

Exploring non-linear inversions: a 1D magnetotelluric example

Seogi Kang, Lindsey J. Heagy, Rowan Cockett, and Douglas W. Oldenburg

July 19, 2017

At some point in many geophysical workflows, an inversion is a necessary step for answering the geoscientific question at hand: whether it is recovering a reflectivity series from a seismic trace in a deconvolution problem, finding a susceptibility model from magnetic data, or recovering conductivity from an electromagnetic survey. This is particularly true when working with data sets where it may not even be clear how to plot the data: 3D direct current resistivity and induced polarization surveys (it is not necessarily clear how to organize data into a pseudosection) or multi-component data, such as electromagnetic data (we can measure three spatial components of electric and/or magnetic fields through time over a range of frequencies). Inversion is a tool for translating these data into a model we can interpret. The goal of the inversion is to find a “model”, some description of the earths physical properties, that is consistent with both the data and geologic knowledge.

In a general inverse problem, we start from a forward problem, of the form $\mathcal{F}[\mathbf{m}] = \mathbf{d}$, where \mathcal{F} is the forward operator (the mathematical description of the physics/problem), \mathbf{d} is our data, and \mathbf{m} is our earth model (an array of numbers that describes the physical properties of the earth). Matt Hall kicked off the discussion of inversions in the Leading Edge in his *Linear Inversion* tutorial (Hall, 2016). He walked through how to solve the classic linear inverse problem, where the forward simulation takes the form $\mathcal{F}[\mathbf{m}] = \mathbf{G}\mathbf{m} = \mathbf{d}$. The example he demonstrated is a deconvolution problem, in that case, \mathbf{G} is a convolution matrix, \mathbf{m} is the reflectivity series and \mathbf{d} is a seismic trace. He introduced the concepts of an underdetermined problem, motivated the need for regularization, formulated the inversion in terms of an optimization problem and solved the linear inverse problem (in true polyglot fashion, using Python, Lua, Julia, and R!). In this tutorial, we will pick up from there and explore a *nonlinear* forward problem, of the form $\mathcal{F}[\mathbf{m}] = \mathbf{d}$; in this case, our forward operator is a function of the model. In the attached notebooks, we use SimPEG (<http://simpeg.xyz>) for the implementation in Python of the physics simulations, optimization, and structure necessary to perform an inversion (Cockett et al., 2015).

1 Magnetotellurics

We will explore the 1D magnetotelluric (MT) survey technique, which is a natural source electromagnetic method. In magnetotellurics, plane-wave source-fields are generated by solar wind (giving us low frequency signals < 1 Hz) and lightning strikes worldwide (giving us higher frequency signals > 1 Hz). In magnetotellurics, the model \mathbf{m} is a description of the earths electrical conductivity σ and $\mathcal{F}[\mathbf{m}]$ solves Maxwells equations — giving the electric field \vec{E} and the magnetic field \vec{H} — for

a plane wave source. In the continuous world, Maxwells equations are:

$$\begin{aligned}\nabla \times \vec{E} + i\omega\mu\vec{H} &= 0 \\ \nabla \times \vec{H} - \sigma\vec{E} &= 0\end{aligned}\tag{1}$$

with boundary conditions

where μ is magnetic permeability, generally taken to be the permeability of free space μ_0 , and $\omega = 2\pi f$ is the angular frequency (f is frequency in Hz). When we discretize Maxwell's equations so that they can be solved numerically (see the first notebook), we obtain a matrix system of the form

$$\mathbf{A}(\mathbf{m})\mathbf{u} = \mathbf{b}\tag{2}$$

where $\mathbf{A}(\mathbf{m})$ is a matrix capturing the physics, \mathbf{u} is vector of the electric and magnetic fields everywhere in our simulation domain, and \mathbf{b} includes the boundary conditions that describe the plane wave source. This problem is nonlinear because this matrix \mathbf{A} depends on the conductivity model. When we solve equation 2 for \mathbf{u} , we obtain the electric and magnetic fields everywhere on our mesh. Our data consist of samples of the electric and magnetic fields where we have receivers. For the MT problem, the measured data, \mathbf{d} , are impedances; for the 1D problem, the impedance Z_{xy} at a single frequency is given by

$$Z_{xy} = -\frac{E_x}{H_y}.\tag{3}$$

Note that the impedance is a complex number, consisting of real and imaginary parts. Impedance is a non-intuitive quantity; often, we instead consider apparent resistivity¹ ρ_a and phase ψ , given by

$$\rho_a = \frac{1}{\mu_0\omega}|Z_{xy}|^2, \quad \psi = \tan^{-1}\left(\frac{\text{Im}(Z_{xy})}{\text{Re}(Z_{xy})}\right).\tag{4}$$

