Linear and Nonlinear Models for Inversion of Electrical Conductivity Profiles in Field Soils from EM-38 Measurements

by

Alex C. Hilgendorf

Submitted in Partial Fulfillment

of the Requirements for the Degree of

Master of Science in Mathematics with Operations Research and Statistics

Option

New Mexico Institute of Mining and Technology

Socorro, New Mexico

August 5, 1997

ACKNOWLEDGEMENT

I would like to thank Dr. Jan Hendricks of the New Mexico Tech Hydrology department for allowing me to research in soil hydrology, as well as Dr. Brian Borchers for guiding me through this odyssey. I would like to extend my thanks to Drs. Hossain, Schaffer, and Stone for their support on my committee. Also, this research would not have been possible without the approval of Drs. Rhoades, Corwin, and Lesch, associated with the US Salinity Laboratory, who painstakingly gathered the data for their own research. These researchers are associated with the US Salinity Laboratory. Finally, I cannot forget the faculty, staff, students, and department mascots that bring honor to New Mexico Institute of Mining and Technology.

This thesis was typeset with LaTeX[1] by the author.

[1] LaTeX document preparation system was developed by Leslie Lamport as a special version of Donald Knuth's TeX program for computer typesetting. TeX is a trademark of the American Mathematical Society. The LaTeX macro package for the New Mexico Institute of Mining and Technology thesis format was adapted by Gerald Arnold from the LaTeX macro package for The University of Texas at Austin by Khe-Sing The.

ABSTRACT

The EM-38 is an instrument used to measure conductivity in the soil, and thus estimate soil salinity. An alternating current is sent through a transmitting coil. This creates an alternating magnetic field that induces current flow in the underlying soil. These currents create secondary magnetic fields. The combination of fields induces a secondary voltage in the receiving coil of the EM-38. The instrument measures the relative strength of the secondary magnetic fields which is a function of the apparent electrical conductivity of the underlying soil.

The response function of the instrument can be represented with a linear or nonlinear model, both of which are presented. A forward comparison of the two models is presented on the 14 sites. The nonlinear model always outperforms the linear model at predicting the EM-38 measurements from measured electrical conductivity (ECa) profiles. The difference is substantial when ECa readings are high (over 500 $\frac{mS}{m}$). Next, the model's ability to invert ECa readings is investigated. With the linear model it is difficult to predict conductivities below 1.5 meters. The nonlinear model is much more computationally expensive but is generally more sensitive to ECa trends and can yield a better solution.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction to Ill-Posed Inverse Problems

1.1 Forward, Inverse, and Identification Problems

The existence of forward and inverse problems follows from the cause and effect relationships found in nature. A specific mathematical model representing some physical system could be represented by K. Then the current state of the system could be represented by x. The present state x may cause some natural unique effect via the system modeled by K. Consider this effect or set of effects to be y. This cause and effect relationship can be generally expressed as

$$Kx \rightarrow y. \tag{1.1}$$

Three important problems are defined in terms of what is known and unknown about this relationship [Groetsch, 1993].

A *forward problem* is the most natural of the three. In this case the system K is well understood and the cause x is well defined. It is left to predict the effect y. Often a cause cannot be directly measured. It is left to reconstruct the cause x, by defining K, and then measuring the effect. This is an *inverse problem*. Both inverse and forward problems assume complete knowledge of K. When K has not been defined, it is left to repeatedly observe cause and effect combinations until K is completely known and all feasible effects can be predicted. This is a *model identification problem*.

Note the above assumption that predicted effects in the forward problem are unique. This does not imply a unique cause for each effect. This interesting observation leads to the discussion of *ill-posed inverse problems*.

1.2 A Prelude to Ill-Posed Problems

A problem is *well-posed* if and only if the following are true [Zauderer, 1983]:

i) the solution exists,

ii) the solution is uniquely determined,

iii) the solution depends continuously on the initial and/or boundary data.

A problem is *well-conditioned* if points (i) and (ii) from above hold as well as a modification of point (iii): a small deviation in problem parameters results in a small deviation in the solution.

The negation of any point implies that the problem is *ill-posed* (or *ill-conditioned*). Of course we desire that any problem have a solution as in point (i). In the practical solution of inverse problems this is not a major issue. An effect is always preceded by a cause. The inverse problem not having a solution would indicate a problem with the model. Consider again that a forward problem is assumed to produce a unique effect. Of course there could be multiple causes inducing the same effect. In this case the corresponding inverse problem of reconstructing the cause is ill-posed by condition (ii).

This thesis will be primarily concerned with point (iii). Any problem is ill-posed if small deviations in the problem induce discontinuous deviations in the solution. This feature is largely prevalent in inverse and identification

problems. In fact, ill-posedness in this sense is extremely problematic: any measurement error is essentially a small parameter deviation. The implications will be discussed at the end of chapter 1. We proceed by examining three problems that are either ill-posed or ill-conditioned.

1.3 Ill-Posed Inverse Problems

These examples are closely related to the inversion problem stated in chapter 3 and solved in the remainder of the thesis. The first has simple intuitive appeal. The last two demonstrate ill-conditioning and ill-posedness in linear algebra and calculus.

1.3.1 A Graphical Example

Consider two lines in a plane. Assume at first, that they are nearly perpendicular and not vertical. Also note that four numbers define the system: two slopes and two y-intercepts. For this system the solution is the point (x, y) satisfying the equations of both lines. Finally we consider a tiny alteration in the system. Allow a small deviation in either of the y-intercepts. This small change results in a correspondingly small change in the solution as shown in figure 1.1.

Now again consider two lines in a plane but assume they are nearly *parallel*. The solution of this system of equations is still the intersection. But a small deviation in either y-intercept results in a remarkably large change in the solution. This is demonstrated in figure 1.2.

In the first case the solution was insensitive to small system devia-

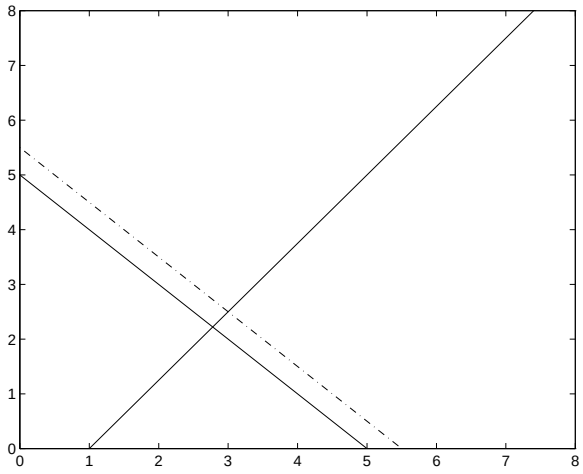

Figure 1.1: A Well-Posed System of Equations.

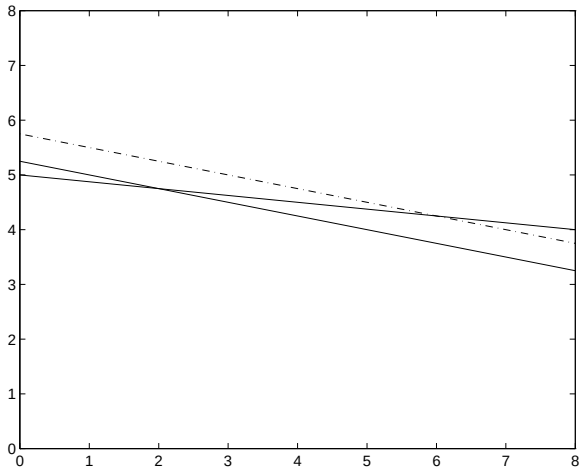

Figure 1.2: An Ill-Conditioned System of Equations.

tions. In the second case the solution to the system of equations is very sensitive to changes in the intercept. This system is *ill-conditioned.*

1.3.2 A Linear Algebra Example

We will look now at an adaptation of an example by Hansen [Hansen, 1992]. Consider the system of equations $Ax = b_0$, where

$$A = \begin{pmatrix} 0.16 & 0.10 \\ 0.17 & 0.11 \\ 2.02 & 1.29 \end{pmatrix}, \text{ and } b_0 = \begin{pmatrix} 0.26 \\ 0.28 \\ 3.31 \end{pmatrix} \tag{1.2}$$

There exists an exact solution $x_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. As in 1.2.1 we consider the effect of a small perturbation of b_0.

$$\text{Let } b_1 = \begin{pmatrix} 0.16 & 0.10 \\ 0.17 & 0.11 \\ 2.02 & 1.29 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 0.01 \\ -0.03 \\ 0.02 \end{pmatrix} = \begin{pmatrix} 0.27 \\ 0.25 \\ 3.33 \end{pmatrix} \tag{1.3}$$

This results in the over determined system of equations $Ax_1 = b_1$. Though there is no exact solution, it is natural to seek a least squares solution, that is, to find

$$\min_x \|Ax - b_1\|_2$$

where $\|\cdot\|_2$ is the Euclidean 2-norm. This solution is $x_{LSQ} = \begin{pmatrix} 7.01 \\ -8.40 \end{pmatrix}$. A relatively small change in a system parameter, b_0, induced a huge change in the solution x.

Recall that if we take A to be acting on x to yield a measurable result b, then finding x for a given b is an *ill-conditioned inverse problem.* Section 2.1 demonstrates how an ill-posed inverse problem can be discretized into an ill-conditioned inverse problem.

1.3.3 Systems of Linear Equations and Condition Numbers

As shown by Forsythe [Forsythe et al., 1977], one can conveniently define the condition number of a matrix so as to develop a relationship between the relative change in b and the relative error in the solution for x. The definition can be based on the 1, 2, or ∞ norms, denoted $\|\cdot\|_1, \|\cdot\|_2,$ or $\|\cdot\|_\infty$ respectively.

Given a square matrix A and a non-zero vector x, let $b_x = Ax$. The *condition number of* A is defined as

$$\text{cond}(A) = \frac{M}{m}, \text{ where} \tag{1.4}$$

$$M = \max_x \frac{\|Ax\|}{\|x\|} = \max_x \frac{\|b_x\|}{\|x\|} \tag{1.5}$$

$$\text{and } m = \min_x \frac{\|Ax\|}{\|x\|} = \min_x \frac{\|b_x\|}{\|x\|} \tag{1.6}$$

By definition, M is equivalent to the conditional matrix norm of A with respect to the given vector norm and is denoted by, $M = \|A\|$. This can be used to find an alternate definition of $\text{cond}(A)$ for nonsingular square matrices. Starting with the definition of m, we have

$$m = \min_x \frac{\|Ax\|}{\|x\|} \tag{1.7}$$

$$= \min_x \frac{\|AA^{-1}x\|}{\|A^{-1}x\|} \tag{1.8}$$

$$= \min_x \frac{\|x\|}{\|A^{-1}x\|} \tag{1.9}$$

$$= \frac{1}{\max_x \frac{\|A^{-1}x\|}{\|x\|}} \tag{1.10}$$

$$= \frac{1}{\|A^{-1}\|} \tag{1.11}$$

Taking $\text{cond}(A)$ to be the ratio of M to m, we are left with

$$\text{cond}(A) = \frac{M}{m} = \|A\| \frac{1}{\frac{1}{\|A^{-1}\|}} = \|A\|\|A^{-1}\| \tag{1.12}$$

This approach only works if A is both nonsingular and square.

Now if b is perturbed by an amount Δb, then the inverse solution of $Ax = b$ is perturbed by an amount Δx. Then we have the new system of equations

$$A(x + \Delta x) = b + \Delta b. \tag{1.13}$$

It follows that $A(\Delta x) = \Delta b$. By the definitions of m and M we have both

$$m\|\Delta x\| \leq \|\Delta b\| \tag{1.14}$$

$$\text{and } M\|x\| \geq \|b\| \tag{1.15}$$

Assuming $m \neq 0$ these results may be combined to form the relationship

$$\frac{\|\Delta x\|}{\|x\|} \leq \text{cond}(A) \frac{\|\Delta b\|}{\|b\|} \tag{1.16}$$

On the right we see that $\text{cond}(A)$ is multiplied by $\frac{\|\Delta b\|}{\|b\|}$. This fraction is essentially the relative change in the right hand side of $Ax = b$. This product bounds $\frac{\|\Delta x\|}{\|x\|}$ which is the relative error caused by this change. With this relationship it quickly becomes apparent that a small condition number of A implies that resulting inverse solutions of a perturbed problem would contain

little error. A large condition number provides a large upper bound on the relative error. Such a system is said to be an *ill-conditioned system of equations.*

1.3.4 A Fredholm Integral Equation of the First Kind

A *Fredholm Integral Equation of the First Kind* is an equation of the following form [Wing, 1991]

$$\int_0^1 k(s,t)\,x(t)\;dt = y(s), \quad 0 \le s \le 1. \tag{1.17}$$

Of course if the functions y and k are known, then the determination of x is an inverse problem. Assume that k is known and that x_0 is a unique inverse solution for a given y_0. Whereas in section 1.2.2 we began with a small perturbation on the vector b_0, we now consider a large perturbation of x_0 of the form

$$m(t) = M\sin(n\pi t) \tag{1.18}$$

Then if follows that

$$\begin{aligned}\int_0^1 k(s,t)[x_0(t)+m(t)]\;dt &= \int_0^1 k(s,t)[x_0(t)+M\sin(n\pi t)]\;dt \\ &= \int_0^1 k(s,t)x_0(t)\;dt + \int_0^1 Mk(s,t)\sin(n\pi t)\;dt\end{aligned}$$

By the Riemann-Lebesgue Lemma, if $k(.,t)$ is square integrable, then

$$\int_0^1 k(\cdot,t)\sin(n\pi t)\;dt \to 0 \text{ as } n \to \infty \tag{1.19}$$

Note that this result is independent of M's magnitude. Then in a forward sense, as n increases, large deviations in x cause small deviations in the projected y. In the context of inverse problems, *small alterations in y can cause discontinuous deviations in the inverse solutions of x* [Groetsch, 1993, Hansen, 1992].

Many problems in the physical sciences, from tomography to geomagnetic prospecting, can be expressed as Fredholm integral equations of the first kind [Parker, 1994, Twomey, 1997, Wing, 1991]. The ill-posedness of these equations is particularly problematic to those interested in solving problems in these fields.

1.4 Why Solving These Problems is Important

An understanding of ill-posed problems is of practical importance to all those doing mathematical work in the physical sciences. Recall from the opening discussion that a problem has three parts, a system model, K, a cause or system state, x, and an effect, y. Most scientific problems are either inverse problems or model identification problems. In the inverse problem the effect is somehow measured and the cause, of actual interest, is induced. There is always some sort of error involved in the measurement of y. We would like the corresponding error involved in determining x to be small as well. The same sorts of things can be said of the model identification problem. Methods for dealing with this ill-conditioned behavior, producing a physically realistic inverse solution for x, are the topic of chapter two.

Chapter 2

Discretization and Regularization: Solving the Ill-Posed Inverse Problem

Section 1.2.3 introduced the Fredholm Integral Equation of the First Kind,

$$\int_0^1 k(s,t)x(t)\,dt = y(s). \tag{2.1}$$

Section 1.2.2 introduced over determined systems of equations. In these cases ill-posed and ill-conditioned behavior was demonstrated. This chapter will focus on three general methods for solving ill-conditioned systems of equations. First, however, we will recall the continuous Fredholm integral equation and demonstrate its relationship to linear systems of equations.

2.1 Discretizing the Fredholm Integral Equation of the First Kind

In many physical applications $x(t)$ and $y(t)$ are continuous univariate functions. To solve the inverse problem, one observes y and deduces x. It is usually impractical to observe y in a continuous sense. For example t may represent time and y may represent some time dependent quantity such as position that is measured at certain points in time. Because y is not completely known, at best we can only discretely approximate x.

Consider the discretization of the linear Fredholm integral equation [Groetsch, 1993, Wing, 1991]. We measure y at points $s_0, s_1, ..., s_M$. This yields

a set of measurements $y_0, y_1, ..., y_M$. Now (2.1) becomes a set of $M+1$ equations predicting y_i:

$$\int_0^1 k(s_0,t)x(t)\ dt = y_0 \tag{2.2}$$

$$\int_0^1 k(s_1,t)x(t)\ dt = y_1 \tag{2.3}$$

and so on up to

$$\int_0^1 k(s_M,t)x(t)\ dt = y_M \tag{2.4}$$

Consider the i^{th} equation. It may be broken up as follows:

$$\int_0^1 k(s_i,t)x(t)dt = \int_0^{P_1} k(s_i,t)x(t)dt + \int_{P_1}^{P_2} k(s_i,t)x(t)dt + \cdots + \int_{P_{N-1}}^1 k(s_i,t)x(t)dt$$

Now let x_j be a constant function approximating $x(t)$ on the interval $[P_{j-1}, P_j)$. The above equation becomes

$$\int_0^1 k(s_i,t)x(t)dt \simeq x_1 \int_0^{P_1} k(s_i,t)dt + x_2 \int_{P_1}^{P_2} k(s_i,t)dt + \cdots + x_N \int_{p_{N-1}}^1 k(s_i,t)dt$$

Note that $\int_{P_{j-1}}^{P_j} k(s_i,t)dt$ may be calculated using a simple integration routine. It functions, then, as a scalar multiple of x_j.

Performing these operations on each equation, the set of equations becomes a linear system in the piecewise constant approximations x_j of x. Finally we have $Kx = y$, where y is a column M-vector containing discrete approximations to $y(t)$, x is a column N-vector containing discrete approximations to x, and K is an $N \times M$ matrix defined as

$$k_{ij} = \int_{P_{j-1}}^{P_j} k(s_i,t)dt \tag{2.5}$$

.

Discretizing the integral equation cures the practical problem of continuity. Ill-conditioned behavior, however, is inherited by the system of linear equations [Groetsch, 1993]. In fact Groetsch shows that decreasing the discretization parameter or increasing the size of the partition of $[0,1]$ previously described, creates a less conditioned system.

Consider Fox and Goodwins' example as cited by Groetsch(1993). The integral equation

$$\int_0^1 (s^2+t^2)^{\frac{1}{2}} f(t)dt = \frac{1}{3}[(1+s^2)^{\frac{3}{2}} - s^2], \quad 0 \leq s \leq 1 \tag{2.6}$$

is discretized into a discrete system of linear equations, $Kf = g$ where K is an $n \times n$ matrix. MATLAB is used to find the condition number of K for different discretization parameters:

n	cond(n)
2	9.9
10	1.9×10^{10}
20	3.2×10^{17}
50	3.3×10^{19}

Recall from equation (1.16) that cond(A) is a general measure of ill-posedness. As the discretization parameter increases, the upper bound on the relative solution error increases dramatically. We must therefore turn to the techniques of regularization to temper this problem.

2.2 Regularization

A direct solution of an ill-conditioned inverse problem is poor because it fails to approximate the true solution in the presence of uncertainty. Perhaps

where the true solution x is almost flat, the inverse solution has a large slope. Or maybe the true solution has little curvature but the inverse solution has a sharp curvature.

