XIX International Conference on Water Resources
CMWR 2012
University of Illinois at Urbana-Champaign
June 17-22,2012

LINEARIZED FUNCTIONAL MINIMIZATION FOR INVERSE MODELING

Brendt Wohlberg[*], Daniel M. Tartakovsky[†], and Marco Dentz[‡]

[*]Theoretical Division
Los Alamos National Laboratory
Los Alamos, NM 87545, USA
Email: brendt@lanl.gov
[†]Department of Mechanical and Aerospace Engineering
University of California, San Diego
La Jolla, CA 92093, USA
Email: dmt@ucsd.edu
[‡]Department of Geosciences
Institute of Environmental Assessment and Water Research (IDAEA-CSIC)
c/ Jordi Girona, 18, 08034 Barcelona, Spain
Email: marco.dentz@idaea.csic.es

Key words: Inverse modeling, Total Variation regularization, Discontinuous coefficients

Summary. Heterogeneous aquifers typically consist of multiple lithofacies, whose spatial arrangement significantly affects flow and transport. The estimation of these lithofacies is complicated by the scarcity of data and by the lack of a clear correlation between identifiable geologic indicators and attributes. We introduce a new inverse-modeling approach to estimate both the spatial extent of hydrofacies and their properties from sparse measurements of hydraulic conductivity and hydraulic head. Our approach is to minimize a functional defined on the vectors of values of hydraulic conductivity and hydraulic head fields defined on regular grids at a user-determined resolution. This functional is constructed to (i) enforce the relationship between conductivity and heads provided by the groundwater flow equation, (ii) penalize deviations of the reconstructed fields from measurements where they are available, and (iii) penalize reconstructed fields that are not piece-wise smooth. We develop an iterative solver for this functional that exploits a local linearization of the mapping from conductivity to head. This approach provides a computationally efficient algorithm that rapidly converges to a solution. A series of numerical experiments demonstrates the robustness of our approach.

1 Introduction

Heterogeneous aquifers typically consist of multiple lithofacies, whose spatial arrangement significantly affects flow and transport in the subsurface. The identification of lithofacies and their hydraulic properties from hydrologic measurements is complicated

by the sparsity of data and by the lack of clear correlation between identifiable geologic indicators and attributes. Recent advances in subsurface applications of inverse modeling explicitly dealt either with heterogeneous environments consisting of one lithofacies[1] or with homogeneous constitutive lithofacies.[2]

The difficulty in accounting for these two scales of heterogeneity (large-scale heterogeneity stemming from uncertain spatial arrangements of multiple lithofacies and small-scale heterogeneity in hydraulic/transport properties of an individual lithofacies) stems from the fact that most modern inverse modeling techniques are optimized to deal with the one scale or the other. For example, level set methods[2] and support vector machines (SVMs)[3] work best for facies delineation, since they are designed to deal with the identification of sharp boundaries. The performance of regression SVMs to infer parameter values (*i.e.*, to deal with small-scale heterogeneity) is less spectacular.[3]

In this study, we explore the use of total variation regularizations as a tool for dealing with both scales heterogeneity simultaneously in a computationally efficient manner. Our goal is to reconstruct conductivity fields—in the presence of multiple geologic facies—from both system-parameter data $K_i = K(\mathbf{x}_i)$ and system-state data $h_i = h(\mathbf{x}_i, t)$. Without loss of generality, we assume that both data sets are collected at the same N locations $\mathbf{x}_i = (x_i, y_i)^T$, where $i \in \{1, \ldots, N\}$. Such problems are ubiquitous in subsurface hydrology since the geologic structure of the subsurface plays a crucial role in fluid flow and contaminant transport.

