A practical strategy to interrogate resolution, uncertainty and value of information in geophysical inverse problem

Partha S. Routh

Upstream Research Company, ExxonMobil, Houston, TX, formerly Dept. of Geosciences, Boise State University, ID, USA.

Abstract

A practical strategy to compute uncertainty in geophysical images is presented in this paper. The methodology is applicable to large scale inverse problems with the goal of obtaining target oriented uncertainty estimates. The first step is to determine the regions of data influence or illumination domains in the model space that are sensitive to the available data in presence of noise. In the second step the image is interpreted to determine the region of interest. The interpretation of the region of interest depends on the type of prior information (regularization operator and constraints) to solve the ill-posed inverse/imaging problem. In the third step, the uncertainty of the average value in the region of interest is determined by solving a minimization-maximization optimization problem to determine the bounds on the average value. If the upper and lower bound are determined for several scales, the functional form of the bounds resembles a funnel function introduced by Oldenburg, 1983. One of the advantages of this approach is that the uncertainty measure depends on the scale. Thus the overall strategy is to reduce the model space using illumination criteria and obtain the uncertainty estimates in the reduced domain. I reduce the model domain by using resolution measure of point spread functions to determine the regions of data influence. Within this region, funnel functions (upper and lower bounds) determine the uncertainty of the model parameter estimates. The min-max problem to determine the bounds are formulated with a constrained optimization procedure and it is solved using interior-point technique. The choice of interior-point method provides a general framework to incorporate different types of constraints including physical bounds on the model within an iterative nonlinear inversion algorithm. Synthetic and field examples from crosswell tomography demonstrate the strategy discussed in this paper.

Introduction

Practical interpretation of inverted models requires that we examine the resolution and uncertainty in the model. Quantifying uncertainty of model parameter estimates leads to a better understanding of the spatial/temporal variability in the subsurface. With a more accurate and complete description of the subsurface, a more reliable cost/benefit analysis is possible that eventually help to make better, more effective policy decisions.

Geophysical data are typically inadequate to probe the subsurface at many scales- temporal and spatial band-limitations are inherent in the geophysical data. Inadequate sampling of the subsurface with geophysical acquisition produces a non-trivial null-space in the data that leads to inadequate sampling of model space. This leads to well known ill-posed inverse problem commonly encountered in most geophysical inversion/imaging applications. For tractable forward problems with fast computation time, stochastic sampling of the model space using global optimization procedure can also provide measures of uncertainty (Sen and Stoffa, 1995). For iterative linearized methods, several types of stabilizers in the form of regularization operators (or commonly known as prior information) can be used to fill this null-space. However, it is often unclear what part of the inverted model is a consequence of the data acquired by the geophysical experiment and what parts came from the prior information? A clear distinction of the information content from these two sources is often hard to determine from the inverted model/image. The inverted model is a mixture of the two quantities i.e. information from inadequate noisy data and the prior information.

In the next section on methodology, I discuss the procedure to reduce the model space and estimate the uncertainty. I provide some theoretical background of the approaches to solve the problem. The uncertainties of the model parameter estimates are determined within the region of data influence using funnel functions. Following this section, I provide 1D synthetic example to demonstrate the strategy and quantify the value of information with uncertainty reduction. Synthetic and field examples with crosswell tomography demonstrate the application to generic geophysical inverse problem.

Methodology

After obtaining an inverted model, if we ask the question about uncertainty, it is often unclear and confusing- what we mean by "uncertainty"? In many cases, some parts of the model are not affected by data, i.e., there are regions in the model beyond the illumination zone of the geophysical data. This has practical consequence for uncertainty estimation. For example, by changing the prior information in the inverse problem we can obtain a completely different set of models that may not be geologically meaningful. However, all of those models perfectly satisfy the geophysical data to the same extent as the "meaningful" models. Therefore a qualified question can be *what is the uncertainty among the "meaningful class" of models that satisfies the data?* The definition of "meaningful" class is subjective and can depend on the geological scenarios or intuition of the subsurface. I recognize that this "meaningful" class is problem dependent.

For example depending on the geological scenario we can ask for a *class of smooth models* or a *class of models with sharp boundaries* or a *class of layered models* or a *class of training models/images depicting possible geological scenarios* or a combination of these classes subject to fitting the geophysical data. Effectively, this "meaningful" class can be defined by choosing the "appropriate" prior for the application. Choosing the meaningful class is the first step towards reducing the model space and can lead to practical utility of uncertainty measures. Ignoring this step would make the goal of determining uncertainty a practically unrealistic problem. Theoretically, we can formulate a solution either within a deterministic or probabilistic framework by ignoring this step, but the practical utility of those solutions for real world problem can be complicated and confusing with the limited scope for risk analysis.

Having chosen the prior that suits our objective the next task is to determine what regions of the model are sensitive to the data acquired by the geophysical experiment? There are several ways to determine this measure depending on the complexity of the problem. Oldenburg and Li (1999) proposed *depth of investigation* (DOI) method to determine the regions that are sensitive to data for a general nonlinear problem. Oldenborger et. al (2007) developed volume of investigation method for large scale 3D problems. Resolution measures such as *point spread functions* (PSF) computed from *model resolution matrix* is another approach to determine the parts of the model sensitive to geophysical data (Backus and Gilbert, 1970; Alumbaugh and Newman, 2000; Routh et. al, 2005; Routh and Miller, 2006; Miller and Routh, 2007; Oldenborger and Routh, 2009). Miller and Routh (2007) showed the link between two approaches by deriving the depth of investigation index as a function of model resolution matrix. Their key result was region of data influence (RDI) obtained from PSF's are slightly different from DOI measures. Regardless, these resolution measures provide an interpretive tool to determine the parts of the inverted model/ image that are sensitive to geophysical data. As expected, the regions that are not sensitive to the data are dominantly

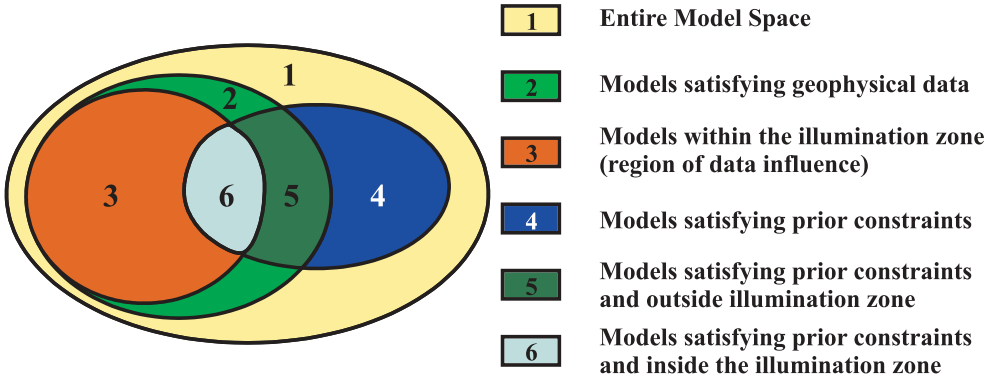

Fig.1 Model space partition showing the regions that satisfy geophysical data, region of data influence i.e. illumination zone, models that satisfy prior constraints. The region shown in light blue is the intersection of models within the illumination zone that satisfying prior constraints. The uncertainty in the model parameters within this region (light blue) are computed using funnel function approach.

