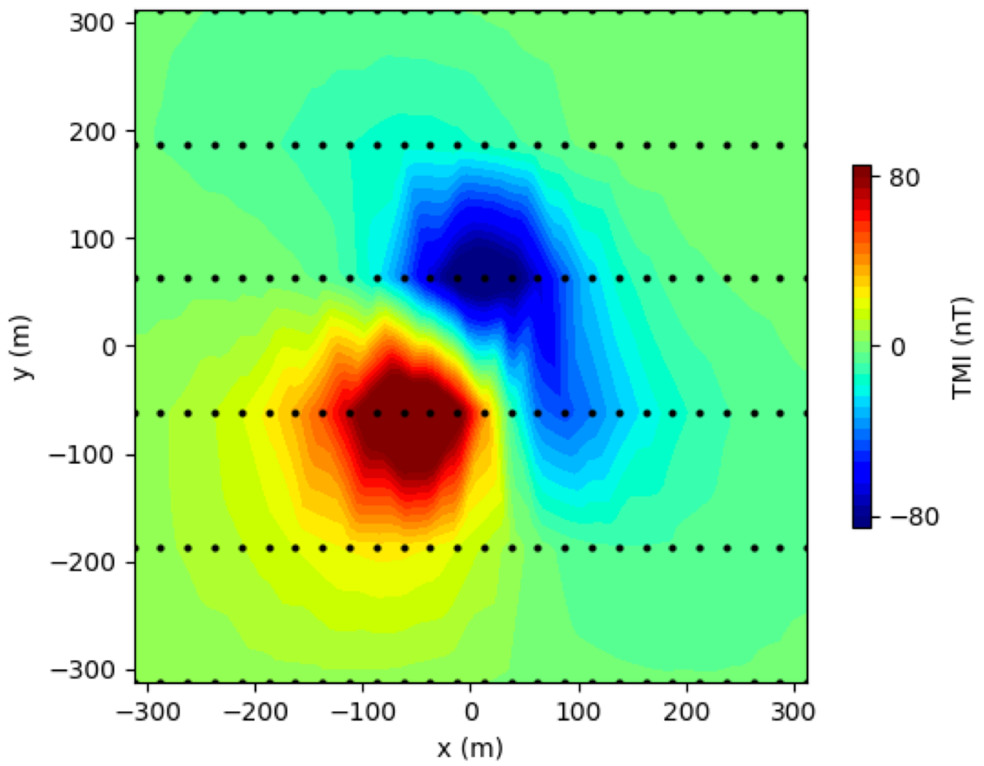
- Earth field
  - Inclination: 30°
  - Declination: 45 °
  - $|B_0|$ = 50,000 nT

- Susceptible block
  - 100m x 100m x 100m block
  - Block susceptibility = 0.5
  - Block top = 50m

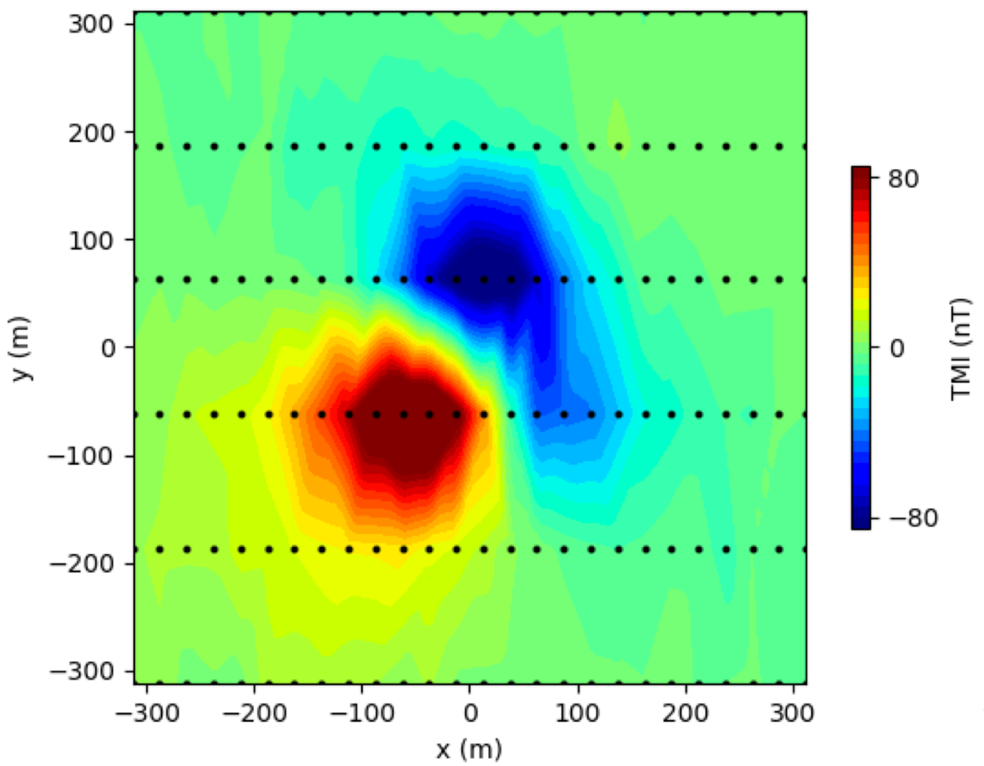
Synthetic survey

Survey parameters:
- 100 m line spacing.
- 25 station spacing.
- N=156 (elevation= 2m)

Contoured “perfect” data



Contoured data with noise = 2 nT



Solving inverse problem

Model objective function

$$\phi_m = \alpha_s \int_v w_s \Big(\kappa - \kappa_{\text{ref}}\Big)^2 dv + \alpha_x \int_v \left(\frac{d\kappa}{dx}\right)^2 dx + \alpha_y \int_v \left(\frac{d\kappa}{dy}\right)^2 dy + \alpha_z \int_v \left(\frac{d\kappa}{dz}\right)^2 dz$$