2 Inverse Modeling as a Nonlinear Optimization Problem

We consider inverse modeling in the context of steady-state saturated flow, subject to

$$\nabla \cdot (K \nabla h) = 0 \tag{1}$$

with the parameter K representing hydraulic conductivity and the system state h representing hydraulic head. When represented by samples on discrete grids, K and h will be denoted by vectors $\mathbf{k}$ and $\mathbf{h}$ respectively. The linear operators representing measurement by selecting a subset of $\mathbf{k}$ and $\mathbf{h}$ grid points will be denoted M_k and M_h respectively, and the corresponding measurements will be denoted by $\hat{\mathbf{k}}$ and $\hat{\mathbf{h}}$ respectively (*i.e.* $M_k \mathbf{k} = \hat{\mathbf{k}}$ and $M_h \mathbf{h} = \hat{\mathbf{h}}$). We estimate $\mathbf{k}$ and $\mathbf{h}$ from measurements $\hat{\mathbf{k}}$ and $\hat{\mathbf{h}}$ by minimizing a functional that penalizes the mismatch with the model connecting $\mathbf{k}$ and $\mathbf{h}$, the error in the estimates with respect to the measurements, and suitable regularization terms expressing our prior knowledge or expectation of the properties of the solution fields. Defining F as the function implicitly defined by the diffusion equation (together with boundary conditions) connecting $\mathbf{k}$ and $\mathbf{h}$, so that $\mathbf{h} = F(\mathbf{k})$, this problem can be expressed as

$$\arg\min_{\mathbf{k},\mathbf{h}} \frac{1}{2} \|F(\mathbf{k}) - \mathbf{h}\|_2^2 + \frac{\gamma_1}{2} \|M_k \mathbf{k} - \hat{\mathbf{k}}\|_2^2 + \frac{\gamma_2}{2} \|M_h \mathbf{h} - \hat{\mathbf{h}}\|_2^2 + \gamma_3 R_k(\mathbf{k}) + \gamma_4 R_h(\mathbf{h}), \tag{2}$$

or, replacing the penalty $\|F(\mathbf{k}) - \mathbf{h}\|_2^2$ by the constraint $\mathbf{h} = F(\mathbf{k})$,

$$\arg\min_{\mathbf{k}} \frac{\gamma_1}{2}\|M_k\mathbf{k} - \hat{\mathbf{k}}\|_2^2 + \frac{\gamma_2}{2}\|M_h F(\mathbf{k}) - \hat{\mathbf{h}}\|_2^2 + \gamma_3 R_k(\mathbf{k}) + \gamma_4 R_h(F(\mathbf{k})). \tag{3}$$

Since we expect that the estimated $\mathbf{k}$ fields may contain discontinuities, the Total Variation (TV)[4] (*i.e.* the ℓ_1 norm of the gradient) is a reasonable choice of regularization terms. (We note that similar choices of regularization term have previously been applied to nonlinear problems in different application areas, but these did not address the inverse problem involving the diffusion equation, and proposed very different algorithms.[5,6]) While a reasonable choice for the regularization for $\mathbf{h}$ would be the ℓ_2 norm of the gradient (since it is smoother than $\mathbf{k}$), we have found in practice that regularization of $\mathbf{h}$ does not have a significant effect, and slightly improved results are obtained when using TV regularization for $\mathbf{k}$ together with an additional term, with smaller parameter, penalizing the ℓ_2 norm of the gradient. While we do not claim that it is inherently superior, the work reported here utilizes the constrained form Eq. (3) above.

If the relationship between $\mathbf{h}$ and $\mathbf{k}$ were linear, this problem could be solved using minor variations on standard algorithms for TV regularization.[7,8] However, F in Eq. (3) is highly nonlinear, representing a functional relationship between $\mathbf{k}$ and $\mathbf{h}$ provided by the diffusion equation (1) and corresponding boundary conditions. We address this problem by utilizing a local linearization[9] of the function that $\mathbf{h} = F(\mathbf{k})$. Given some vector $\mathbf{k}_j$,

$$F(\mathbf{k}) \approx F(\mathbf{k}_j) + J_F(\mathbf{k}_j)(\mathbf{k} - \mathbf{k}_j), \tag{4}$$

where $J_F(\mathbf{k}_j)$ is the Jacobian of function F evaluated at $\mathbf{k}_j$. This linearization allows the nonlinear problem to be solved by iterative solution of a linear subproblem, described in detail in the following section.