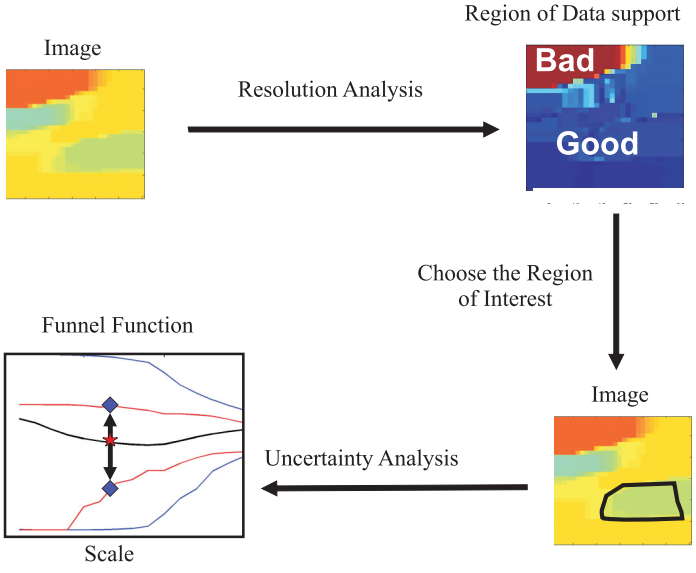

Fig.2 The practical strategy to compute scale dependent uncertainty using funnel functions. The flow shows the various steps to determine the upper and lower bounds of the average model estimates within the region of interest.

biased towards the prior information. *Therefore it is meaningless to interrogate uncertainty of the model parameters in the region that are not sensitive to geophysical data*. This is an important practical step since it eliminates the parts of the model where the uncertainty is unbounded. The conceptual model space shown in Figure 1 shows the basic idea behind model reduction procedure. The region we are interested need to be within the illumination zone of the geophysical data and satisfy the prior constraints.

The first step of choosing the prior is followed by the second step of eliminating the parts of the model that are not in the resolving capacity of the geophysical data. These two steps provide us with a reduced model domain that are sensitive to the data and honors the imposed prior to produce models from a "meaningful" class. I will denote this reduced domain as the *region of data influence (RDI)*. Having obtained the RDI the task of determining the uncertainty of the model parameters within RDI is a "better" posed problem and perhaps more practical. Figure 2 shows the overall approach of the proposed strategy.

Within the reduced model domain (i.e. region of data influence), the model parameters are affected by data noise, errors in the forward model, model parameterization and regularization operators. Therefore uncertainties associated with the value of a model at a single point of interest are still unbounded. This difficulty motivated Backus and Gilbert (1970a,b,c) (BG) to construct averages of the model over some finite region surrounding the point of interest. The finite region can be cell thickness or layer thickness or more generally the *region of interest* (ROI). The averaging functions were designed to be, in some sense, close to a delta function. The BG averaging kernel at a particular point in the model is obtained by taking linear combination of the data kernels, given by

$$A(r, r_0) = \sum_{j=1}^{N} \alpha_j(r_0)\, g_j(r) \,, \qquad (1)$$

where $g_j(r)$ is the data kernel function and $\alpha_j(r_0)$ are the coefficients to be determined so that the averaging function is $A(r,r_0)$ close to a delta function i.e. $\min_{\alpha} \|A(r, r_0) - \delta(r - r_0)\|^2$. The appeal of the BG average is that it has a unique value since the inner product of the averaging kernel with any acceptable model produces the same average.

$$\langle m(r_0) \rangle = (A(r, r_0), \hat{m}(r)) \qquad (2)$$

Therefore, to interpret the BG average we need to consider the average value $\langle m(r_0) \rangle$ and the associated averaging function $A(r,r_0)$. Unfortunately, if the averaging kernel is not peaked at the location of interest and/or has side lobes, then the particular BG average may not be very useful. Thus, the interpretability of the BG average value depends on the nature of the averaging kernel which is *not* in our control, since it is a linear combination of the data kernels shown in equation (1).

Instead of producing the BG unique average where the averaging function is a linear combination of the data kernels, we can *alternatively formulate the problem by choosing an averaging function or window (such as box car function or any linear functional) and determine the average value within this window of choice*. By doing so, the scale information is explicitly introduced. This is the key advantage since we are free to choose this window- i.e. the desired scale where we like to determine the average value in the region of interest (ROI). The disadvantage is that the average value obtained with the window is no longer unique; since the chosen window (i.e. averaging function) cannot be constructed using a finite number of data kernels- a limitation inherent in all geophysical acquisition. However, the average value of the model over the chosen window (or ROI) can be constrained by the data and other available information. As demonstrated by Oldenburg (1983), the upper and lower bounds of the average value can be obtained by solving a maximization and minimization problem for the average value subject to the available constraints. For a chosen ROI, these bounds on the model average are of practical value since they provide a measure of uncertainty for the chosen scale. As the region of interest increases, the uncertainty associated with the average value decreases giving a funnel like shape. Thus the basic idea in the funnel function approach is to determine the upper and lower bounds of the average values in a specified model region.

With this analysis, *scale dependent upper and lower bounds of the average value of the model in ROI can be obtained by solving maximization and minimization problem*. Following Routh and Oldenburg (2007), I calculate these bounds using an interior point method and impose physical bounds of the model in the optimization procedure.

Linearized Resolution Analysis

Consider a linear inverse problem, $d^{obs} = Gm + \varepsilon$. The linear problem has been studied in detail in previous literature (Backus and Gilbert, 1970; Oldenburg, 1983; Snieder, 1990; Parker, 1994). Here I present necessary equations to illustrate the various components of the resolution measure. In a linear problem the objective function to be minimized can be expressed as a combination of data misfit and model norm given by,

$$\phi = \left\| W_d \left(d^{obs} - Gm \right) \right\|^2 + \beta \left\| W_m \left(m - m_0 \right) \right\|^2 \qquad (3)$$

where $G \in R^{N \times M}$ is forward modeling operator, $W_d \in R^{N \times N}$ is the data weighting matrix, $W_m \in R^{M \times M}$ is the model weighting matrix, m_0 is the reference model and β is the regularization parameter. Denoting the Hessian by

$H = \left(G^T W_d^T W_d G + \beta W_m^T W_m\right)$, the recovered model is given by

$$\hat{m} = H^{-1} G^T W_d^T W_d G\, m + H^{-1} G^T W_d^T W_d\, \varepsilon + \beta\, H^{-1} W_m^T W_m\, m_0 \quad (4)$$

