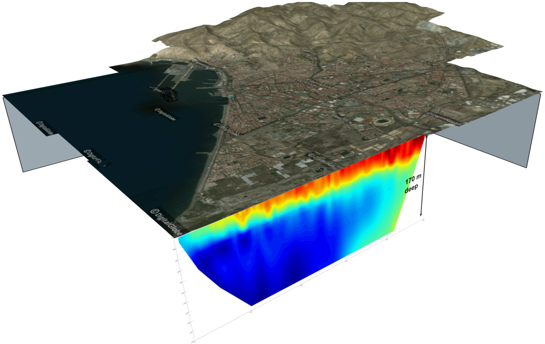
The exploration manager asks a geologist, an engineer, and a geophysicist what 2 + 2 is.



Nguyen et al., 2009

Frederic Nguyen, Cargese 2018 summer school

LIÈGE université

Figure: J. Hadamard (1865-1963)

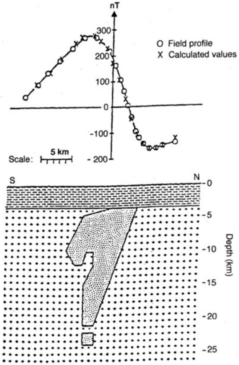


Figure: Inversion of magnetic data (source:?)

from Zhdanov (2015)

Figure: A.N. Tikhonov (1906-1993)

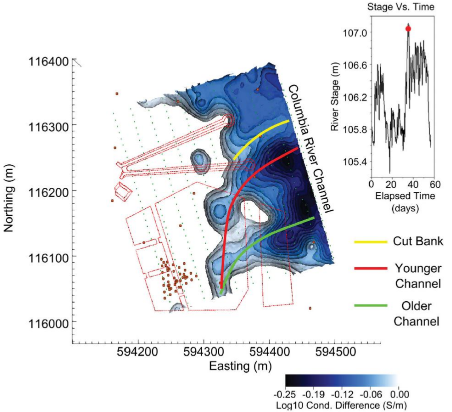


Figure: 4D river water intrusion Johnson et al., (2015)

from Zhdanov (2015)

Outline

Least-squares

Our objective if to find a model **m** which is capable of reproducing our observable data **d**. However, in practice, it is impossible to find a **m** that solve exactly **g(m)=d**. As an alternative, one can minimize the **data residuals** (measure of the misfit) to some tolerance δ:

$$\mathbf{r=d\text{-}g(m)} \tag{1}$$

We generally use the L_p norm to quantify the misfit:

$$||\mathbf{r}||_p = (\sum_{i=1}^{m} |r_i|^p)^{\frac{1}{p}} \leq \delta \tag{2}$$

SVD

Moore-Penrose inverse

If the problem is linear $g(\mathbf{m}) = \mathbf{Gm}$ then we can use the SVD

$$\mathbf{G} = [\mathbf{U}_p, \mathbf{U}_0] \begin{pmatrix} \mathbf{S}_p & 0 \\ 0 & 0 \end{pmatrix} [\mathbf{V}_p, \mathbf{V}_0]^T \tag{3}$$

where **U** and **V** an m by m and n by n orthogonal matrix, respectively, and **S** an m by n diagonal matrix with nonnegative diagonal elements called singular values arranged in decreasing size $(s_1 \geq s_2 \geq \ldots \geq s_{min(m,n)})$.

to compute a generalized inverse of the matrix **G**, called the Moore-Penrose pseudoinverse:

$$\mathbf{G}^\dagger = \mathbf{V}_p \mathbf{S}_p^{-1} \mathbf{U}_p^T \tag{4}$$

Definition

We define the pseudoinverse solution $m_\dagger$ to the inverse problem $\mathbf{Gm} = \mathbf{d}$:

$$m_\dagger = \mathbf{V}_p \mathbf{S}_p^{-1} \mathbf{U}_p^T \mathbf{d} = \sum_{i=1}^{p} \frac{\mathbf{U}_{:,i}^T (\mathbf{d} + \epsilon)}{s_i} \mathbf{V}_{:,i} \tag{5}$$

which **always** exist.

Resolution and covariance of the pseudoinverse solution

Definition

$$Cov(\mathbf{m}_\dagger) = \mathbf{G}^\dagger Cov(\mathbf{d}(\mathbf{G}^\dagger)^T \tag{6}$$

$$= \sigma^2 \sum_{i=1}^{p} \frac{V_{:,i} V_{:,i}^T}{s_i^2} \tag{7}$$

Note that as p increases, the variance increases.