Data misfit $\phi_d = \sum_{j=1}^{N} \left(\frac{G_{ij}\kappa_j - d_j^{obs}}{\epsilon_j}\right)$

Choose $\kappa_{ref} = 0, \alpha_s = 0.0001, \alpha_x = \alpha_y = \alpha_z = 1$

$L_x = \sqrt{\frac{\alpha_x}{\alpha_s}} = 100$

The Inverse problem is:

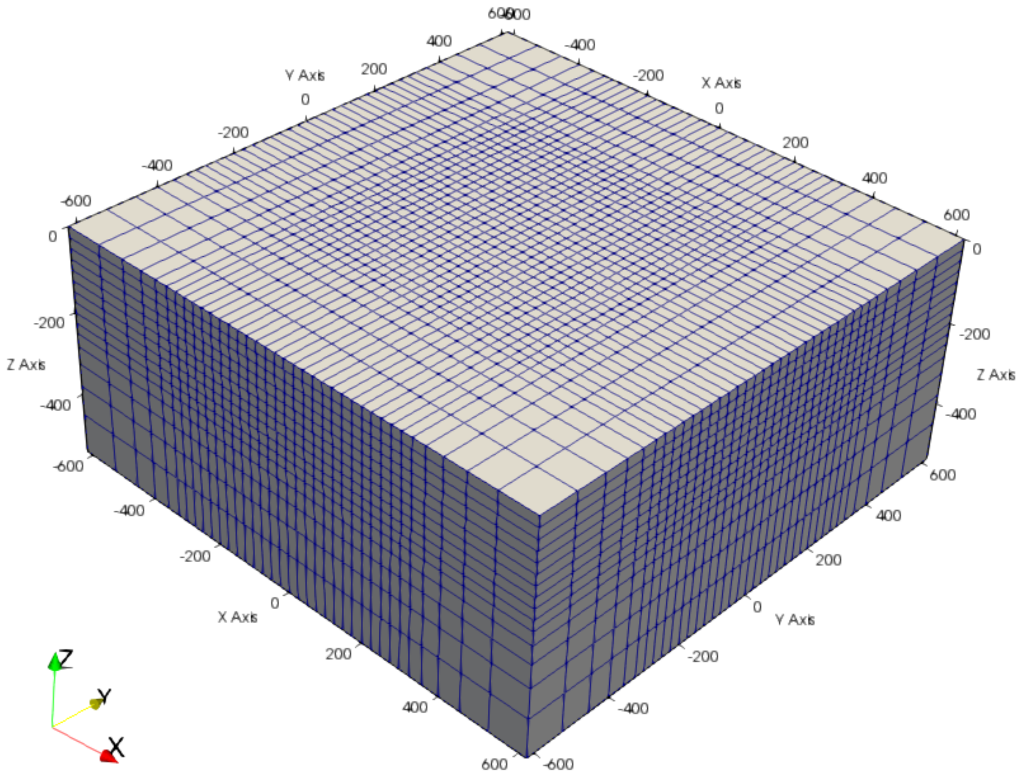
$$\text{minimize} \quad \phi = \phi_d + \beta\phi_m$$

$$\text{find } \beta \text{ such that } \phi_d = \phi_d^* \text{ where } \phi_d = N$$