To handle these types of problems we must have some method of measuring their extremeness. Consider the unscaled 0^{th}, 1^{st}, and 2^{nd} discrete unscaled derivative operators:

$$L_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2.7}$$

$$L_1 = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}, \tag{2.8}$$

$$L_2 = \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{pmatrix}. \tag{2.9}$$

Measures of size, slope, and curvature for some inverse solution $x = (x_1,\ x_2,\ x_3,\ x_4)^{\mathrm{T}}$ are, respectively,

$$\|L_0 x\|_2 = \| \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \| \tag{2.10}$$

$$= \|(x_1,\ x_2,\ x_3,\ x_4)^{\mathrm{T}}\|_2 \tag{2.11}$$

$$\|L_1 x\|_2 = \| \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \| \tag{2.12}$$

$$= \|(x_1 - x_2,\ x_2 - x_3,\ x_3 - x_4)^{\mathrm{T}}\|_2 \tag{2.13}$$

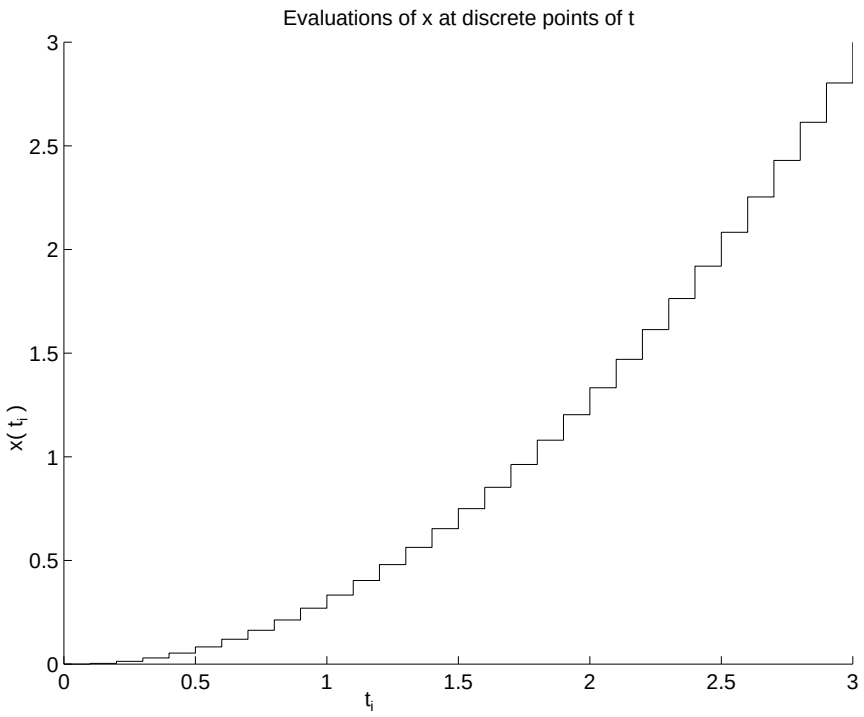

Figure 2.1: Step Function Representation of x(t).

$$\|L_2 x\|_2 = \| \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \| \tag{2.14}$$

$$= \|(x_1 - 2x_2 + x_3,\ x_2 - 2x_3 + x_4)^{\mathrm{T}}\|_2 \tag{2.15}$$

Consider the following example. Some quantity, x, depends on t as $x(t) = t^2/3$. Letting t range from 0.1 to 3 in steps of 0.1 we generate the set of 30 x values:

$$0.0033,\ \ 0.0133,\ \ 0.0300,\ \ 0.0533,\ ...,\ 2.8033,\ \ 3.0000.$$

The step function in figure 2.1 displays this vector of x values against

the corresponding vector of t values. This parabola has a small amount of curvature on (0,3] whereas the relative slope is quite large. These qualities can be demonstrated with the previously described discrete derivative operators.

The norm of the *zeroth* discrete derivative of x, or more plainly the norm of x, is defined by $(\|L_0 x\|_2)^2 = (\|Ix\|_2)^2 = (\|x\|)^2 = (0.0033)^2 + (0.0133)^2 + (0.0300)^2 + (0.0533)^2 \cdots + (2.8033)^2 + (3.0000)^2$ so that $\|L_0 x\| = 7.6551$.

The norm of the *first* discrete derivative of x is defined by $(\|L_1 x\|_2)^2 = (0.0033 - 0.0133)^2 + (0.0133 - 0.0300)^2 + (0.0300 - 0.0533)^2 + \cdots + (\ldots - 2.8033)^2 + (2.8033 - 3.0000)^2$ so that $\|L_1 x\| = 0.6324$.

The norm of the *second* discrete derivative of x is defined by $(\|L_2 x\|_2)^2 = (0.0033 - 2(0.0133) + 0.0300)^2 + (0.0133 - 2(0.0300) + 0.0533)^2 + \cdots + (\ldots - 2(2.8033) + 3.0000)^2$ so that $\|L_2 x\|_2 = 0.0353$.

The first discrete unscaled derivative is much larger than the second discrete unscaled derivative. Also, the norm of x is larger than either unscaled derivative. When it is known that any one of these norms should be small for a good inverse solution, one can impose constraints on these norms to ensure that quality. The methods of regularization seek to accomplish two goals [Hansen, 1992, Hanke and Hansen, 1993, Engl, 1993]:

- To implement side constraints so as to ensure stable solutions.
- To determine proper weights to place on these side constraints.

The L-curve Criterion, Cross Validation, and the Discrepancy Principle, are all methods that accomplish these goals [Hansen, 1992].

2.2.1 L-curve Principle

Because a linear Fredholm integral equation may be discretized, we have at present the general problem

$$\min_x \|Ax - b\|_2^2 \tag{2.16}$$

where x and b are column vectors and A is a matrix.

To smooth out the ill-conditioned tendencies we append a penalty function, referred to as the *regularization term*, involving a discrete derivative operator [Hansen, 1992]:

$$\min_x \{ \|Ax - b\|_2^2 + \lambda^2 \|L_i x\|_2^2 \}. \tag{2.17}$$

The choice of derivative operator is important. Solutions could be sensitive to this choice. Of course this requires some assumption on the solution shape. In the presence of no information one might try several derivative operators and compare resulting solutions.

A more difficult subproblem is choosing λ^2, the weight placed on the regularization term. *Note that any good inverse solution, x^*, will impose a small residual error $\|Ax^* - b\|_2^2$ and a small regularization term $\|L_i x^*\|_2^2$.*

The L-curve principle is as follows [Hansen, 1992]. Repeatedly solve the modified minimization problem with several values of λ^2. Then for each value of λ^2, plot the regularization term against the residual term for the corresponding inverse solution, x_λ^*. In other words plot the set of points

$$(\|Ax_\lambda - b\|, \|L_i x_\lambda\|). \tag{2.18}$$

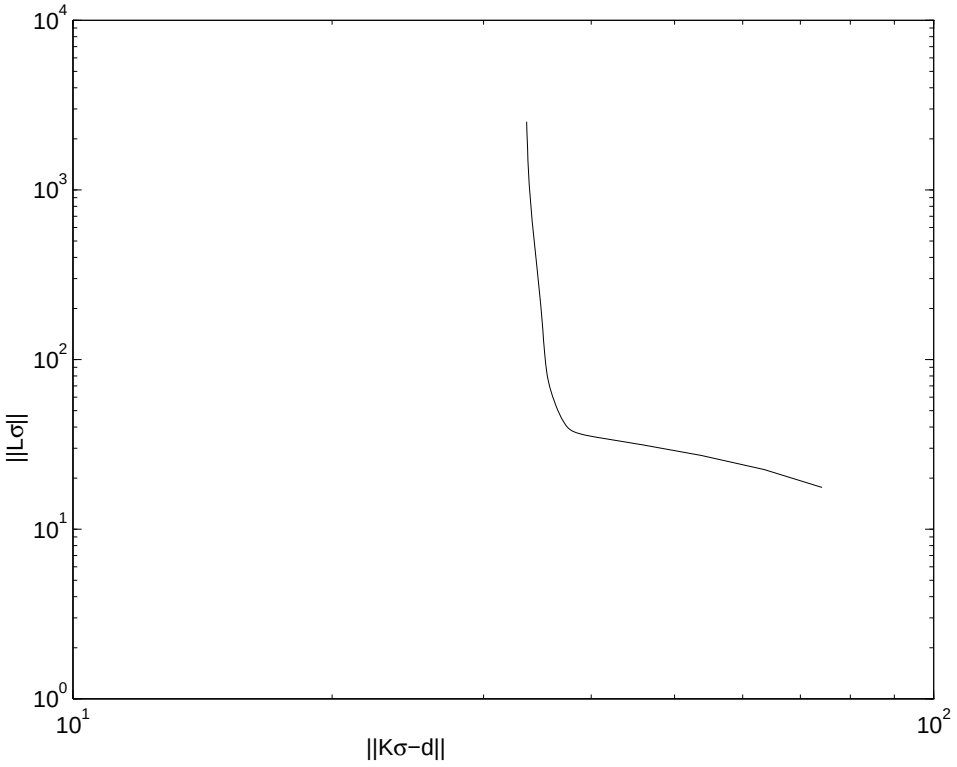

Figure 2.2: Example of an L-curve.

This process will trace out an "L-curve" as in figure 2.2.

In practice this curve tends to be an "L" with a definite corner. As λ^2 increases, flow moves down the curve. This is because a large value of λ^2 is associated with a large weight on the regularization term $\|L_i x\|_2^2$. Then to properly minimize $\|Ax - b\|_2^2 + \lambda^2 \|L_i x\|_2^2$, x must be such that $\|L_i x\|_2^2$ is very small. The residual term $\|Ax - b\|_2^2$ is allowed to grow. Such a point tends to be far from the y-axis and close to the x-axis.

The best choice of λ^2 will produce a solution corresponding to the L-curve corner. To see why it is best, one can view the corner of figure(2.2) and imagine a small improvement being made on either term. If one wants a solution that is slightly more smooth than the corner solution, then a large sacrifice in sum squared error must be made. If one wants a slightly smaller sum squared error, then a large sacrifice must be made in smoothness.

The L-curve criterion can, however, fail to produce a distinct corner. In this case one may turn to cross validation or the discrepancy principle.

2.2.2 Cross-Validation

As with the L-curve criterion, ordinary cross validation (OCV) is a method of choosing a good value of λ^2 with respect to (2.17)

$$\min_x \|Ax - b\|_2^2 + \lambda^2 \|L_i x\|_2^2.$$

But for notational convenience this section will discuss the identical problem:

$$\min_x \sum_{i=1}^{M} (a_i x - b_i)^2 + \lambda^2 \sum_{i=1}^{M} (l_{i(p)} x)^2. \tag{2.19}$$

Here a_i is the i^{th} row of A and $l_{i(p)}$ is the i^{th} row of the p^{th} discrete derivative operator L_p. In other words, $l_{i(p)}x$ is the i^{th} element of $L_p x$. Again $\|Ax - b\|_2^2$ measures the amount of solution error and $L_i x$ is a quality desired to be small.

OCV uses the idea that if some element of b, say b_i, is eliminated from b then a good value of λ^2 will still yield a good solution [Hansen, 1992, Wahba, 1990]. It will be good in that the resulting solution x_λ^* will predict a vector b' that can be used to interpolate the missing element, b_i. Of course the misfit of this prediction is $|b_i - b_i'|$.

Given a *fixed* λ^2, one could perform this prediction process M times and sum the squares of the misfits. Then we have

$$\sum_{i=1}^{M}(a_1 x_{\lambda,i}^* - b_i)^2 \tag{2.20}$$

where $x_{\lambda,i}^*$ minimizes

$$\sum_{j=1,j\neq i}^{M}(a_j - b_j)^2 + \lambda^2 \sum_{j=1,j\neq i}^{M}(l_{j(p)}x)^2 \tag{2.21}$$

Letting λ^2 vary we have the function

$$V(\lambda) = \sum_{I=1}^{M}(a_i x_{\lambda,i}^* - b_i)^2 \tag{2.22}$$

Finally we pose an "external" minimization problem,

$$\min_\lambda V(\lambda). \tag{2.23}$$

OCV provides an alternative method of finding a good regularization weight. It can be, however, computationally expensive. Each instance requires that (2.17) be solved M times.

2.2.3 Discrepancy Principle

The *discrepancy principle* is of particular interest when there is no L-curve corner and when OCV is computationally difficult to implement. Here we still assume that a good solution x, produces a small $\|Ax-b\|_2$ and $\|L_i x\|_2$. We go further to assume an explicit bound on the size of $\|Ax-b\|_2$, say ϵ. We are left with a second rendition of the original minimization problem [Hansen, 1992]:

$$\min_x \|L_i x\|_2$$

$$\text{s.t. } \|Ax - b\|_2 \leq \epsilon. \tag{2.24}$$

This method has intuitive appeal over the L-curve criterion and OCV. The column vector b usually consists of measured values. An experimenter often has some knowledge of total error involved in the measurements. The norm $\|Ax - b\|_2$ is a measure of fitted error so we take ϵ to be the estimated total error in measuring b. Solutions tend to be, however, very sensitive to the choice of ϵ. Furthermore, the measurement error is itself a random variable. If the actual measurement error is less than the expected measurement error, then the solution will be overregularized. If the actual measurement error is too large, then the solution will be underregularized resulting in a noisy solution.

2.3 Non-Linear Problems

Chapter 2 has, so far, outlined the process of discretizing linear Fredholm integral equations and solving the resulting ill-conditioned linear least squares problems. The second half of this thesis will compare a linear and non-linear inverse problem. A word should be said of non-linear inverse problems.

As long as a non-linear forward problem maps vectors x onto vectors y, a minimization problem such as (2.24) may be set up. Any of the methods described in this chapter may be used to pick a regularization parameter.

For example, let $F : x \in R^n \rightarrow y \in R^m$. Then we have the nonlinear L-curve problem

$$\min_x \{ \|F(x) - b\|_2 + \lambda^2 \|L_i x\|_2 \} \tag{2.25}$$

and the nonlinear discrepancy problem

$$\min_x \|L_i x\|_2$$

$$\text{s.t. } \|F(x) - b\|_2 \leq \epsilon. \tag{2.26}$$

In the nonlinear case there can be an added difficulty of multiple local minima causing optimization routines to stall before finding global minima.

Chapter 3

The Geonics EM-38 Problem

3.1 The Inversion Problem Defined

Having examined the techniques of regularization and the need for them, chapter 3 provides the physical background for a soil hydrology problem. Chapter 3 concludes with a discussion of two different predictive models. In the remaining chapters we apply the introductory material to this problem and discuss the results.

A soil's *salinity* is, by definition, the level of dissolved inorganic solutes in the soil. There are many instances in which knowing soil salinity is important. For example, if irrigated agriculture is to remain sustainable, the salinity within these soils must remain at a tolerable level [Rhoades, 1993].

These solutes tend to be conductive. When the electrical conductivity of soil, referred to as ECa, can be directly measured, salinity can be immediately determined. Invasive measurements are, however, both costly and time consuming. This brings us to the focus of the second half of the thesis: *the non-invasive inversion of soil conductivity from above ground measurements.*

Consider an instrument containing two coils of wire. An alternating current is sent through the transmitting coil. This creates an alternating magnetic field that induces current flow in the underlying soil. These currents create secondary magnetic fields. Finally the combination of fields induces a

secondary voltage in the receiving coil of the EM-38. The instrument would measure the relative strength of these secondary fields. [Borchers et al., 1997, Wait, 1982, McNeill, 1985, McNeill, 1990].

Geonics Limited markets a device called the *EM-38 Soil Electrical Conductivity Meter*. This machine will be referred to as the EM-38. It is essentially a hand held bar separating two coils of wire. The coils are fixed with respect to each other and are exactly 1 meter apart. The coils can be positioned either vertically or horizontally. It should be noted that the coil orientation drastically affects the measurements taken. This study uses *both* vertical and horizontal measurements at any given height [Borchers et al., 1997, Rhoades, 1993].

For future reference, the problem of predicting EM-38 measurements from a vertical electrical conductivity profile will be referred to as the *forward EM-38 problem*. Inverting the profile from the EM-38 measurements is then the *inverse EM-38 problem*. The remainder of this chapter will discuss a linear and nonlinear model for the forward and inverse problems.

3.2 The Linear Model

The *linear model* can be stated as follows [Borchers et al., 1997, McNeill, 1980]

$$m^V(h) = \int_0^\infty \phi^V(z+h)\,\sigma(z)\,dz, \tag{3.1}$$

$$m^H(h) = \int_0^\infty \phi^H(z+h)\,\sigma(z)\,dz, \tag{3.2}$$

where

$$\begin{aligned} m^H(h) &\leftarrow \text{predicted horizontal measurement at height } h, \\ m^V(h) &\leftarrow \text{predicted vertical measurement at height } h, \\ \sigma(z) &\leftarrow \text{soil conductivity } z \text{ units below the surface}, \\ \phi^H(\cdot),\, \phi^V(\cdot) &\leftarrow \text{sensitivity functions.} \end{aligned}$$

The sensitivity functions are defined as

$$\phi^V(z) = \frac{4z}{(4z^2+1)^{\frac{3}{2}}}, \tag{3.3}$$

$$\text{and} \quad \phi^H(z) = 2 - \frac{4z}{(4z^2+1)^{\frac{1}{2}}}, \tag{3.4}$$

and represent the relative weight placed on the conductivity at depth z. They are, in fact, linear approximations derived from the nonlinear model discussed in the next section. Figure 3.1 displays these functions. While the coils are in a horizontal position, the instrument is highly sensitive to conductivities near the surface. Measurements taken in a vertical position are, however, more sensitive to conductivities slightly beyond the surface. In either case $\lim_{z\to\infty} \phi^\bullet(z) = 0$.

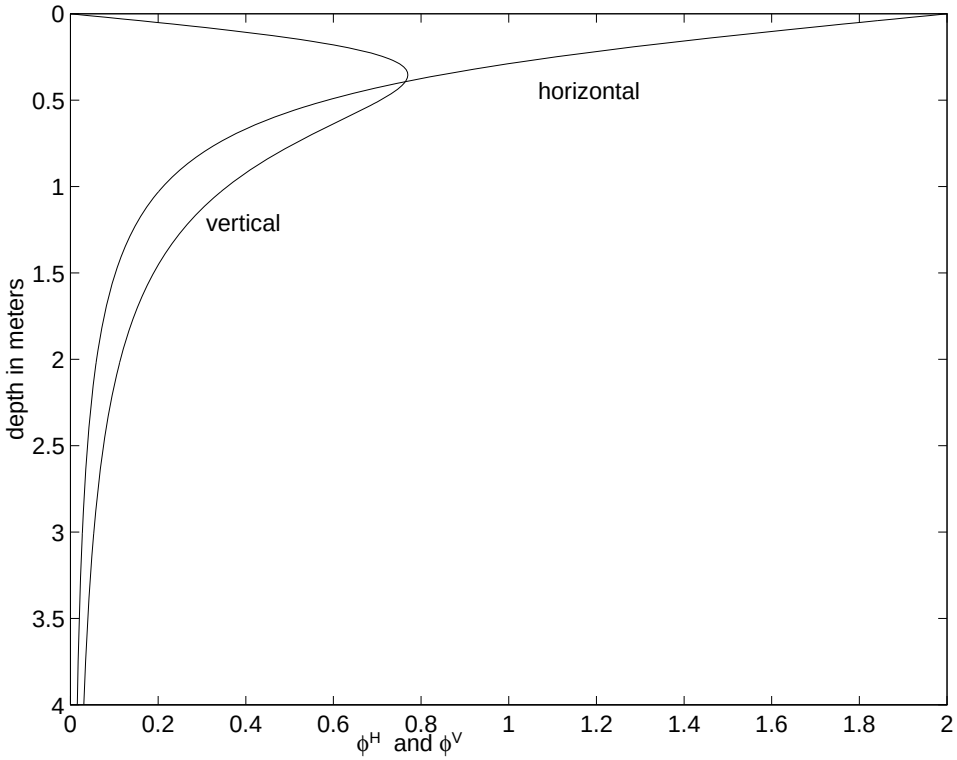

Figure 3.1: Sensitivity Functions ϕ^V and ϕ^H.

When the EM-38 device is held at height h, the additional displacement must be taken into account. The kernels become $\phi^H(z+h)$ and $\phi^V(z+h)$. Finally, the equations are similar in form to the Fredholm integral equation of the first kind (2.1).

In fact they are discretized in the same fashion yielding the discrete forward and inverse problem $K\sigma = b$ and $\min_\sigma \|K\sigma - b\|,\ \sigma \geq 0$ respectively. Since conductivities are non-negative, we constrain σ to be greater than or equal to 0. Appending the weighted penalty term from chapter 2, we are left with the problem:

$$\begin{aligned} \min_\sigma \quad & \{\|K\sigma - d\| + \lambda^2 \|L_i \sigma\|\} \\ \text{s.t.} \quad & \sigma \geq 0. \end{aligned}$$

In this thesis we attempt to solve 14 different instances of this problem. Each instance is based on a data set taken by Rhoades, Corwin, and Lesch [Rhoades et al., 1990]. These data sets will be introduced in chapter 4. In each case, the L-curve criterion was used to find a "good" value of λ^2. All minimization is performed in MATLAB using the *nnls* command. This command implements the non-linear least squares algorithm, *NNLS* [Lawson and Hanson, 1974].

It is widely known (and will be shown in subsequent chapters) that this model breaks down in the presence of high conductivities. This flaw induces the need for the following nonlinear model.

3.3 The Non-Linear Model

The linear model maps continuous profile functions $\sigma(z)$ into continuous measurement functions, $m^V(h)$ and $m^H(h)$. Due to its form it is easily

discretized using chapter 1 techniques. On the other hand, the nonlinear model must initially assume a discretized soil profile. It then maps these conductivity vectors into a discrete function describing EM-38 measurements at different heights. Assume then, that the soil is discretized into M layers where the M^{th} layer is semi-infinite. Let d_i represent the thickness of the i^{th} layer and let σ_i be the conductivity of this layer. This discretization is seen is figure (3.2).