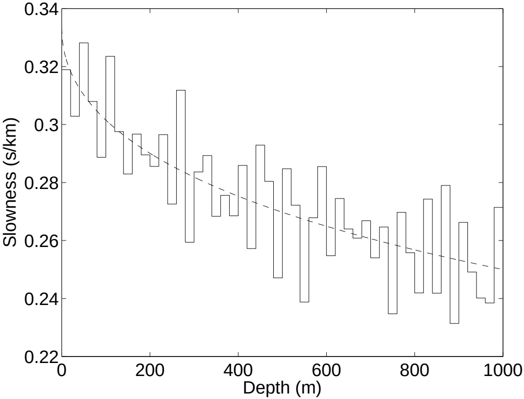
Definition

The **model resolution matrix $\mathbf{R}_m$** is defined by:

$$\mathbf{m}_\dagger = \mathbf{G}^\dagger d \simeq \mathbf{G}^\dagger \mathbf{G} \mathbf{m}_{true} \tag{8}$$

$$\mathbf{R}_m = \mathbf{G}^\dagger \mathbf{G} = \mathbf{V}_p \mathbf{V}_p^T \tag{9}$$

Consider a vertical seismic profile (VSP) problem for a smooth velocity profile, the full SVD solution yields :



Aster et al., 2012

Tikhonov and the L-curve

We will examine another way of stabilizing the solution using the Tikhonov regularization, one of the most widely used technique to solve inverse problems (Tikhonov and Arsenin, 1977). It can be stated as follows:

$$\min ||\mathbf{m}||_2 \tag{10}$$
$$||\mathbf{Gm}\text{-}\mathbf{d}||_2 \leq \delta \tag{11}$$

so that unneeded features will not appear in the regularized solution.

Draw Lcurve.

Note that this is equivalent to:

$$\min ||\mathbf{Gm}\text{-}\mathbf{d}||_2 \tag{12}$$
$$||\mathbf{m}||_2 \leq \epsilon \tag{13}$$

Applying the method of the Lagrange multiplier we have:

$$\min ||\mathbf{Gm\text{-}d}||_2 + \alpha^2 ||\mathbf{m}||_2 \tag{14}$$

where α is a **regularization parameter**. By computing the derivative of this functional with respect to **m** and using the SVD of **G** we can find the solution $\mathbf{m}_\alpha$

$$\mathbf{m}_\alpha = \sum_{i=1}^{k} \frac{s_i^2}{(s_i^2 + \alpha^2)} \frac{\mathbf{U}_{:,i}^T \mathbf{d}}{s_i} \mathbf{V}_{:,i} \tag{15}$$

The quantities $\frac{s_i^2}{(s_i^2+\alpha^2)}$ are called filter factors. They control the instabilities induced by the small s_i

Higher-order Tikhonov regularization

The Tikhonov term $||\mathbf{m}||_2$ can be written as $||\mathbf{Lm}||_2$ where $\mathbf{L} = \mathbf{I}$ or the zeroth order derivative operator. Instead of using the zeroth order regularization we can choose to the use the first, second, third...derivative:

$$\mathbf{L}_1 = \begin{bmatrix} -1 & 1 & & \\ & -1 & 1 & \\ & & -1 & 1 \\ & & & \end{bmatrix} \tag{16}$$

and the inverse problem becomes:

$$\min ||\mathbf{Gm\text{-}d}||_2 + \alpha^2 ||\mathbf{L}_p\mathbf{m}||_2 \tag{17}$$

and can be solved using the generalized singular value decomposition $(\mathbf{G}, \mathbf{L})$.

Contribution to the functional

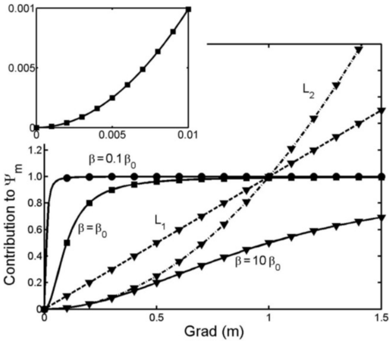
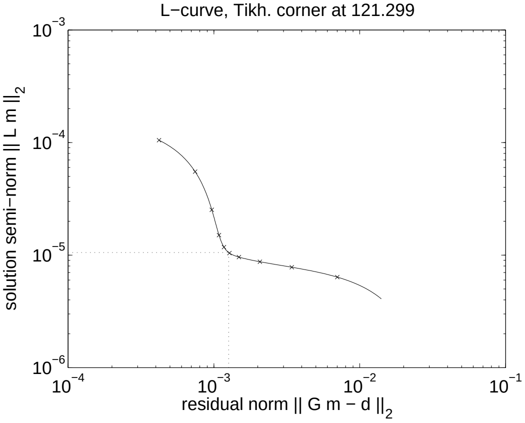
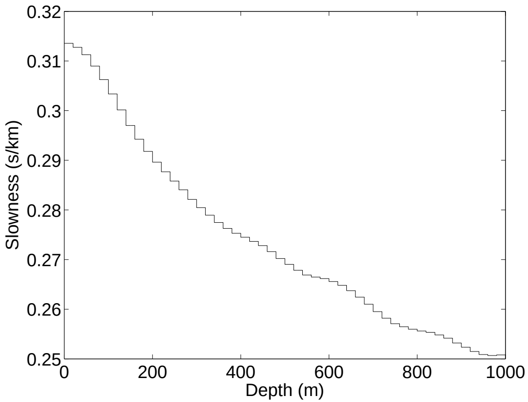


