
The University of Queensland School of Earth and Environmental Sciences

Inversion Using Finite Elements

Lutz Gross
E-mail : l.gross@uq.edu.au

Outline

- Formulation of the inversion problem
 - Gravity
 - MT
- Cost Function Gradient
- Solver: BFGS method
- FEM implementation

Inversion Problem

Optimization problem for property function m

Find $\arg \min_m \Phi(m)$

$$\text{cost function : } \Phi(m) = \underbrace{\Phi_d(m)}_{\text{misfit}} + \underbrace{\mu \cdot \Phi_r(m)}_{\text{regularization}}$$

Positive regularization parameter

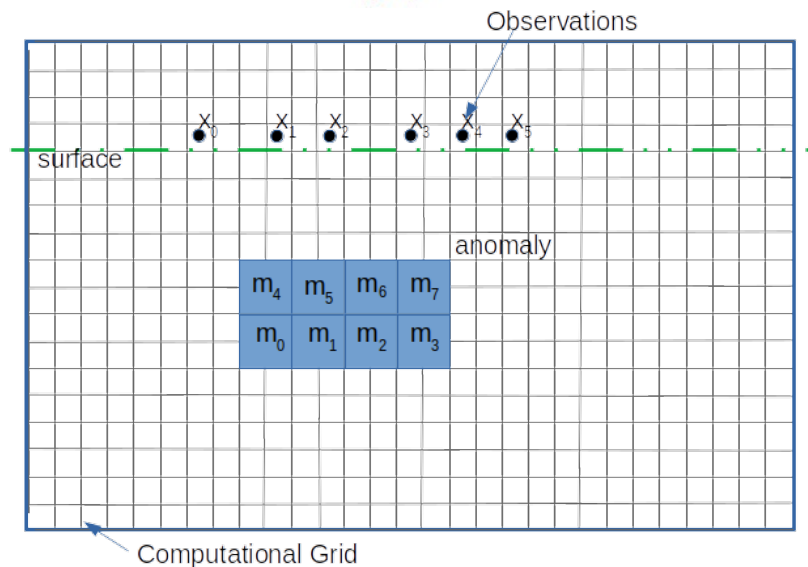
Property Function m

Before:

unknown is a vector of real values

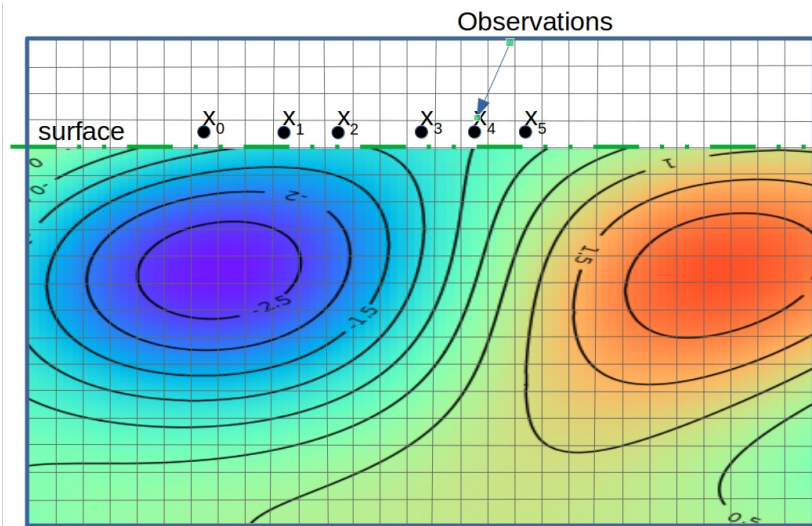
$$\mathbf{m} = [m_0, \dots, m_{N_p-1}]^T$$

$$m = \sum_{k=0}^{N_p-1} m_k \chi_k$$



Now:

unknown m is a function of $\mathbf{x}=(x_0, x_1)$



Regularization

- Now we use integrals and gradients

$$\Phi_r(m) = \frac{1}{2} \int_{\Omega} v_1 \|\nabla m\|^2 + v_0 |m|^2 dx$$

Before:

$$\Phi_r(\mathbf{m}) = \frac{1}{2} \mathbf{m}^T \mathbf{m}$$

weighting factors: $v_1, v_0 \geq 0$

recall: $\|\nabla m\|^2 = \left(\frac{\partial m}{\partial x_0}\right)^2 + \left(\frac{\partial m}{\partial x_1}\right)^2 + \left(\frac{\partial m}{\partial x_2}\right)^2$

Misfit

- Now we use integrals and weighting functions:

$$\Phi_d(m) = \frac{1}{2} \int_{\Omega} w_d \cdot |d - d^{obs}|^2 dx$$

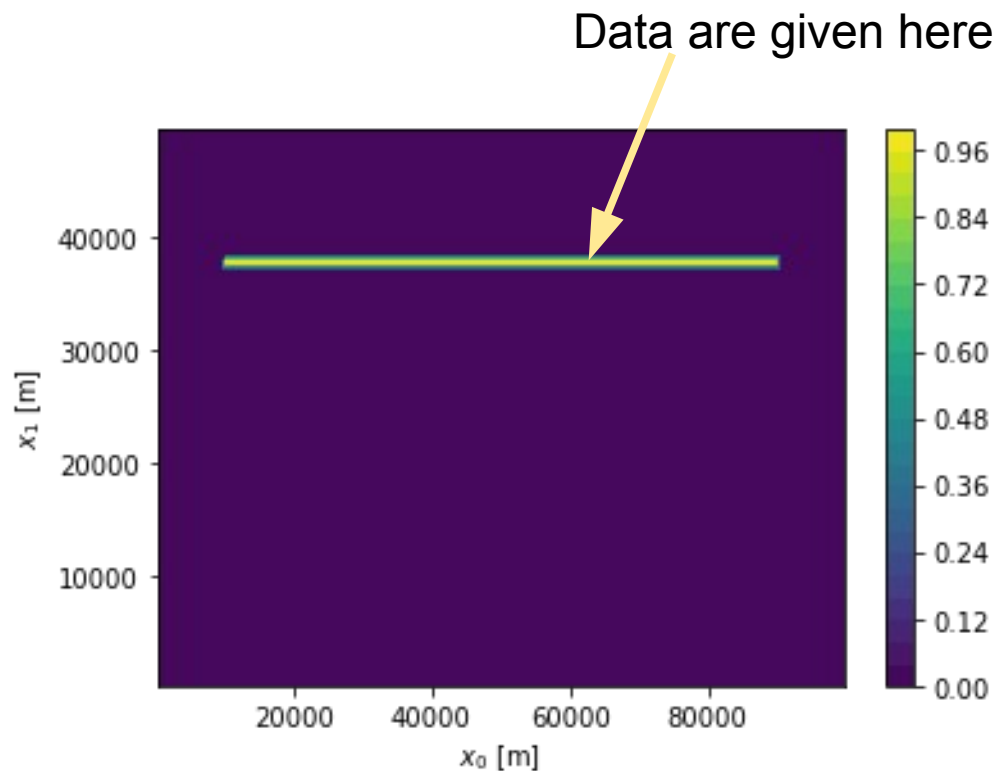
Before:

$$\Phi_d(\mathbf{m}) = \frac{1}{2} (\mathbf{d} - \mathbf{d}^{obs})^T \mathbf{W}_d (\mathbf{d} - \mathbf{d}^{obs})$$

- Forward model: $d = F(m)$
- Weighting function w_d

Weighting Function

$$w_d(\mathbf{x}) = \begin{cases} > 0 & \text{data } d^{obs}(\mathbf{x}) \text{ at } \mathbf{x} \\ 0 & \text{no data} \end{cases}$$



normalize:

$$\int_{\Omega} w_d dx = 1$$

Forward Model: Gravity

- Parametrization of density

$$\rho = \rho' \cdot m + \rho_{ref}$$

- Gravity potential u from PDE:

$$-\nabla^T \nabla u = -4\pi G \rho$$

- Predicted data

$$d = F(m) = g_z = -e_z^T \nabla u$$

Inversion Problem

cost function : $\Phi(m) = \Phi_d(m) + \mu \cdot \Phi_r(m)$

Find $m^* = \arg \min_m \Phi(m)$

We need a critical point condition for minimum m^*

Towards a Critical Point Condition

- For any increment δm for the property function:
- The function

$$\varphi_{\delta m}(\alpha) = \Phi(m^* + \alpha \cdot \delta m) \text{ for } \alpha \in \mathbb{R}$$

has a minimum at $\alpha=0$.