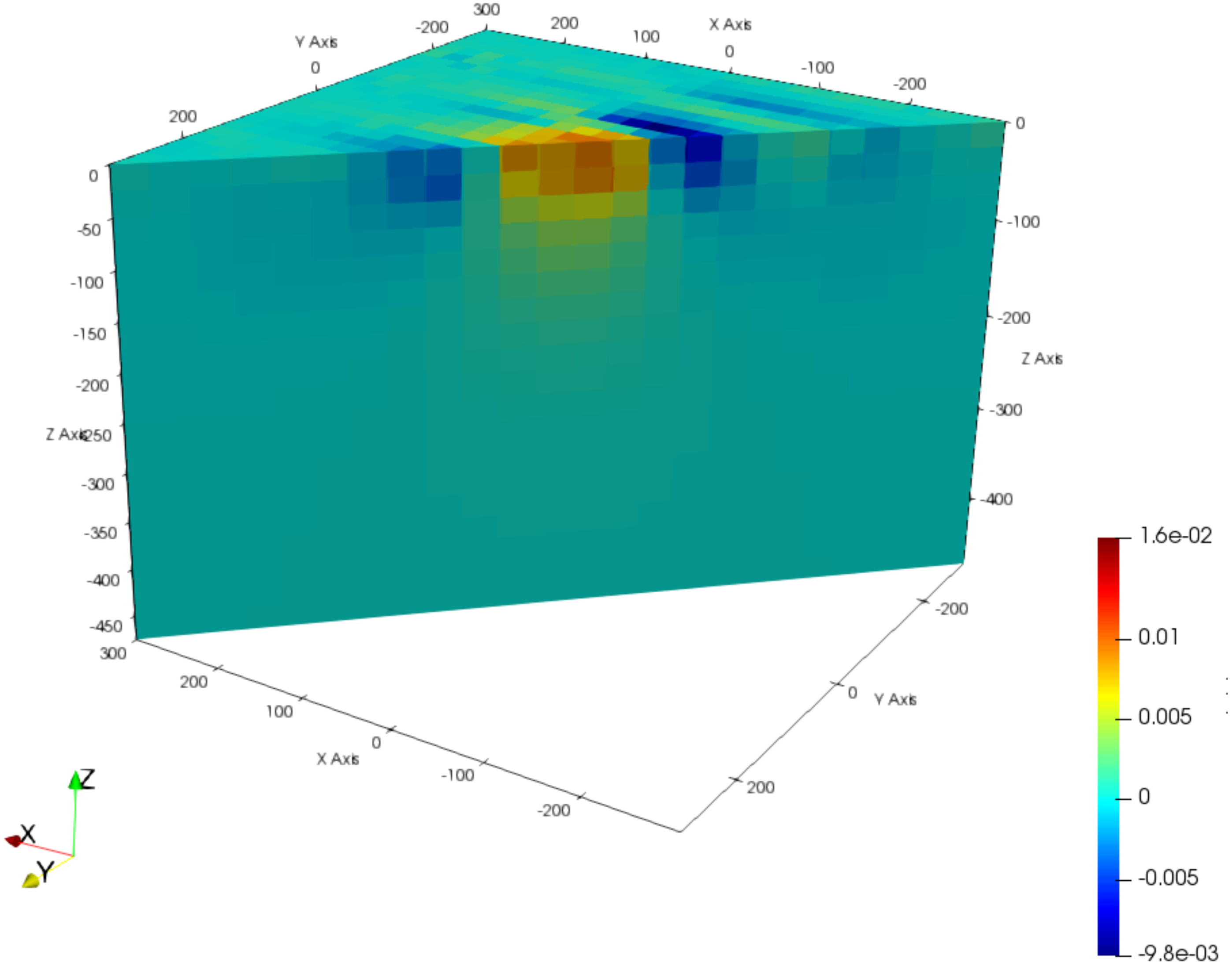
Discretization

- Earth model for inversion:
  - dx=dy=dz=25m
  - N/S and E/W padding = 300m
  - Number of cells (M) = 19440

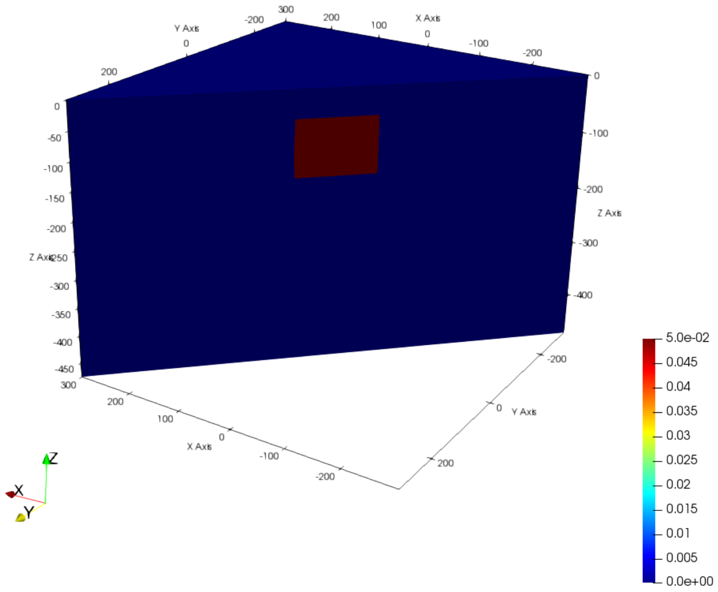
- Therefore:
  - No. data is N = 176
  - No. unknowns is M = 11,492



Inversion results



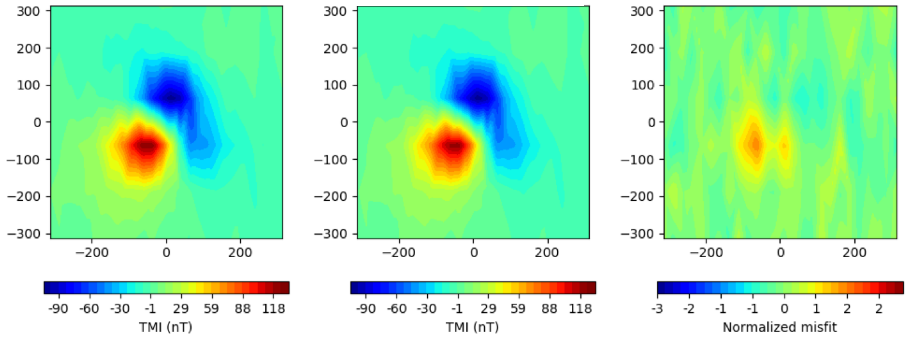
True model



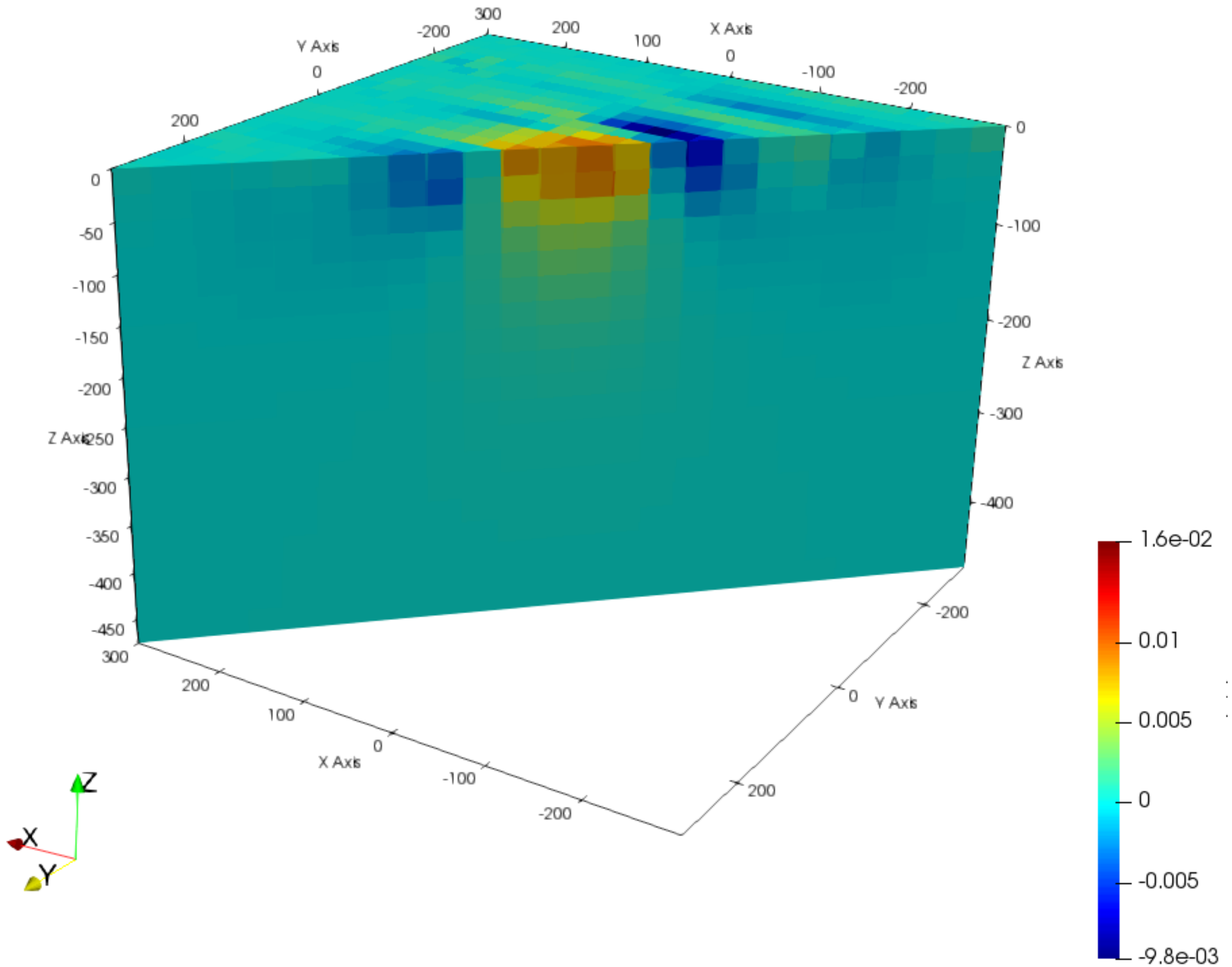
Obs.