The measurement at height h, is described by the following equations [Wait, 1982].

$$m^V(h) = \mathrm{Im}(1 + B^3 T_0) \quad \text{and} \quad m^H(h) = \mathrm{Im}(1 + B^2 T_2) \tag{3.5}$$

$$T_0 = \int_0^\infty -R_0(\frac{gB}{r})\, g^2\, e^{\frac{-2gh}{\delta}}\, J_0(gB)\, dg, \tag{3.6}$$

$$T_2 = \int_0^\infty -R_0(\frac{gB}{r})\, g\, e^{\frac{-2gh}{\delta}}\, J_1(gB)\, dg, \quad \text{and} \tag{3.7}$$

$m^V(h) \leftarrow$ height h measurement with EM-38 in vertical orientation,

$m^H(h) \leftarrow$ height h measurement with EM-38 in horizontal orientation,

$Im(z) \leftarrow$ imaginary part of a complex number, z,

$f \leftarrow$ EM-38 frequency(14.6 khz),

$\omega \leftarrow 2\pi f$,

$r \leftarrow$intercoil spacing: 1 meter for the EM-38,

$\mu_0 \leftarrow$magnetic permeability of air: $4\pi 10^{-7} henry/m$,

$\mu_i \leftarrow$ magnetic permeability of the i^{th} layer,

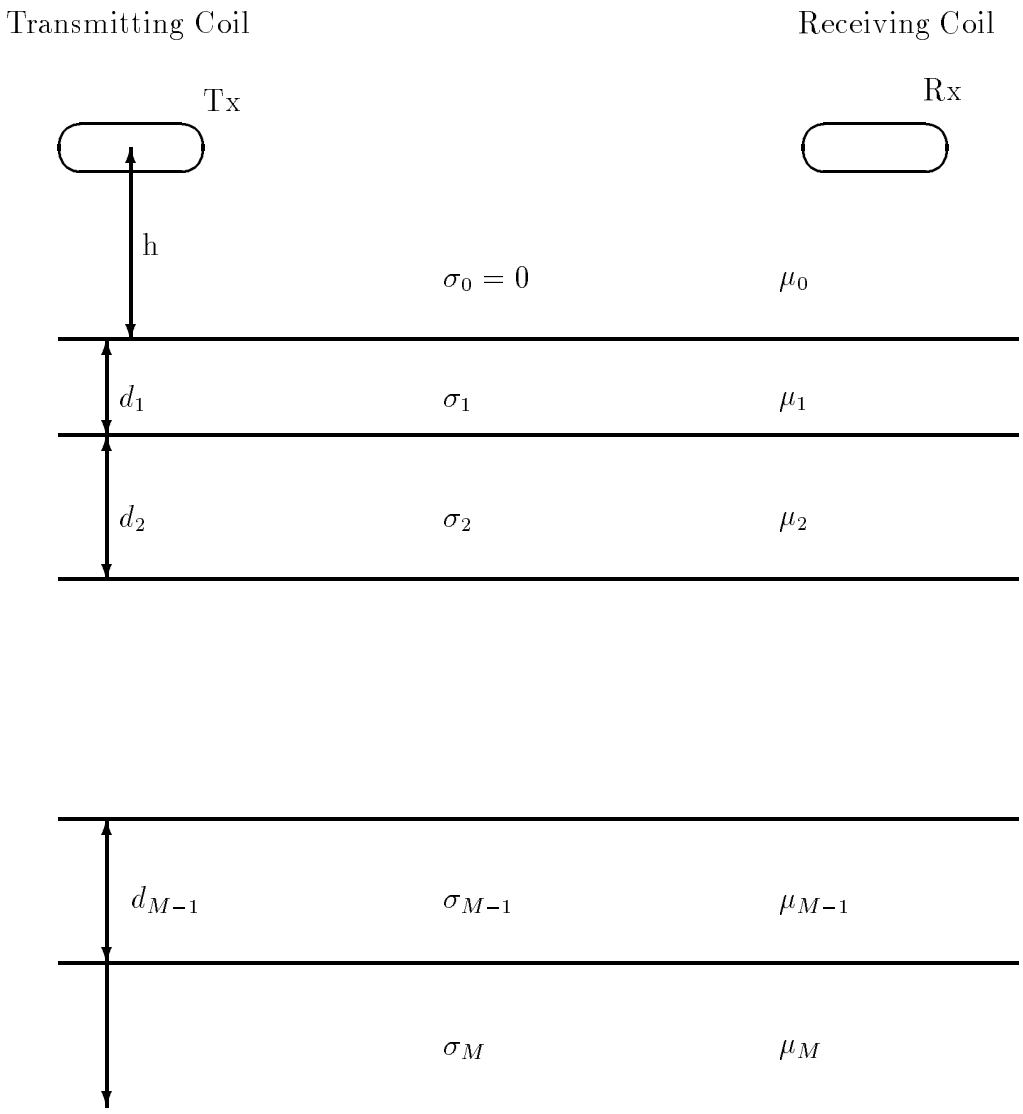

Figure 3.2: Soil Discretization for Nonlinear Model.

$\delta \leftarrow \sqrt{\frac{2}{\sigma_1 \mu_0 \omega}}$,

$B \leftarrow \frac{r}{\delta}$,

$J_0,\ J_1 \leftarrow$Bessel functions of orders 0 and 1.

Finally we define R_0 as follows

$$R_0(\lambda) = \frac{N_0 - Y_1}{N_0 + Y_1} \tag{3.8}$$

where

$$Y_1 = N_1 \left(\frac{Y_2 + N_1 \tanh u_1 d_1}{N_1 + Y_2 \tanh u_1 d_1} \right), \tag{3.9}$$

$$Y_2 = N_2 \left(\frac{Y_3 + N_2 \tanh u_2 d_2}{N_2 + Y_3 \tanh u_2 d_2} \right), \tag{3.10}$$

and so on to

$$Y_{M-2} = N_{M-2} \left(\frac{Y_{M-1} + N_{M-2} \tanh u_{M-2} d_{M-2}}{N_{M-2} + Y_{M-1} \tanh u_{M-2} d_{M-2}} \right), \tag{3.11}$$

$$Y_{M-1} = N_{M-1} \left(\frac{N_M + N_{M-1} \tanh u_{M-1} d_{M-1}}{N_{M-1} + N_M \tanh u_{M-1} d_{M-1}} \right), \tag{3.12}$$

and where N_k and u_k are given by

$$N_k = \frac{\sqrt{\lambda^2 + i\sigma_k \mu_k \omega}}{i\mu_k \omega}, \qquad i = \sqrt{-1} \tag{3.13}$$

$$u_k = \sqrt{\lambda^2 + i\sigma_k \mu_k \omega}. \tag{3.14}$$

Here Y_k is the *surface admittance* at the top of the i^{th} layer and N_k is the *characteristic admittance* of the i^{th} layer [Nabighian, 1987, Wait, 1982]. It should be noted that the magnetic permeability of soil at the specific sites in question, and in most soils, is essentially equal to the magnetic permeability of air. Thus we make the simplifying assumption $\mu_i = \mu_0$.

The nonlinear model is much more computationally intensive than the preceding linear model. Assuming that a discrete conductivity profile is known, each integral must be numerically computed N times to calculate N forward predictions. This was not true for the linear model which could be integrated by hand. To make practical use of this model, an efficient method of calculating the integrals must be used.

When B is taken to be a variable, equations (3.6) and (3.7), become Hankel transforms with kernels,

$$-R_0(\frac{gB}{r})\, g\, e^{\frac{-2gh}{\delta}} \quad \text{and} \quad -R_0(\frac{gB}{r})\, g^2\, e^{\frac{-2gh}{\delta}}. \tag{3.15}$$

Although B depends only on σ_1, the kernels of the Hankel transforms depend upon h, the height at which the EM-38 is displaced above the ground. For each value of h, we calculate the Hankel transform at the designated value of B. A FORTRAN code was published in 1979 by Walter L. Anderson that does this [Anderson, 1979]. Finally, it should be noted that the vast majority of computational time in solving the nonlinear inverse problem is spent in this integration routine. It is no trivial task.

The techniques of regularization that we have previously discussed, may be used to condition this problem. Let vectors d and σ represent the thicknesses and conductivities of some proposed profile. Then let $F(d,\sigma)$ be the set of M predicted measurements given this profile. We are finally left with the discrete inverse problem:

$$\begin{aligned} \min_\sigma \quad & \|F(d,\sigma) - m_e\| + \lambda^2 \|L_i \sigma\| \\ \text{s.t.} \quad & \sigma \geq 0 \end{aligned}$$

where m_e is an actual set of EM-38 measurements.

Whereas MATLAB was a suitable environment to run the linear forward and inverse problems, a FORTRAN code had to be constructed to solve the related nonlinear problems. The nonlinear least squares portion was solved with a double precision IMSL subroutine called DBCLSF. To guard against the possibility of finding a local minimum point a multistart approach was used.

First a set of logarithmically spaced λ^2s was chosen. For each value of λ^2, three initial guesses were chosen. After the three runs were completed the "best" solution was taken to be the solution with the lowest function value. Finally an L-curve was constructed. This study constructed L-curves by fitting cubic splines around the points $(\|K\sigma_\lambda - \text{data}\|,\ \|L\sigma_\lambda\|)$. Then a new plot was made for each site of the calculated curvature vs. the value of λ^2. The highest point on this curve is the point of greatest curvature or the "L-curve corner".

ECa readings are known to rarely fall at or above 3000 mS/m and never below 0 mS/m within the first few meters of earth. Occasionally this "L-curve corner" yields a solution that is underregularized in that the maximum and minimum conductivities are either zero or too large. With this in mind an additional rule was adopted: if a solution reaches either 0 or 3000 mS/m before 3 meters, then λ^2 is increased until a solution within these constraints is found.

The only detail left to define is the choice of initial guesses for the multistart method. For the sites discussed in chapter 3, the largest ECa reading is about twice the largest EM-38 reading. Thus one initial guess is taken to be a constant vector of the largest reading and another initial guess is taken to

be the twice the largest guess. The third initial guess is taken to be the best solution from the last set of three runs.

Chapter 4

Forward Results

Chapters 4 and 5 compare the performance of the two preceding models using data sets from 14 different sites. This data is supplied to us through a technical report by Rhoades, Corwin, and Lesch [Rhoades et al., 1990]. We will begin chapter 4 with a discussion of this data. Then this data will be used to compare forward predictions. In chapter 5 we will compare inverse solutions.

4.1 The Rhoades, Corwin, and Lesch Data Sets

A total of 14 sites throughout California were chosen. At each site detailed soil electrical conductivity, ECa, and Geonics EM-38 measurements were taken.

Before actually taking ECa readings, the top soil was leveled off and dry mulch was removed. Readings were taken at depths of 5, 10, 15, 20, 50, 75, 100, 125, 150, 175,..., 275, and 300 cm. At each depth, 9 readings were taken. These 9 readings were spaced on a 60cm$\times$60cm grid. Measurements down to 20 cm were taken by vertically inserting a "Martek bedding probe." A Martek "Rhoades-probe" was inserted horizontally for depths 25-200 cm. To make this possible a trench was excavated in the E-W direction. To sample the final meter, the top 2 meters were stripped off and a "Rhoades-probe" was inserted vertically.

At each site, EM-38 measurements were taken at heights 0, 10, 20, 30, 40, 50, 75, 100, and 150 cm. At each height, 8 measurements were taken. For both horizontal and vertical orientations of the EM-38, a measurement was taken with the instrument facing directions (N, NE, E, SE).

4.2 Discussion of Forward Plots

The forward plots in appendix A were created in the following manner. At each depth, the measured soil conductivities were averaged. This profile was then used to interpolate the conductivities of 10 cm layers down to 3 meters with a semi-infinite layer beginning at the 3 meter mark. It was assumed that the conductivity of the semi-infinite layer is the conductivity of the last thin layer. It was found that the forward predictions are sensitive to this assumption. More will be said of this later. Finally, using the linear MATLAB code and the nonlinear FORTRAN code of the previous chapter, forward EM-38 predictions were made at heights 0, 10, 20, 30, 40, 50, 75, 100, and 150 cm.

Then for each site, two plots were made. One assuming a vertical orientation of the instrument and one assuming a horizontal orientation. Each contains two dotted lines representing predictions of the linear and nonlinear models, as well as a solid line representing the average EM-38 measurement at each height. There are a total of 28 forward plots.

The following observations are a visual interpretation of the 28 plots.

Observation 1: Without exception, the nonlinear model outperforms the linear model. In every plot the nonlinear predictions lie closer to the EM-38 measurements than the linear predictions.

Observation 2: The linear model always over predicts the EM-38 measurements. In fact they are usually over predicted by a substantial amount. It is not unusual for the linear predictions to be $50\% - 100\%$ larger than the actual measurement.

Observation 3: The nonlinear model almost always over predicts EM-38 measurements. But it usually over predicts by less than 30% of the actual measurements. In some cases the nonlinear model's predictions are impressively close to the actual measurements. Examples include sites 4, 6, 7, and 10.

Observation 4: At sites with high soil conductivities, the nonlinear model tends to make large improvements on the forward predictions. At sites with low conductivities, the nonlinear model usually makes only subtle improvements.

To demonstrate observation 4, we will partition the 14 sites into those with "high" conductivities and those with "low" conductivities. The partition will be based on the largest conductivity measurement. Comparing the set of 14 maximum conductivities, it is natural to define a *high conductivity site* to be a site with a maximum reading of at least $500\frac{\text{mS}}{\text{m}}$. A *low conductivity site* will be any site with a maximum reading of under $500\frac{\text{mS}}{\text{m}}$. Refer to table (4.1).

Finally consider an alternative partitioning of the 14 sites. This will be based on the extent to which the nonlinear model outperforms the linear model. On site 2, 9, 11, 12, and 13, the nonlinear model makes modest improvements. This is generally consistent with observation 4 because 3 of the 5 sites have low conductivities. An exception is site 2 which has been categorized as a high conductivity site. In examining the measured conductivity profile (figure

Site	Max Conductivity	Site Type
1	1200 mS/m	High
2	1000 mS/m	High
3	250 mS/m	Low
4	300 mS/m	Low
5	2000 mS/m	High
6	2000 mS/m	High
7	2000 mS/m	High
8	1000 mS/m	High
9	250 mS/m	Low
10	700 mS/m	High
11	1000 mS/m	High
12	300 mS/m	Low
13	300 mS/m	Low
14	200 mS/m	Low

Table 4.1: Site Type Based on Max Soil Conductivity

4.1), we see that the conductivity drops from 1000 $\frac{\text{mS}}{\text{m}}$ to 500 $\frac{\text{mS}}{\text{m}}$ in the first 20 centimeters and continues to decrease. Site 2 might be better categorized as a *low* conductivity site. The nonlinear model only makes a small improvement on the linear model's predictions for site 2 and thus this site is still consistent with observation 4.

The nonlinear model makes a large improvement on sites 1, 3, 4, 5, 6, 7, 8, 10, and 14. This is also generally consistent with observation 4 because 6 of the 9 sites have high conductivities.

Also the computational difficulties of the nonlinear model were touched upon in chapter 3. Given these difficulties and observation 4, the nonlinear model is best fit for sites with high conductivities.

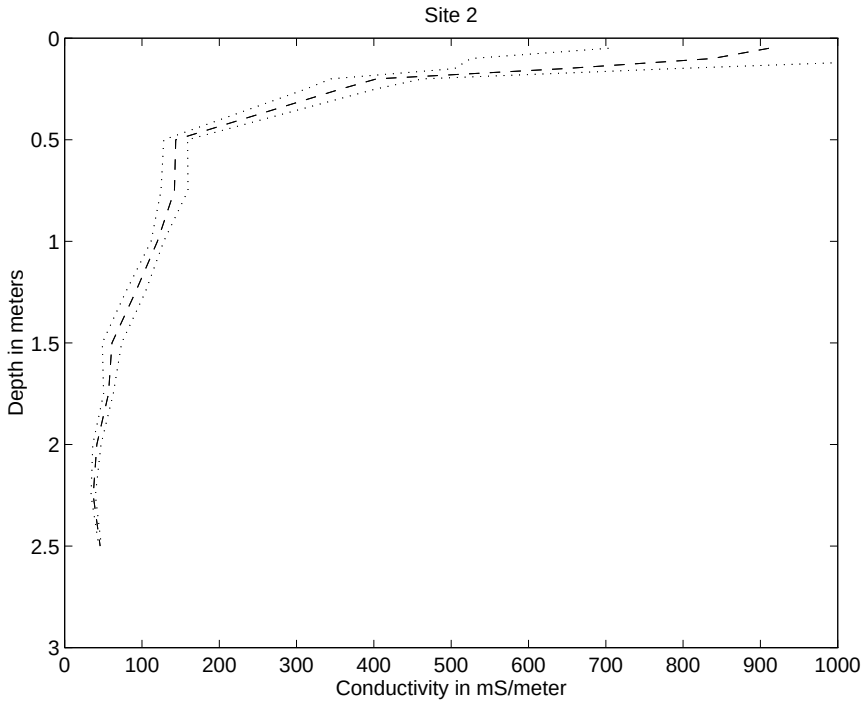

Figure 4.1: Site 2: Measured Conductivity Profile.

Chapter 5

Inverse Results

5.1 Discussion of Inverse Plots

The inverse problem has been solved for each of the 14 Rhoades, Corwin, and Lesch sites using each of the two models. Appendix B contains one plot for each site. Each plot contains a solution for each model and the average measured conductivity at each depth with one standard deviation bars.

The technical details of these runs were outlined in the previous chapter. It should be noted that the L-curve criterion was used to pick the regularization parameter λ^2. Attempts to use cross-validation on the linear model were made. The resulting solutions were by no means better than the L-curve solutions. Given the computationally expensive nature of cross-validation, application to the nonlinear model is currently infeasible. Also, an attempt to use the discrepancy principle on the linear model was made. This attempt was also not successful. The solution is far too sensitive to changes in the upper bound of $\|K\sigma - d\|_2^2$. The L-curve method is inexpensive to implement and relatively stable.

Chapter 4 partitioned the 14 sites into high and low conductivity profiles. It was concluded that the nonlinear model was more appropriate than the linear model for high conductivity sites. We will continue this line of reasoning here.

Again take sites 2, 9, 11, 12, and 13 to be low conductivity sites. For sites 2, 11, and 12 the linear and nonlinear solutions are nearly identical. The nonlinear solution for site 13 is less regularized than the linear solution. It is more regularized for site 9. But for both sites the solution quality is similar.

Now consider the remaining sites. The quality of the solutions for sites 3, 4, 8, 10, and 14 are nearly identical. The nonlinear solution for sites 5, 6, and 7 are excellent whereas the linear solutions fails after 1.5 meters. Here the nonlinear model is more sensitive than the linear model to lower depths. This also will be discussed in the next section. These sites are interesting as the measurements all increase with depth. Nonlinear success for site 7 is also due in part to the second order regularization. This regularization is most advantageous when a profile has little curvature.

Figure (B.8) of the site 8 inverse solutions demonstrates an important point about the solution process: *the choice of* λ^2 *is critical.* The linear solution is able to pick up trends in the first 1.5 meters because it is somewhat "underregularized" and allowed to bend. Perhaps with a good value of λ^2, the nonlinear models could yield an better result. For this site the nonlinear L-curve resulted in large λ^2 resulting in a solution with no curvature.

Chapter 6

Possible Explanations of Unexplained Behaviors

This discussion of influences not included in these models has been reserved for a separate chapter. Note that every influence listed below could potentially effect both forward and inverse calculations.

6.1 Semi-infinite Layer

Both models assume a semi-infinite layer beginning at the 3 meter mark. By semi-infinite we mean that the conductivity is uniform from 3 meters down toward the center of the earth. Some assumption on this conductivity must be made for the forward predictions. The plots included in this paper assume that the conductivity of the bottom layer is the final interpolated conductivity. A set of similar plots was created assuming this conductivity to be zero. This had a large effect on the linear model's predictions and a small effect on the nonlinear model's predictions. We conclude that the linear model is sensitive to conductivities below 3 meters.

6.2 Magnetic Permeability

In the chapter 3 description of the nonlinear model, μ_i represented the magnetic permeability of the i^{th} layer of soil. Also μ_0 was taken to be the magnetic permeability of air. This study has consistently assumed that $\mu_i = \mu_0$

for all i. But for soils containing magnetic material such as magnetite this is not the case. It is known, however, that very little of these materials existed at the Rhoades, Corwin, and Lesch sites.

6.3 Discretization

Both models assume a vertical discretization of the soil profile. Though this cannot be avoided in practice, it always induces some level of error.

6.4 Instrument Calibration

Calibration problems with the instruments used to collect the data could affect the results. This could be the case with either the EM-38 or the $MartekProbe$.