- Therefore

$$\left. \frac{d \varphi_{\delta m}}{d \alpha} \right|_{\alpha=0} = 0 \text{ for all } \delta m$$

Gradient of cost function

- Directional derivative of Φ at m along δm :

$$\langle \nabla \Phi(m) | \delta m \rangle = \left. \frac{d \Phi(m + \alpha \cdot \delta m)}{d \alpha} \right|_{\alpha=0}$$

- Linear function of δm
- Gradient of cost function Φ at m : $\nabla \Phi(m)$

Critical Point Condition

- Find m^* with

$$\langle \nabla \Phi(m^*) | \delta m \rangle = 0$$

for any increment δm

We need to calculate this gradient!?

Increments

- We can see $\delta\Phi$ as a change of Φ due to a `tiny` change δm to m :

$$\delta\Phi = \langle \nabla\Phi(m) | \delta m \rangle$$

- Idea is to track the changes through from the property function m all the way to the cost function Φ

$m \rightarrow \text{density} \rightarrow \text{potential} \rightarrow \text{gravity} \rightarrow \text{misfit} \rightarrow \text{cost function}$

←————— Derive in backwards direction —————→

Increments (cont)

- Some simple rules:

$$f = f(m) \quad : \delta f = f' \cdot \delta m = \frac{df}{dm} \cdot \delta m$$

$$f = f(g(m)) \quad : \delta f = f' \cdot \delta g = f' \cdot g' \cdot \delta m$$

$$f = \nabla g \quad : \delta f = \nabla \delta g$$

Gradient Of Regularization

$$\Phi_r(m) = \frac{1}{2} \int_{\Omega} v_1 \|\nabla m\|^2 + v_0 m^2 d\mathbf{x}$$

$$\delta \Phi_r = \frac{1}{2} \int_{\Omega} v_1 \cdot \delta \|\nabla m\|^2 + v_0 \cdot \delta m^2 d\mathbf{x}$$

$$\delta m^2 = 2m \cdot \delta m$$

$$\delta \|\nabla m\|^2 = \delta (\nabla^T m \nabla m) = 2 \nabla^T m \nabla \delta m$$

$$\delta \Phi_r = \langle \nabla \Phi_r(m) | \delta m \rangle = \int_{\Omega} (v_1 \cdot \nabla^T m \nabla \delta m + v_0 \cdot m \cdot \delta m) d\mathbf{x}$$

Representation of the Gradient

- Identify coefficients:

$$\langle \nabla \Phi_r(m) | \delta m \rangle = \int_{\Omega} \left(\underbrace{v_1 \cdot \nabla^T m}_{=X} \nabla \delta m + \underbrace{v_0 \cdot m}_{=Y} \delta m \right) d\mathbf{x}$$

- Represent gradient by $X=X(m)$, $Y=Y(m)$:

$$\nabla \Phi_r(m) \Leftrightarrow (X, Y)$$

- Then

$$\langle \nabla \Phi_r(m) | \delta m \rangle = \int_{\Omega} (X^T \nabla \delta m + Y \cdot \delta m) d\mathbf{x}$$

Gradient of Misfit

- This is a bit harder!

$$\Phi_d = \frac{1}{2} \int_{\Omega} w_d \cdot |d - d^{ops}|^2 dx$$

$$\delta \Phi_d = \frac{1}{2} \int_{\Omega} w_d \cdot \delta |d - d^{ops}|^2 dx = \int_{\Omega} w_d \cdot (d - d^{ops}) \cdot \delta d \, dx$$

$$d = g_z = -e_z^T \nabla u \quad \longrightarrow \quad \delta d = \delta g_z = -e_z^T \nabla \delta u$$

$$\delta \Phi_d = \int_{\Omega} w_d \cdot (d^{ops} - d) \cdot e_z^T \nabla \delta u \, dx$$



We need a δm here!

Gradient of Misfit

- PDE needs to be solved to connect u with m .
- We solve this in weak form:

$$\rho = \rho' \cdot m + \rho_{ref}$$



$$\delta \rho = \rho' \cdot \delta m$$

$$\int_{\Omega} \nabla^T v \nabla u \, d\mathbf{x} = -4\pi G \int_{\Omega} v \rho \, d\mathbf{x}$$

δ PDE



$$\int_{\Omega} (\nabla^T v \nabla \delta u) \, d\mathbf{x} = \int_{\Omega} \gamma (v \delta m) \, d\mathbf{x}$$

$$\gamma = -4\pi G \rho'$$

$$\rho' = \frac{d\rho}{dm}$$

Adjoint problem

- The magic idea: Solve the adjoint equation for the adjoint potential u^*
- How to get the adjoint equation?
 - collect all δu terms of δPDE
 - Set this equal $\delta \Phi_d$
 - Replace $v \rightarrow u^*$, $\delta u \rightarrow v$
 - Solve adjoint equation for any v to get u^*

Adjoint problem (cont)

- collect all δu terms of δPDE : $\int_{\Omega} (\nabla^T v \nabla \delta u) d\mathbf{x}$
- Set this equal $\delta \Phi_d$

$$\int_{\Omega} \left(\nabla^T \underbrace{v}_{\rightarrow u^*} \nabla \underbrace{\delta u}_{\rightarrow v} \right) d\mathbf{x} = \int_{\Omega} w_d \cdot (d^{ops} - d) \cdot e_z^T \nabla \underbrace{\delta u}_{\rightarrow v} dx$$

- Replace $v \rightarrow u^*$, $\delta u \rightarrow v$:

$$\int_{\Omega} (\nabla^T u^* \nabla v) d\mathbf{x} = \int_{\Omega} w_d \cdot (d^{ops} - d) \cdot e_z^T \nabla v dx \text{ for all } v$$

What is the gradient of ϕ_d then?

- Adjoint problem: set $v = \delta u$:

$$\int_{\Omega} (\nabla^T u^* \nabla \delta u) d\mathbf{x} = \int_{\Omega} w_d \cdot (d^{ops} - d) \cdot e_z^T \nabla \delta u dx = \delta \Phi_d$$

- in δ PDE: set $v = u^*$:

$$\int_{\Omega} (\nabla^T u^* \nabla \delta u) d\mathbf{x} = \int_{\Omega} \gamma(u^* \delta m) d\mathbf{x}$$

- So we have: $\delta \Phi_d = \int_{\Omega} \gamma(u^* \delta m) d\mathbf{x}$

Gradient of cost function

- We put this all together:

$$\langle \nabla \Phi | \delta m \rangle = \int_{\Omega} \left(\mathbf{v}_1 \cdot \nabla^T m \nabla \delta m + (\gamma u^* + \mathbf{v}_0 \cdot m) \cdot \delta m \right) d\mathbf{x}$$

- Representation for an implementation:

$$\nabla \Phi = (X, Y) = (\mu \mathbf{v}_1 \cdot \nabla m, \gamma u^* + \mu \mathbf{v}_0 \cdot m)$$

How to find the minimum?