3 Iterative Local Linearization

Since our algorithm can be applied to any inverse problem of this type, we switch to a notation that is not domain specific, and pose the problem in terms of vectors $\mathbf{u}$ and $\mathbf{v}$, subject to $\mathbf{v} = F(\mathbf{u})$, and with measurement operators P and Q such that $P\mathbf{u} = \mathbf{s}$, and $Q\mathbf{v} = \mathbf{t}$. The problem formulated in the previous section corresponds to $\mathbf{u} = \mathbf{k}$, $\mathbf{v} = \mathbf{h}$, $P = M_k$, $Q = M_h$, $\mathbf{s} = \hat{\mathbf{k}}$, and $\mathbf{t} = \hat{\mathbf{h}}$. Eq. (3) can be rewritten as

$$\arg\min_{\mathbf{u}} \frac{\alpha}{2}\|P\mathbf{u} - \mathbf{s}\|_2^2 + \frac{\beta}{2}\|QF(\mathbf{u}) - \mathbf{t}\|_2^2 + \gamma \left\| \sqrt{\mathcal{D}(\mathbf{u})} \right\|_1 + \frac{\delta}{2} \left\| \sqrt{\mathcal{D}(\mathbf{u})} \right\|_2^2 , \tag{5}$$

where $\mathcal{D}(\mathbf{u}) = (D_x\mathbf{u})^2 + (D_y\mathbf{u})^2$, D_x and D_y are discrete derivative operators in the horizontal and vertical directions respectively, and scalar operations, *e.g.* $\cdot^2$ and $\sqrt{\cdot}$, applied to a vector denote element-wise operation.

3.1 Linear Subproblem

Consider some initial values $\mathbf{u}_j$ and $\mathbf{v}_j = F(\mathbf{u}_j)$, define $\mathbf{u} = \mathbf{u}_j + \mathbf{w}$ and, for small $\|\mathbf{w}\|_2^2$, linearize about $\mathbf{u}_j$ so that $\mathbf{v} \approx \mathbf{v}_j + A\mathbf{w}$. To ensure that $\mathbf{w}$ is small, *i.e.* that the penalties on $\mathbf{u}$ and $\mathbf{v}$ are satisfied, we add an additional term penalizing the magnitude of $\mathbf{w}$. (This term plays essentially the same role as the damping term in the Levenberg-Marquardt method.[10]) This gives

$$\begin{aligned} \arg\min_{\mathbf{w}} \frac{1}{2} \|\mathbf{w}\|_2^2 + \frac{\alpha}{2} \|P(\mathbf{u}_j + \mathbf{w}) - \mathbf{s}\|_2^2 + \frac{\beta}{2} \|Q(\mathbf{v}_j + A\mathbf{w}) - \mathbf{t}\|_2^2 + \\ \gamma \left\| \sqrt{\mathcal{D}(\mathbf{u}_j + \mathbf{w})} \right\|_1 + \frac{\delta}{2} \left\| \sqrt{\mathcal{D}(\mathbf{u}_j + \mathbf{w})} \right\|_2^2 . \end{aligned} \tag{6}$$

We minimize this functional using the Alternating Direction Method of Multipliers (ADMM) approach.[11] Introducing auxiliary variables $\mathbf{x} \approx D_x(\mathbf{u}_j + \mathbf{w})$ and $\mathbf{y} \approx D_y(\mathbf{u}_j + \mathbf{w})$, and *scaled dual* variables $\mathbf{b}_x$ and $\mathbf{b}_y$, and splitting the resulting optimization problem into two subproblems, we obtain

$$\begin{aligned} \arg\min_{\mathbf{w}} \frac{1}{2} \|\mathbf{w}\|_2^2 + \frac{\alpha}{2} \|P\mathbf{w} - (\mathbf{s} - P\mathbf{u}_j)\|_2^2 + \frac{\beta}{2} \|QA\mathbf{w} - (\mathbf{t} - Q\mathbf{v}_j))\|_2^2 + \\ \frac{\lambda}{2} \|D_x\mathbf{w} - (\mathbf{x} - D_x\mathbf{u}_j - \mathbf{b}_x)\|_2^2 + \frac{\lambda}{2} \|D_y\mathbf{w} - (\mathbf{y} - D_y\mathbf{u}_j - \mathbf{b}_y)\|_2^2 , \end{aligned} \tag{7}$$