For an earth that is a half-space, the apparent resistivity equals the true resistivity, and the phase is 45° . When we implement the computation of the data, we define a method that (a) selects the values of the electric and magnetic fields at the surface of the earth from the fields that we computed everywhere in our domain \mathbf{u} , and (b) computes their ratio to provide us with impedance data, that is: $\mathbf{P}(\mathbf{u}) = \mathbf{d}$.

In summary, to implement the forward simulation for the MT problem ($\mathcal{F}[\mathbf{m}] = \mathbf{d}$), we break it into two steps:

1. Solve $\mathbf{A}(\mathbf{m})\mathbf{u} = \mathbf{b}$
2. Compute the impedance data $\mathbf{d} = \mathbf{P}(\mathbf{u})$

In the first notebook, we provide details on how each step is performed using a finite difference approach, and if you are looking for more on numerical discretization, we wrote a tutorial on finite volume methods (Cockett et al., 2016).

The inversion aims to solve $\mathcal{F}^{-1}[\mathbf{d}]$ for a model. Just as in the linear problem, we require regularization in order to select a model from the infinitely many that can fit the data. Before we tackle this *ill posed* inverse problem, let's explore an example of nonuniqueness: how can different models give us the same data?

¹Note that resistivity is the inverse of conductivity $\rho = 1/\sigma$

2 Go forwards

A classic example that demonstrates the non-uniqueness of MT data is the equivalence of the conductivity-thickness product (conductance) of a thin layer. If we start with a layer that has a conductivity of σ , halve its thickness and double its conductivity, the resulting data will be similar. In Figure 1, we show apparent resistivity and phase data for five models, each of which has the same conductance. In all of the simulations, the data shows a decrease in apparent resistivity and an increase in phase starting at $\sim 10\text{Hz}$. Thus, in all of the data we have evidence of a conductive layer, and the frequency range at which it appears is an indicator of the depth of the layer (you can explore by changing the `depth` variable in the model setup of the second notebook). However, all scenarios produce similar data. Even with a small amount of noise, we cannot expect an inversion code to separate the conductivity and thickness of a conductive unit without incorporating additional information. When setting up the inverse problem and defining regularization (next up!), it is important to realize that the choices we make there will influence the character of the model we recover, as the data alone do not provide us with a unique model.

3 Go backwards

There are many models that can fit the data, so we need a means of *regularizing* our inversion so that we can select a single, reasonable model from the (infinitely!) many that agree with the data. We will come back to how the regularization ϕ_m is defined, but for now, consider it as a measure of how “reasonable” the model is based upon our prior knowledge about the earth. Formally, we pose the inversion as an optimization problem:

$$\begin{aligned} & \underset{\mathbf{m}}{\text{minimize}} && \phi_m(\mathbf{m}) \\ & \text{subject to} && \mathcal{F}[\mathbf{m}] = \mathbf{d}_{\text{obs}} \end{aligned} \tag{5}$$

Essentially we are saying “find the model \mathbf{m} that best fits the assumptions we are making in the regularization $\phi_m(\mathbf{m})$, and that agrees with the observed data \mathbf{d}_{obs} .” In practice, our data are noisy, so there is no sense in fitting them exactly. Rather, we pose the optimization problem as a trade-off between fitting the data and fitting the regularization, so our inverse problem can be stated as:

$$\underset{\mathbf{m}}{\text{minimize}} \quad \phi(\mathbf{m}) = \phi_d(\mathbf{m}) + \beta\phi_m(\mathbf{m}) \tag{6}$$

where $\phi_d(\mathbf{m})$ is the data misfit, a measure of how far our simulated data are from the observed data; $\phi_m(\mathbf{m})$ is the regularization; and β is a trade-off parameter that weights the relative importance of the data misfit and regularization in the optimization. A larger β says that we want our model to do a good job minimizing the regularization, while a smaller β turns down the volume of the regularization and says that fitting the data is more important in the inversion.