- Quasi-Newton Scheme

- Iterative process creating sequence of property function approximations m^k $k=1,2,3,4,\dots$
- second-order approximation of cost function at m^k

$$\Phi(m^k + p) \approx \Phi(m^k) + \langle \nabla \Phi(m^k) | p \rangle + \langle B^k p | p \rangle$$

- with an approximation B^k of the Hessian of Φ
- Apply gradient with respect to p :
$$\nabla \Phi(m^k + p) = \nabla \Phi(m^k) + B^k p$$

Iterative Scheme

- Iterative scheme:

- Solve for search direction p^k :

$$B^k p^k = -\nabla \Phi(m^k)$$

- Apply line search

$$\alpha_k = \arg \min_{\alpha} \Phi(m^k + \alpha \cdot p^k)$$

- Update solution:

$$m^{k+1} = m^k + s^k \quad \text{with } s^k = \alpha^k \cdot p^k$$

- Calculate

$$g^k = \nabla \Phi(m^{k+1}) - \nabla \Phi(m^k)$$

- Update Hessian:

$$\text{such that: } B^{k+1} s^k = g^k$$

- Next step:

$$k \leftarrow k+1$$

Inversion of Hessian

- Gradient and hence difference of gradients are represented by tuple $g=(X,Y)$:

solving $B^k p = g$ means :

$$\langle B^k p | q \rangle = \int_{\Omega} (X^T \nabla q + Y \cdot q) d\mathbf{x} \text{ for all } q$$

- See weak form of PDE template

BFGS

- Broyden–Fletcher–Goldfarb–Shanno algorithm
 - Special update of approximate Hessian

$$B^{k+1} = B^k + \frac{1}{a_k} |g^k\rangle\langle g^k| - \frac{1}{b_k} |B^k s^k\rangle\langle B^k s^k|$$

$$a_k = \langle g^k | s^k \rangle = \int_{\Omega} \left(X_g^T \nabla s^k + Y_g \cdot s^k \right) d\mathbf{x} \text{ with } g^k = (X_g, Y_g)$$

$$b_k = \langle B^k s^k | s^k \rangle = \int_{\Omega} \left(X_B^T \nabla s^k + Y_B \cdot s^k \right) d\mathbf{x} \text{ with } B^k s^k = (X_B, Y_B)$$

- Inverse $(B^k)^{-1}$ of B^k can be calculated explicitly.

Remarks

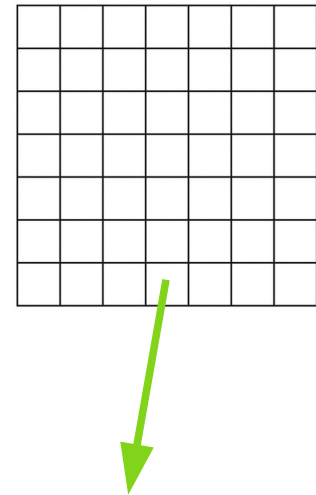
- Implemented with
 - truncation of terms in the Hessian
 - restart to throw away bad approximations
- Inverse of the initial Hessian approximation B^0 is required
 - In inversion: Use the Hessian of the regularization

get $p = (B^0)^{-1} g$ by solving

$$\int_{\Omega} \left(v_1 \nabla^T q \nabla p + v_0 q p \right) d\mathbf{x} = \int_{\Omega} \left(X_g^T \nabla q + Y_g \cdot q \right) d\mathbf{x} \text{ for all } q$$

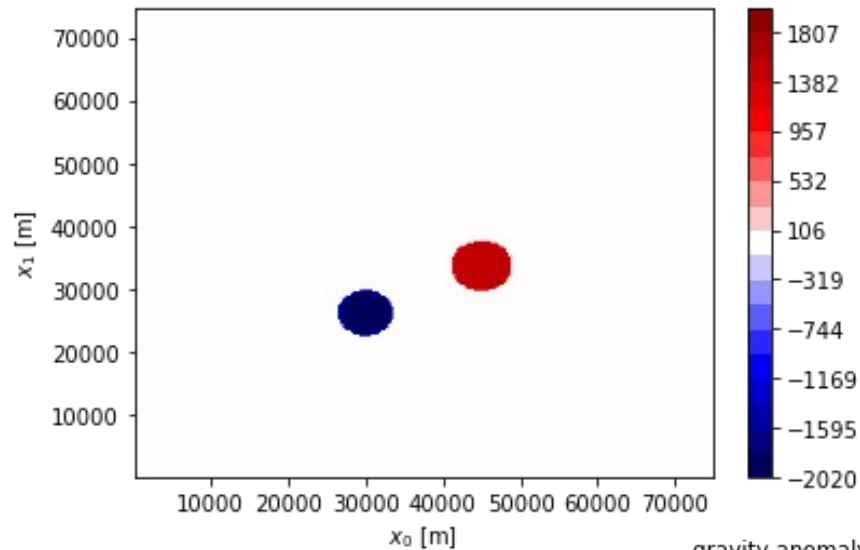
FEM implementation

- Use FEM to solve:
 - Forward problems
 - Adjoint problems
 - Hessian approximation
- Values
 - Potential $u \rightarrow$ FEM nodes
 - Adjoint solution $u^* \rightarrow$ FEM nodes
 - Cost function gradients $(X,Y) \rightarrow$ integration points

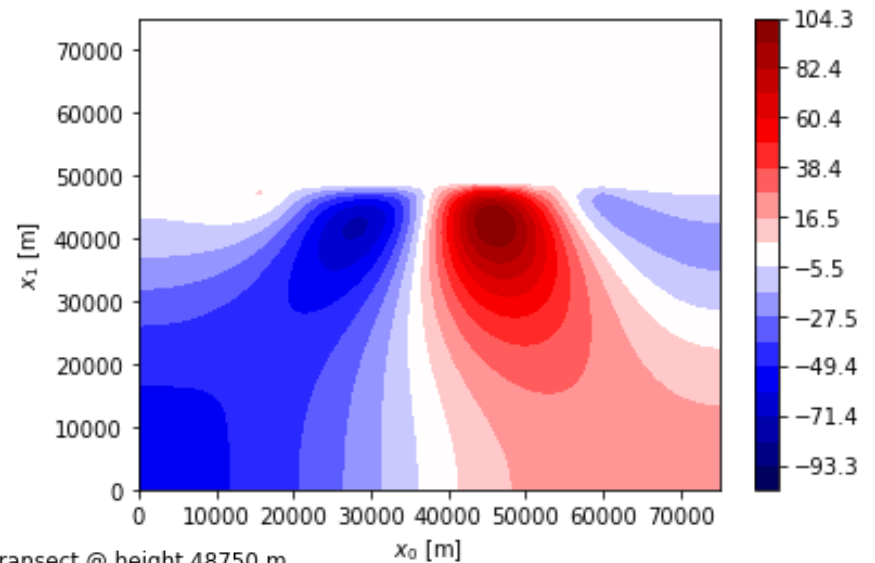


Gravity

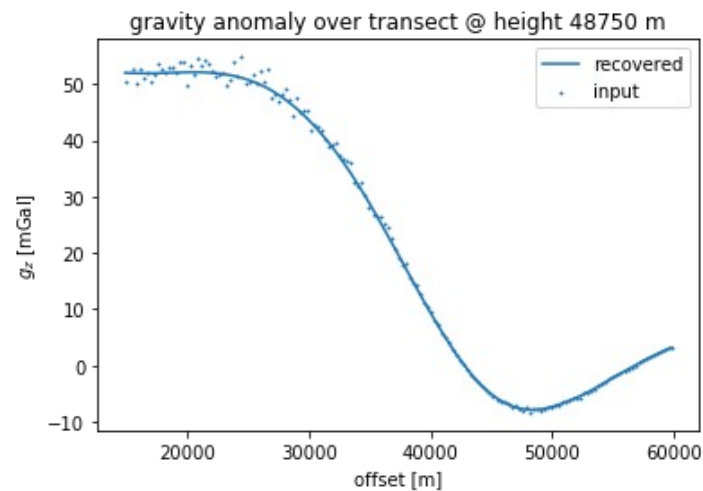
True Density



Recovered



$$\mu = 10^{-6}$$



FEM inversion

Inversion MT-TE

- Conductivity model (needs to be non negative)