Equation (4) can be represented in a compact form $\hat{m} = Rm + \eta + \gamma$, where R is the resolution matrix, η is the noise contribution term and γ is the bias due to reference model, given by

$$\begin{aligned} R &= \left(G^T W_d^T W_d G + \beta W_m^T W_m\right)^{-1} G^T W_d^T W_d G \\ \eta &= H^{-1} G^T W_d^T W_d\, \varepsilon \\ \gamma &= \beta H^{-1} W_m^T W_m\, m_0 \end{aligned} \quad (5)$$

R provides a linear mapping between the estimated model parameters, $\hat{m}$ and the true model parameters, m. For a nonlinear problem $d^{obs} = Fm + \varepsilon$, the estimated solution for the model perturbation is given by

$$\delta\hat{m} = H^{-1} G^T W_d^T W_d G\, \delta m + H^{-1} G^T W_d^T W_d \left(Q(\delta m) + \varepsilon\right) + \beta\, H^{-1} W_m^T W_m\, \delta m_0 \quad (6)$$

where G is the sensitivity matrix, and $Q(\delta m)$ is the higher order terms typically neglected in the linearized analysis. When the higher order terms are not negligible, it introduces a bias into the solution of the estimated model in addition to the bias introduced by the reference model. For linearized analysis we typically neglect the higher order terms. A key aspect of the resolution operator is that it depends on the physics, forward modeling errors, acquisition geometry, parameterization of the model, regularization operator or prior information, regularization parameter and the noise in the data.

Region of Data Influence Using Point Spread Function

Point spread functions (PSF) are columns of the resolution matrix. A PSF describes how a delta-like perturbation in the model is broadcasted in the inverted result. If a particular PSF is wide and/or has large side lobes, then the corresponding model is poorly resolved at that location and the resolution 'width' is broader than the cell dimension. This introduces the idea that spread of these PSF's can be used as a measure to assess the resolving capability of the data and the prior information. *Note that the prior information can effectively be treated as data in the inverse problem, so the PSF and hence the spread will be influenced by the prior information*. For each cell, this spread number can be used as a diagnostic measure to compare the resolving capability of the inverse operator in different regions of the model. Oldenborger and Routh (2009) provide details on how these PSF's are computed for large scale 3D inverse problems and their implication on model resolution. Among possible choices for the spread I chose the following definition (Routh and Miller, 2006)

$$S(r_k) = \left[\frac{\int W(r, r_k)\left(p(r, r_k) - \Delta(r, r_k)\right)^2 dr}{\alpha + \int p^2(r, r_k)\, dr} \right]^{1/2}, \quad (7)$$

where $W(r, r_k)$ is a distance weighting matrix that increases the spread value when side lobes exist away from the region of interest. The weights in the matrix can be chosen using a distance rule given by $W(r, r_k) = 1 + \left|r - r_k\right|^{\lambda}$ where $\lambda \geq 1$ (Routh et. al, 2005; Miller and Routh, 2007). For examples presented in this paper, I choose $\lambda = 1$, and $\Delta(r, r_k) = 1$ when $r = r_k$ and $\Delta(r, r_k) = 0$ for $r \neq r_k$, i.e., discrete delta function. The point spread function $p(r, r_k)$ is for the cell whose center is at r_k, and α is a small threshold parameter chosen so that the denominator in the spread equation is not equal to zero when the energy of the PSF is zero. This can occur in the regions where there is no illumination. If the PSF peak is shifted from the region of interest, the chosen spread criterion will generate large spread values. Another alternative way to compute RDI is use DOI (Oldenburg and Li, 1999) or volume of investigation method (Oldenborger and Routh, 2009). Miller and Routh, (2007) showed that an approximate relation exists between the rows of the resolution matrix and the DOI or volume of investigation measures.

Funnel Function Method: Uncertainty and Value of Information

Funnel functions introduced by Oldenburg (1983) are obtained by solving a minimization and maximization problem for the average model value. Oldenburg (1983) used linear programming (LP) approach to solve the problem since LP is a natural framework to solve a linear, constrained optimization problem. Following Routh and Oldenburg (2007) I use a nonlinear, constrained optimization method to solve the minimization and maximization problem using the interior point method (Nocedal and Wright, 2006). Thus for each region of interest, I obtain a lower and upper bound of the average value. Narrower the bound, the smaller is the uncertainty of the average value in the chosen region. The interior point formulation allows incorporation of other types of prior information that maybe available from well-logs or other sources of geological information. Therefore the bounds determined by the funnel functions have the ability to honor different types of information i.e. geophysical data with errors, constraints on the model and physical bounds on the model. It provides the mechanism to introduce constraints in a progressive manner to examine how uncertainty is reduced by adding new information- i.e. it provides a metric to measure the *value of information*. Although we can impose various types of constraints on the model, one of the goals in funnel function analysis is to use minimal constraints, and thereby

determine the resolving power of the data. However, the framework is general and allows the integration of other types of constraints.

Denoting the average value of model in a region of interest (ROI) as $\langle m \rangle_{ROI}$ the minimization problem (Routh and Oldenburg, 2007) can be written as:

$$\min \langle m \rangle_{ROI} = \frac{1}{V_{ROI}} \int_{V_{ROI}} m \, dv$$
$$\phi_d = \left\| W_d \left(d^{obs} - g(m) \right) \right\|^2 = \phi_d^* \qquad (8)$$
$$subject \ \ to \ m^{MIN} \le m \le m^{MAX}$$

where d^{obs} are the observed data, m^{min}, m^{max} are the physical bounds on the model, g is the forward modeling operator and w_d is the data weight matrix that contains the standard deviations of data errors and ϕ_d^* is the target data misfit. Similarly the maximization problem can be written as:

$$\max \langle m \rangle_{ROI} = \frac{1}{V_{ROI}} \int_{V_{ROI}} m \, dv$$
$$subject \ to \ \phi_d = \left\| W_d \left(d^{obs} - g(m) \right) \right\|^2 = \phi_d^* \qquad (9)$$
$$subject \ to \ m^{MIN} \le m \le m^{MAX}$$

The maximization problem can be transformed to a minimization problem where the reciprocal of the model in the ROI is minimized. Since equation (8) and (9) are solved using a nonlinear optimization method, this transformation allows us to use the same minimization algorithm with slight modifications. The transformation is not arbitrary since the reciprocal of many physical properties have a physical meaning (e.g., conductivity-resistivity; velocity-slowness etc.). However, if the model values approach zero then a small threshold parameter can be added to the model before taking the reciprocal. Thus the maximization problem in equation (9) is transferred to a minimization problem (Routh and Oldenburg, 2007), given by

$$\max \langle m \rangle_{ROI} = \max \frac{1}{V_{ROI}} \int_{V_{ROI}} m \, dv = \min \frac{1}{V_{ROI}} \int_{V_{ROI}} (1/m) dv. \qquad (10)$$

The minimization in equation (8) and maximization problem in equation (9) is solved using interior point method where we represent the discrete form of the integral in equation (8) by

$$\min \langle m \rangle_{ROI} = \min \frac{1}{V_{ROI}} \int_{V_{ROI}} m \, dv = \min S^T m. \qquad (11)$$