The data misfit is often taken to be a weighted ℓ_2 norm:

$$\phi_d(\mathbf{m}) = \frac{1}{2} \|\mathbf{W}_d(\mathcal{F}(\mathbf{m}) - \mathbf{d}_{\text{obs}})\|_2^2 \tag{7}$$

where \mathbf{W}_d captures the noise model (typically it is a diagonal matrix containing the standard deviation of each datum).

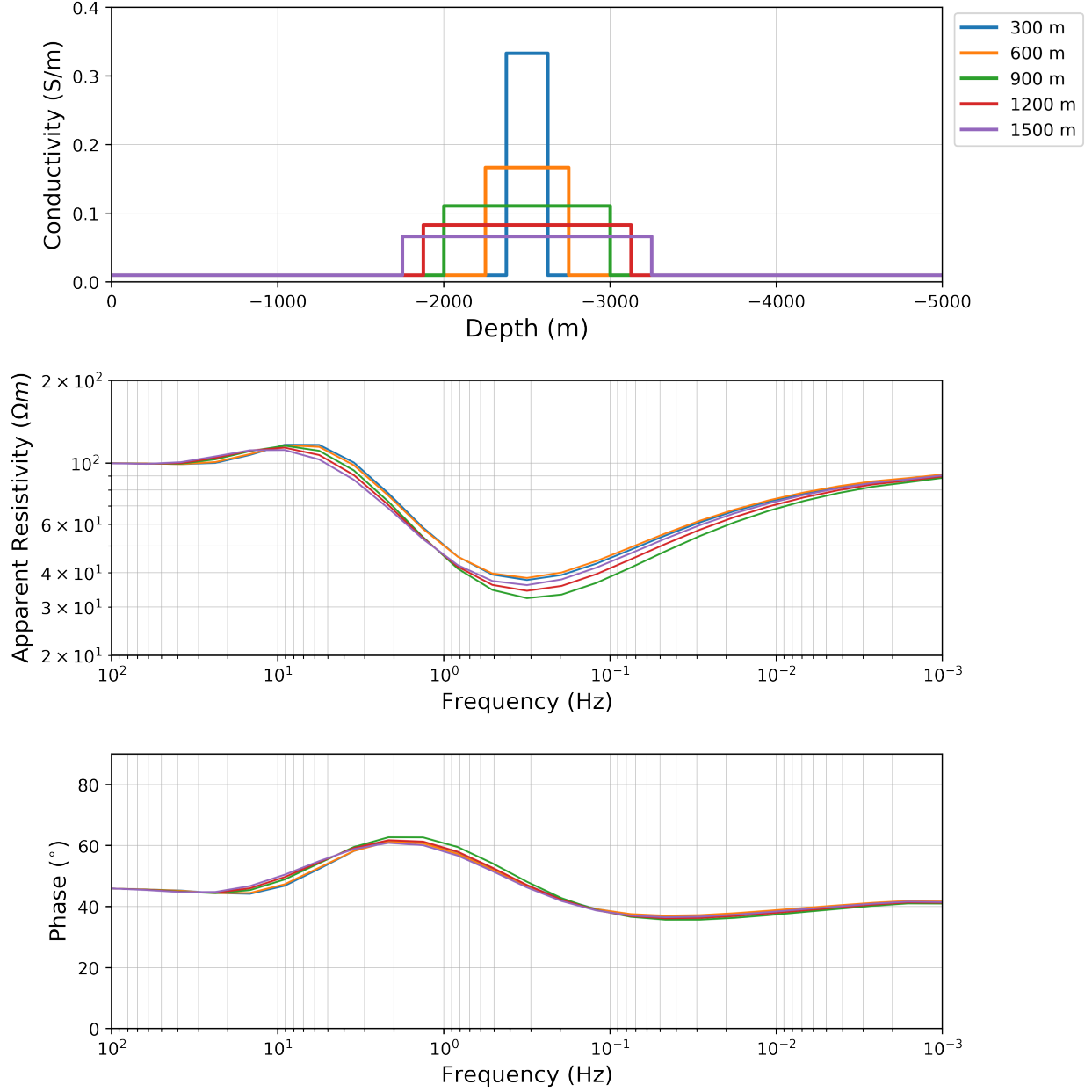


Figure 1: Magnetotelluric responses from five models, each having an equivalent conductivity-thickness product for the conductive layer. (a) Conductivity models, (b) apparent resistivity (ρ_a), and (c) phase (ϕ).