$$\sigma = \sigma_{ref} e^m$$

- Electric field E_x solution of PDE

$$-\nabla^T \nabla E_x + j \mu_0 \omega \sigma E_x = 0$$

- Predicted data

$$d = F(m) = Z_{xy} = j \omega \mu_0 E_x \cdot \left(\frac{\partial E_x}{\partial x_1} \right)^{-1}$$

Gradient of MT Cost Function?

$$\Phi_d = \frac{1}{2} \int_{\Omega} w_d \cdot |Z - Z^{ops}|^2 dx$$

$$\delta |a|^2 = \delta(\bar{a} a) = a \delta \bar{a} + \delta a \bar{a} = 2 \Re[\delta a \bar{a}]$$

$$\delta \Phi_d = \int_{\Omega} w_d \cdot \Re[(\overline{Z - Z^{ops}}) \delta Z] dx$$

$$Z = j \omega \mu_0 u \cdot \left(\frac{\partial u}{\partial x_1} \right)^{-1}$$

$$\delta Z = j \omega \mu_0 \left(\delta u \frac{\partial u}{\partial x_1} - u \frac{\partial \delta u}{\partial x_1} \right) \cdot \left(\frac{\partial u}{\partial x_1} \right)^{-2}$$

$$\delta \Phi_d = \Re \int_{\Omega} w_d \cdot (Y_d \delta u + \mathbf{X}_d^T \nabla \delta u) d\mathbf{x} \text{ with}$$

$$Y_d = j \omega \mu_0 \overline{(Z - Z^{ops})} \left(\frac{\partial u}{\partial x_1} \right)^{-1} \quad \mathbf{X}_d = -j \omega \mu_0 \overline{(Z - Z^{ops})} \cdot u \cdot \left(\frac{\partial u}{\partial x_1} \right)^{-2} \mathbf{e}_1$$

Adjoint Problem

- Weak form of forward model:

$$\int_{\Omega} \nabla^T v \nabla u + j \omega \sigma u v d\mathbf{x} = 0 \text{ for all } v$$

- Apply δ :

$$\int_{\Omega} \nabla^T v \nabla \delta u + j \omega \mu_0 (u \delta \sigma + \sigma \delta u) v d\mathbf{x} = 0 \text{ for all } v$$

- Adjoint problem: collect δu :

$$\int_{\Omega} (\nabla^T u^* \nabla v + j \mu_0 \omega \sigma u^* v) d\mathbf{x} = \int_{\Omega} w_d \cdot (Y_d v + \mathbf{X}_d^T \nabla v) d\mathbf{x} \text{ for all } v$$

Gradient MT-TE Mode

- Finally we have:

$$\begin{aligned}\delta \Phi_d &= -\Re \int_{\Omega} \mathbf{j} \mu_0 \omega u u^* \delta \sigma d\mathbf{x} = -\Re \int_{\Omega} (\mathbf{j} \mu_0 \omega u u^* \sigma') \delta m d\mathbf{x} \\ \delta \sigma &= \frac{d\sigma}{dm} \delta m = \sigma' \delta m\end{aligned}$$

$= \int_{\Omega} \underbrace{(-\Re \mathbf{j} \mu_0 \omega u u^* \sigma')}_{=Y} \delta m d\mathbf{x}$

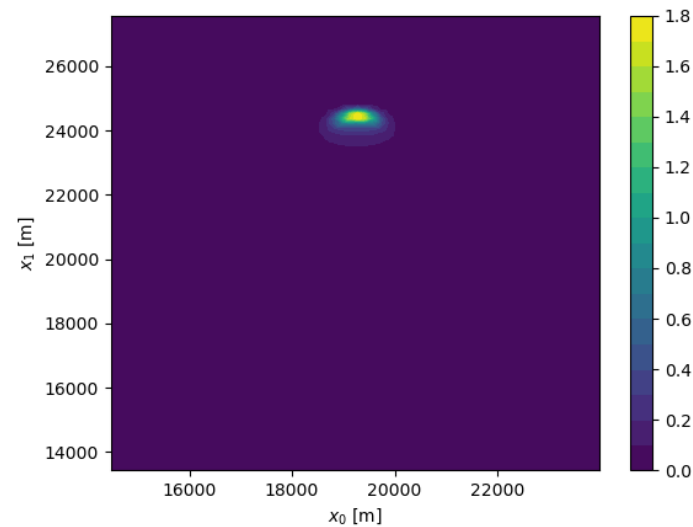
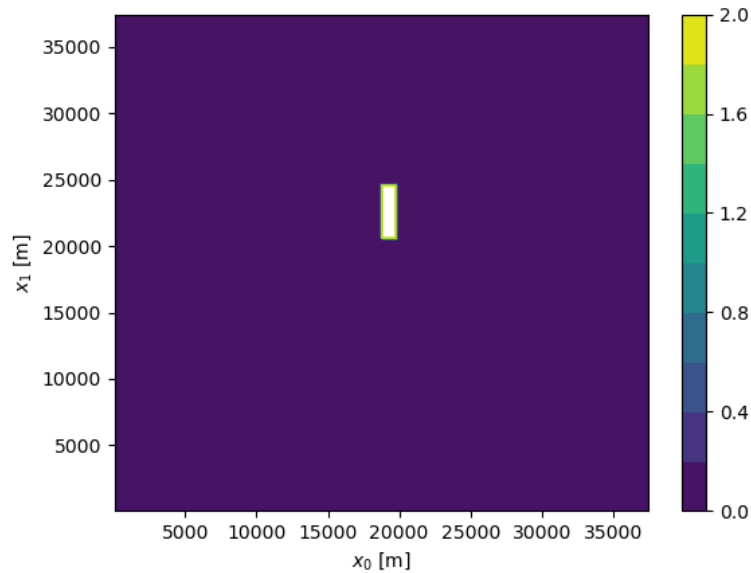
Property function m is real

- Put is all together:

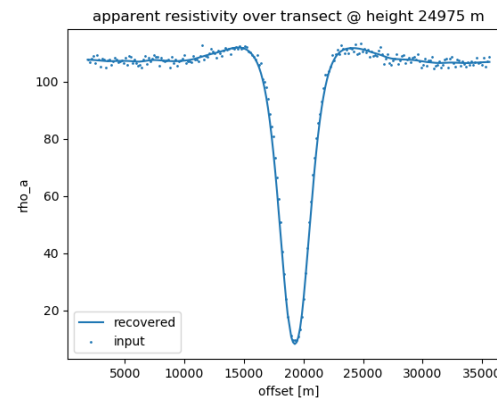
$$\langle \nabla \Phi | \delta m \rangle = \int_{\Omega} \left(\mathbf{v}_1 \cdot \nabla^T m \nabla \delta m + (\mathbf{v}_0 \cdot \mathbf{m} - \Re \mathbf{j} \mu_0 \omega u u^* \sigma') \cdot \delta m \right) d\mathbf{x}$$

- Ready for the BFGS

MT Example



$f=5\text{Hz}$, $\mu=10^{-2}$



Joint Inversion

- Idea: invert for two or more physical properties simultaneously.
 - Reduce the `degree of ill-posedness` - really???
 - Create consistent distributions of physical properties.
- How to connect two inversions?
 - For instance Gravity and MT.