Figure: Contribution of the functional and β Blascheck et al., 2008

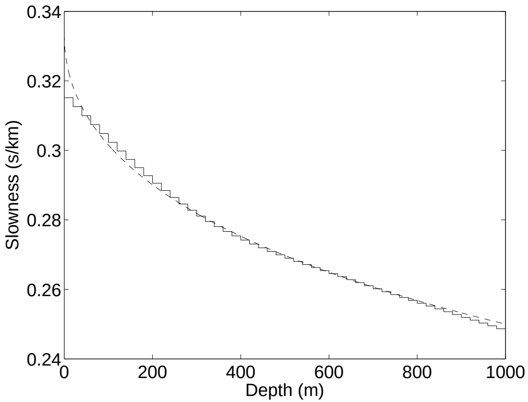
Using the first order Tikhonov solution with α^2 chosen according to the L-curve criteria:



Using the first order Tikhonov solution with α^2 chosen according to the L-curve criteria:



Using the second order Tikhonov solution with α^2 chosen according to the L-curve criteria:



Gauss-Newton, Levenberg-Marquardt and Occam

Gauss-Newton

Consider problems of the form $\mathbf{d} = g(\mathbf{m})$ where $\mathbf{m}$ is a n by 1 model parameter vector and $\mathbf{d}$ is a m by 1 data vector. As for the linear problem, we want to find the minimum of:

$$f(\mathbf{m}) = \sum_{i}^{m} \underbrace{(\frac{g(\mathbf{m})_i - d_i}{\sigma_i})^2}_{f_i(\mathbf{m})} \tag{18}$$

$$f(\mathbf{m}) = \sum_{i}^{m} f_i(\mathbf{m})^2 \tag{19}$$

To compute the increment on the model we had to solve:

$$\nabla f(\mathbf{m}^0) = -\frac{1}{2}\nabla^2 f(\mathbf{m}^0)\Delta\mathbf{m} \tag{20}$$

which becomes (re-introducing the iteration index k)

$$\mathbf{J}(\mathbf{m}^k)^T\mathbf{J}(\mathbf{m}^k)\Delta\mathbf{m}^k = -\mathbf{J}(\mathbf{m}^k)^T\mathbf{f}(\mathbf{m}^k) \tag{21}$$

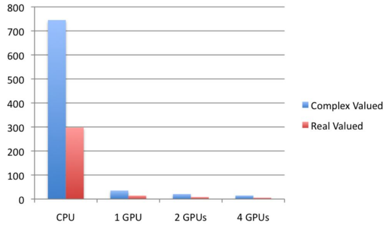
Where the $\mathbf{J}^T\mathbf{J}$ term approximates the Hessian $\nabla^2 f(\mathbf{m}^0)$. Note the similarity between this equation and the normal equations from the linear problem case:

$$\mathbf{G}^T\mathbf{G}\mathbf{m} = \mathbf{G}^T\mathbf{d} \tag{22}$$

Jacobian

The cost is mainly associated with the computation of the Jacobian:

- Analytical expression.
- Finite-differences.
- Adjoint state methods (Plessix, 2006).
- Mesh adaptation. (Rucker et al., 2006)
- Approximation or compression of **J** (Broyden, 1965; Li et al., 2003)
- GPU computation (Borsic et al., 2012)



Borsic et al., 2012

This method will fail if the product $\mathbf{J}(\mathbf{m}^k)^T\mathbf{J}(\mathbf{m}^k)$ us close to singularity. To ensure convergence, one can add a positive term on the diagonal of this product resulting in the **Levenberg-Marquardt** method:

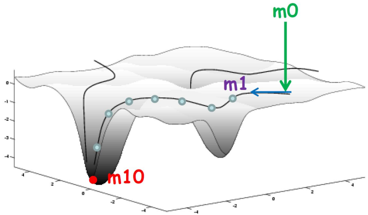
$$(\mathbf{J}(\mathbf{m}^k)^T\mathbf{J}(\mathbf{m}^k) + \lambda\mathbf{I})\Delta\mathbf{m}^k = -\mathbf{J}(\mathbf{m}^k)^T\mathbf{f}(\mathbf{m}^k) \qquad (23)$$

where λ must be adjusted in order to ensure convergence. Note the similarity with the Tikhonov regularized problem:

$$(\mathbf{G}^T\mathbf{G} + \alpha^2\mathbf{I})\mathbf{m} = \mathbf{G}^T\mathbf{d} \qquad (24)$$

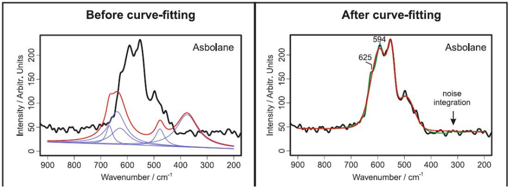
Levenberg-Marquardt method

1. Choose $\mathbf{m}^0$
2. Pick a λ value as small as possible that minimizes the misfit (downhill steps)
3. solve for $\Delta\mathbf{m}^k$
4. Let $\mathbf{m}^{k+1} = \mathbf{m} + \Delta\mathbf{m}^k$
5. Let $k = k + 1$
6. Check convergence criteria



Step 6 of the algorithm correspond to the establishment of the convergence of the solution. We can consider the three following stopping criteria:

- $\nabla f(\mathbf{m}^k) \approx 0$
- The successive models $\mathbf{m}^k$ do not evolve
- The residuals $f(\mathbf{m})$ do not change
- Am I trapped in a local minima ? (yes...) $\rightarrow$ hybrid methods combined with LM (accepts uphill steps)



Mule et al., 2016

Occam principle

In science, Occam razor is used as a heuristic (rule of thumb) to guide scientists in the development of theoretical models rather than as an arbiter between published models, in our case, it was translated as: "**In doubt, smooth**" or more formaly

$$\min ||\mathbf{Lm}||_2 \quad (25)$$

$$||\mathbf{g}(\mathbf{m}) - \mathbf{d}||_2 \leq \delta \quad (26)$$

or

$$\min ||\mathbf{g}(\mathbf{m}) - \mathbf{d}||_2^2 + \lambda ||\mathbf{Lm}||_2^2 \quad (27)$$

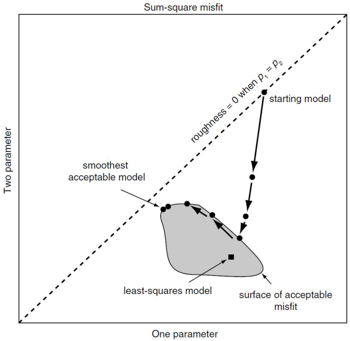
Letting

$$\hat{\mathbf{d}}(\mathbf{m}^k) = \mathbf{d} - \mathbf{g}(\mathbf{m}^k) + \mathbf{J}(\mathbf{m}^k)\mathbf{m}^k \quad (28)$$

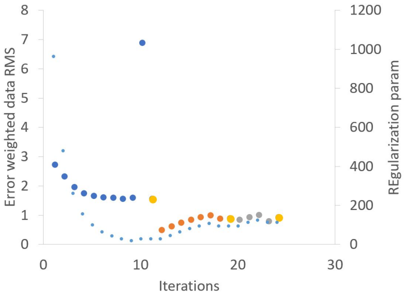
The Gauss-Newton update becomes:

$$\mathbf{m}^k + \Delta\mathbf{m}^k = (\mathbf{J}(\mathbf{m}^k)^T\mathbf{J}(\mathbf{m}^k) + \lambda^2\mathbf{L}^T\mathbf{L})^{-1}\mathbf{J}(\mathbf{m}^k)^T\hat{\mathbf{d}}(\mathbf{m}^k) \quad (29)$$

Occam algorithm: maximizing the constraint



Constable et al. (2015)

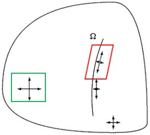


CRTomo (Kemna, 2000)

Regularization freedom

And the geophysicist replied: Which model do you want? (Constable et al., 2015)

Consider that we have access to borehole geological log $\rightarrow$ boundaries



Kaipio et al., 1999)

$$||W_d(\mathbf{d} - f(\mathbf{m})||^2 + \lambda(||W_m(\mathbf{m} - \mathbf{m}_0)||^2 + \alpha||\mathbf{m} - \mathbf{m}_0||^2)$$