6.5 Temperature Effects

Soil conductivity is affected by temperature. There is always a certain danger in invasive experimentation. The removal of top soil or side soils could have an effect on ECa readings.

6.6 Vertical Effects

Though the sensitivity of the instrument decays with depth, an extremely high conductivity below 3.0 meters could affect results. A water table could, for example, cause a sharp increase in conductivity. Second order regularization would have a difficult time picking out the sharp increase. Perhaps the solution would try to compensate by over predicting at depths directly above the water table.

6.7 Depth Sensitivity

Both models have some difficulty inverting profiles below 1.5 meters. This sensitivity matter is a particular problem for the linear model. This is not surprising. As shown by figure(3.1) the instrument is most sensitive to shallow depths.

6.8 Lateral Homogeneity

Both models make an important assumption: that conductivity varies only with depth. This assumption does not hold in practice. At nearly every depth at nearly every site, there is considerable variability in the measured ECa values. Furthermore at some sights there are strong lateral trends in ECa. Though a thorough sensitivity analysis has not yet been conducted, it seems reasonable that EM measurements would be sensitive to surface or subsurface anomalies beyond the 60 cm by 60 cm square in which the ECa was actually measured. A larger range of data is not currently available.

A forward sensitivity analysis can, however, be performed. Recall that at each depth, 9 measurements were taken. Figures (6.1) and (6.2) depict forward nonlinear predictions for site 1. The dashed line represents forward predictions using only the largest measured conductivity at each depth. Here we use the conservative assumption that the semi-infinite layer is at the last measured conductivity. The dash/dot line represents the predictions using only the lowest conductivity at each depth. Here we use the liberal assumption that the semi-infinite layer is 0 $\frac{mS}{m}$. In neither plot do the EM-38 readings lie within these predictions. It seems that lateral variability in the ECa measurements

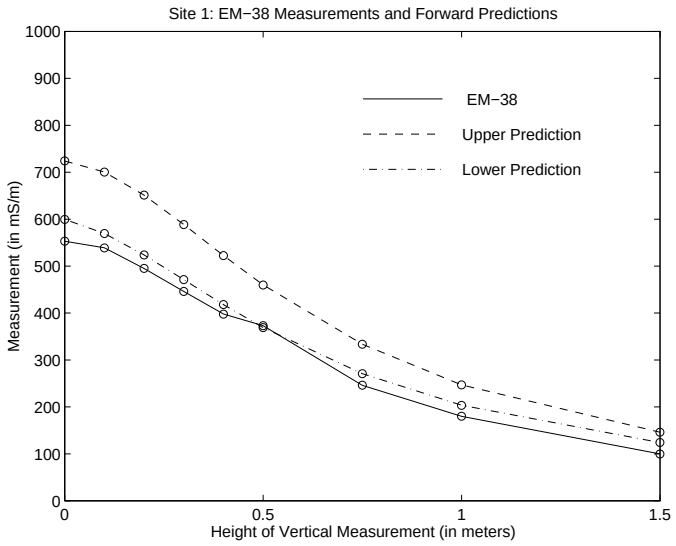

Figure 6.1: Site 1: Nonlinear Vertical Sensitivity Analysis.

taken within the 60 cm $\times$ 60cm grid is not enough to explain the consistent over predictions.

It was shown in the previous chapter that site 1 is the only high conductivity site in which the nonlinear model fails to produce a reasonable solution. The authors of the data set produced a different type of sensitivity analysis that can help explain this result [Rhoades et al., 1990]. Recall that at each depth, ECa readings were taken in a grid pattern. A linear regression can be performed on each layer. Of course the ECa is regressed onto the forward and lateral positions of the data. If the result is a small set of coefficients then we can have confidence that the soil conductivity is truly laterally homogeneous. For site 1, such a regression was significant at the 0.01 level for three depths. The resulting regressions are listed in table (6.1).

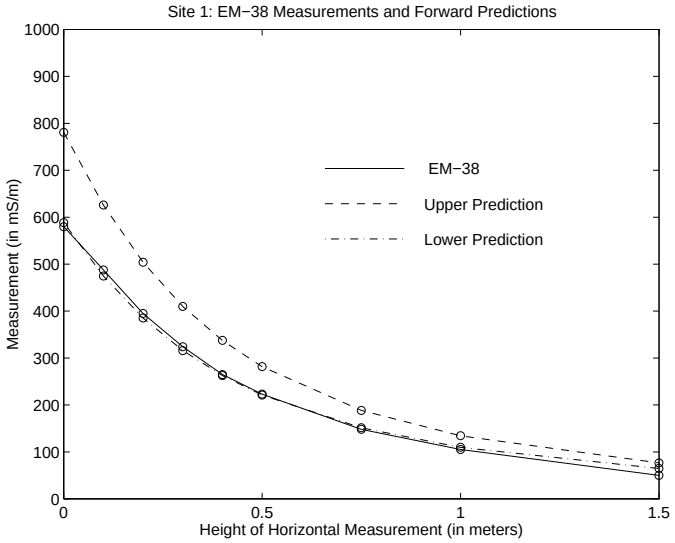

Figure 6.2: Site 1:Nonlinear Horizontal Sensitivity Analysis.

Depth(m)	Model: ECa in $\frac{mS}{m}$, x,y in meters	R^2
1.00	ECa = 965.8 - 42.8x - 254.4y	0.946
1.75	ECa = 973.8 + 149.4x + 193.9y	0.825
3.00	ECa = 567.9 - 131.1x - 46.1y	0.856

Table 6.1: Site 1 Regressions Significant at α=0.01

For example, at 1.75 meters below the earth, a move of 1 meter in the y direction could result in an ECa increase of almost 200 mS/m. With this sort of change it is possible that varying conductivities outside of the measurement grid could have a strong impact on the EM-38 measurements and thus the inverse solutions.

Chapter 7

Summary and Conclusions

The EM-38 inverse problem is only successfully solved as a least squares problem when EM-38 measurements can be accurately predicted from a measured conductivity profile. With this in mind, an extensive effort to compare forward linear and nonlinear predictions has been made. It was found that both models over predicted the EM-38 measurements. The nonlinear model without exception outperformed the linear model in forward predictions. The difference was substantial for sites with high soil conductivities. In fact, the nonlinear model generally produced satisfactory results for our purposes.

In choosing a regularization parameter, the L-curve principle should be used. The discrepancy principle is too sensitive to estimations of ϵ and cross-validation is too computationally expensive.

In some cases the nonlinear model does not outperform the linear model. The linear model, on the other hand, only outperforms the nonlinear model for site 1. The general failure of both models on site 1 can be, however, partially attributed to a lack of lateral homogeneity.

Considering the computational expense of solving the nonlinear model, the linear model may be appropriate for low conductivities and for depths down to 1.5 meters.

This study introduces a comparison of two models and an overview of

how they might be solved. External influences are not completely understood. Further research is needed to understand the influences of lateral anomalies. Perhaps the range of ECa readings could be extended to a 3.0 meter square and a more thorough sensitivity analysis could be performed. Finally, this study has inverted only one dimensional profiles. Two and three dimensional models are being examined [Alumbaugh et al., 1996, Newman and Alumbaugh, 1997, Alumbaugh and Newman, 1997].

Appendix A

Forward Plots

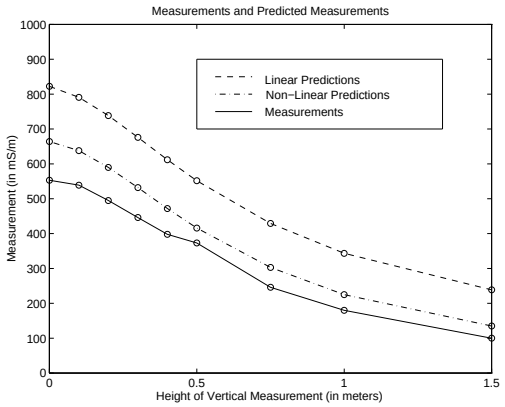

Figure A.1: Site 1: Vertical Orientation.

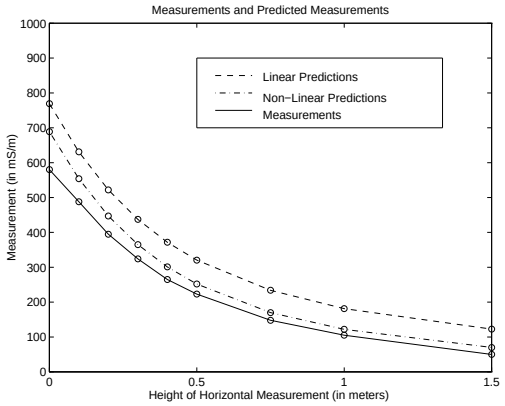

Figure A.2: Site 1: Horizontal Orientation.

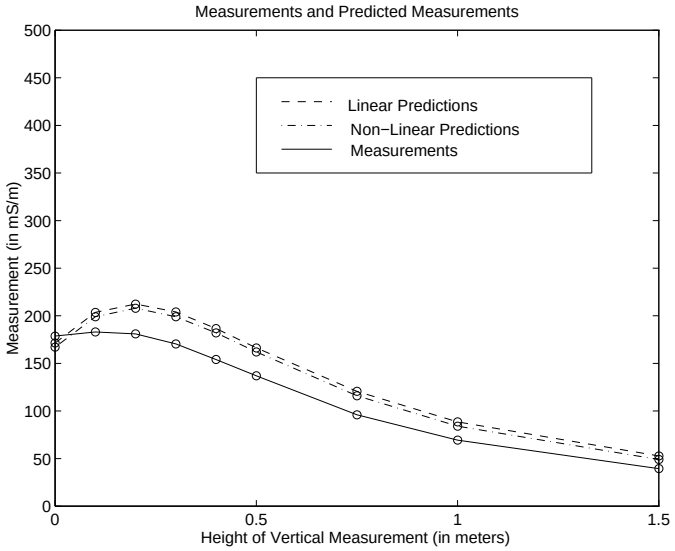

Figure A.3: Site 2: Vertical Orientation.

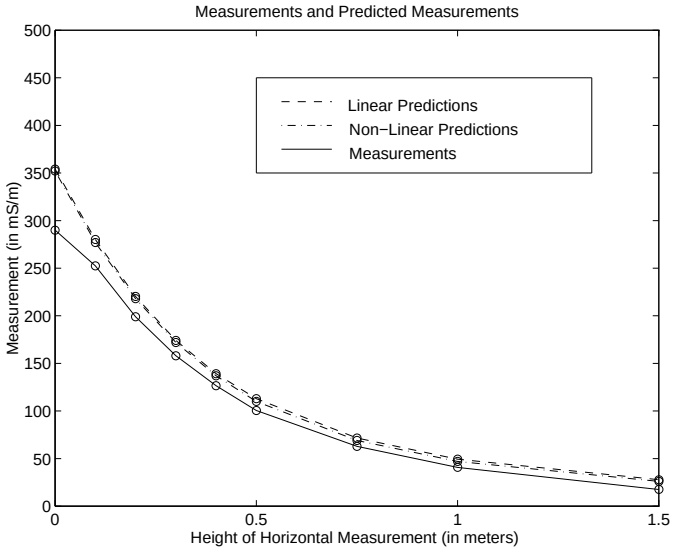

Figure A.4: Site 2: Horizontal Orientation.

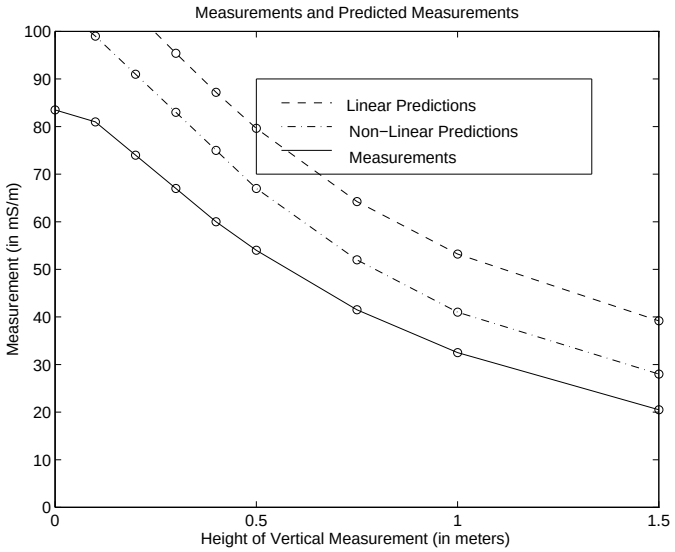

Figure A.5: Site 3: Vertical Orientation.

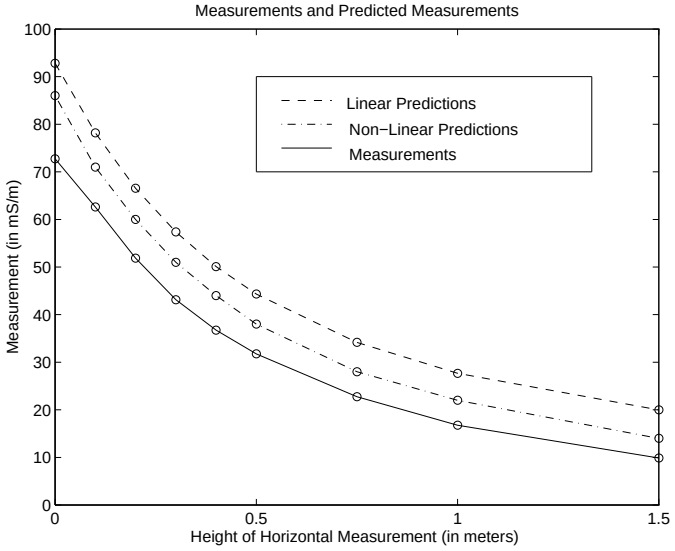

Figure A.6: Site 3: Horizontal Orientation.

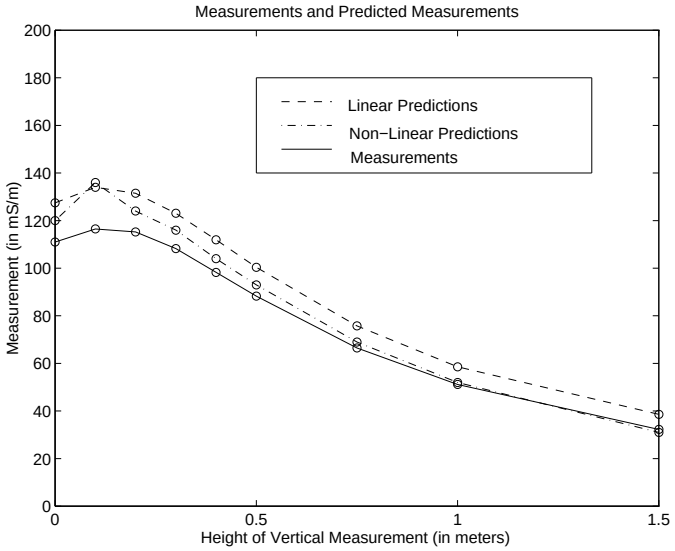

Figure A.7: Site 4: Vertical Orientation.

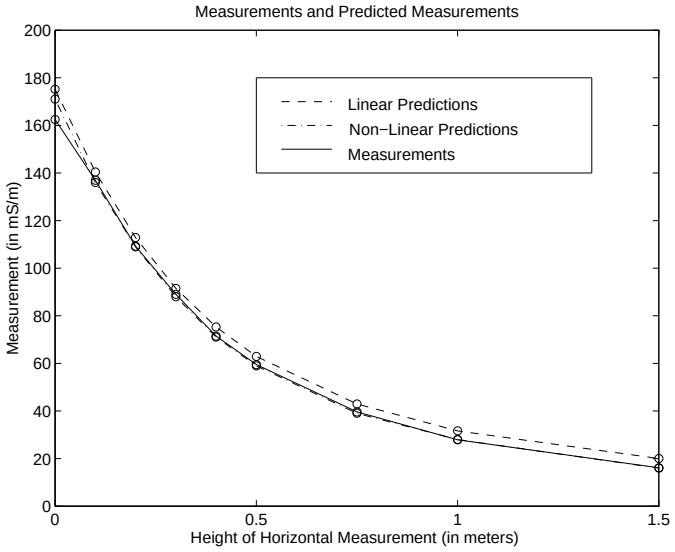

Figure A.8: Site 4: Horizontal Orientation.

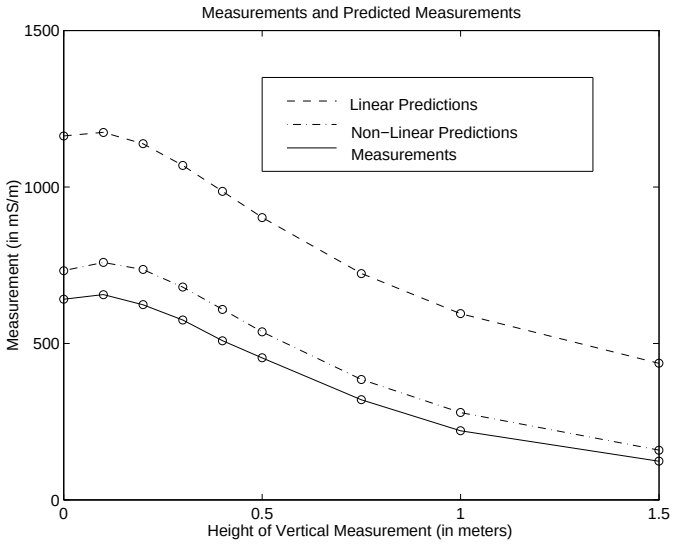

Figure A.9: Site 5: Vertical Orientation.

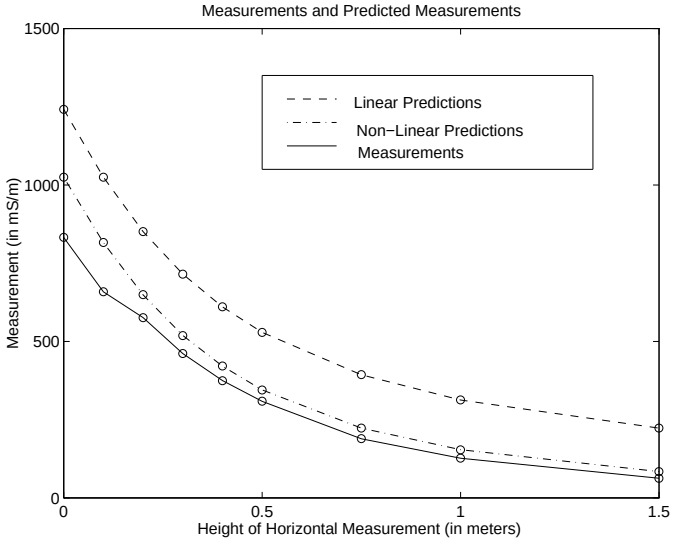

Figure A.10: Site 5: Horizontal Orientation.

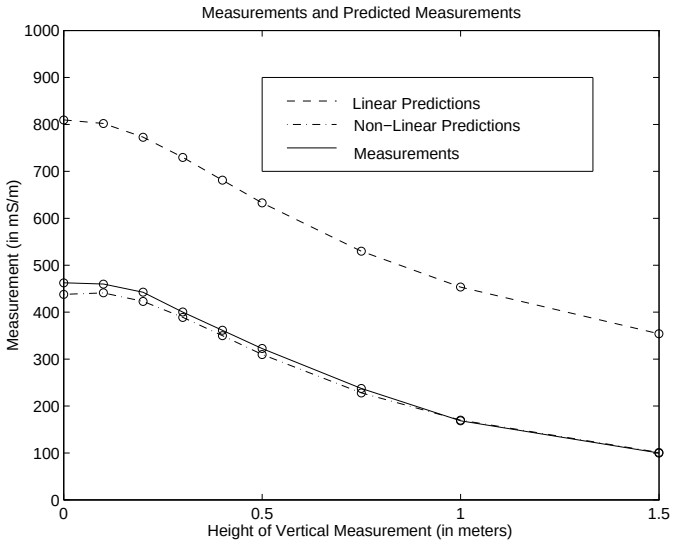

Figure A.11: Site 6: Vertical Orientation.