Pred.

Misfit



Inversion results



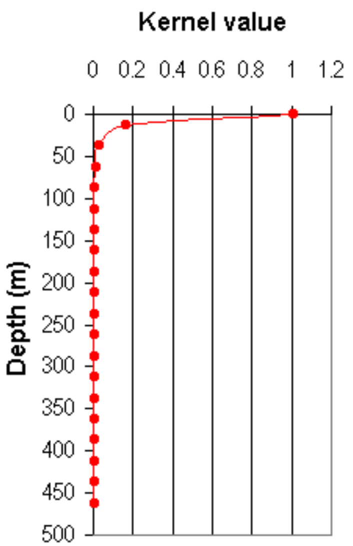
- Two primary difficulties:
  - Concentration of susceptibility is near the surface
  - Regions of negative susceptibility

What went wrong?

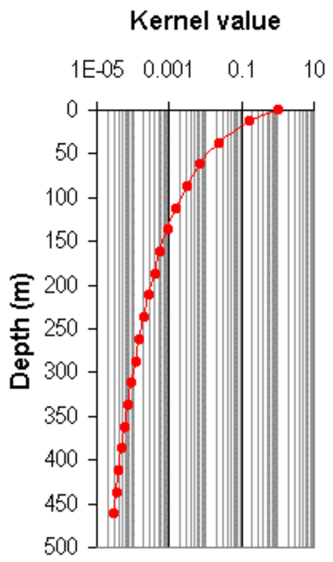
Fundamental non-uniqueness of all potential fields:

- As a consequence of Green's third identity ...
  - an observed magnetic field can be caused by a thin layer of susceptible material at any arbitrary depth

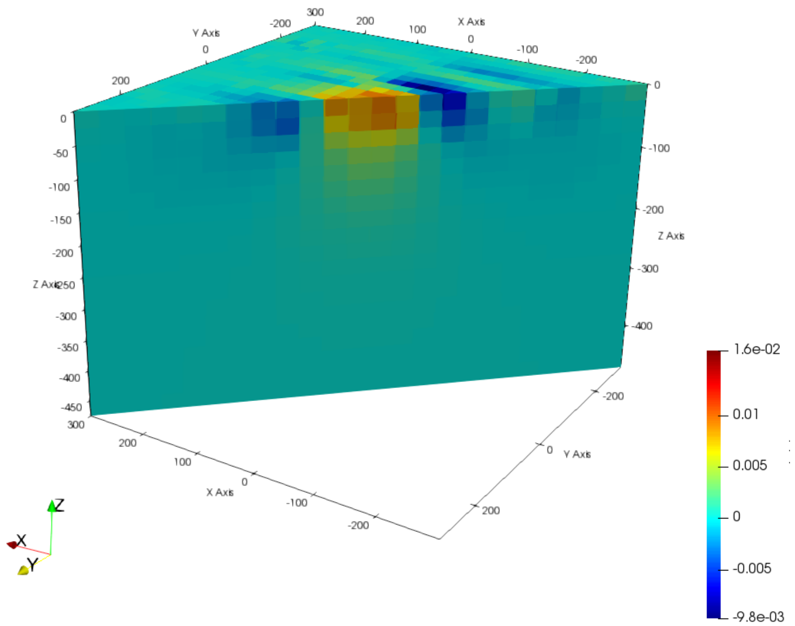
- The rapid decay of our kernels causes a concentration of κ near the surface to be a preferred solution.