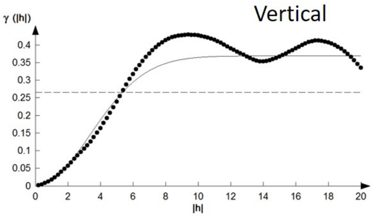
with

$$W_m^T W_m = \beta_t W_t^T W_t + \beta_n W_n^T W_n$$

We could also have a borehole geophysical log $\rightarrow$ variogram and reference values

$$||W_d(\mathbf{d} - f(\mathbf{m})||^2 + \lambda||C_m^{-0.5}(\mathbf{m} - \mathbf{m}_0)||^2$$

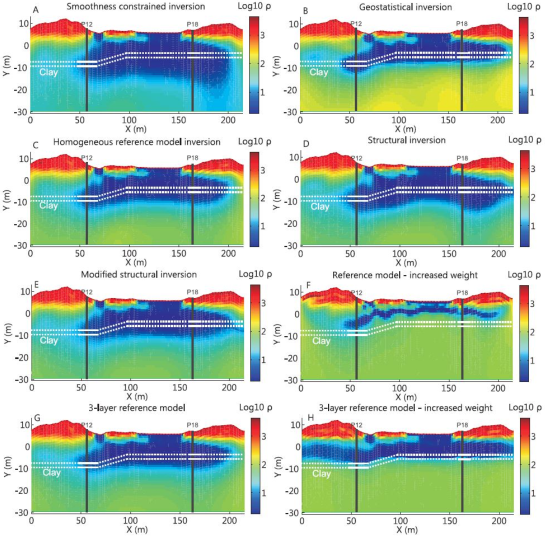
with C_m the covariance matrix obtained from a vertical variogram but assuming some horizontal range.



Chasseriau and Chouteau, 2003; Hermans et al., 2012

Least-squares
SVD
Tikhonov and the L-curve
Gauss-Newton, Levenberg-Marquardt and Occam
Regularization freedom
Image appraisal
Time-lapse dedicated regularization

Including borehole geological log



Caterina et al., 2014

Least-squares
SVD
Tikhonov and the L-curve
Gauss-Newton, Levenberg-Marquardt and Occam
Regularization freedom
Image appraisal
Time-lapse dedicated regularization

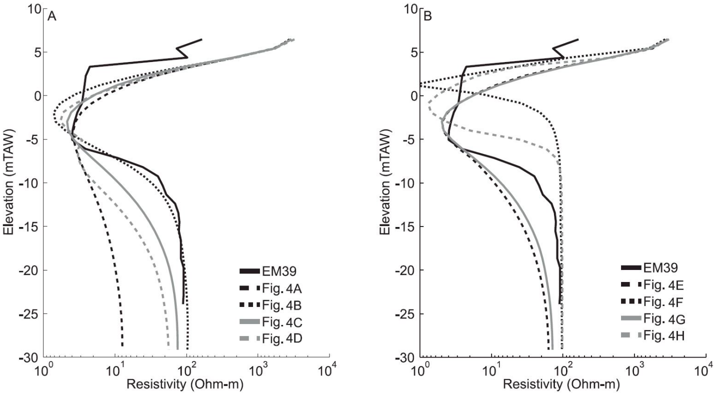


Figure: P18 borehole validation

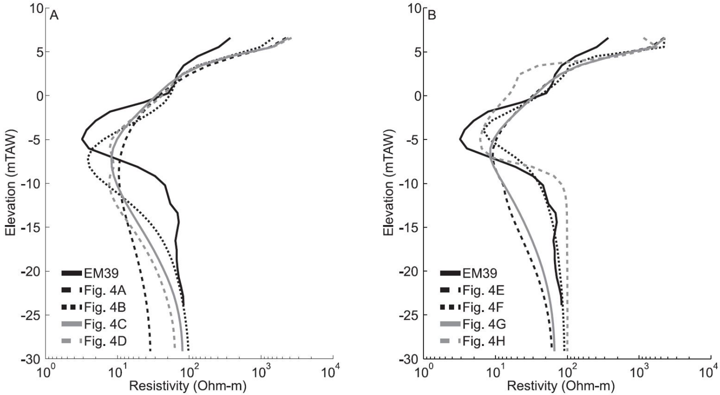


Figure: P12 borehole validation

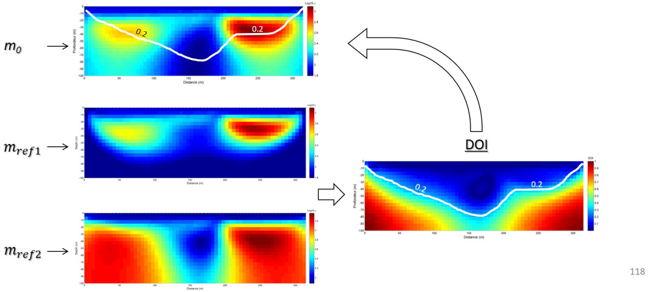
Image appraisal

DOI index

Least-squares
SVD
Tikhonov and the L-curve
Gauss-Newton, Levenberg-Marquardt and Occam
Regularization freedom
Image appraisal
Time-lapse dedicated regularization