$$\begin{aligned} \arg\min_{\mathbf{x},\mathbf{y}} \gamma \left\| \sqrt{\mathbf{x}^2 + \mathbf{y}^2} \right\|_1 + \frac{\delta}{2} \left\| \sqrt{\mathbf{x}^2 + \mathbf{y}^2} \right\|_2^2 + \\ \frac{\lambda}{2} \|\mathbf{x} - (D_x(\mathbf{u}_j + \mathbf{w}) + \mathbf{b}_x)\|_2^2 + \frac{\lambda}{2} \|\mathbf{y} - (D_y(\mathbf{u}_j + \mathbf{w}) + \mathbf{b}_y)\|_2^2 . \end{aligned} \tag{8}$$

The $\mathbf{w}$ subproblem is solved by setting the gradient of the $\mathbf{w}$ functional to zero, giving the linear system

$$\begin{aligned} &\left(I + \alpha P^T P + \beta A^T Q^T Q A + \lambda D_x^T D_x + \lambda D_y^T D_y\right) \mathbf{w} = \alpha P^T (\mathbf{s} - P\mathbf{u}_j) \\ &+ \beta A^T Q^T (\mathbf{t} - Q\mathbf{v}_j) + \lambda D_x^T (\mathbf{x} - D_x\mathbf{u}_j - \mathbf{b}_x) + \lambda D_y^T (\mathbf{y} - D_y\mathbf{u}_j - \mathbf{b}_y) . \end{aligned} \tag{9}$$

The $\mathbf{x}, \mathbf{y}$ subproblem is solved by generalizing the shrinkage function used in a Split Bregman (equivalent to ADMM) solution for TV regularization.[8] Defining the function

$$\sigma(s, t, \alpha, \beta) = \frac{\max\left(0, \sqrt{s^2 + t^2} - \alpha\right)}{(1 + \beta)\sqrt{s^2 + t^2}} , \tag{10}$$

we can write solution of this subproblem in closed-form as

$$\mathbf{x} = (D_x(\mathbf{u}_j + \mathbf{w}) + \mathbf{b}_x)\, \sigma\left(D_x(\mathbf{u}_j + \mathbf{w}) + \mathbf{b}_x, D_y(\mathbf{u}_j + \mathbf{w}) + \mathbf{b}_y, \gamma/\lambda, \delta/\lambda\right) \tag{11}$$

$$\mathbf{y} = (D_y(\mathbf{u}_j + \mathbf{w}) + \mathbf{b}_y)\, \sigma\left(D_x(\mathbf{u}_j + \mathbf{w}) + \mathbf{b}_x, D_y(\mathbf{u}_j + \mathbf{w}) + \mathbf{b}_y, \gamma/\lambda, \delta/\lambda\right) , \tag{12}$$

where the function $\sigma(\mathbf{s}, \mathbf{t}, \alpha, \beta)$ with vector arguments $\mathbf{s}$ and $\mathbf{t}$ denotes element-wise operation on corresponding elements of $\mathbf{s}$ and $\mathbf{t}$.

The full algorithm for solving the linear subproblem consists of iteratively solving both subproblems (*i.e.* at iteration k, first solving the linear system to obtain $\mathbf{w}^{k+1}$, and then using $\mathbf{w}^{k+1}$ within the shrinkage function to obtain $\mathbf{x}^{k+1}$ and $\mathbf{y}^{k+1}$), followed by scaled dual variable updates[11]

$$\mathbf{b}_x^{(k+1)} = \mathbf{b}_x^{(k)} + D_x \left(\mathbf{u}_j + \mathbf{w}^{(k+1)}\right) - \mathbf{x}^{(k+1)} \tag{13}$$

$$\mathbf{b}_y^{(k+1)} = \mathbf{b}_y^{(k)} + D_y \left(\mathbf{u}_j + \mathbf{w}^{(k+1)}\right) - \mathbf{y}^{(k+1)} \ . \tag{14}$$