Constitutive Relationships

- Explicit relationships are known
 - typically empirical
 - for instance: density as function of electric conductivity

$$\rho = F(\sigma)$$

Some lateritic soil

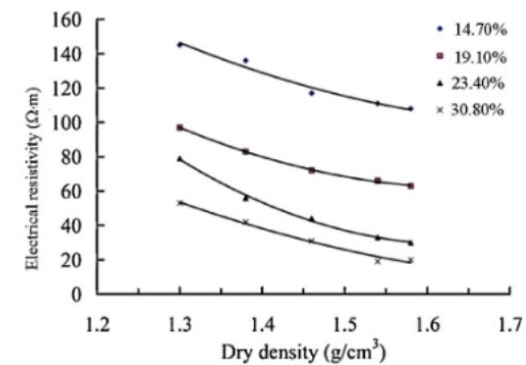


Fig. 12. Relationship between dry density and electrical resistivity at different water contents.

<https://doi.org/10.1016/j.jrmge.2013.07.003>

How to build the inversion?

- Unknown: property function m
 - Density and conductivity distribution are both parameterized with m :

$$\begin{aligned}\sigma &= \sigma_{ref} e^m \\ \rho &= F(\sigma_{ref} e^m)\end{aligned}$$

- Cost function to minimize:

$$\Phi(m) = \underbrace{\alpha^{(1)} \Phi_d^{(1)}(m)}_{\text{gravity}} + \underbrace{\alpha^{(2)} \Phi_d^{(2)}(m)}_{\text{MT}} + \underbrace{\mu \cdot \Phi_r(m)}_{\text{regularization}}$$

Tweak factors to adjust balance : $\alpha^{(1)}, \alpha^{(2)}$

- BFGS ready!

Without Constitutive Relationships?

- Difficulties: It is an optimistic approach!
 - Relationships are unknown
 - Or can only be obtained for shallow regions
 - Relationships may be valid in certain regions but the extends of the regions are unknown
- Assumption: There is a relationship between density and conductivity but the actual relationship is unknown.

$$\rho = F(\sigma) \text{ with unknown } F$$

Joint Inversion Set-up

- Unknowns: $\mathbf{m}=(m_0, m_1)$ $m_0=\rho$ $m_1=\ln\left(\frac{\sigma}{\sigma_{ref}}\right)$
- If there is relationship between σ and ρ then also for the respective property functions m_0, m_1

$$G(m_1, m_2)=0$$

- Problem: Function G is not known!
- Let's assume it not depending on location \mathbf{x}

Gradient Correlation

- Apply spatial gradient:

$$a_0 \cdot \nabla m_0 + a_2 \nabla m_1 = 0 \quad \text{with } a_0 = \frac{\partial G}{\partial m_0} \quad a_1 = \frac{\partial G}{\partial m_1}$$

- This is a system of linear equations for (a_0, a_1) :
 - Two equations in 2D
 - Three equations in 3D
- Normal equation must have non-zero solution.

Gradient Correlation (cont.)

- Multiply by ∇m_0 and by ∇m_1 :

$$\begin{bmatrix} \|\nabla m_0\|^2 & \nabla^T m_0 \nabla m_1 \\ \nabla^T m_1 \nabla m_0 & \|\nabla m_1\|^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- For non-trivial solution: determinant has to be zero:

$$C(m_0, m_1) = \|\nabla m_0\|^2 \|\nabla m_1\|^2 - (\nabla^T m_0 \nabla m_1)^2 = 0$$

- Notice: always $C(m_0, m_1) \geq 0$
 - Also known as Gram determinant
- We want to minimize $C(m_0, m_1)$ in the inversion

Cost Function

- Cost function to minimize over $\mathbf{m}=(m_0, m_1)$:

$$\Phi(\mathbf{m}) = \underbrace{\alpha^{(1)}(\Phi_d^{(1)}(m_0) + \mu_1 \Phi_r^{(1)}(m_0))}_{\text{gravity}} + \underbrace{\alpha^{(2)}(\Phi_d^{(2)}(m_1) + \mu_1 \Phi_r^{(2)}(m_1))}_{\text{MT}} + \theta \underbrace{\Phi_c(m_0, m_1)}_{\text{correlation}}$$

$$\Phi_c(m_1, m_2) = \frac{1}{2} \underbrace{\int_{\Omega} \left(\frac{1}{\|\nabla m_0\|^2} + \frac{1}{\|\nabla m_1\|^2} \right)}_{\text{balance vs. regularization}} \cdot C(m_0, m_1) d\mathbf{x}$$

<https://doi.org/10.1093/gji/ggz134>

Cross Gradient Form

- Lagrange's identity:

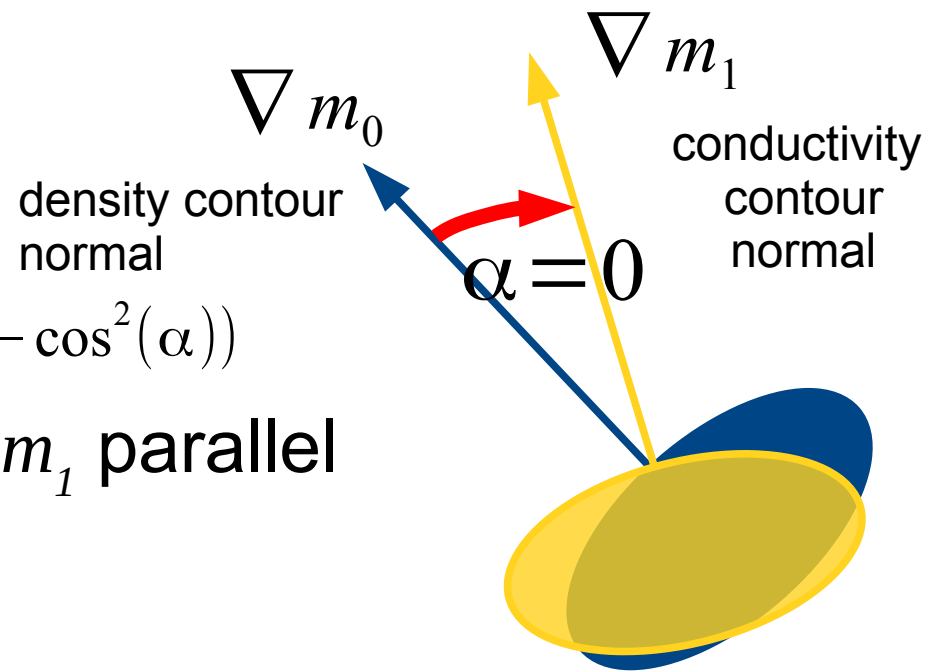
$$C(m_0, m_1) = \|\nabla m_0\|^2 \|\nabla m_1\|^2 - (\nabla^T m_0 \nabla m_1)^2 = \|\nabla m_0 \times \nabla m_1\|^2$$