The regularization is one place where *a priori* information about the geologic setting can be brought in. There are a variety of regularization functionals that can be chosen, but one of the most widely used is Tikhonov regularization, which again uses ℓ_2 norms:

$$\phi_m(\mathbf{m}) = \frac{1}{2} (\alpha_s \|\mathbf{W}_s(\mathbf{m} - \mathbf{m}_{\text{ref}})\|_2^2 + \alpha_z \|\mathbf{W}_z(\mathbf{m})\|_2^2) \quad (8)$$

The first term is often referred to as the "smallness" as it measures the "size" of the model (in the ℓ_2 sense). The matrix \mathbf{W}_s is generally taken to be a diagonal matrix that may contain information about the length scales of the model or be used to weight the relative importance of various parameters in the model. The scalar α_s weights the relative importance of this term in the regularization. Notice that we include a reference model, \mathbf{m}_{ref} . Often this is defined as a constant value, but if more information is known about the background, that can be used to construct a more intricate reference model. Here, we will not delve too far into how the reference model impacts the recovered results, but you are encouraged to change \mathbf{m}_{ref} in the notebooks and investigate its impact.

The second term is often referred to as the "smoothness". The matrix \mathbf{W}_z approximates the derivative of the model with respect to depth, and is hence a measure of how "smooth" the model is. The term α_z weights the relative importance of smoothness in the regularization.

From this setup, we see that there are quite a number of choices to make: defining uncertainties on the data (\mathbf{W}_d), selecting a reference model (\mathbf{m}_{ref}), choosing the importance of smallness and smoothness (α_s and α_z), and selecting a trade-off parameter (β). Lets start by assuming a known noise model, fix α_s and α_z , and explore the impact of the trade-off parameter β . Our forward problem depends upon the electrical conductivity. For the inverse problem, however, we are free to use any function of the conductivity as a parameter. The electrical conductivity of earth materials varies by many orders of magnitude and is strictly positive. Thus it is advantageous to use $\log(\sigma)$ as the model in the inverse problem. For a nonlinear problem, we also have the additional choice of the initial model \mathbf{m}_0 at which to start the inversion. Although we will not discuss the choice of \mathbf{m}_0 , you are encouraged to change the initial model in the notebooks and examine the impact it makes because it can be significant!

3.1 The β knob

If the noise is Gaussian, then the sum of squares (our data misfit) is a Chi-squared distribution, which has an expected value of N_{data} (in our case, we divide this by two to match our definition of ϕ_d). Thus, the ideal choice of β is one that gives us $\phi_d^* \approx \frac{1}{2}N_{\text{data}}$. To demonstrate the effect of β , we consider a five layer model, originally shown in Whittall and Oldenburg (1992), and will demonstrate inversions when we achieve the target misfit, underfit the data and overfit fit the data. The conductivity model used is the solid black line in Figure 2a. For these inversions we fix the regularization parameters to $\alpha_s = 10^{-2}$, $\alpha_z = 1$ and set $\mathbf{m}_{\text{ref}} = 10^{-2}S/m$, and the initial model, $\mathbf{m}_0 = \mathbf{m}_{\text{ref}}$ (feel free to change them in the notebook!). We start the inversion with a large β and decrease its value to plot the trade-off or Tikhonov curve (Figure 2b). In blue, we show the inversion that is stopped when the data misfit approximately equals the target misfit (the star in Figure 2b). Figures 2c and 2d show the data as apparent resistivity and phase, which is a visualization of our complex-valued impedance data. The blue line in Figure 2a shows the recovered model, which identifies the general structure and conductivity values of the five layers. In this case, we are employing a smooth regularization, thus we expect to recover smoothly varying structures.

If we instead choose a larger β , reducing the contribution of the data misfit to the objective function, we underfit the data, as is shown in orange in Figure 2. Although we still see evidence of two conductive structures, we do not recover their amplitudes, and do a poor job resolving the location and widths of the conductive layers (if you had to pick the top of the first layer, where should it be?). Examining the plots in 2c and d, there is more insight about the subsurface conductivity that can be learned by pushing the inversion to extract more from the data.