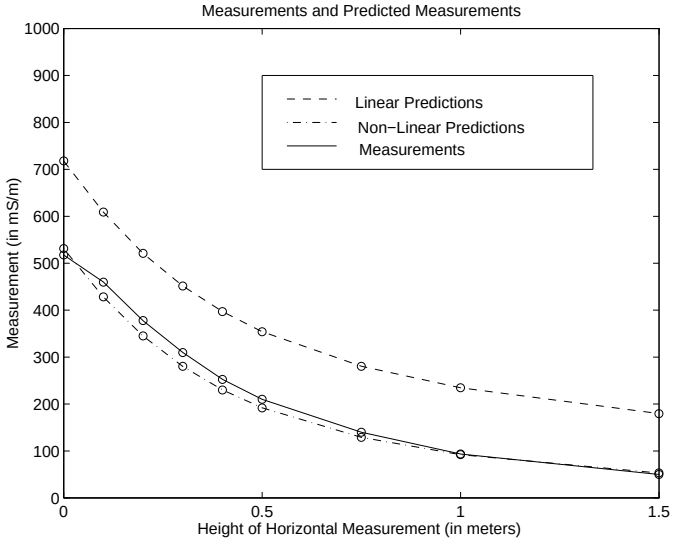

Figure A.12: Site 6: Horizontal Orientation.

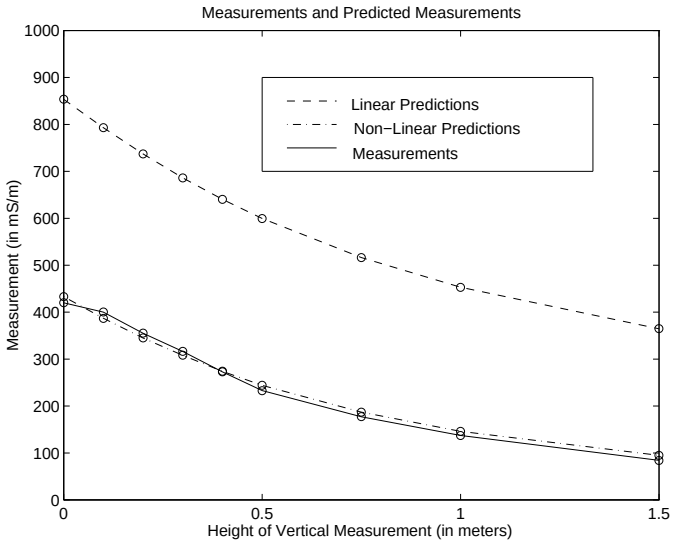

Figure A.13: Site 7: Vertical Orientation.

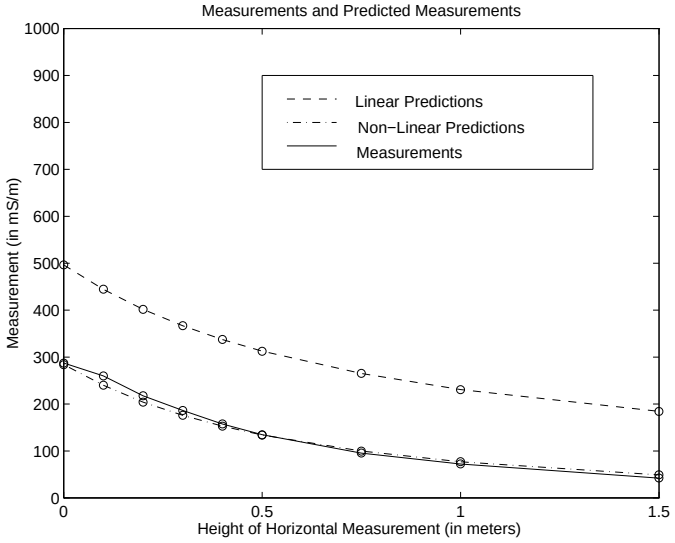

Figure A.14: Site 7: Horizontal Orientation.

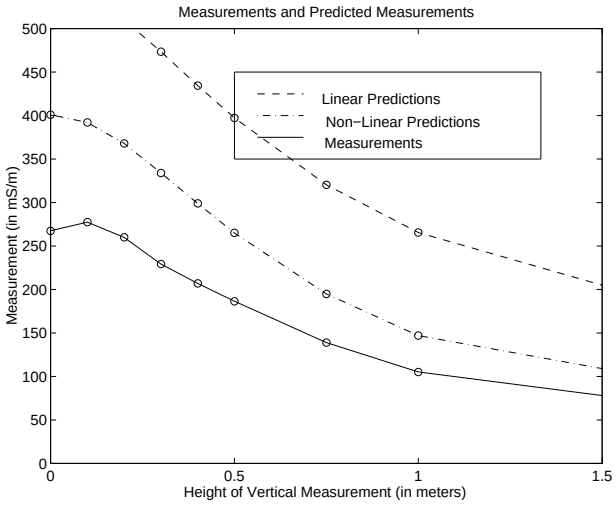

Figure A.15: Site 8: Vertical Orientation.

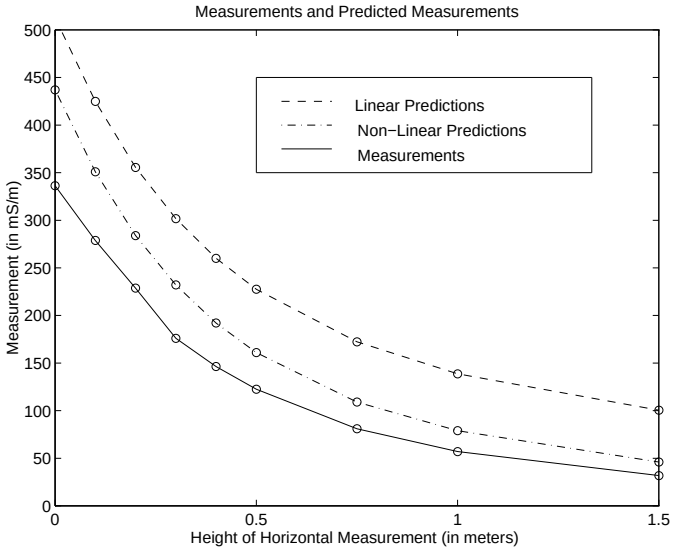

Figure A.16: Site 8: Horizontal Orientation.

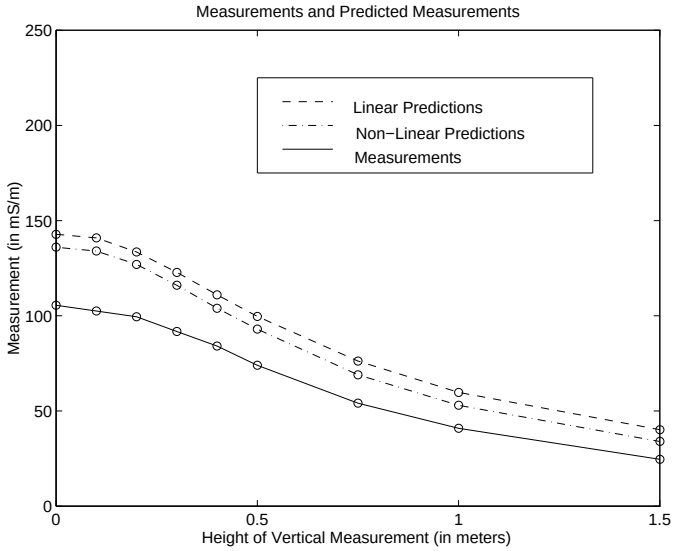

Figure A.17: Site 9: Vertical Orientation.

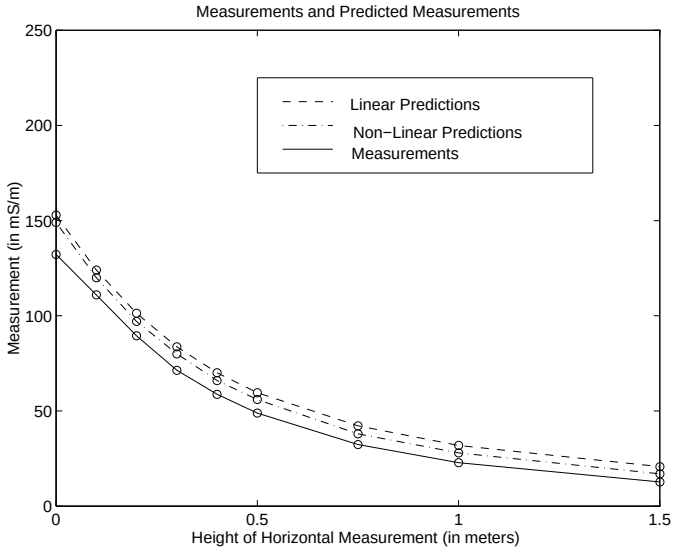

Figure A.18: Site 9: Horizontal Orientation.

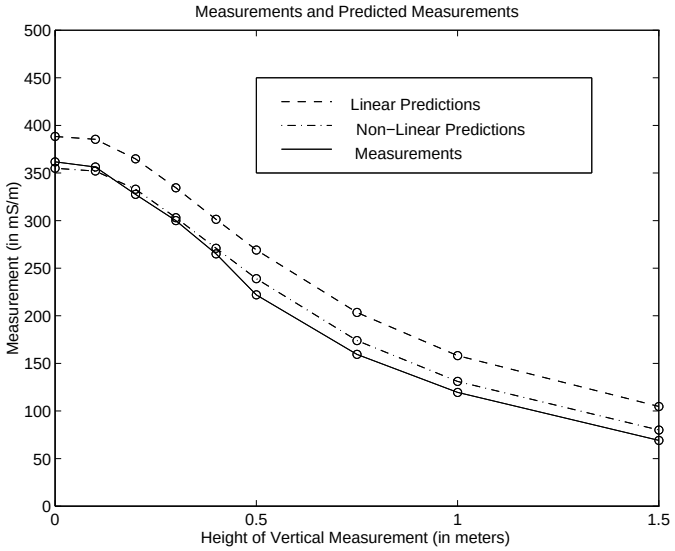

Figure A.19: Site 10: Vertical Orientation.

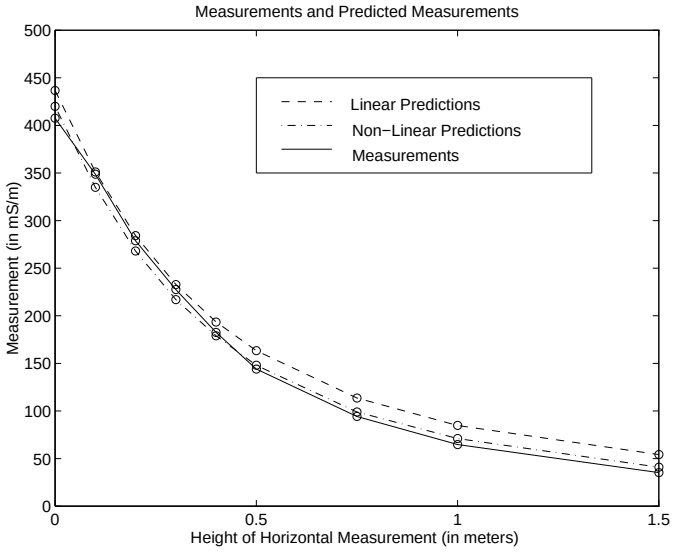

Figure A.20: Site 10: Horizontal Orientation.

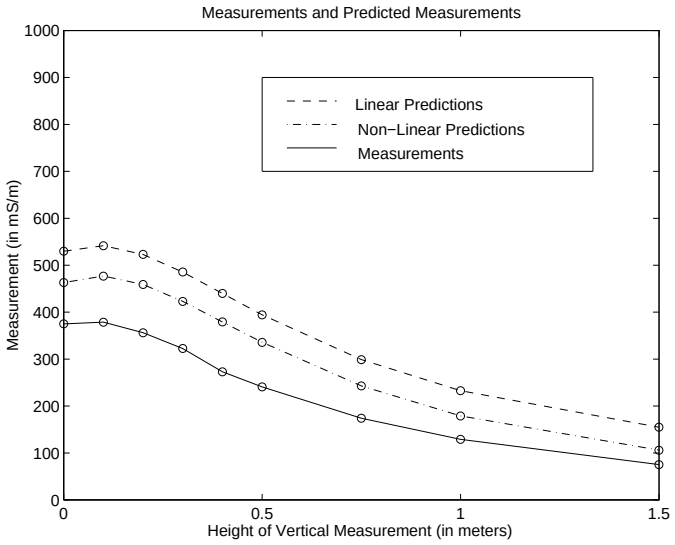

Figure A.21: Site 11: Vertical Orientation.

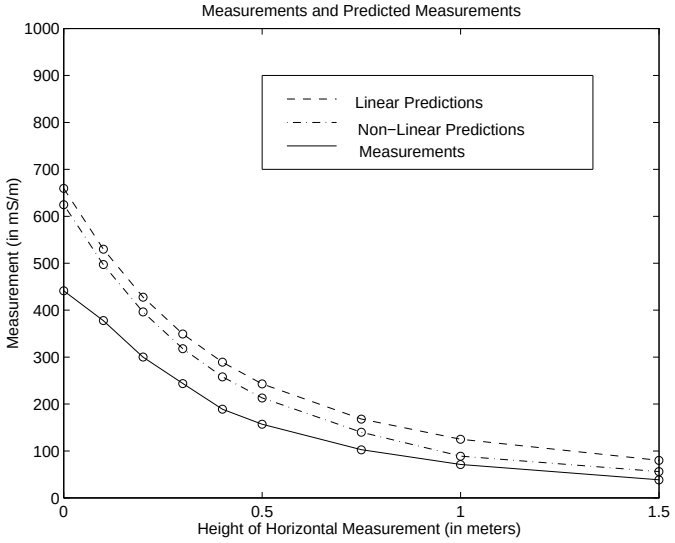

Figure A.22: Site 11: Horizontal Orientation.

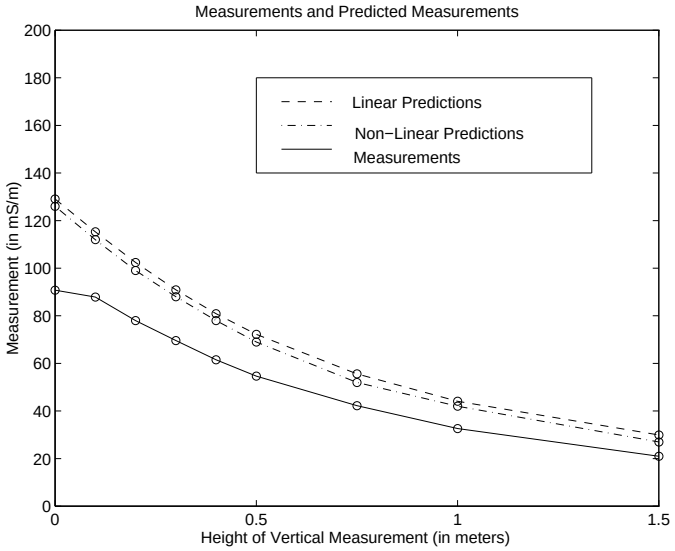

Figure A.23: Site 12: Vertical Orientation.

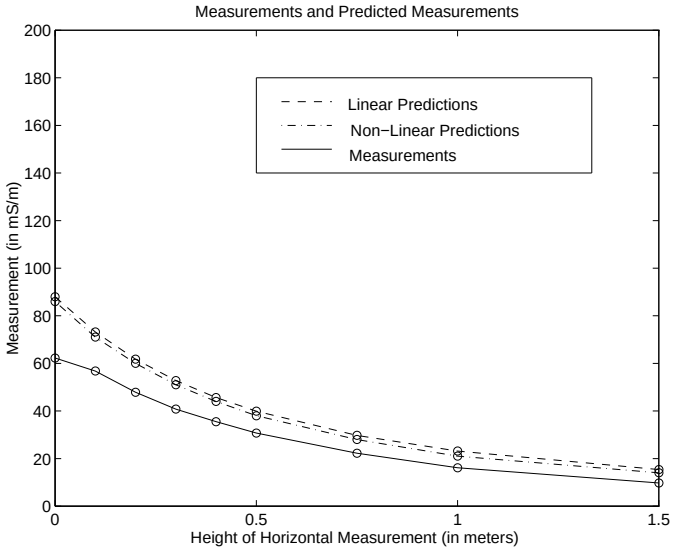

Figure A.24: Site 12: Horizontal Orientation.

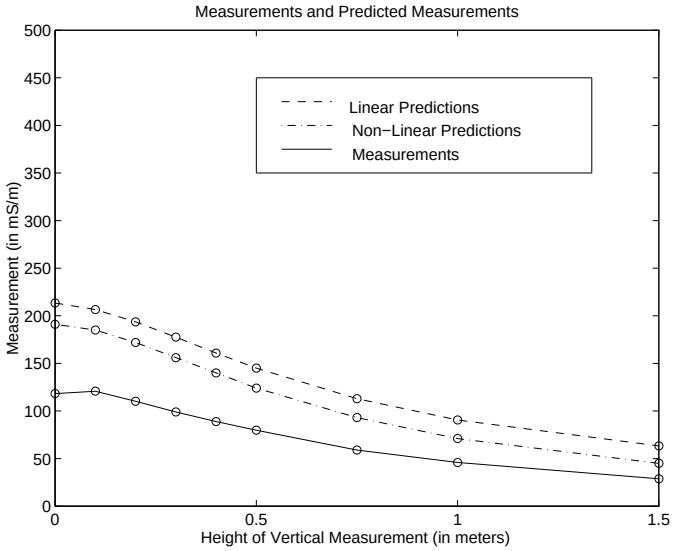

Figure A.25: Site 13: Vertical Orientation.

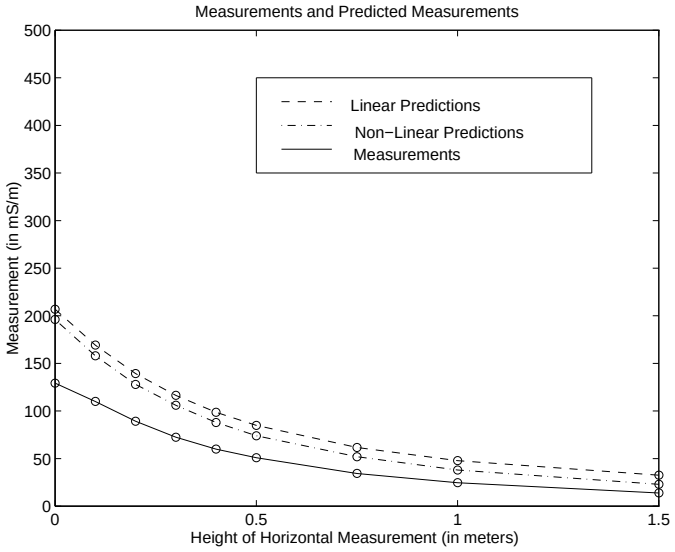

Figure A.26: Site 13: Horizontal Orientation.

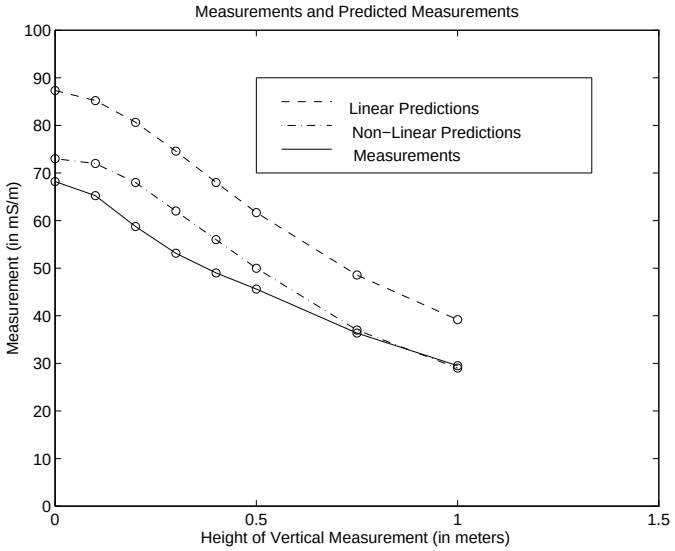

Figure A.27: Site 14: Vertical Orientation.

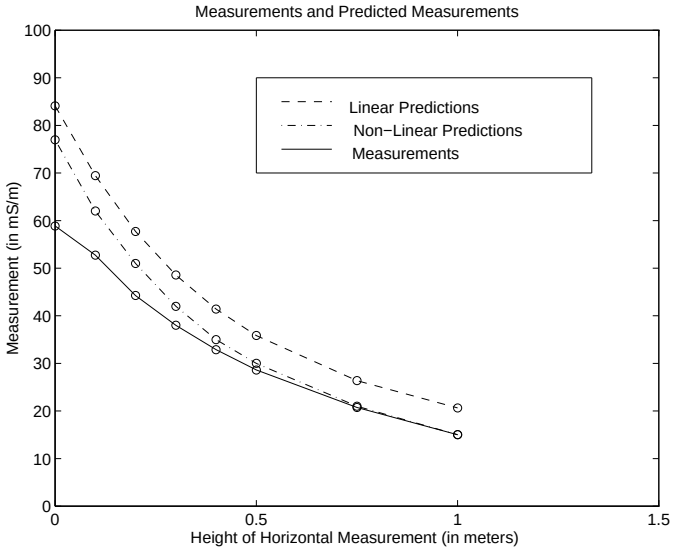

Figure A.28: Site 14: Horizontal Orientation.