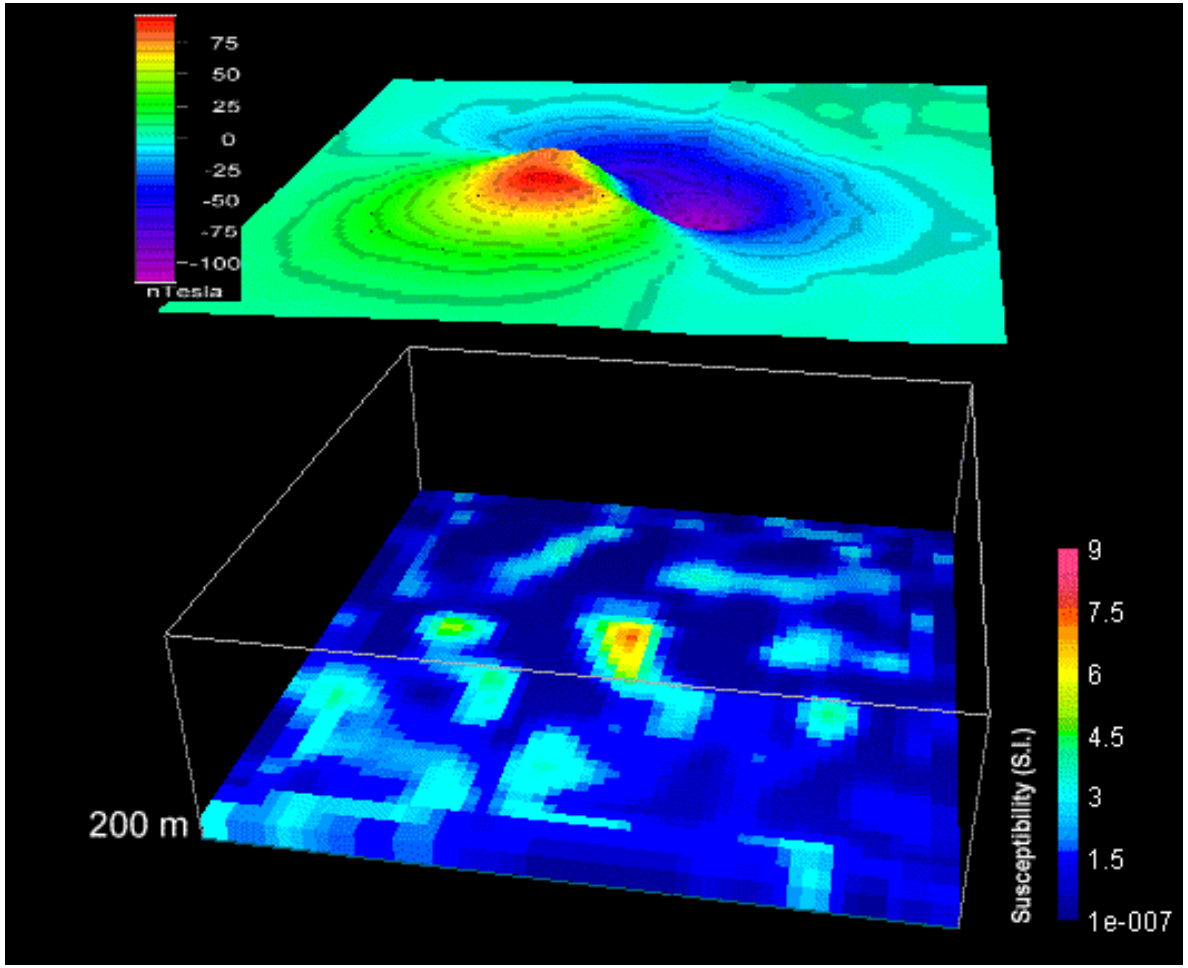
linear



log



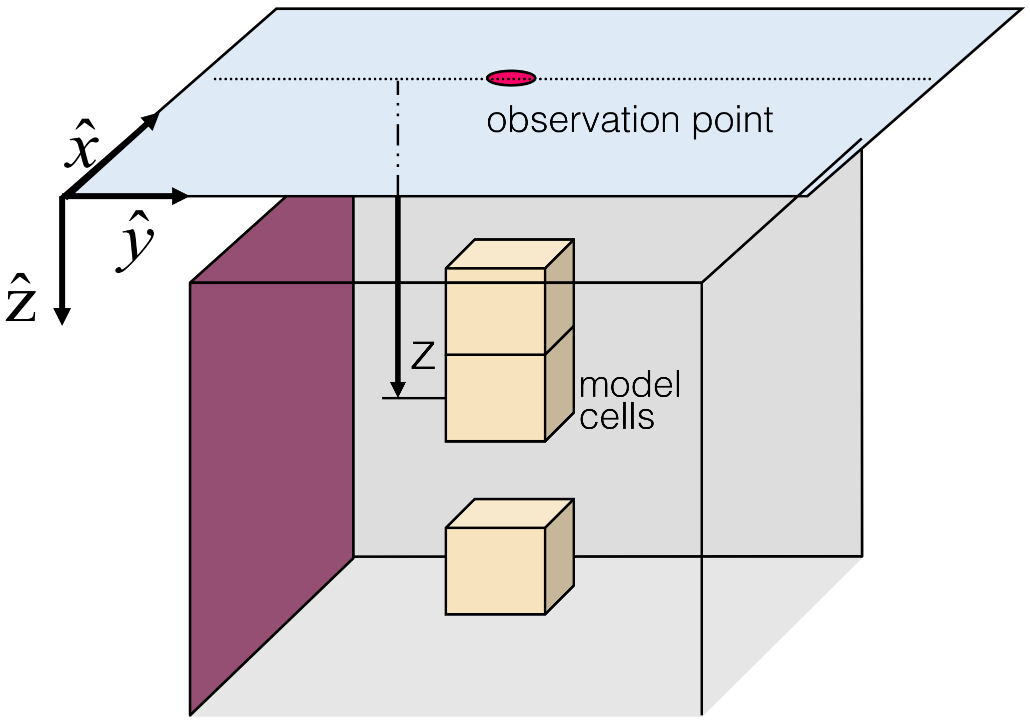
Non-uniqueness

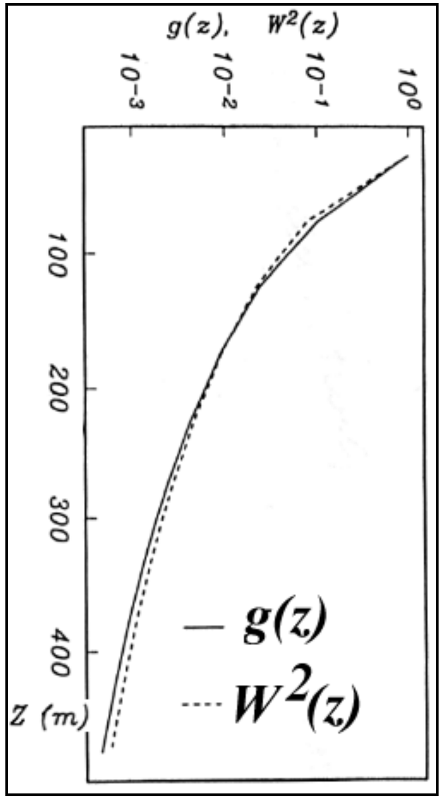


- Equivalent layer models which reproduce the data

Sensitivity weighting

- Decays with depth: $w(z) = \dfrac{1}{(z+z_o)^{3/2}}$
- z_0: by least-squares fit between g(z) and $w^2(z)$





Inversion with sensitivity weighting

Model objective function