- Known as “cross-gradient” form

- Geometrical interpretation

$$\|\nabla m_0 \times \nabla m_1\| = \|\nabla m_0\| \cdot \|\nabla m_1\| (1 - \cos^2(\alpha))$$

- Target: make ∇m_0 and ∇m_1 parallel



Implementation

- Solution by BFGS method
 - Gradient:

$$\langle \nabla \Phi(\mathbf{m}) | \delta \mathbf{m} \rangle = \alpha^{(1)} \langle \nabla \Phi^{(1)}(m_0) | \delta m_0 \rangle + \alpha^{(2)} \langle \nabla \Phi^{(2)}(m_1) | \delta m_1 \rangle + \theta \langle \nabla \Phi_c(\mathbf{m}) | \delta \mathbf{m} \rangle$$

$$\langle \nabla \Phi_c(\mathbf{m}) | \delta \mathbf{m} \rangle = \int_{\Omega} \left((a_0 \nabla m_0 - a \nabla m_1)^T \nabla \delta m_0 + (a_1 \nabla m_1 - a \nabla m_0)^T \nabla \delta m_1 \right) d\mathbf{x}$$

with some a_0, a_1, a depending $\nabla m_0, \nabla m_1$

<https://doi.org/10.1093/gji/ggz134>

Check Correlation

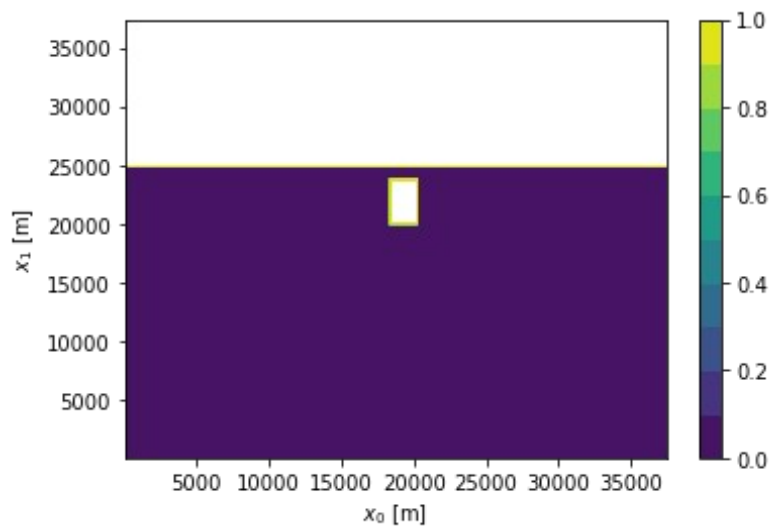
- Measure of correlation:

$$R = \frac{\nabla^T m_1 \nabla m_0}{\|\nabla m_1\| \|\nabla m_0\|} \in [-1, 1]$$

- Good correlation: $R^2 \approx 1$
- Weak correlation: $R^2 \ll 1$
 - Marks regions in which G is changing with x .
 - Interpretation: interface of geological units

Test Case

- $f=5\text{Hz}$, $\sigma_1=2$, $\rho_1=1500$
- $\sigma_{bg}=0.01$, $\rho'=500$
- $\alpha^g=1e-4$, $\alpha^{mt}=0.0$
- $\mu^g, \mu^{mt}=1e6$
- $\theta=1e8$
- 62500 cells.

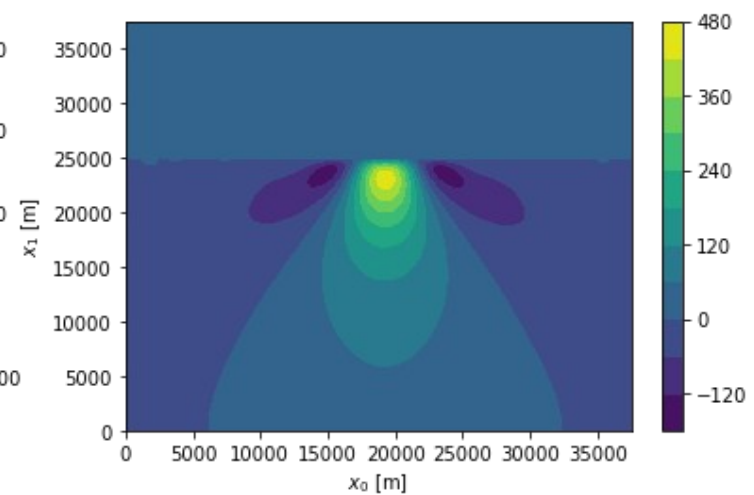
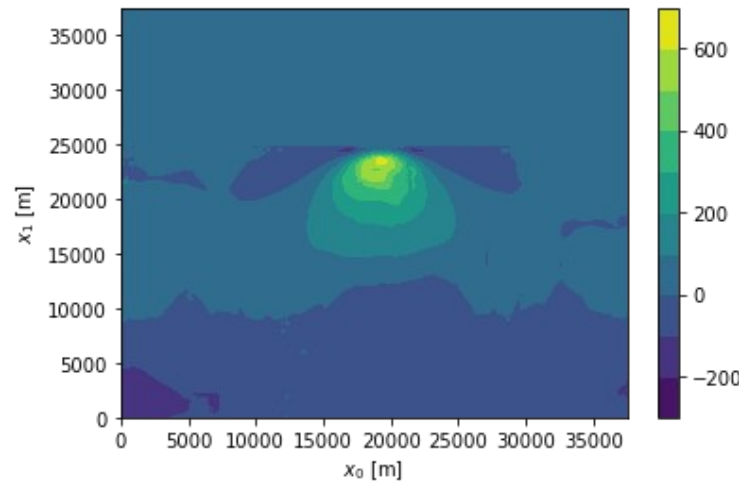


Properties

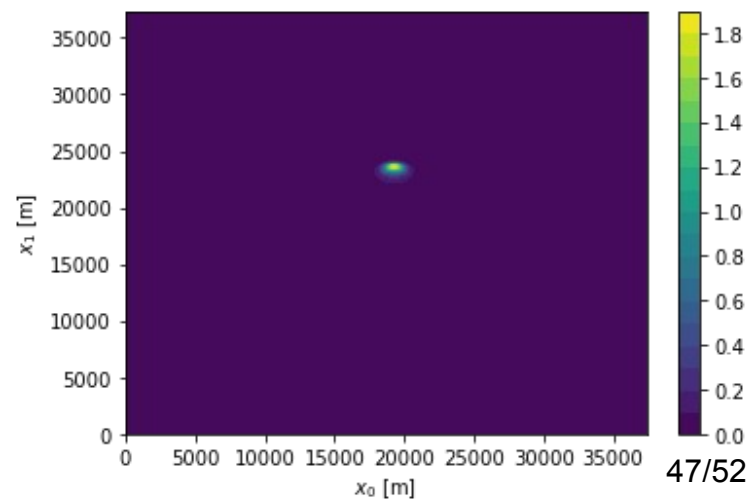
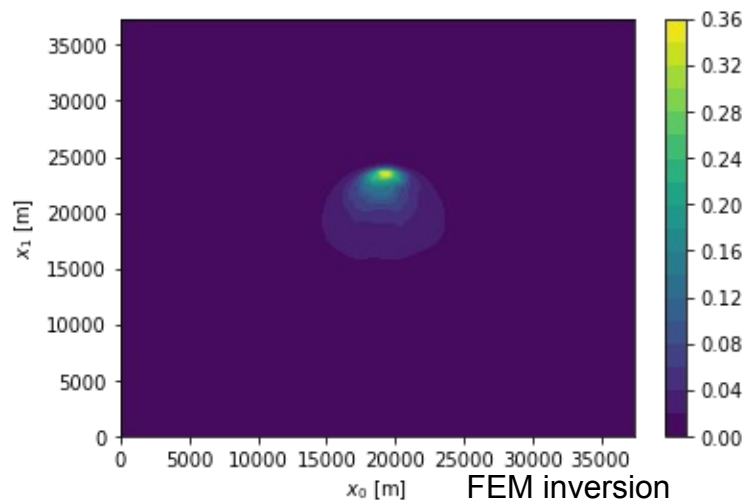
Joint inversion