S is a diagonal screen vector with non-zero entries in the ROI and m is a vector. Note that choice of window is

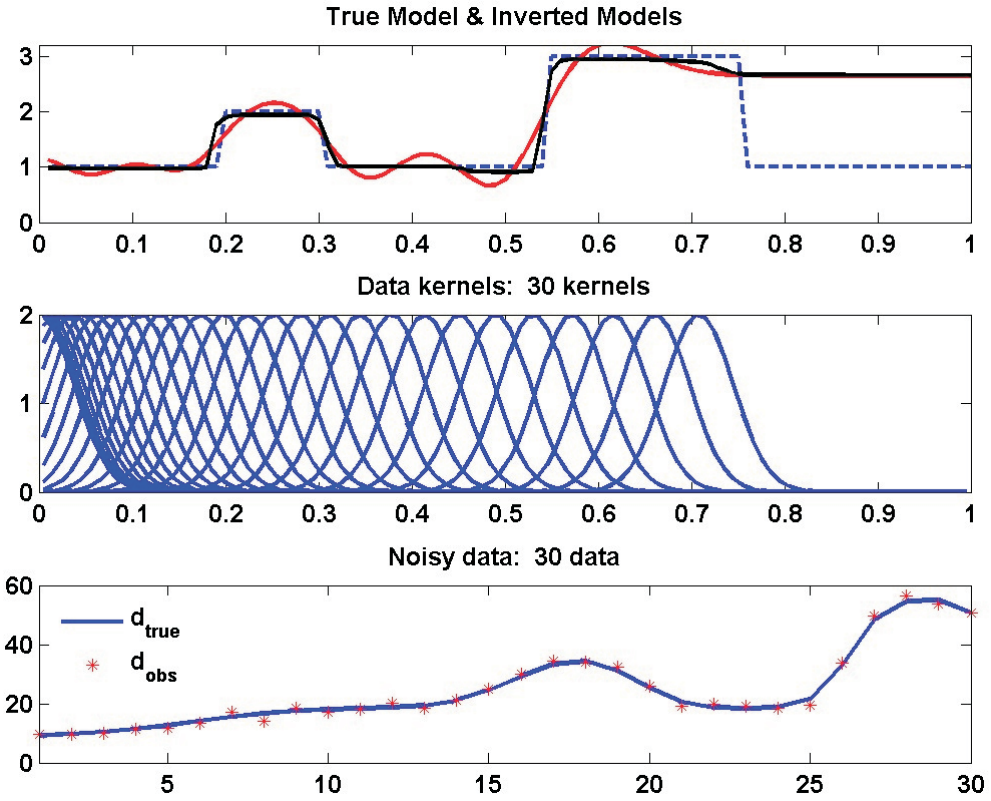

Fig.3 Top panel shows the true model (blue dotted) and inverted models using a smooth regularization (red) and a total variation regularization (black). The kernels are plotted in the middle panel and and data in the bottom panel. The x-axis in the top and middle panel is depth and in the bottom panel are index for data points.

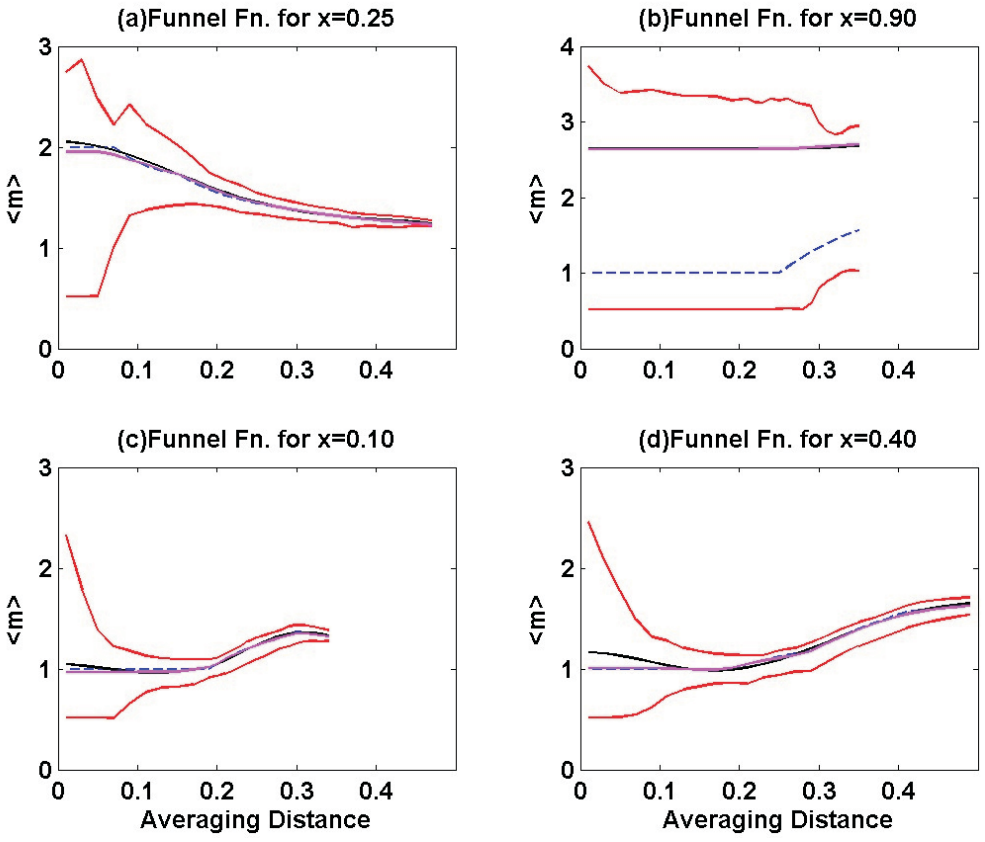

Fig.4 Funnel functions for the 1D example. The dotted blue curve is true average, the black curve is the average from smooth model inversion and purple curve is the average of TV model. The top and bottom red curves in each panel are the upper and lower bound of the model. Panel (a) is funnel functions for x=0.25 and panel (b) for x=0.9.

completely left to user and is a part of the interpretive step. In general, one can choose any linear functional of model as the window of choice. In this paper I choose the screen vector such that $S_k = \Delta v_k / V_{ROI}$ *for* $k \in ROI$ and $S_k = 0$, *for* $k \notin ROI$, k is the cell index of the discretized model. The min-max optimization can be mathematically represented by

$$\min\phi = \beta S^T m + \frac{1}{2}\left\| W_d \left(d^{obs} - g(m) \right) \right\|^2 + \phi_B \qquad (12)$$

$$\min\phi = \beta S^T \left(\frac{1}{m} \right) + \frac{1}{2}\left\| W_d \left(d^{obs} - g(m) \right) \right\|^2 + \phi_B \qquad (13)$$

where $\phi_B = -\lambda \sum_{k=1}^{M} \left(1 - \frac{m_k}{m_k^{MAX}} \right) - \lambda \sum_{k=1}^{M} \left(\frac{m_k - m_k^{MIN}}{m_k^{MAX}} \right)$

The mixed norm problem in equation (12) and (13) is solved using iterative re-weighted least squares (IRLS) method (Farquharson and Oldenburg, 1998; Routh at. al, 2003). At each iteration, the model is updated by solving for a model perturbation using a line search procedure. The regularization parameter β and the barrier parameter is reduced using cooling procedure (Routh, 1999; Routh et., al, 1998; 2005).