3.2 Outer Iterations

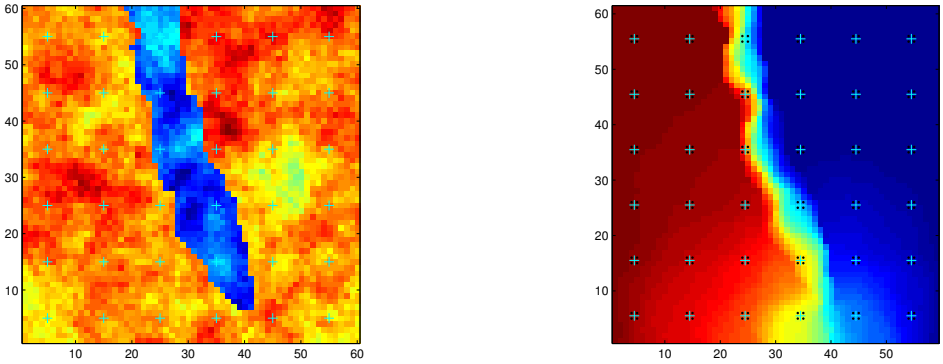
Each solution of the linear subproblem provides a descent direction $\mathbf{w}$ for the outer nonlinear problem. Since the function f is highly nonlinear, the linearization A may only provide a good approximation to the function behavior within a very small region, so a line search is necessary. If $\mathbf{u}_j + \mathbf{w}$ does not reduce the functional value, parameters α, β, γ, and δ are reduced by a factor of 2, and the linear subproblem is repeated. When the update $\mathbf{w}$ is acceptable, the update $\mathbf{u}_{i+1} = \mathbf{u}_j + \mathbf{w}$ is made together with computation of the linearization at $\mathbf{u}_{i+1}$, and the process is repeated. If the initial update $\mathbf{w}$ is found to be acceptable, parameters α, β, γ, and δ are increased by a factor of 2, so that the effective step size is not permanently reduced by a line search from a previous outer iteration.

4 Computational Example

The simulations reported below relied on a Galerkin finite-element solution of Eq. (1) on a rectangular domain, subject to no-flow boundary conditions on the two horizontal boundaries and constant heads on the two vertical boundaries (20 on the left and 0 on the right, with the heads expressed in consistent model units). Our numerical experiments revealed that TV is not an appropriate regularization term for conductivity fields, which often have values spanning many orders of magnitude. The $\ln K$ field, however, has properties appropriate to this regularization. We therefore define function $g(\mathbf{x}) = F(\exp(\mathbf{x}))$, and solve the problem in terms of $\ln K$ and h instead of K and h. The Jacobian for this modified function is calculated by application of the chain rule, giving $J_g(\mathbf{x}) = J_F(\exp(\mathbf{x}))\,\mathrm{diag}(\exp(\mathbf{x}))$, where $\mathrm{diag}(\cdot)$ denotes construction of a diagonal matrix with the diagonal consisting of its vector argument.

We test our algorithm on a synthetic problem constructed from the $\ln K$ field shown in Fig. 1(a). The corresponding reference heads field is presented in Fig. 1(b). We use 25 samples from each field (sampled on a regular grid, at the same position in each field) to reconstruct the fields using our method. We use algorithm parameters $\alpha = 5, \beta = 5, \gamma = 1 \times 10^{-6}$, $\delta = 2 \times 10^{-7}$, and $\lambda = 5$, and choose the initial value for $\mathbf{k}$ as the mean of the conductivity sample values, and the initial value of $\mathbf{h} = F(\mathbf{k})$. We allow 60 outer iterations (corresponding to 79 evaluations of the nonlinear function F, including the linesearch), taking approximately 610s on a multiprocessor Intel Xeon X5570 workstation,

to compute the results displayed in Fig. 2. Root Mean Squared (RMS) errors on the log conductivity and heads fields are 1.76 and 0.41 respectively (the variances of the corresponding reference fields are 7.09 and 18.60 respectively).



(a) Log conductivity reference (b) Heads reference

Figure 1: Reference log conductivity (a) and hydraulic head (b) fields. The + signs represent measurement locations used to reconstruct these fields.