On the other extreme, we can choose a very small β , and try to fit all of the details in the

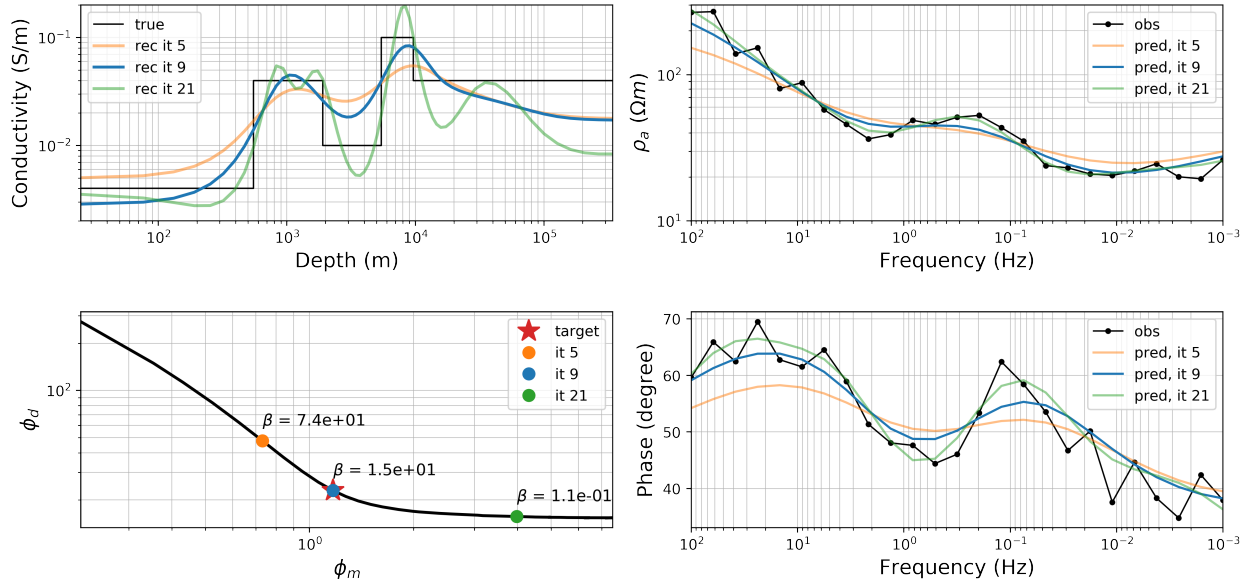


Figure 2: Inversions which fit the data (blue), underfit the data (orange), and overfit the data (green). (a) True (black) and recovered electrical conductivity models. (b) Tikhonov curve showing the trade off between the misfit and regularization, target misfit (red star), and achieved misfit corresponding to each of the inversion results shown (c) observed (black) and predicted apparent resistivity data, (d) observed (black) and predicted phase data.

data. Doing this, we obtain the results shown in green in Figure 2. When we push the inversion to fit the (noisy!) data very closely, we end up fitting the noise. In order to do this, conductivity contrasts are exaggerated and oscillatory and erroneous conductivity structures are introduced in the inversion.

3.2 The α knobs

For the inversions shown in Figure 2, we prescribed the values the α_s , α_z . What impact do they have on the character of the model we recover?

In Figure 3, we compare two inversions with different regularization parameters: (1) a “smooth” inversion (blue line), with $\alpha_s = 10^{-5}$ and $\alpha_z = 1$, and (2) a “small” inversion (orange line): with $\alpha_s = 1$ and $\alpha_z = 10^{-5}$. In both, β was chosen so that a desired target misfit was achieved. The smooth inversion penalizes large gradients; the resulting model has two smooth peaks. Note that we smooth over the resistive third layer, over-estimating its conductivity. The “small” inversion instead favors models that are close to the reference model; this model has more structure. The resistivity of the first layer matches well (the conductivity of the first layer is equivalent to our reference model) and the conductivity of the third layer is closer to its true value, but additional oscillatory structures are introduced at depth. In the third notebook, you can explore the impact of these parameters yourself!

In practice, these parameters are often determined by experimentation; strategies such as examining length scales are often successfully adopted (see page 38 in Oldenburg and Li (2005)).