Value of Information from uncertainties : 1D Synthetic Example

I consider a 1D example where the kernels are Gaussian blurring function and the model is a blocky function shown in Figure 3. The kernels do not extend beyond x=0.8, therefore we lack resolution at depth. We invert the data using a smooth regularization, shown in red and total variation (TV) regularization, shown in black (Figure 3). As expected, the model obtained using TV recovers sharp jumps compared to the smooth model inversion.

To compute the funnel function we consider a ROI around x=0.25. Unit box-cars centered about this point are used as the averaging window. Choosing the bounds to compute funnel function is a crucial part, here we choose the bounds by inspecting the model values from the inversion given by $0.5 \leq m \leq 4$. Figure 4(a) shows the funnel function centered at x=0.25. As the length of averaging window increases the uncertainty, that is the difference between the upper and lower bounds in the average value, is reduced.

Figure 4(a) shows that the true average (blue), smooth model average (black) and TV model average (purple)

have very similar values. This is because the region around x=0.25 is well resolved by the data kernels. Beyond the averaging distance of 0.07 the bounds on the average value for x=0.25 is greater than unity. This is expected since the block has amplitude of 2.0. For x=0.1 or x=0.4 we note that the average value is close to unity as expected. This clearly demonstrates that funnel functions can provide distinct information about the average value in different regions of model. And the uncertainty in the average value decrease as the averaging window becomes larger. This indicates that resolving power of the data increases as the window (or ROI) gets larger. Next I examine funnel functions in the region where the resolving power of data is poor (x=0.9). Comparing the averages between the true and the inverted models clearly shows that the inversion averages do not agree with true averages (Figure 4(b)). This is because the inverted models in Figure 3 depart significantly from the true model at depth- a consequence of the regularization operator. The funnel functions for x=0.9 shows that the upper and lower bounds stay close to the physical bounds implying that resolving power of data is extremely poor. When the averaging window length increases to 0.25 there is a decrease in the uncertainty of the average value.

To examine how the new information leads to better estimates of model average (reduction of uncertainty), I add more kernels at depth in the region of no-coverage zone in Figure 3 shown in Figure 5. The new funnel functions at x=0.9 with the additional kernels show significant reduction in uncertainty. Comparing Figure 4(b) and Figure 6, I can quantify the reduction of uncertainty due to the addition of new information. This is useful in a practical sense since it provides a quantitative measure of *value of information*.

Tomography Example

I consider a synthetic crosswell example to illustrate the analysis with 2D data. 400 data were generated using a slowness model with 2160 cells shown in Figure 7(a). The sources were placed in the left side of the model and receivers on the right. Since sources were placed from depth of 10m the top left part of the model is not illuminated by the rays. The inverted model using total variation regularization is shown in Figure 7(b). The details of the algorithm is given in Routh et. al (2007). The region of data influence is shown in Figure 8. In Figure 8 the region in blue indicates the region where the model is sensitive to the data i.e. better resolved region.

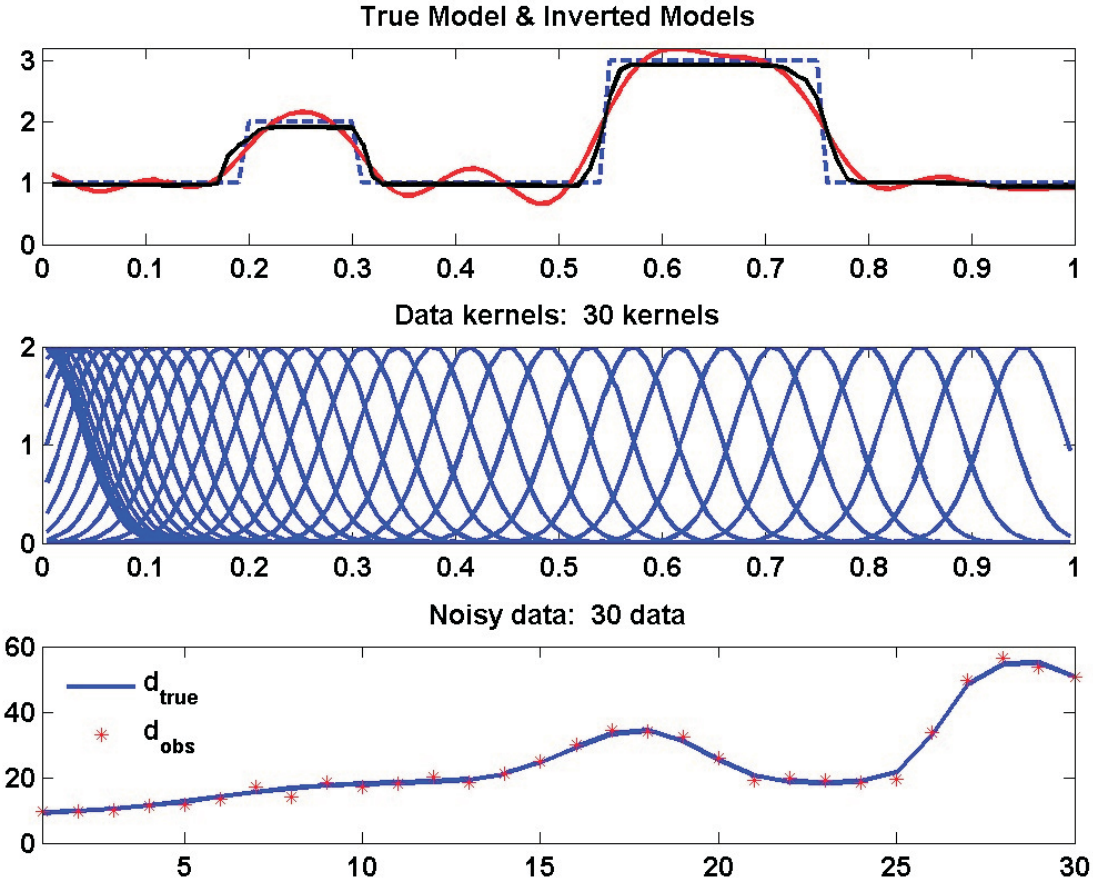

Fig.5 Example with more data kernels to examine the value of information of the data. Top panel shows the true model (blue dotted) and inverted models using a smooth regularization (red) and a total variation regularization (black). The kernels are plotted in the middle panel and and data in the bottom panel. The x-axis in the top and middle panel is depth and in the bottom panel are index for data points.