Inversion $\theta=0$

Density



Conductivity



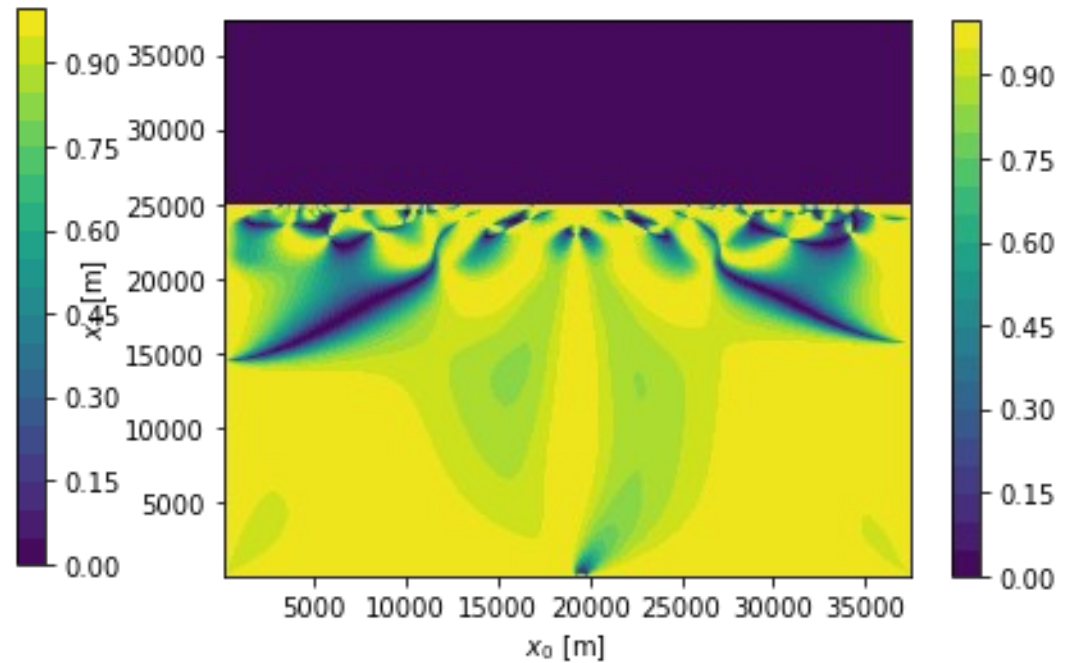
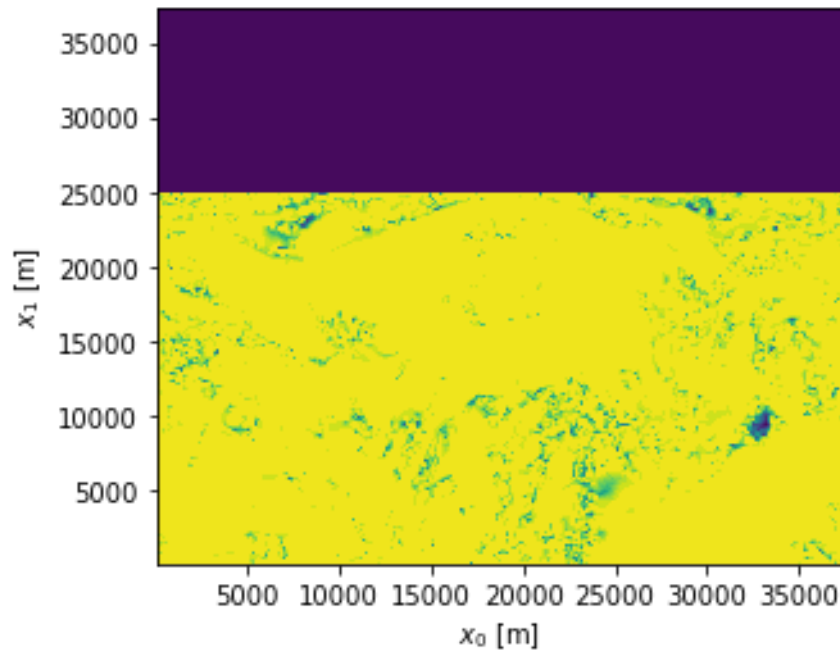
FEM inversion

47/52

Correlation R^2

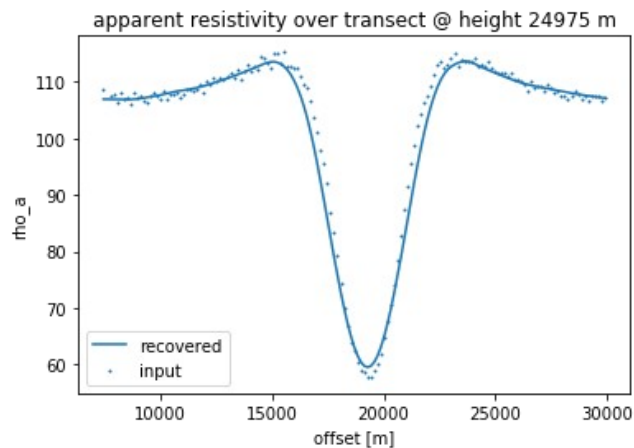
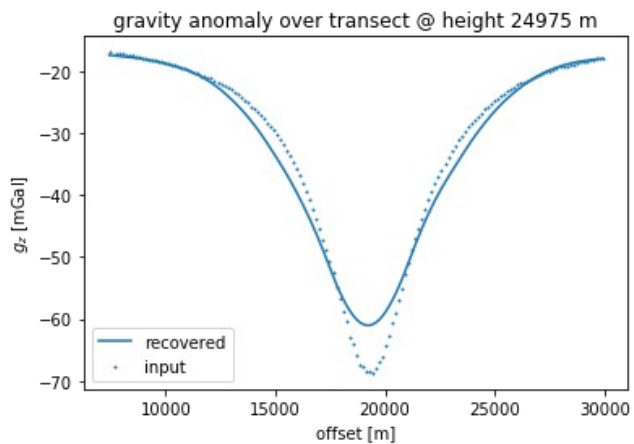
Joint inversion