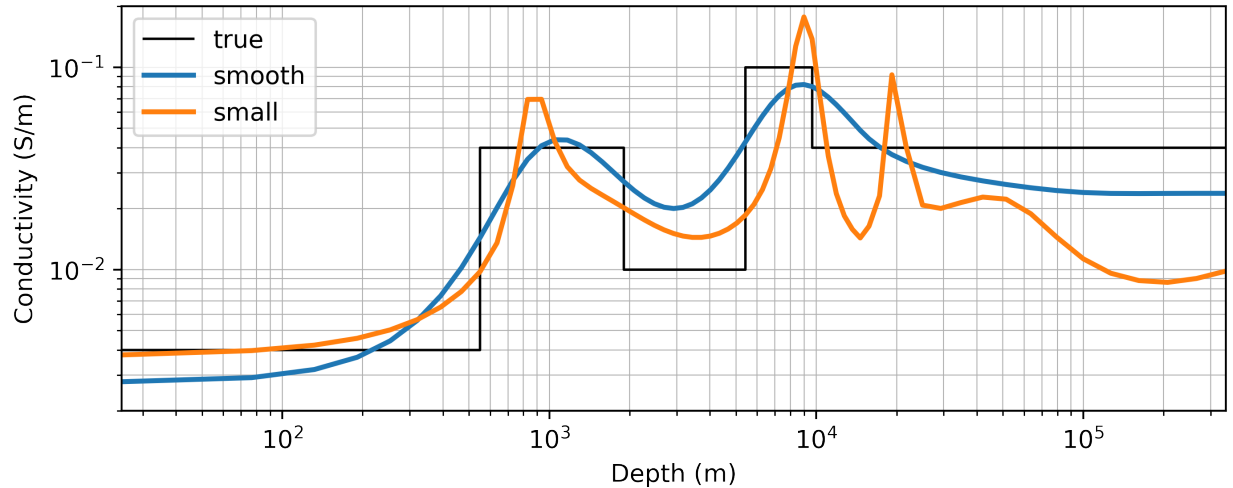


Figure 3: Comparing the use of Smooth regularization versus Small regularization in the inversion.

Changing the relative values of α_s and α_z is one way to bring in a priori information. If we know very little, often starting with a smooth inversion is a good option; this penalizes structure (high gradients) while showing general trends. If more structure is expected, or a reliable reference model can be built from additional data such as physical property measurements, well logs, or additional geophysical/geologic data, then the influence of the smallness term may be increased. There are a few other ways to bring in additional a-priori information. If we are expecting a more “blocky” model, we can choose a different norm (such as an ℓ_1 norm), or if we have structural constraints, we can introduce other weighting structures (e.g. on the smoothness); these are knobs for another tutorial and there is discussion in Oldenburg and Li (2005).

4 Summary

In this tutorial, we have introduced the forward simulation for magnetotellurics and explored a few aspects of the inverse problem. Prior to jumping into an inversion, it is important to know the limitations of the survey and data, and what you can and cannot resolve, even if there is no noise! Forward modelling is a powerful tool for setting realistic expectations of an inversion.

To set up and solve the inverse problem, we posed the inversion as an optimization problem that searches for a model of the earth that minimizes an objective function consisting of a data misfit and a regularization term. There are many choices to be made in defining the various elements of the inverse problem, including how to assign uncertainties, selecting a trade-off parameter, defining the regularization function, and choosing initial and reference models. In this tutorial we explored two of the knobs: (1) the trade-off parameter and (2) the relative importance of smallness and smoothness contributions in Tikhonov regularization. The interactive notebooks that are provided allow you to change parameters and experiment with their impact.

References

- Cockett, R., L. J. Heagy, and D. W. Oldenburg, 2016, Pixels and their neighbors: Finite volume: The Leading Edge, **35**, 703–706.
- Cockett, R., S. Kang, L. J. Heagy, A. Pidlisecky, and D. W. Oldenburg, 2015, SimPEG: An open source framework for simulation and gradient based parameter estimation in geophysical applications: Computers & Geosciences, **85**, 142–154.
- Hall, M., 2016, Linear inversion: The Leading Edge, **35**, 1085–1087.
- Oldenburg, D. D. W., and Y. Li, 2005, Inversion for applied geophysics: A tutorial: Investigations in geophysics, **13**, 1–85.
- Whittall, K., and D. Oldenburg, 1992, Inversion of magnetotelluric data for a one-dimensional conductivity: Society of Exploration Geophysicists.