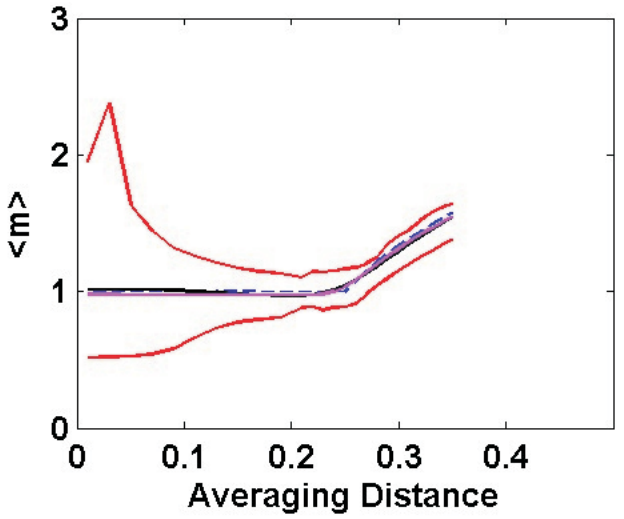

Fig.6 Funnel functions at x=0.9 for the 1D example with 5 extra kernels at depths in addition to kernels in Figure 2.

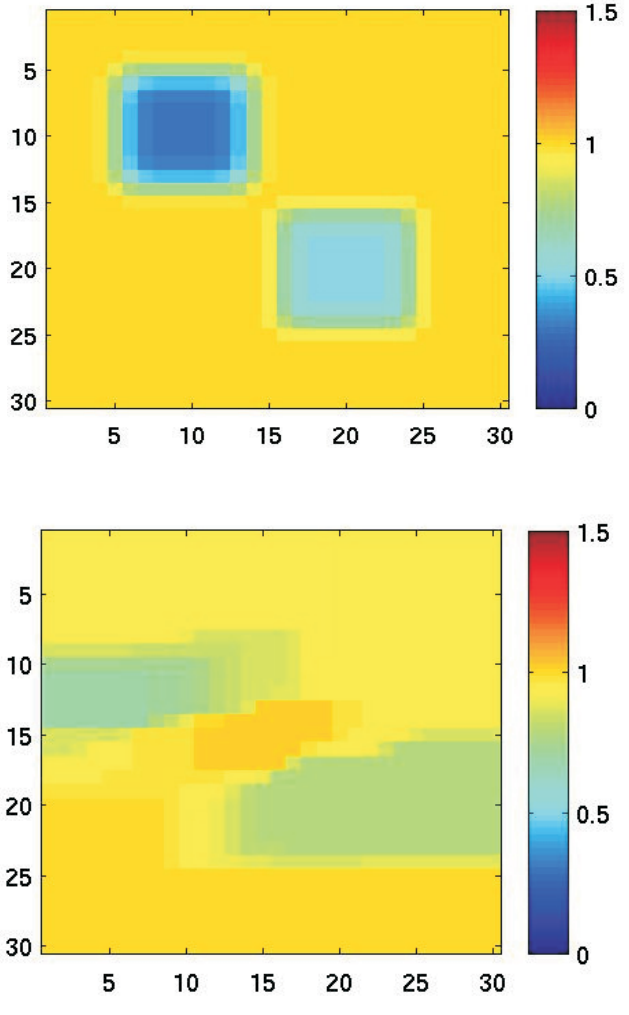

Fig.7 (a) The synthetic velocity model with sources in the left side and receivers on the right. Sources start from 10m depth where as the receivers are placed at equal interval from surface to 30m depth. (b) Inverted model using curved ray tomography with total variation regularization. Velocity units are in km/sec.

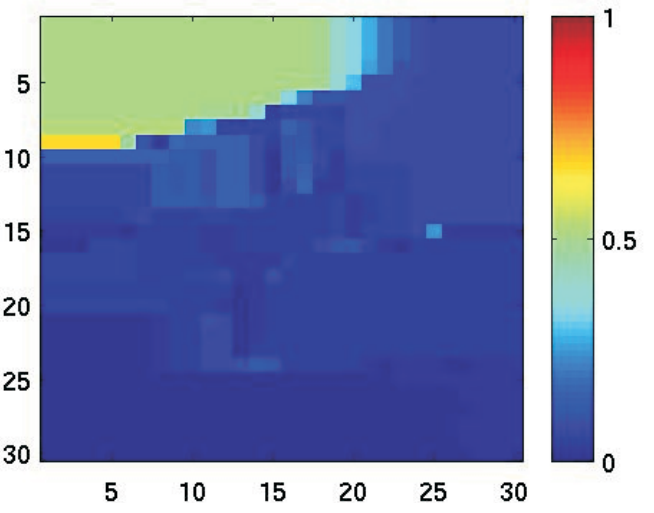

Fig.8 Region of data influence (RDI) computed by applying the spread criteria to point spread functions. The region in blue indicates the region illuminated by the data and region in green are not affected by the data.

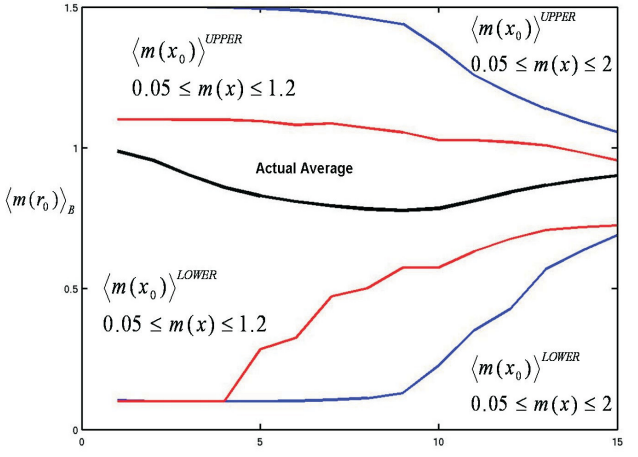

Fig.9 Funnel functions computed for the synthetic tomographic model. Two sets of upper and lower bounds are computed using different physical bounds on the model parameters. The x-axis indicates the increasing scale index. The scale of the ROI in Figure 8 is located at index value of 10 in this figure. The actual model average is shown in solid black line.

Having obtained the RDI the next step is to interrogate the uncertainty of the velocity estimates. Funnel function analysis requires that we choose a window in the model. For this synthetic example I consider several windows with increasing size centered on x=20 and z=20. For each window I solve the maximization and minimization problem. The min-max optimization problem is solved by imposing physical bounds on the velocity. In Figure 9 the funnel functions for two such bounds are presented. The black line indicates the true average value of the velocity for each chosen window (or scale). The red curve indicates the funnel function with tighter bounds $0.05 \le m \le 1.2$ and the blue curve are the funnel function with wider bounds $0.05 \le m \le 2.0$. The funnel functions indicate that as the scale of window