Inversion $\theta=0$

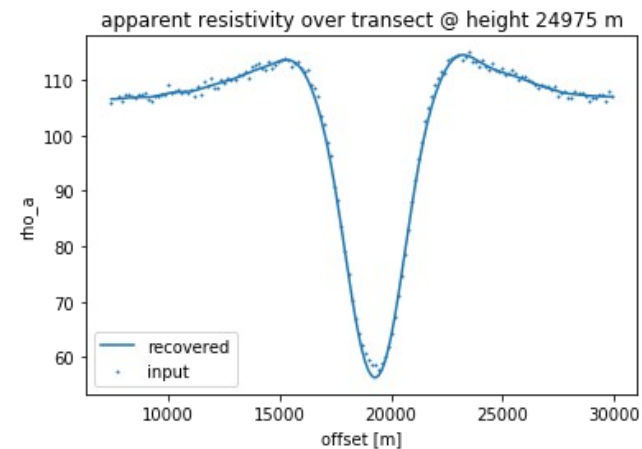
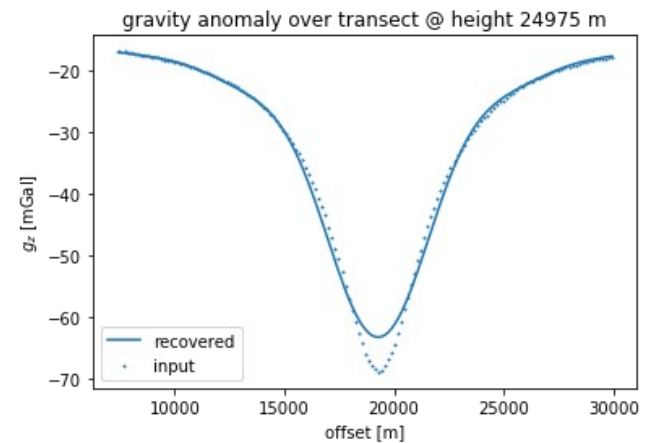


Data Recovery

Joint inversion

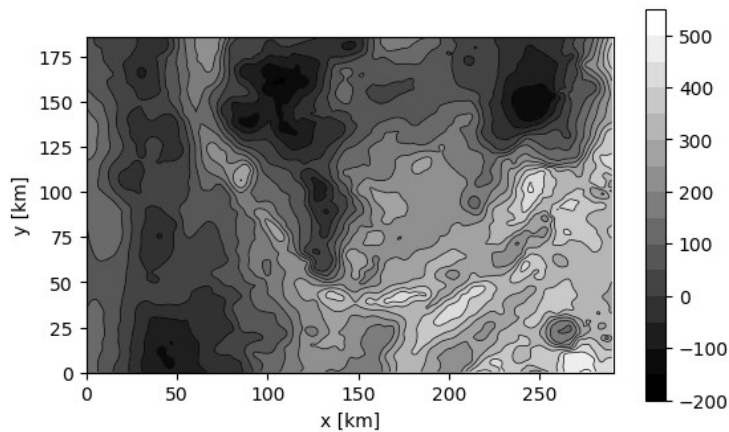


Inversion $\theta=0$

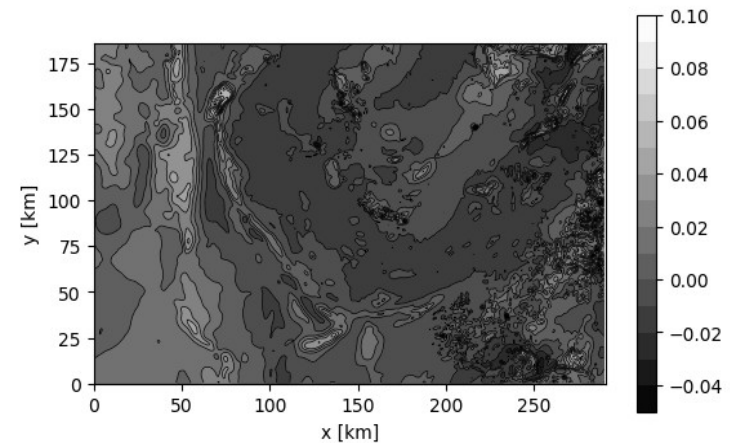


Field Data set from Central QLD

Gravity Anomalies QLD:



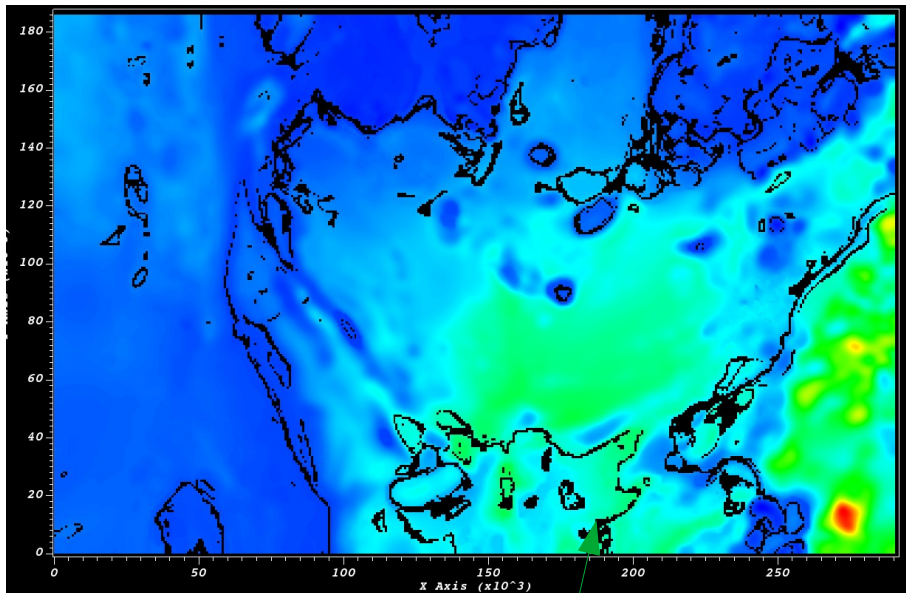
Magnetic field intensity QLD



Joint Inversion with 30 Million cells on ~200 cores
Implementation in python3 + esys-escript

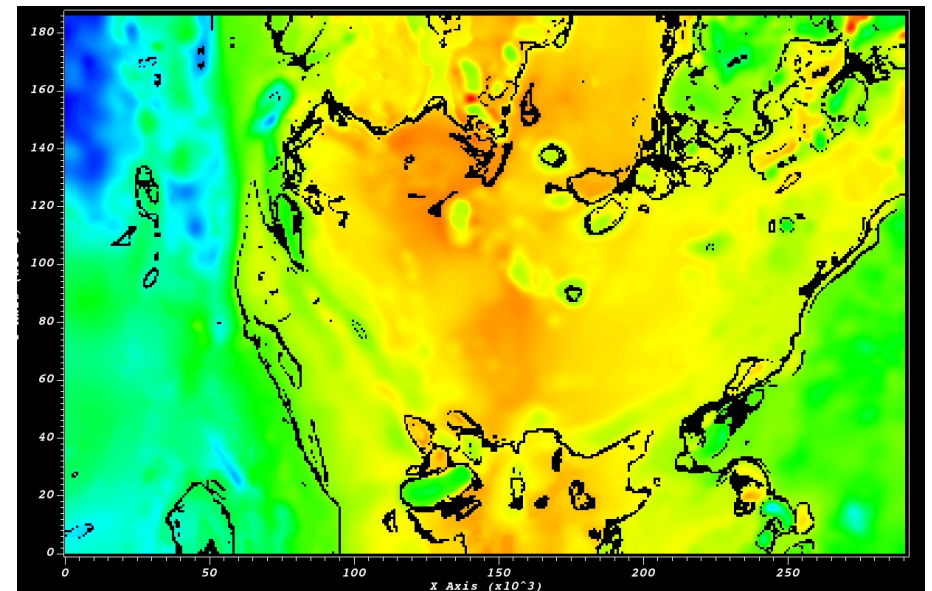
Some Results

Density at depth 500m



Locations with $R^2 < 0.97$
= boundary of geological units

Log of susceptibility at depth 500m



Thanks!

Questions, Suggestions, Share programs:
Please get in contact on e-mail:

l.gross@uq.edu.au

Or WeChat