Although the estimated $\ln K$ field captures the key features of the reference field (*i.e.* the presence of distinct hydrofacies), the RMS error is relatively large. We examine the sensitivity of solutions of Eq. (1) to this error by solving the flow problem with the estimated $\ln K$ fields for two different types of change in driving forces: varying head at the left boundary, and a pumping well with varying rate. Fig. 3 exhibits the RMS error of the difference between the heads corresponding to the reference and estimated $\ln K$ fields as a function of the RMS error of the change in heads solution for the reference field with respect to the original (*i.e.* unmodified boundary conditions) heads field.

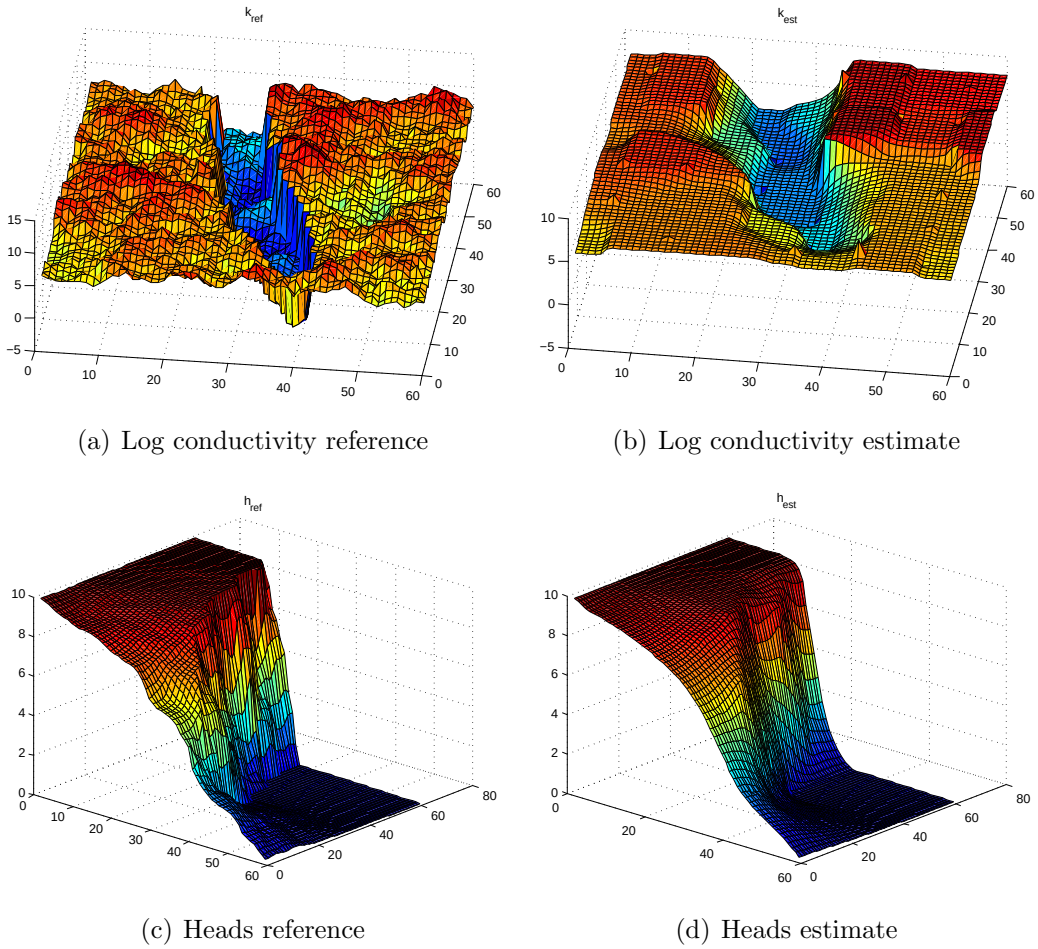
5 Summary and Conclusions

We have proposed a regularization approach to inverse modeling of subsurface processes. Our approach yields the conductivity field estimates that capture the salient features of the field, even when only two samples are available within the low-conductivity inclusion. While the RMS error in the conductivity reconstruction is relatively large, the estimated conductivity field provides robust head predictions for flow regimes with moderate size perturbations from those used in the inverse procedure.