Appendix B

Linear and Nonlinear Inverse Solutions

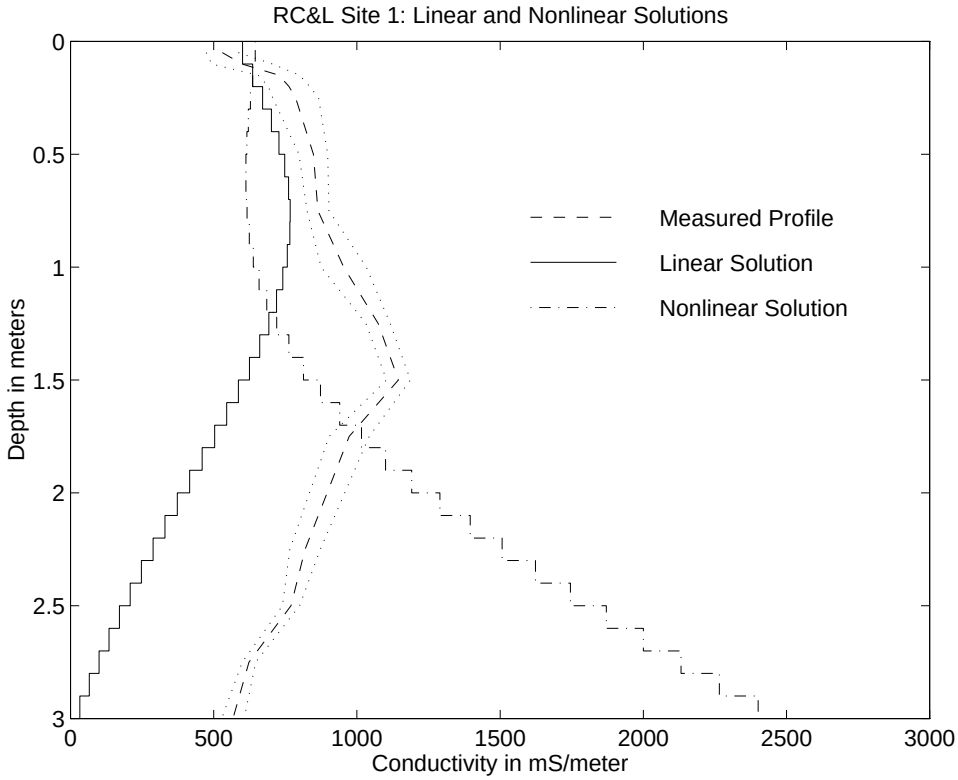

Figure B.1: Site 1: Inverse Solutions.

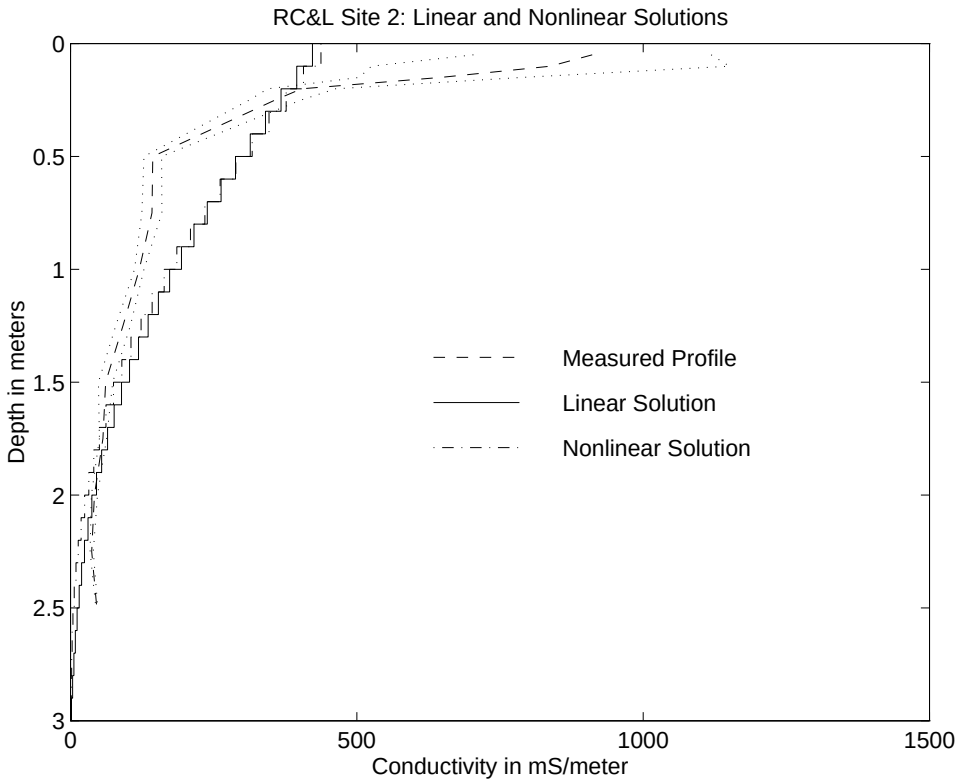

Figure B.2: Site 2: Inverse Solutions.

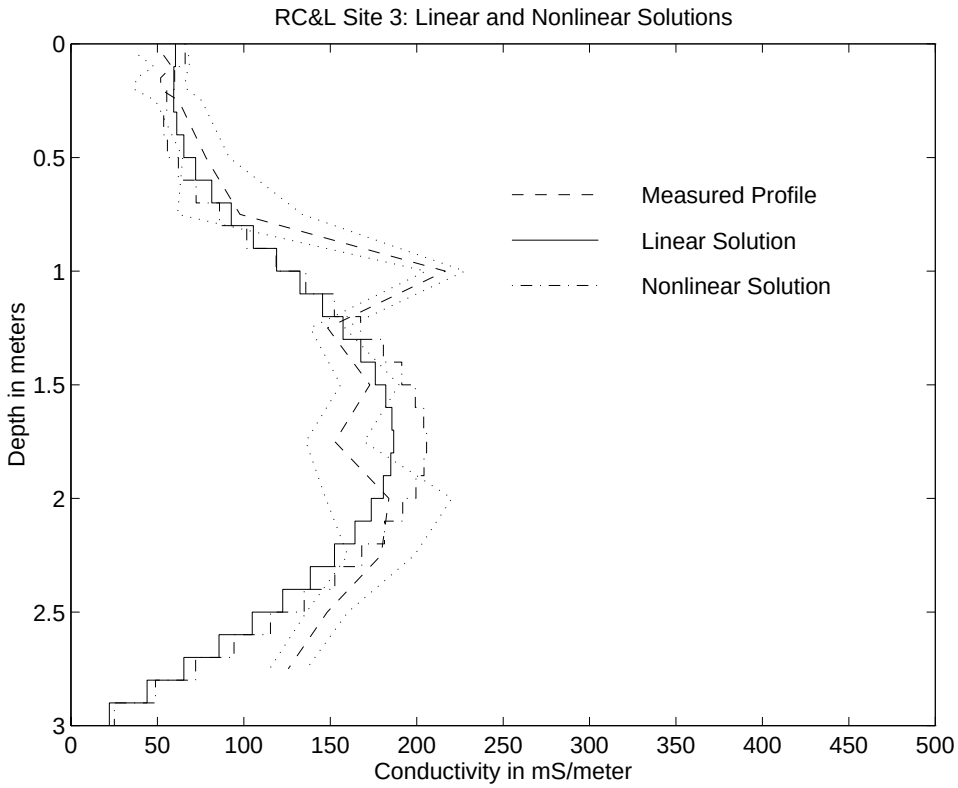

Figure B.3: Site 3: Inverse Solutions.

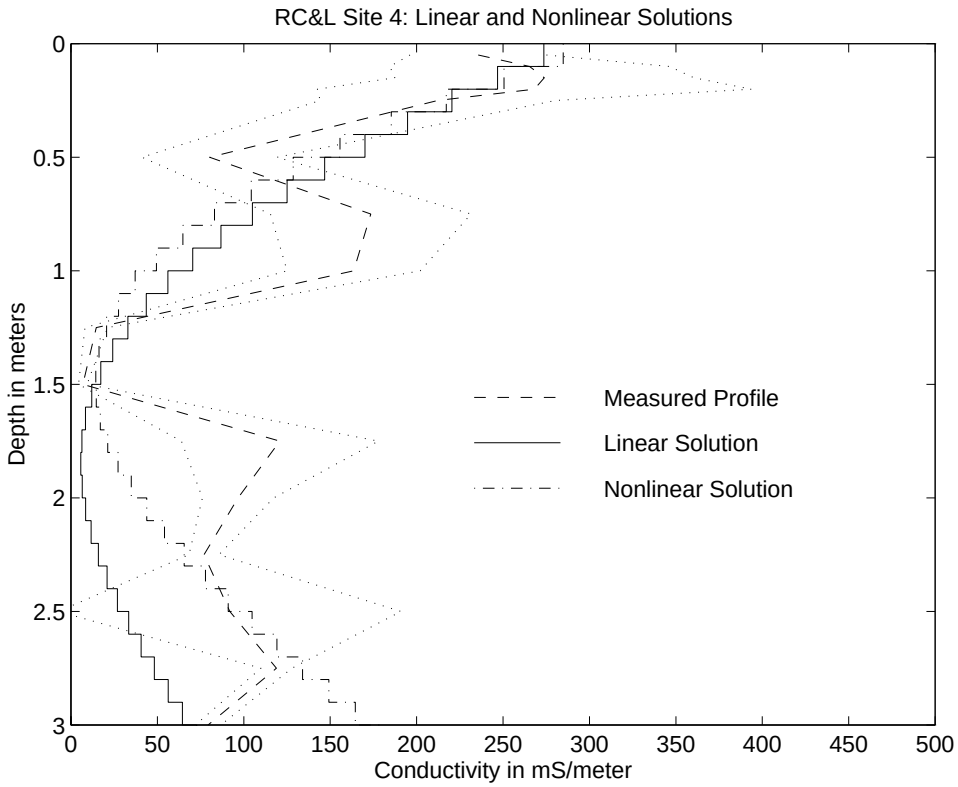

Figure B.4: Site 4: Inverse Solutions.

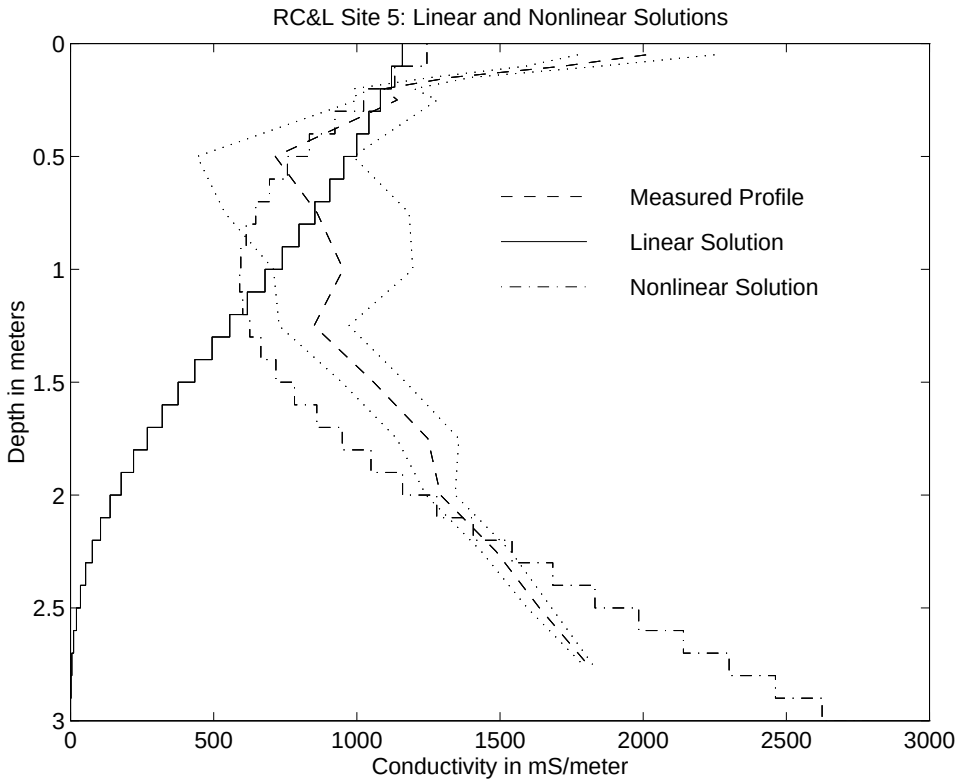

Figure B.5: Site 5: Inverse Solutions.

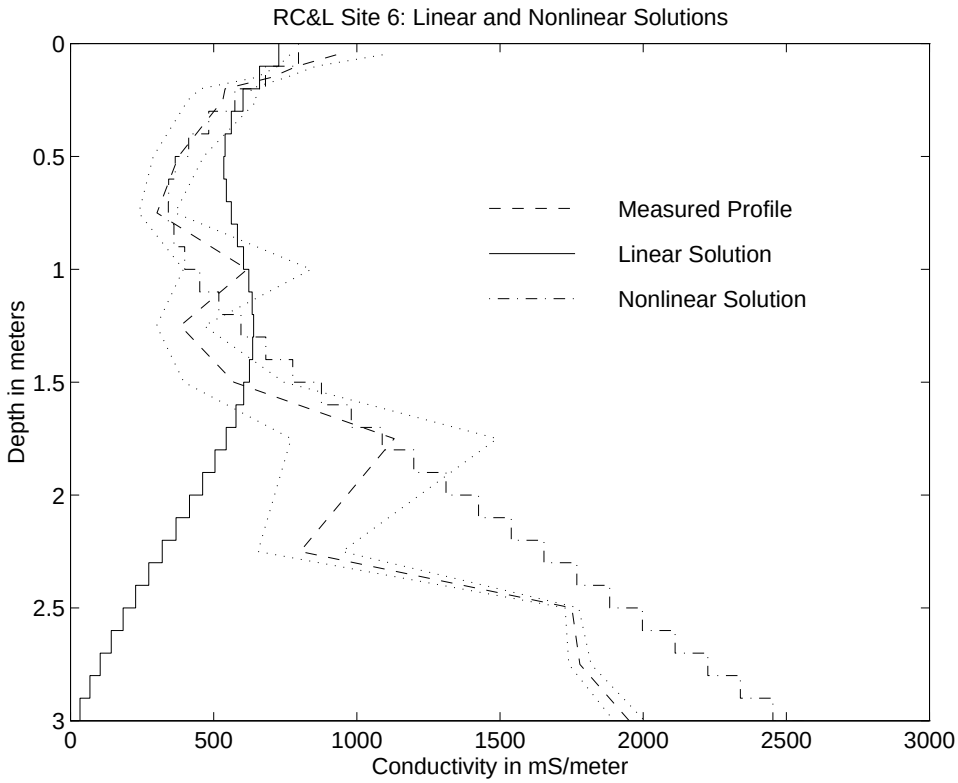

Figure B.6: Site 6: Inverse Solutions.

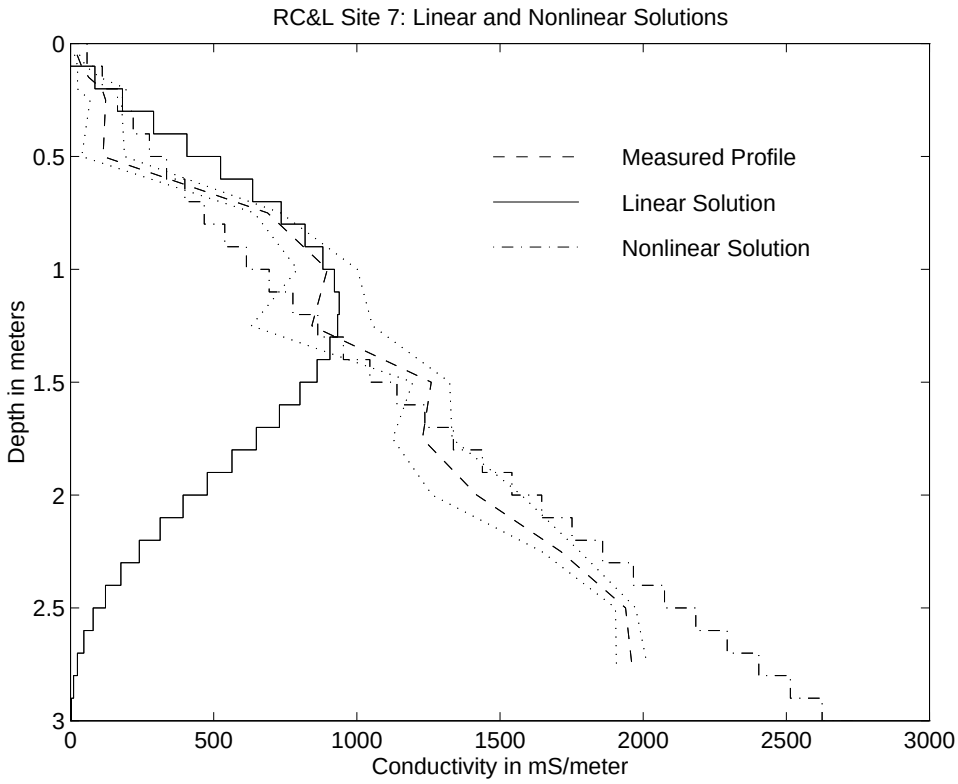

Figure B.7: Site 7: Inverse Solutions.

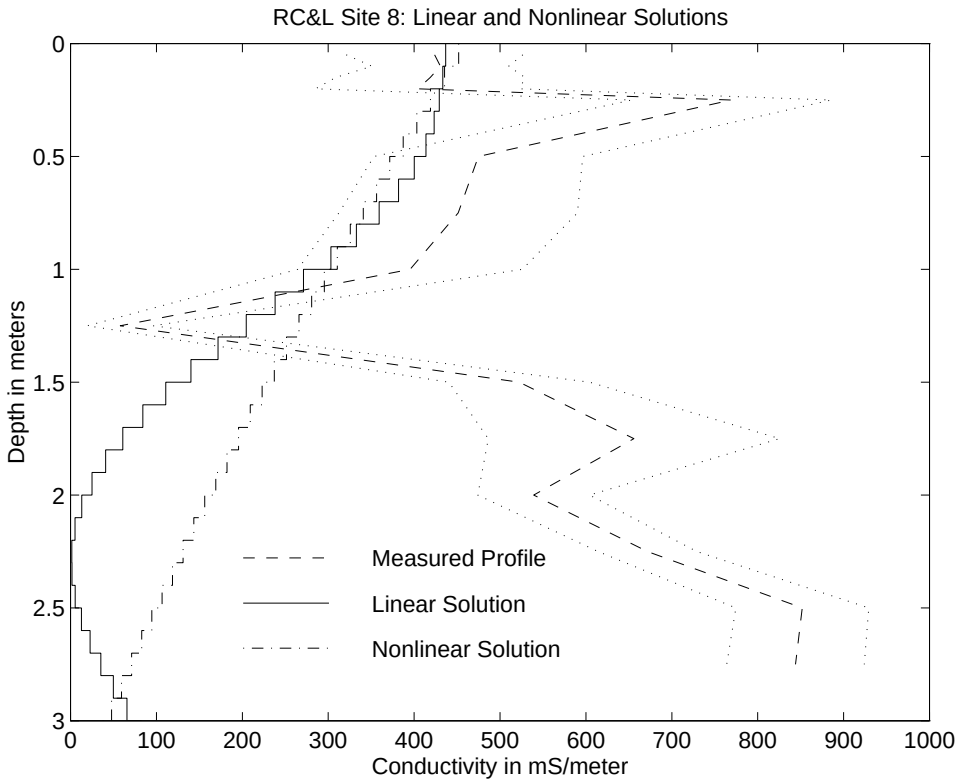

Figure B.8: Site 8: Inverse Solutions.