To appraise the obtained image, it is often useful to use the DOI Index (depth of investigation, Oldenburg and Li, 1999):

$$\frac{\left|\mathbf{m}^1 - \mathbf{m}^2\right|}{\left|\mathbf{m}_{ref1} - \mathbf{m}_{ref2}\right|}$$



Caterina et al., 2013

Resolution and sensitivity

There are several commonly used appraisal tools that one can use to judge the quality of an inverse model (in addition to misfit criteria). One of the difficulty associated with them is to define a threshold.

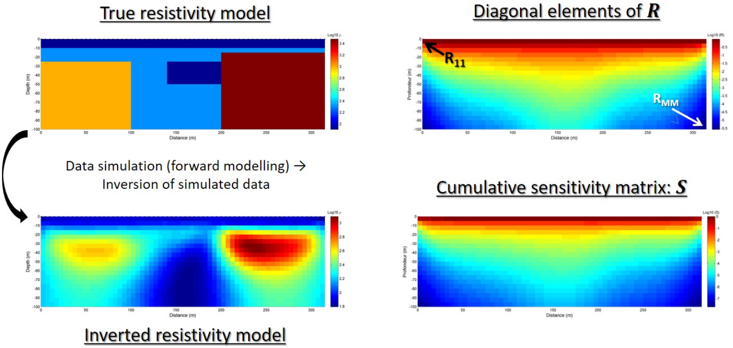
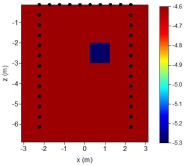


Figure: Resolution and cumulative sensitivity. Caterina et al., 2013

Uncertainty estimation

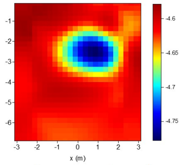


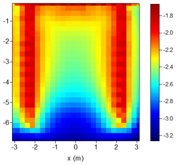
Linear data uncertainty

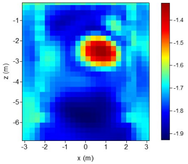
NL data uncertainty

$$\mathbf{C}_d^{-1} = \mathbf{W}^{\mathrm{T}}\mathbf{W}, \quad \mathbf{C}_{m0}^{-1} = \lambda \mathbf{R}^{\mathrm{T}}\mathbf{R} \quad (2)$$

$$\mathbf{C}_m^{(d)} = (\mathbf{J}^{\mathrm{T}}\mathbf{C}_d^{-1}\mathbf{J} + \mathbf{C}_{m0}^{-1})^{-1}\,\mathbf{J}^{\mathrm{T}}\mathbf{C}_d^{-1}\mathbf{J}\,(\mathbf{J}^{\mathrm{T}}\mathbf{C}_d^{-1}\mathbf{J} + \mathbf{C}_{m0}^{-1})^{-1} \quad (4)$$





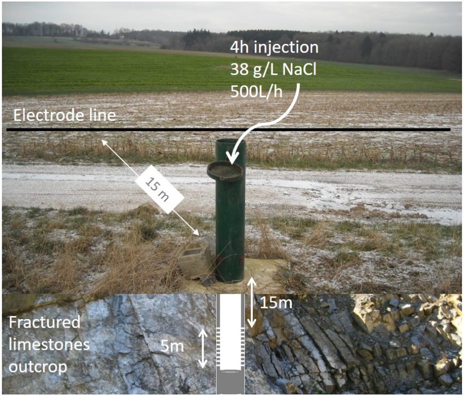


inverted
(mean of 100 realizations)

Figure: Comparing data error propagation. (Kemna et al., 2007; Gubbins, 2004)

Time-lapse dedicated regularization

Example of minimum gradient support



Robert et al., 2012

Consider the following regularization functional (minimum gradient support, Portniaguine and Zhdanov, 1999):

$$\int_V \frac{(\nabla m) \cdot \nabla m)}{(\nabla m) \cdot \nabla m) + \beta^2} dV$$

The choice of the β parameter controls the stability of the gradient $\rightarrow$ here we propose to select based on a line search on the time-lapse residuals (no other prior information required).