Future work includes a systematic comparison of our approach with its state-of-the-art counterparts, and the use of our approach in the hydraulic tomography context.

Acknowledgment

This study was funded by the Office of Science of the U.S. Department of Energy, Advanced Scientific Computing Research program in Applied Mathematical Sciences.



(a) Log conductivity reference

(b) Log conductivity estimate

(c) Heads reference

(d) Heads estimate

Figure 2: Inverse modeling results using the functional minimization approach. The $\ln K$ and h fields were reconstructed from 25 measurements of each. The estimated $\ln K$ field captures the salient properties of the reference, despite only two of the samples having fallen within the low-conductivity inclusion.

REFERENCES

[1] H. J. H. Franssen, A. Alcolea, M. Riva, M. Bakr, N. van der Wiel, F. Stauffer, and A. Guadagnini, "A comparison of seven methods for the inverse modelling of groundwater flow. application to the characterisation of well catchments," *Adv. Water Resour.* **32**(6), pp. 851 – 872, 2009.

[2] M. A. Iglesias and D. McLaughlin, "Level-set techniques for facies identification in reservoir modeling," *Inverse Problems* **27**, p. 035008 (36pp), 2011.

[3] B. Wohlberg, D. M. Tartakovsky, and A. Guadagnini, "Subsurface characterization with Support Vector Machines," *IEEE Trans. Geosci. Remote Sens.* **44**(1), pp. 47 – 57, 2006. doi:10.1109/TGRS.2005.859953.

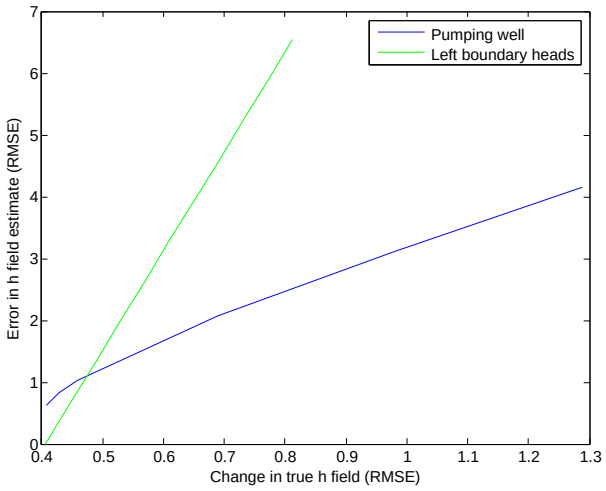


Figure 3: Sensitivity of head prediction accuracy to the error in the conductivity field estimate.

[4] L. Rudin, S. J. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms.," *Physica D. Nonlinear Phen.* **60**, pp. 259–268, November 1992.

[5] C. G. Farquharson and D. W. Oldenburg, "Non-linear inversion using general measures of data misfit and model structure," *Geophys. J. Int.* **134**(1), pp. 213–227, 1998.

[6] U. Ascher and E. Haber, "Computational methods for large distributed parameter estimation problems with possible discontinuities," in *Proc. Symp. Inverse Problems, Design & Optimization*, pp. 201–208, 2004.

[7] P. Rodríguez and B. Wohlberg, "Efficient minimization method for a generalized total variation functional," *IEEE Trans. Image Proc.* **18**, pp. 322–332, Feb. 2009.

[8] T. Goldstein and S. J. Osher, "The Split Bregman method for l1-regularized problems," *SIAM J. Imaging Sci.* **2**(2), pp. 323–343, 2009.

[9] K. P. Bube and R. T. Langan, "Hybrid ℓ^1/ℓ^2 minimization with applications to tomography," *Geophysics* **62**(4), pp. 1183–1195, 1997.

[10] J. Pujol, "The solution of nonlinear inverse problems and the Levenberg-Marquardt method," *Geophysics* **72**(4), pp. W1–W16, 2007.

[11] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.* **3**(1), pp. 1–122, 2010.