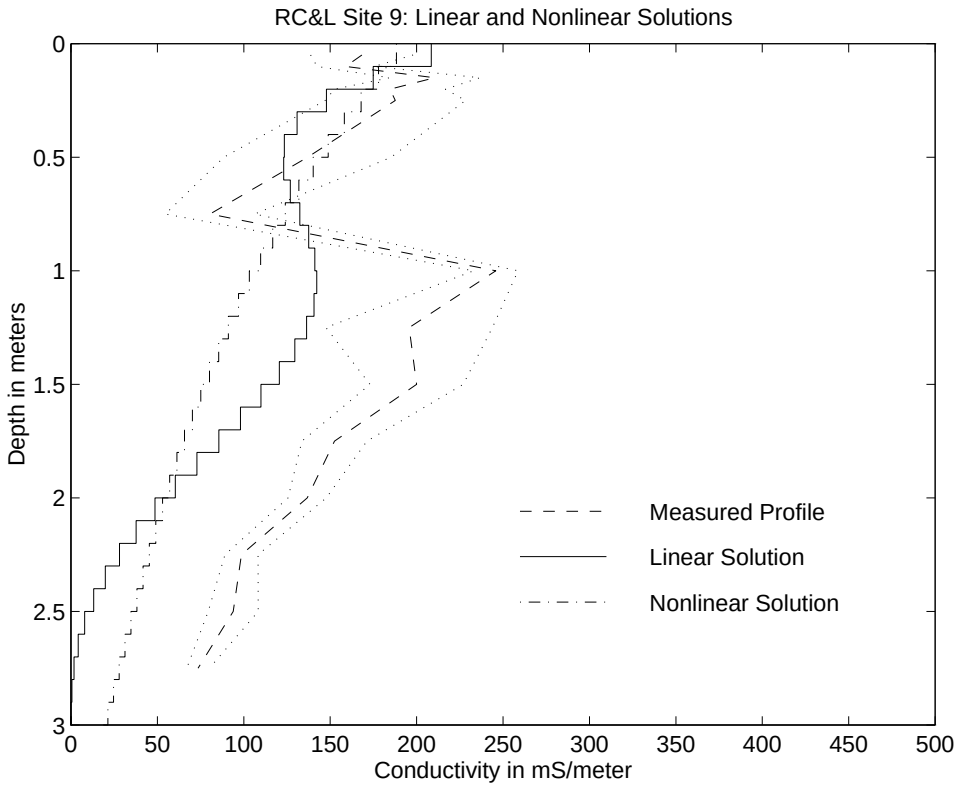

Figure B.9: Site 9: Inverse Solutions.

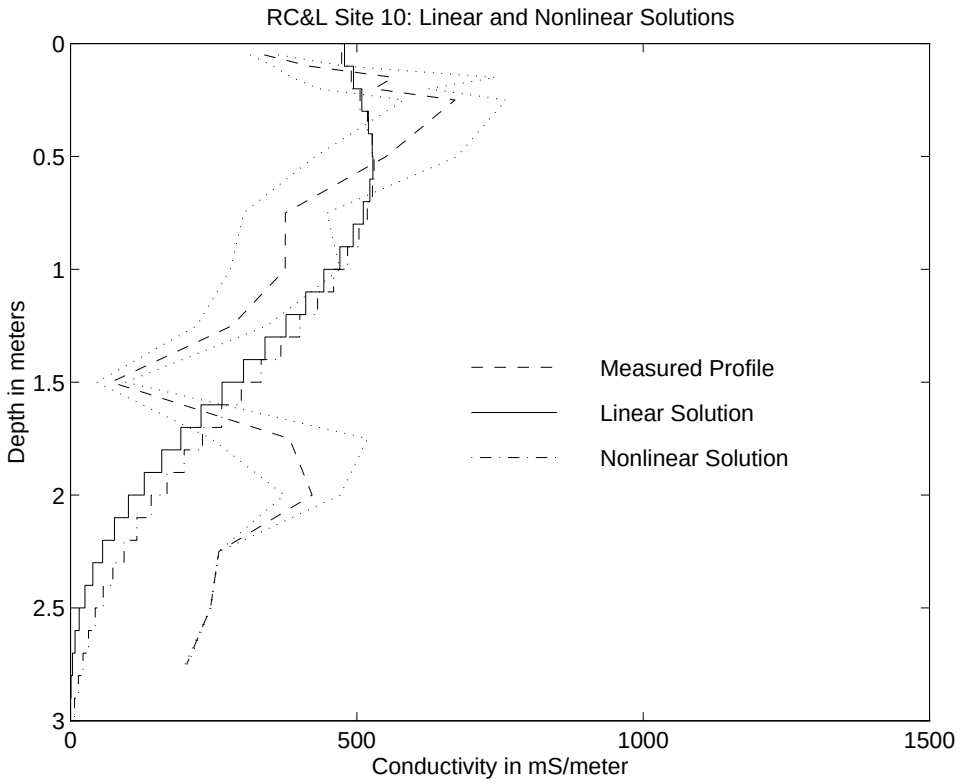

Figure B.10: Site 10: Inverse Solutions.

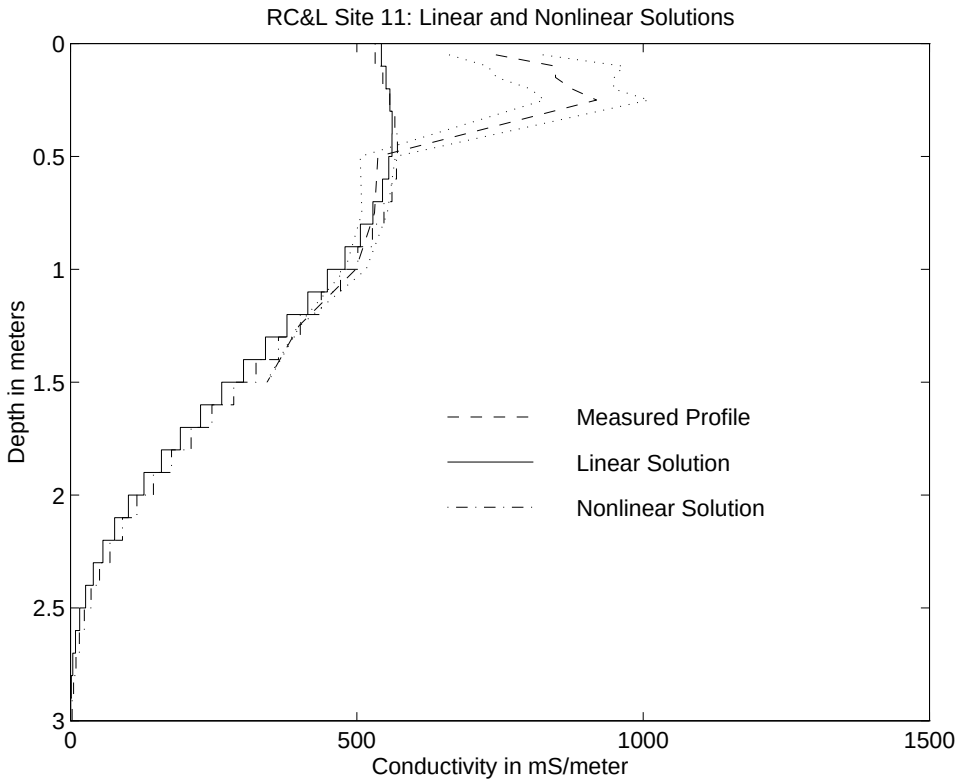

Figure B.11: Site 11: Inverse Solutions.

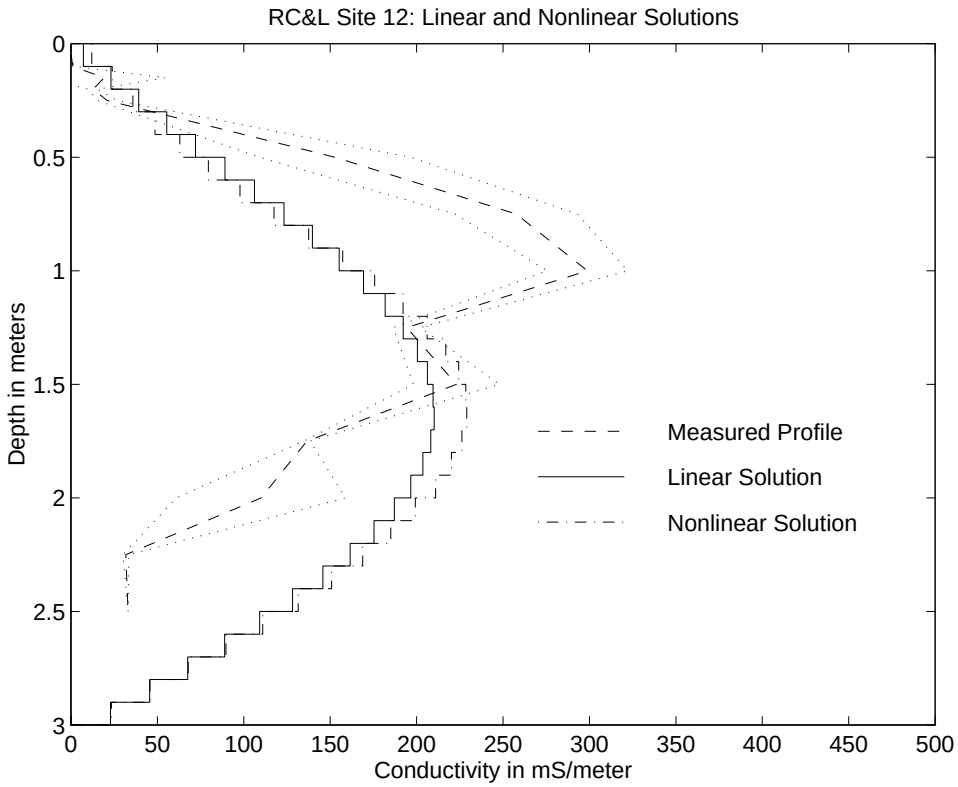

Figure B.12: Site 12: Inverse Solutions.

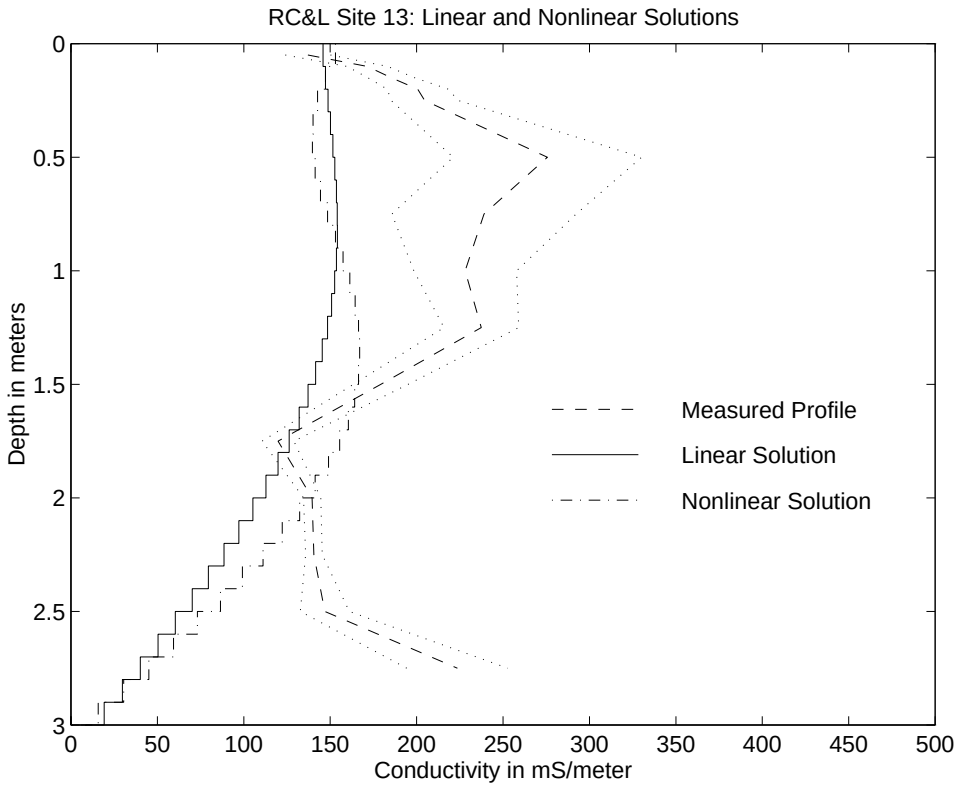

Figure B.13: Site 13: Inverse Solutions.

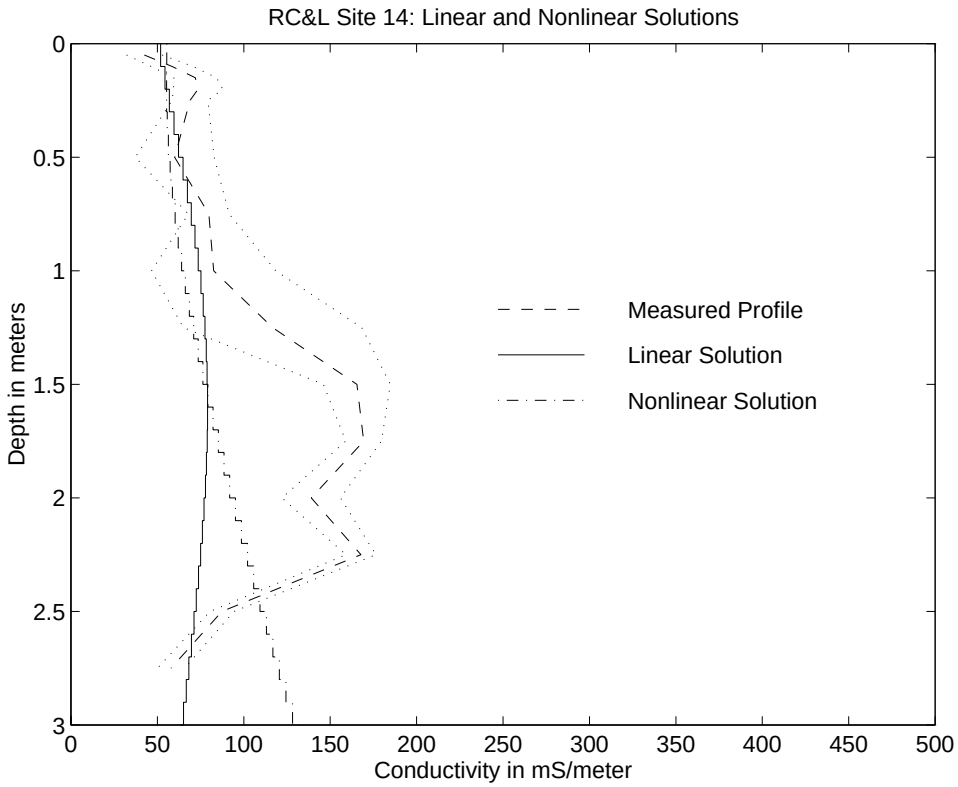

Figure B.14: Site 14: Inverse Solutions.

Appendix C

Linear L-curves

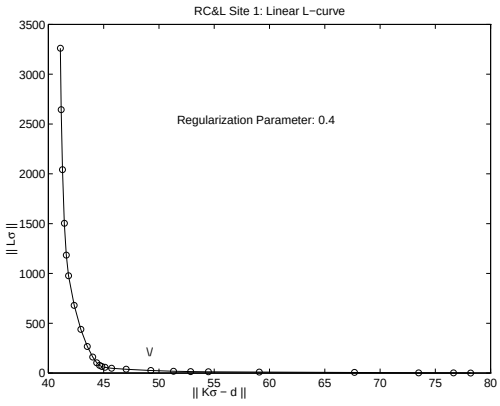

Figure C.1: Site 1: Linear L-curve.

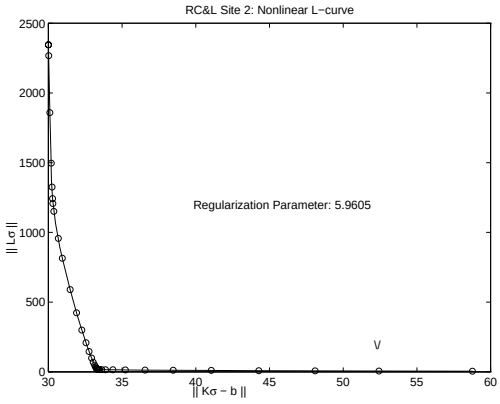

Figure C.2: Site 2: Linear L-curve.

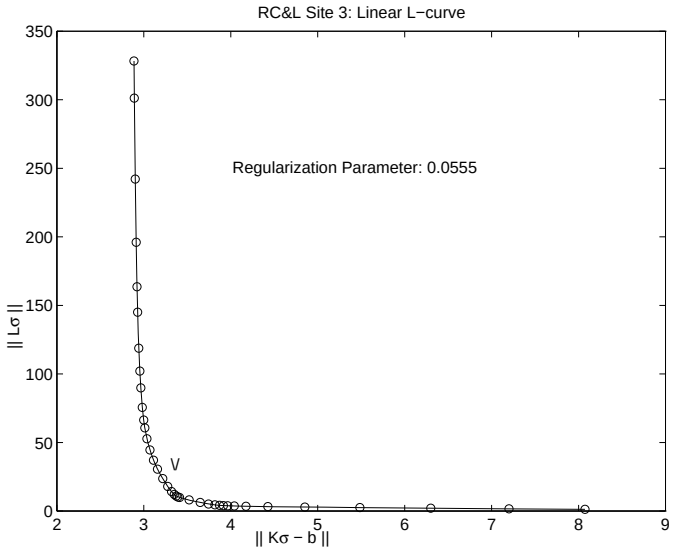

Figure C.3: Site 3: Linear L-curve.

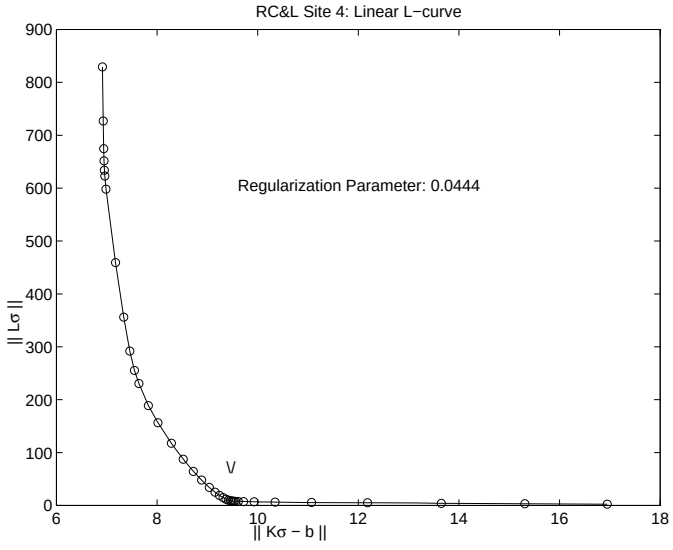

Figure C.4: Site 4: Linear L-curve.

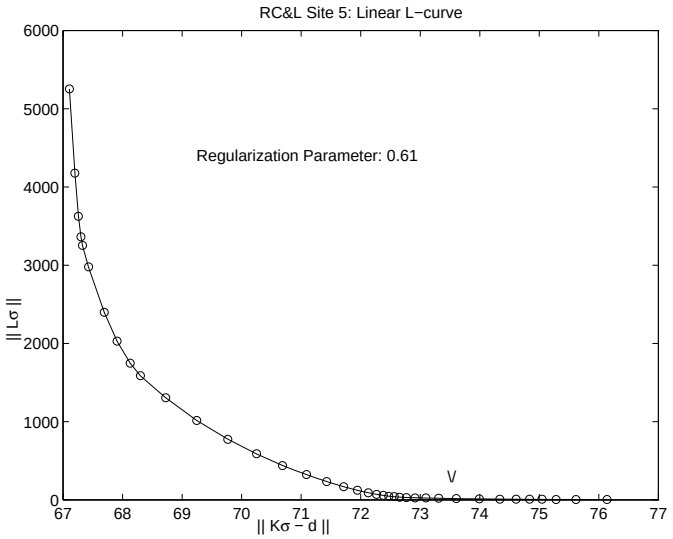

Figure C.5: Site 5: Linear L-curve.

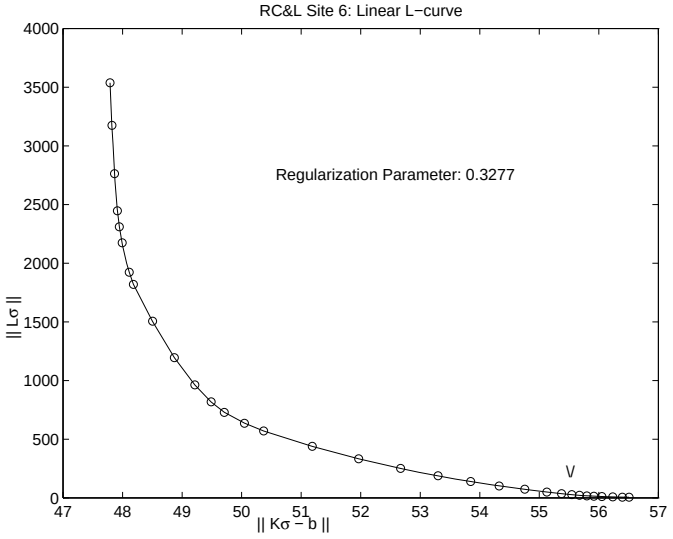

Figure C.6: Site 6: Linear L-curve.