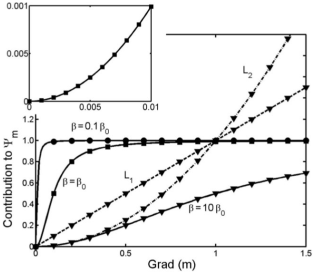


Figure: Contribution of the functional and β

Benchmarking the approach

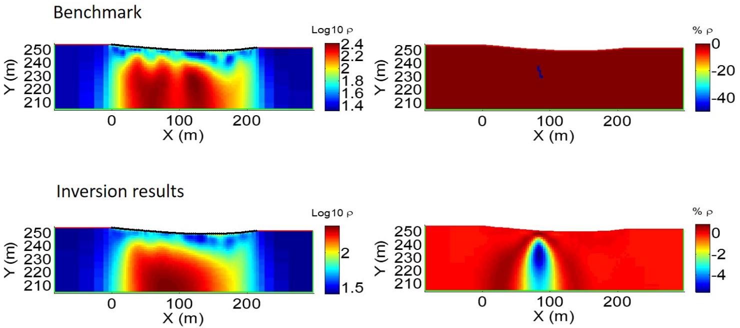
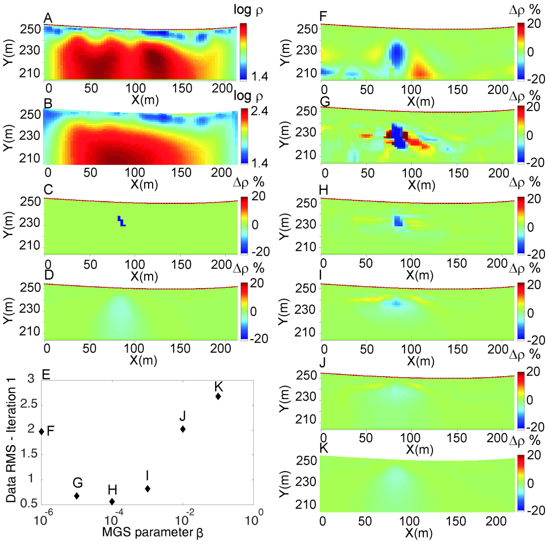


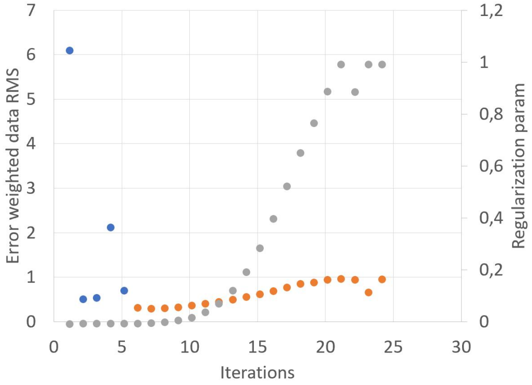
Figure: Smoothness constraint inversion

Nguyen et al., 2016

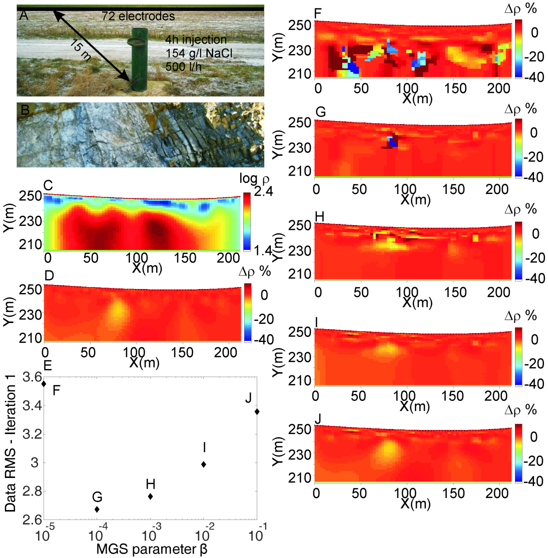
Benchmarking the approach



Optimization behaviour



Testing the approach with field data



Hydrogeophysical inversion

A seawater intrusion model was inverted using ERT data and the Levengerg-Marquardt algorithm in PEST (Doherty., 2015).

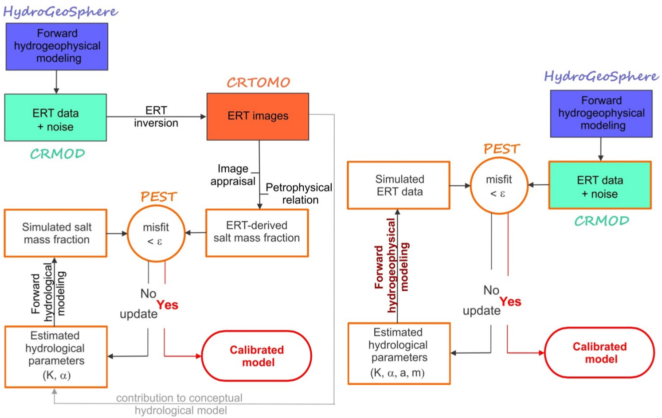


Figure: Sequential and joint inversion approaches (Beaujean et al., 2014)