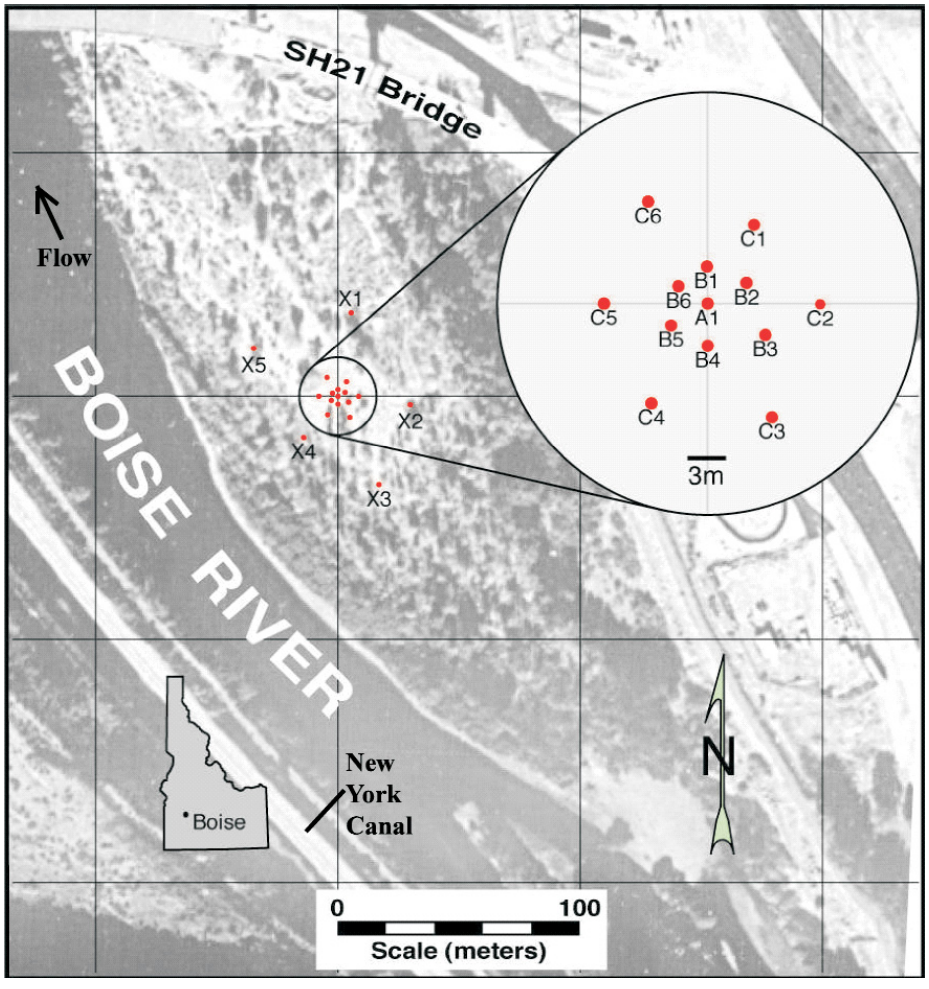

Fig.10 Boise Hydrogeophysical Research Site (BHRS) location along the Boise River near Boise, Idaho (Barrash and Reboulet, 2004). The inset shows the relative locations of the inner 13 boreholes at the site.

increases the upper and lower bounds of the average value for the particular window are narrower. Thus the uncertainty given by the upper and lower bound are scale dependent.

Field example from Nearsurface Geophysical experiment

Next I consider crosswell radar traveltime data from Boise Hydrogeophysical Research Site (BHRS). The BHRS is a multidisciplinary field laboratory situated in a sand and gravel bar adjacent to the Boise River about 15 kilometers southeast of downtown Boise, Idaho (Figure 10), and is designed for testing and developing geophysical and hydrogeologic methods. The BHRS consists of a heterogeneous unconsolidated alluvial aquifer composed of unaltered coarse and fine-grained sediments underlain by a red clay aquitard (Barrash and Reboulet, 2004). There are 18 boreholes in this site, which are cased with 10-centimeter (cm) inside diameter screened PVC pipe, and were drilled to an average depth of about 19 m and penetrate the underlying clay aquitard (Barrash et. al, 2006). I consider one tomogram panel acquired in the B wells. To invert the data I assume standard deviation of errors as 1% of the maximum data amplitude. BHRS is a layered sand and silt depositional environment. The model obtained by using TV regularization better indicates layering compared to the smooth inversion model in Figure 11(a). I consider the layer at z=15m and compute the funnel function (Figure12) at the location indicated by star in 11 (b). Physical bounds with $1 \le m \le 20$ ns/m are imposed during the min-max optimization. The ROI in this example corresponds to an area of 30 m around the location indicated by star in Figure 11 (b). The funnel function suggests that the bounds for this ROI are between $8 \le m_{ROI} \le 13$. The bound determined in this manner has practical utility. For example, if the model is interpreted in terms of lithologic units then the geophysical bounds computed using funnel function for those units can be converted to petrophysical bounds via petrophysical transformation. Consequently these computed petrophysical bounds provide a measure of uncertainty for flow simulation problem. Although flow simulation is not carried out in this example but the overall strategy proposed in this paper provides a path forward to propagate scale dependent geophysical uncertainty to scale dependent petrophysical uncertainty.

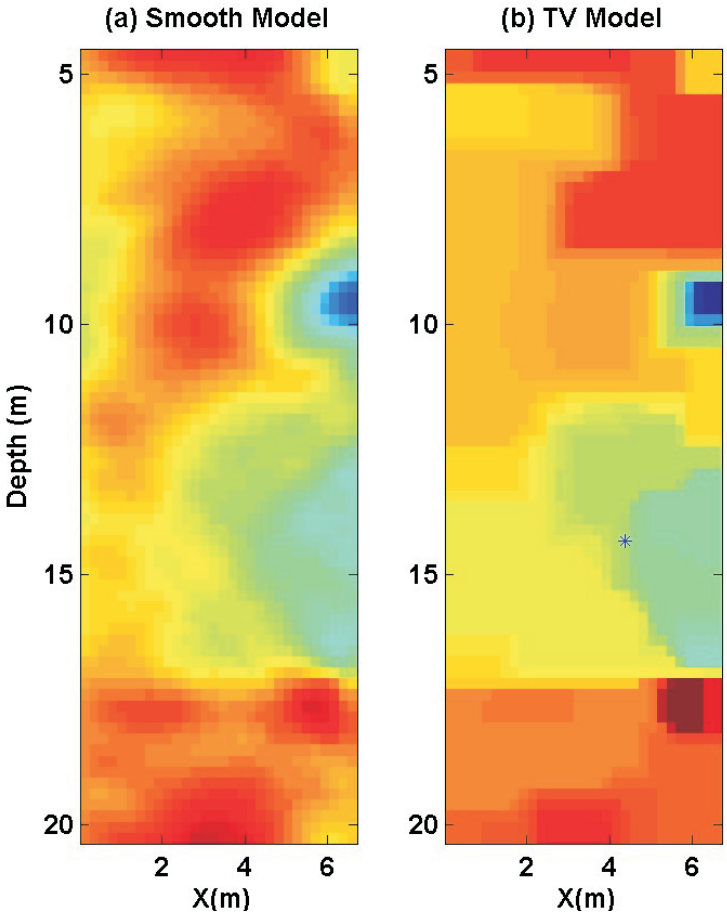

Fig.11 Inversion of BHRS crosswell data for slowness. The color range is 9-13 ns/m from blue to brown. (a) Smooth model inversion and (b) is obtained using TV regularization. The star indicates the position around which the funnel functions are computed by choosing area of increasing dimension.