$$\phi_m = \alpha_s \int_v w_s \Big(\kappa - \kappa_{\text{ref}}\Big)^2 dv + \alpha_x \int_v w_x \Big(\frac{d\kappa}{dx}\Big)^2 dx + \alpha_y \int_v w_y \Big(\frac{d\kappa}{dy}\Big)^2 dy + \alpha_z \int_v w_z \Big(\frac{d\kappa}{dz}\Big)^2 dz$$

Data misfit $\phi_d = \sum_{j=1}^{N} \Big(\frac{G_{ij}\kappa_j - d_j^{obs}}{\epsilon_j}\Big)$

$\{w_s, w_x, w_y, w_z\}$: additional weightings

Choose $\{w_s, w_x, w_y, w_z\} \propto \frac{1}{(z+z_0)^3}$

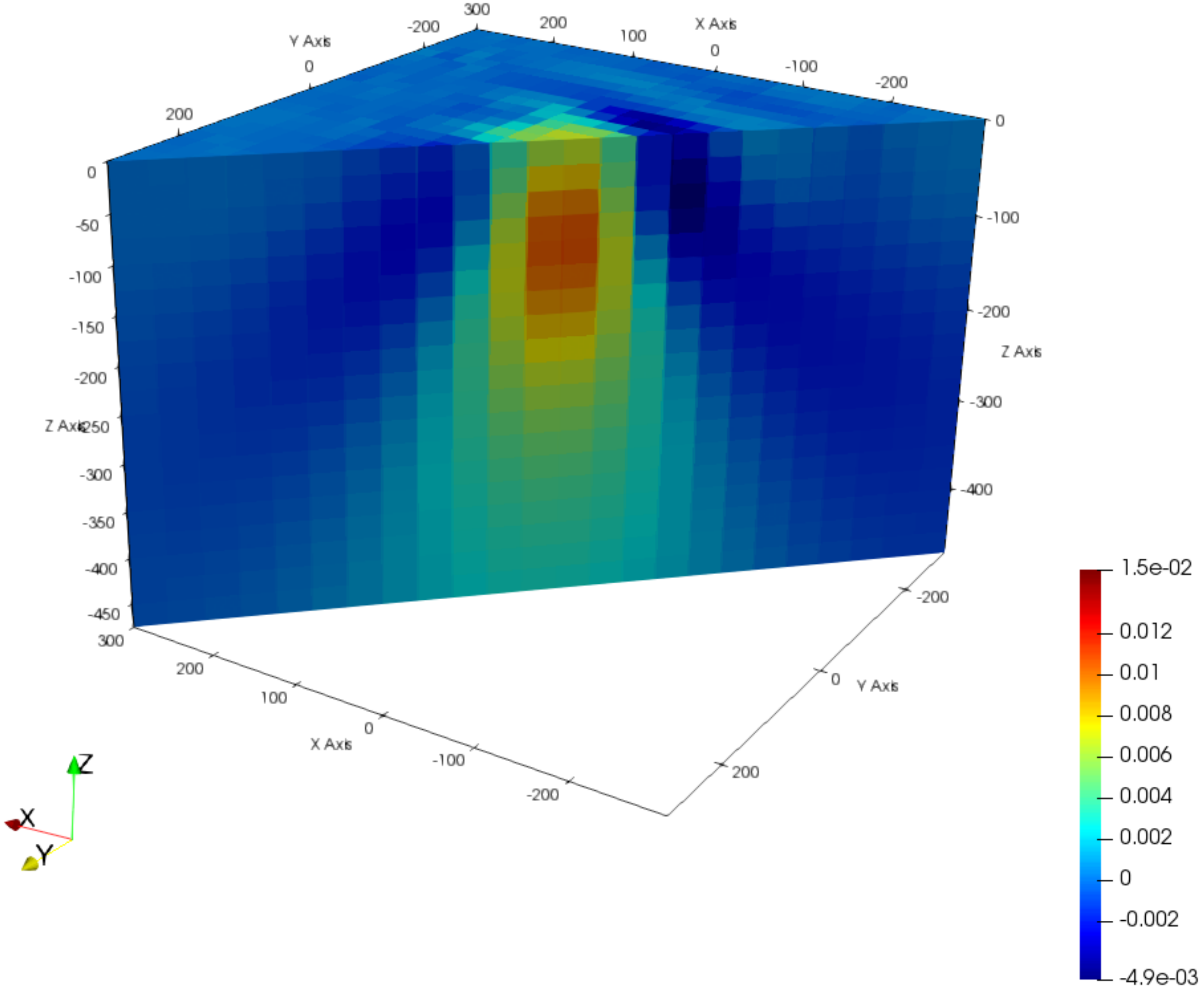
Allows cells at depth to contribute

The Inverse problem is:

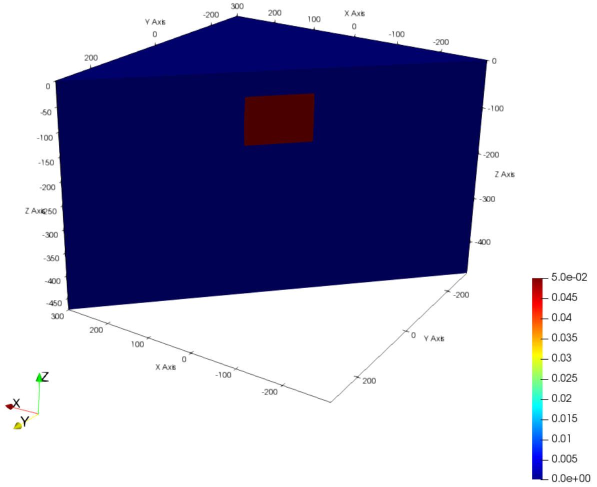
minimize $\quad \phi = \phi_d + \beta\phi_m$

find β such that $\phi_d = \phi_d^*$ where $\phi_d = N$

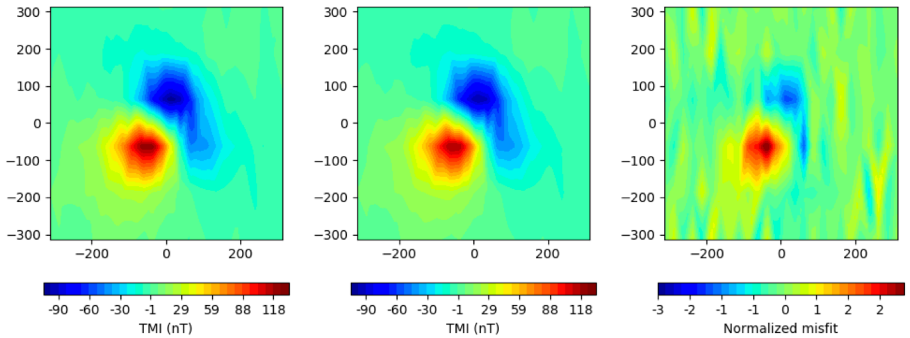
Inversion with sensitivity weighting



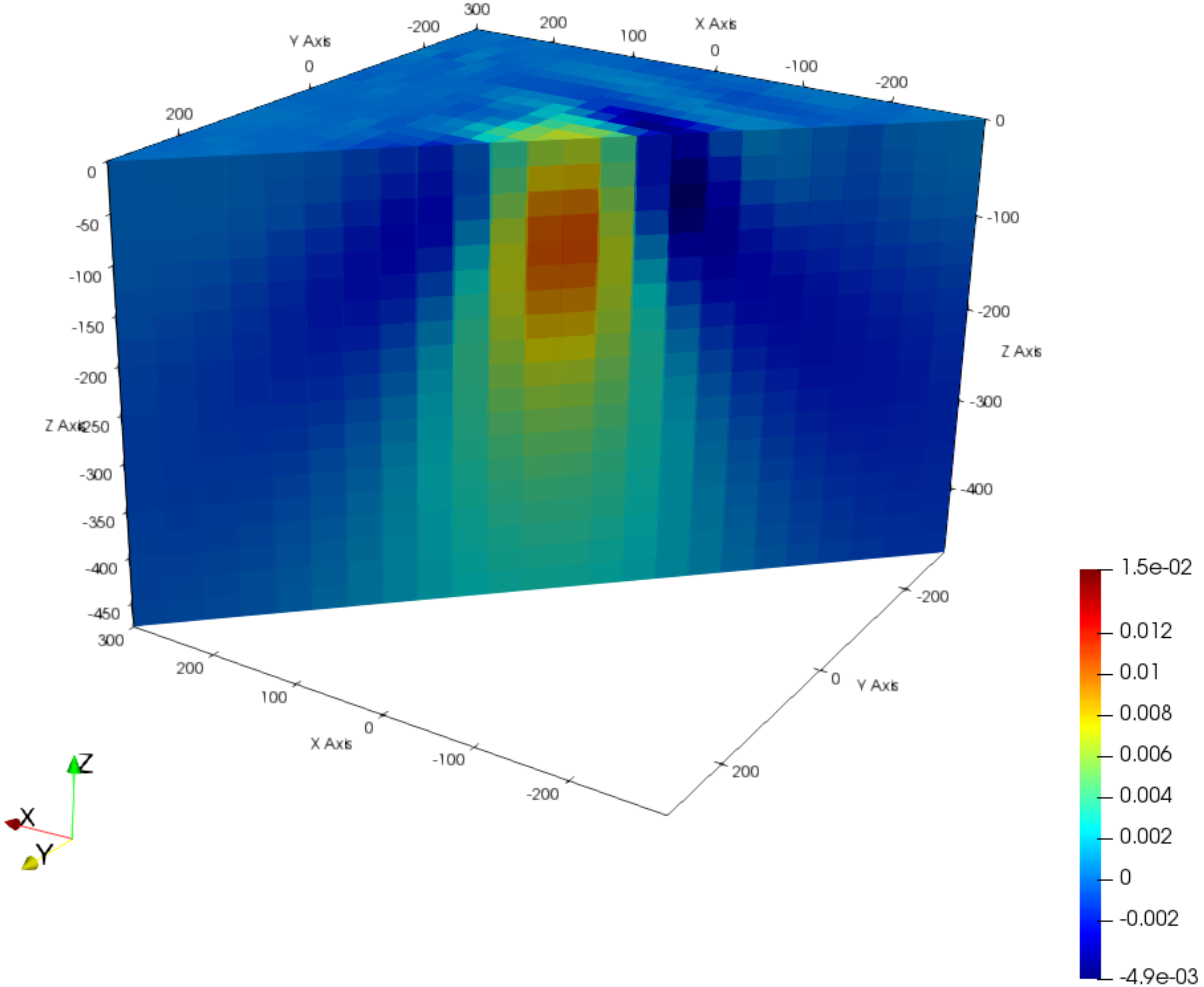
True model



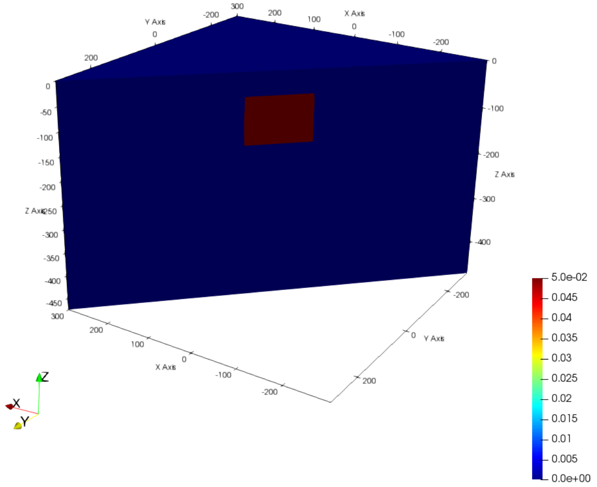
Obs. Pred. Misfit



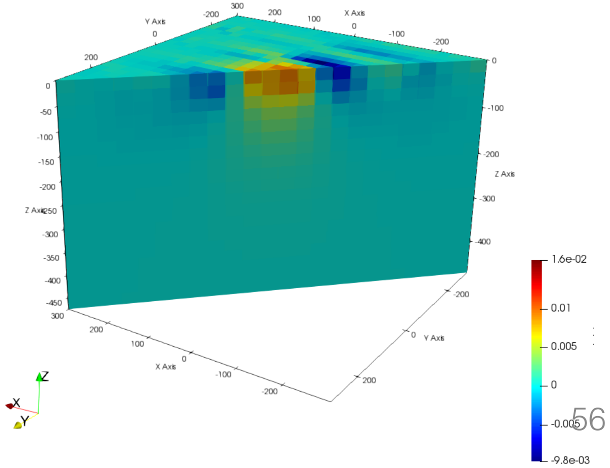
Inversion with sensitivity weighting



True model

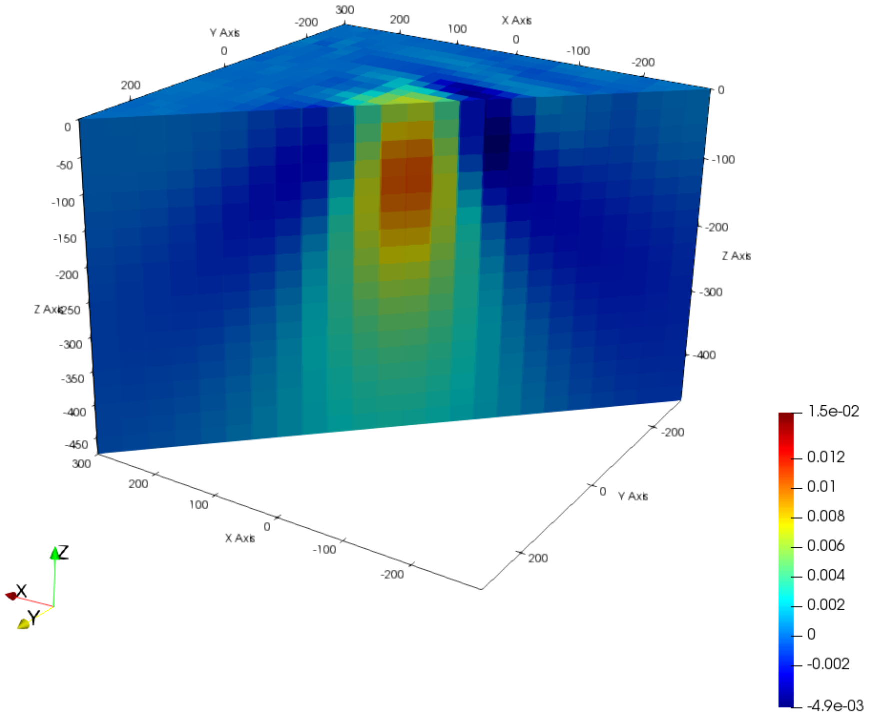


Without sensitivity weight



Summary for sensitivity weighting

- Structure is no longer concentrated at surface.
- Main anomaly is at a reasonable depth.
- BUT:
  - Negative κ persists
  - There is a long tail extending down and out.
- Require positivity



Positivity

- Susceptibility is positive
- The inverse problem is stated as

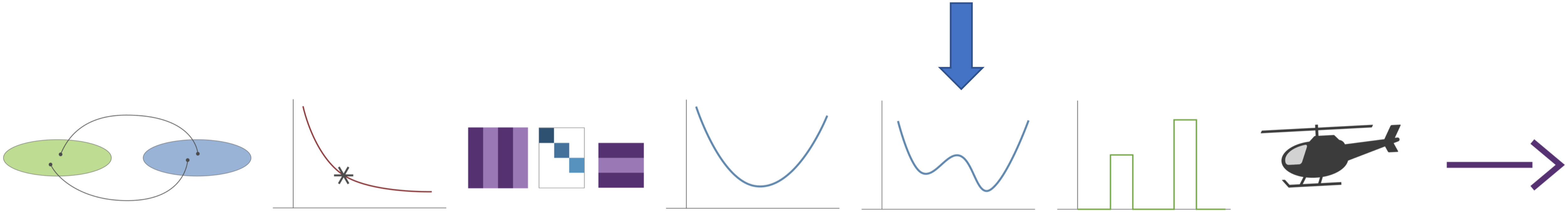
$$\text{minimize} \quad \phi = \phi_d + \beta\phi_m$$
$$\text{subject to } m \geq 0$$

- The problem becomes non-linear. Need to address this.

Outline

- Linear inversion using L2 norms
- 1D inversion app
- 2D cross-well tomography
- Incorporating prior information
  - Alpha parameters
  - Reference model
- 3D magnetics
  - Role for weightings

Next up …



Overview Tikhonov SVD Linear Nonlinear Lp-norms Field scale The future

The end