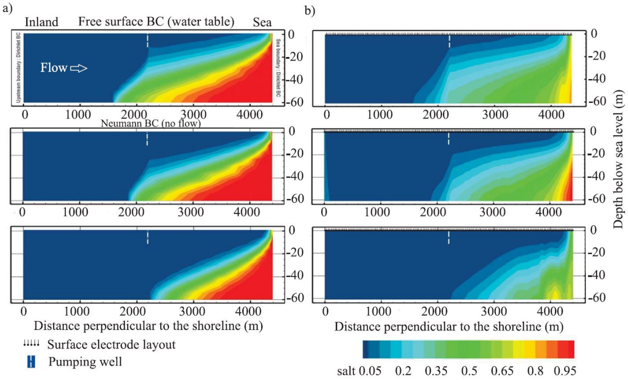


Figure: Homomgeneous model and corresponding ERT. (Beaujean et al., 2014)

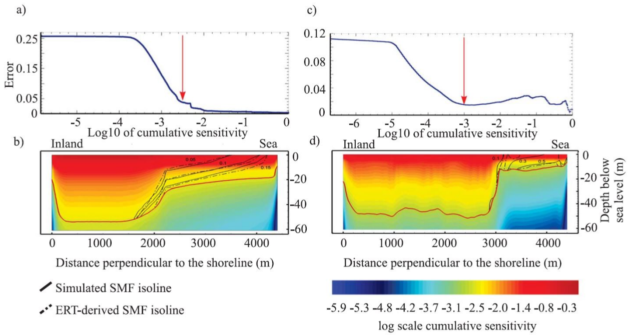


Figure: Filtering ERT before inversion. (Beaujean et al., 2014)

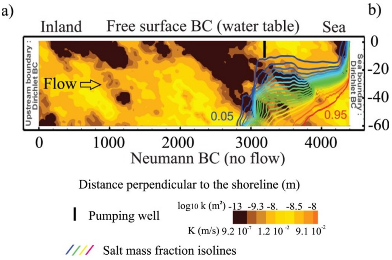


Figure: Simuation

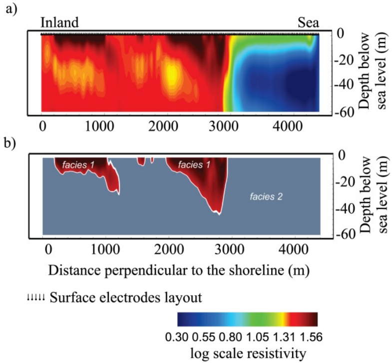


Figure: Conceptualisation

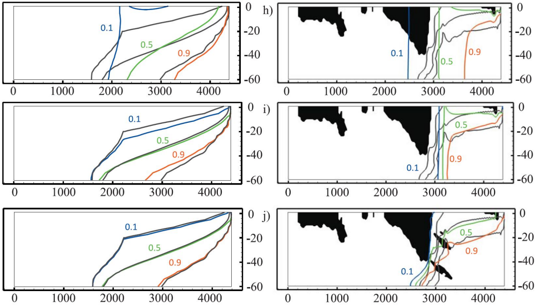


Figure: Inversion results in terms of salt mass fraction isolines using ERT images

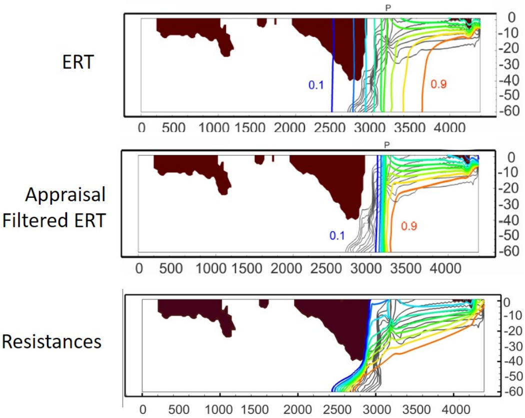


Figure: Inversion results in terms of salt mass fraction isolines using resistances (bottom image)

Questions to take home

- Linear or non-linear ?
- Data uncertainty estimation ?
- Objective function definition ?
- How costly is the CPU on the Jacobian ?
- What type of constraint would be best suited ?
- How do I tune all those λ, α, L_p, W_m, m_0 ?
- How far can I trust my inverted and conceptual model ?
- What type of uncertainty am I accounting for mainly ?
- Should I explore more models ?
- Is Belgium really going to win the world cup ?

THANKS !