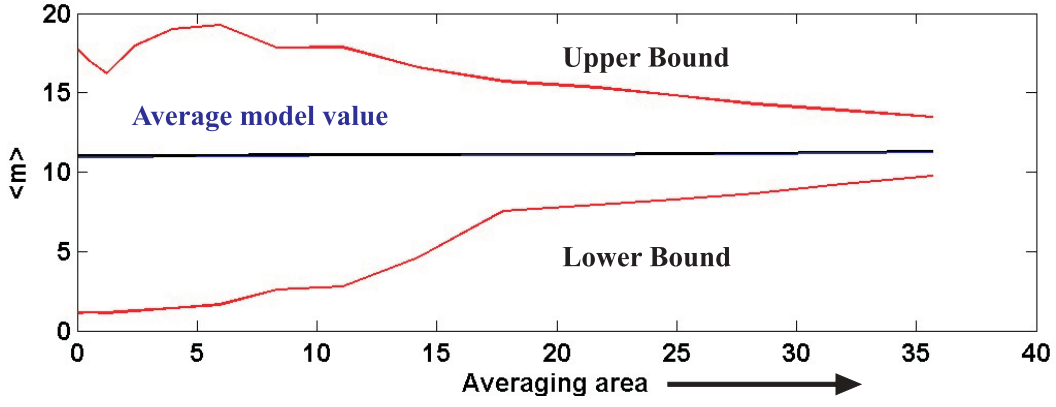

Fig.12 Funnel functions for the BHRS data set about the point indicated by star in Figure 11. Averaging area is in m^2.

Conclusions

The recovered model from an inversion algorithm will show features and the question that arises is whether the features are demanded by the data and whether different portions of the model domain have values that are distinctly different. Funnel function analysis provides a formalism to compute upper and lower bounds to localized averages of the model and hence answer some of the posed questions. In this work the funnel function method is solved by formulating the max-min optimization problem using an interior point method and imposing physical constraints via log barrier function. The bounds for the localized averages of the model depend upon a priori information about the model, but in this work only minimal information was incorporated. It is shown that certain portions of the model domain were distinctly different from other parts. The strategy proposed in this paper provides a general framework to determine scale dependent (target-oriented) uncertainties taking into account the resolution of the geophysical data.

Acknowledgments

I thank Prof. Doug Oldenburg, University of British Columbia for many insights and discussions on inverse problems and its application to geophysics over the years. I thank Tim Johnson, Idaho National Labs and Warren Barrash from Boise State University for providing the tomography data set from BHRS. I thank Upstream Research Company, ExxonMobil for permission to publish this paper.

References

Alumbaugh, D., and Newman, G., 2000, Image appraisal in 2D and 3D electromagnetic inversion, Geophysics, 65, 1455-67.

Backus, G., and Gilbert, F., 1970a, Inference from inadequate and inaccurate data I, National Academy of Science Proceedings, 65, 1-7.

Backus, G., and Gilbert, F., 1970b, Inference from inadequate and inaccurate data II, National Academy of Science Proceedings, 65, 281-287.

Backus, G., and Gilbert, F., 1970c, Inference from inadequate and inaccurate data III, National Academy of Science Proceedings, 67, 282-289.

Backus, G., and Gilbert, F., 1970, Uniqueness in the inversion of gross earth data, Phil. Trans. Roy. Soc. Series A Mathematical and Physical Sciences, 266, 123-192.

Barrash, W., and Reboulet, E.C., 2004. Significance of porosity for stratigraphy and textural composition in subsurface, coarse fluvial deposits: Boise Hydrogeophysical Research Site, *Geological Society of America Bulletin*, 116(9), p.1059-1073.

Barrash, W., Clemo, T., Fox, J.J., and Johnson, T.C., 2006. Field, laboratory, and modeling investigation of the skin effect at wells with slotted casing, Boise Hydrogeophysical Research Site, Journal of Hydrology, 326(1-4), p.181-198

Farquharson, C., and Oldenburg, D. W., 1998, Non-linear inversion using general measures of data misfit and model structure, Geophysical Journal International, 134, 213-227.

Miller, C. R., and Routh, P. S., 2007, Resolution analysis of geophysical images: Comparison of point spread function and region of data influence measures, Geophysical Prospecting, 55, 6, 835-852.

Nocedal, J., and Wright, S., 2006, Numerical Optimization, Springer Verlag Publication.

Oldenborger, G.A., and P.S. Routh, 2009, The point-spread function measure of resolution for the 3D electrical resistivity experiment, Geophysical Journal International, 176, 405-414.

Oldenborger, G.A., P.S. Routh, and M.D. Knoll, 2007, Model reliability for 3D electrical resistivity tomography: Application of the volume of investigation index to a time-lapse monitoring experiment, Geophysics, 72(4), F167-F175.

Oldenburg, D., 1983, Funnel functions in linear and nonlinear appraisal, J. Geophy. Res., B 9 7387-98.

Oldenburg, D., and Li, Y., 1999, Estimating the depth of investigation in dc resistivity and IP surveys, Geophysics, 64, 403-416.

Parker, R., 1994, Geophysical Inverse Theory, Princeton University Press.

Routh, P.S., 1999, Electromagnetic coupling in frequency domain induced polarization data, Ph.D. thesis, University of British Columbia.

Routh, P.S., and Oldenburg, D. W., and Li, Y., 1998, Regularized inversion of spectral IP parameters from complex resistivity data, SEG, Expanded Abstracts, 17, no. 1, 810-813.

Routh, P.S., Anno, P. D., Baumel, R. T., and Chavarria, J. A., 2003, Inversion for source wavelet and AVA parameters from prestack seismic data, SEG Expanded Abstracts, 22, no. 1, 698-701.

Routh, P. S., Oldenborger, G. A., and Oldenburg, D. W., 2005, Optimal survey design using point spread function measure of resolution, SEG, Expanded Abstracts, 1033-1036.

Routh, P. S., Anno, P. D., and Baumel, R. T., 2006, A simultaneous inversion for source wavelet and AVA parameters from prestack seismic data: US Patent # 7072767

Routh, P., and Miller, C., 2006, Image interpretation using appraisal analysis, SAGEEP proceedings, 1812-1820.

Routh, P.S., and Oldenburg, D. W., 2007, Uncertainty estimates using funnel functions, SEG Expanded Abstracts, 26, no. 1, 1187-1191.

Routh, P., Qu, L., Sen M., Anno, P., 2007, Inversion for non-smooth models with physical bounds, SEG expanded abstract, 26, no. 1, 1795-1799.

Sen, M. K., and Stoffa, P. L., 1995, Global optimization method in geophysical inversion, Elsevier Publication.

Snieder, R., 1990, A perturbative analysis of nonlinear inversion, Geophys. J. Int., 101, 545-556.