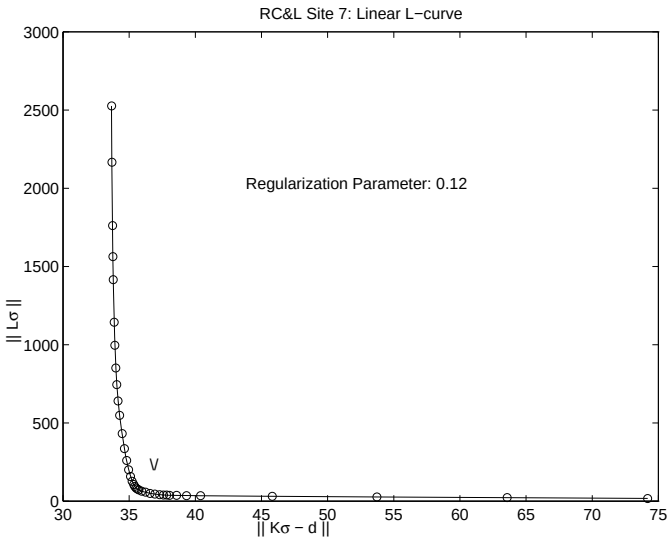

Figure C.7: Site 7: Linear L-curve.

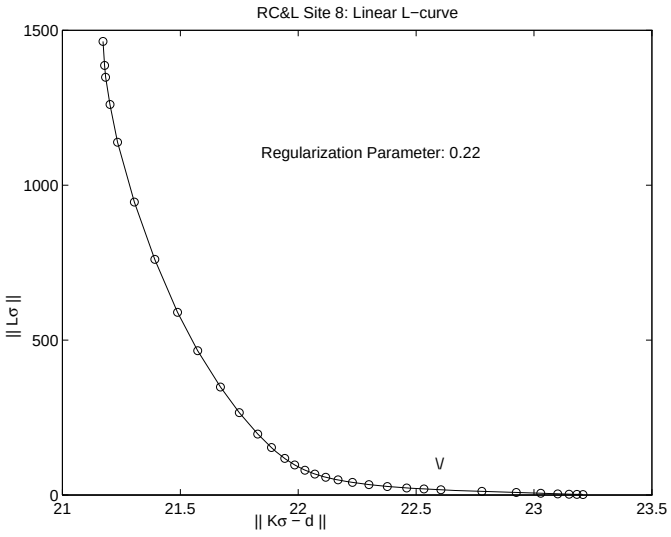

Figure C.8: Site 8: Linear L-curve.

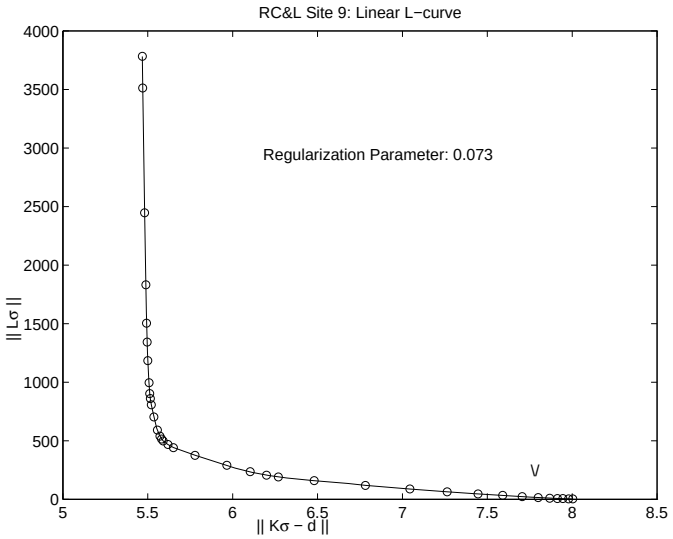

Figure C.9: Site 9: Linear L-curve.

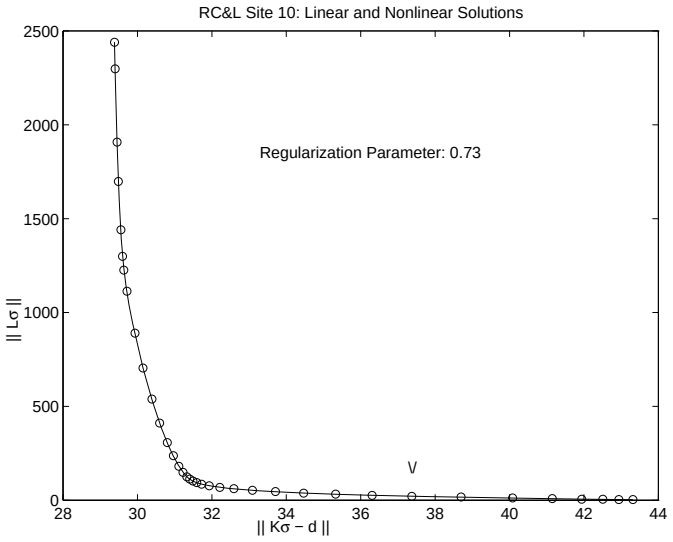

Figure C.10: Site 10: Linear L-curve.

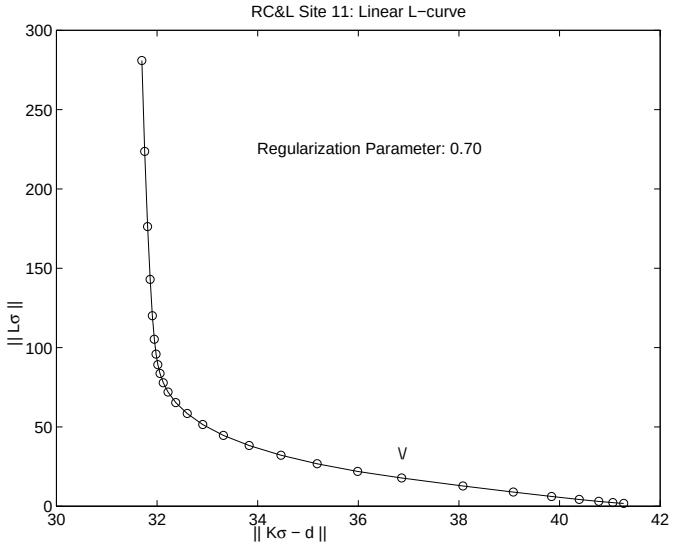

Figure C.11: Site 11: Linear L-curve.

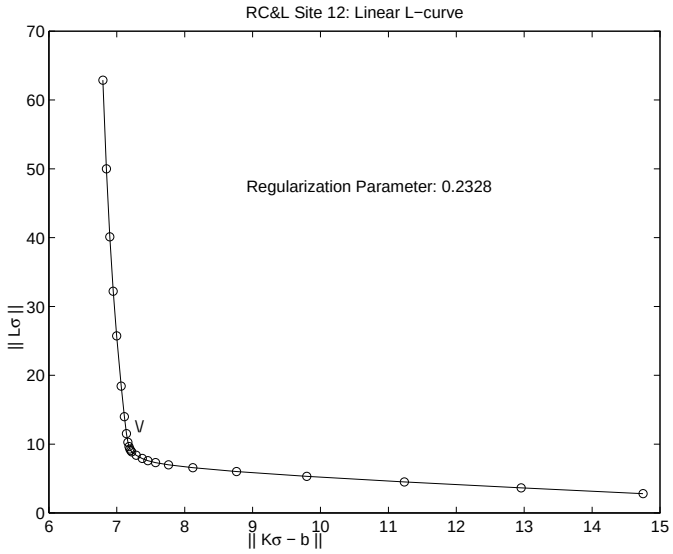

Figure C.12: Site 12: Linear L-curve.

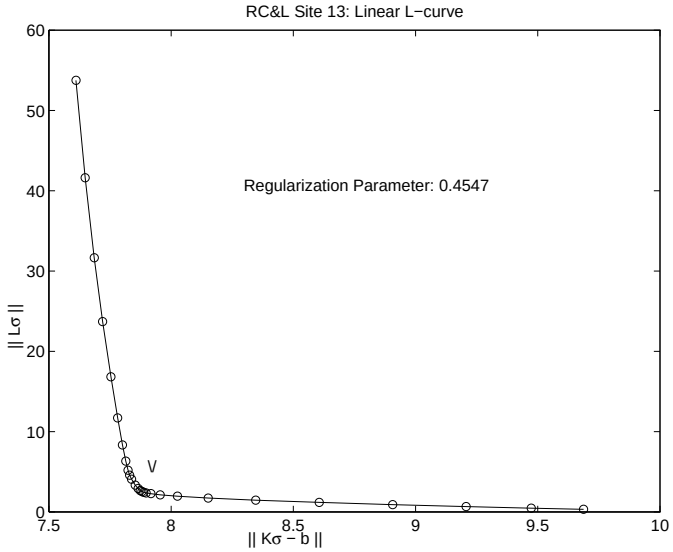

Figure C.13: Site 13: Linear L-curve.

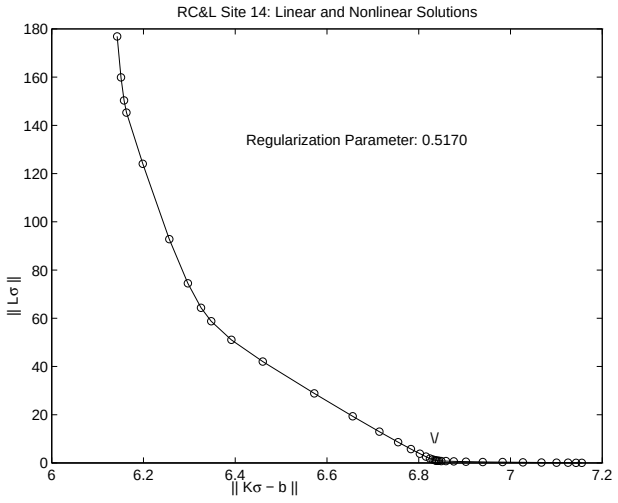

Figure C.14: Site 14: Linear L-curve.

Appendix D

Nonlinear L-curves

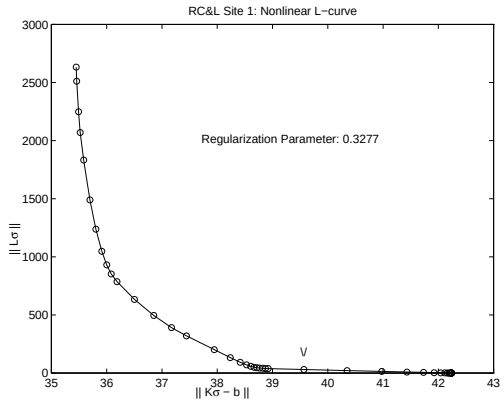

Figure D.1: Site 1: Nonlinear L-curve.

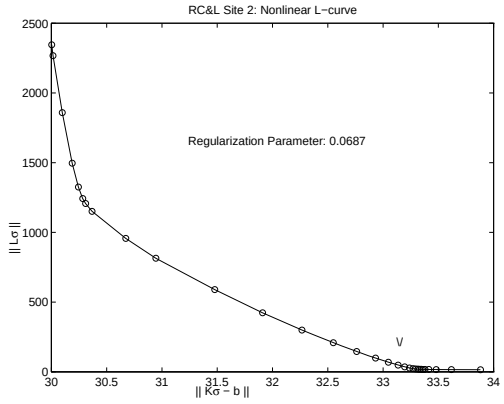

Figure D.2: Site 2: Nonlinear L-curve.

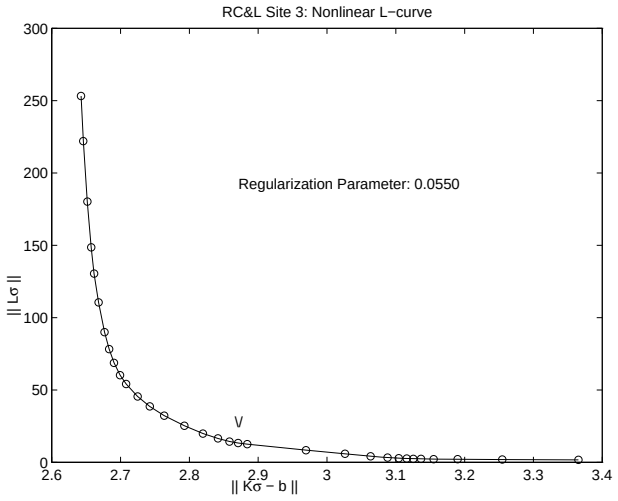

Figure D.3: Site 3: Nonlinear L-curve.

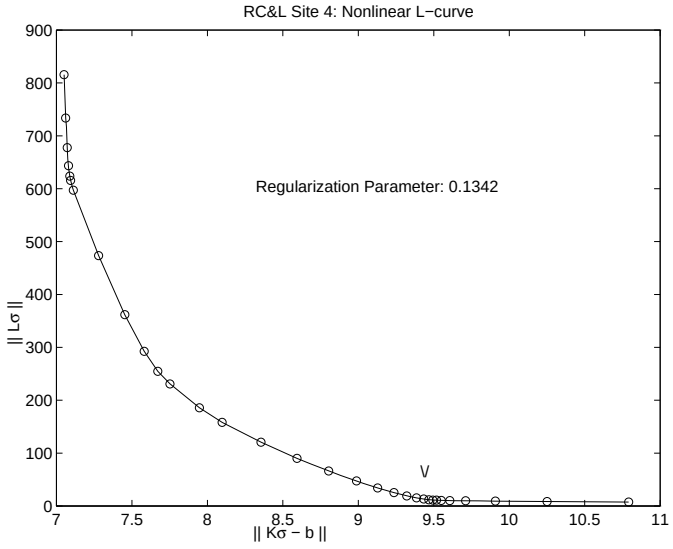

Figure D.4: Site 4: Nonlinear L-curve.

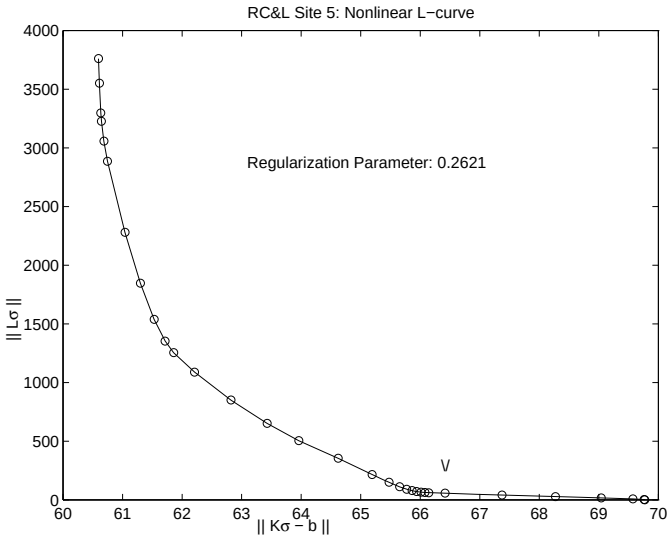

Figure D.5: Site 5: Nonlinear L-curve.

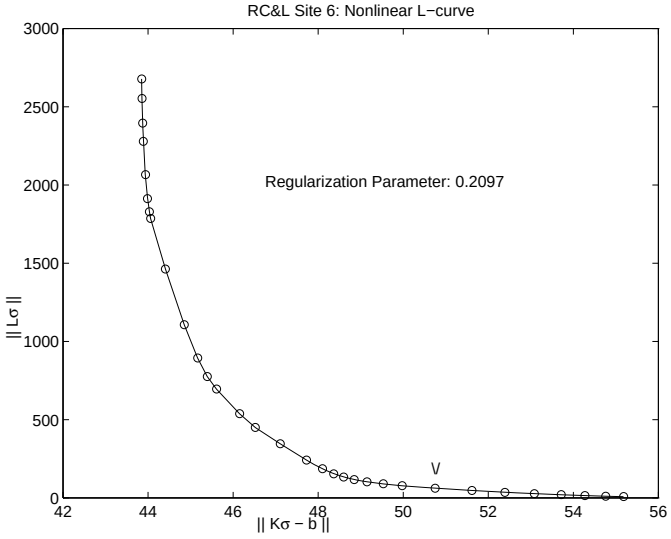

Figure D.6: Site 6: Nonlinear L-curve.

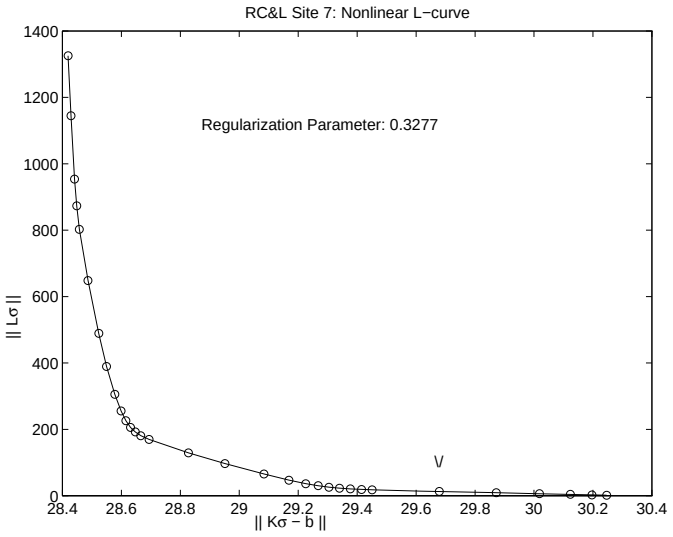

Figure D.7: Site 7: Nonlinear L-curve.

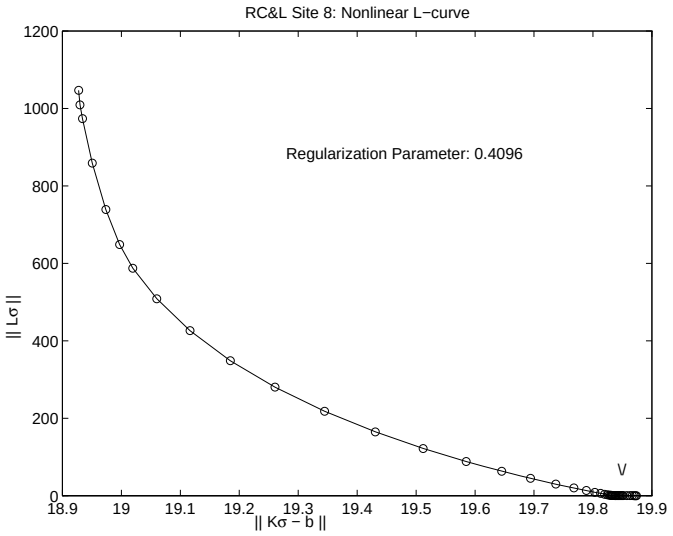

Figure D.8: Site 8: Nonlinear L-curve.

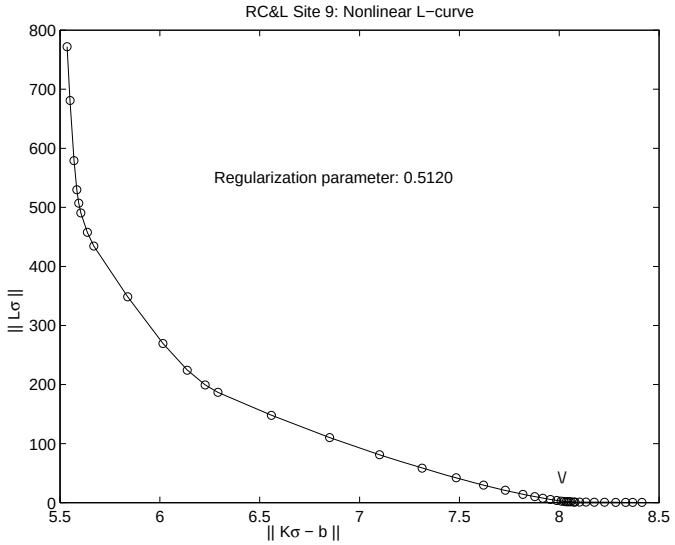

Figure D.9: Site 9: Nonlinear L-curve.

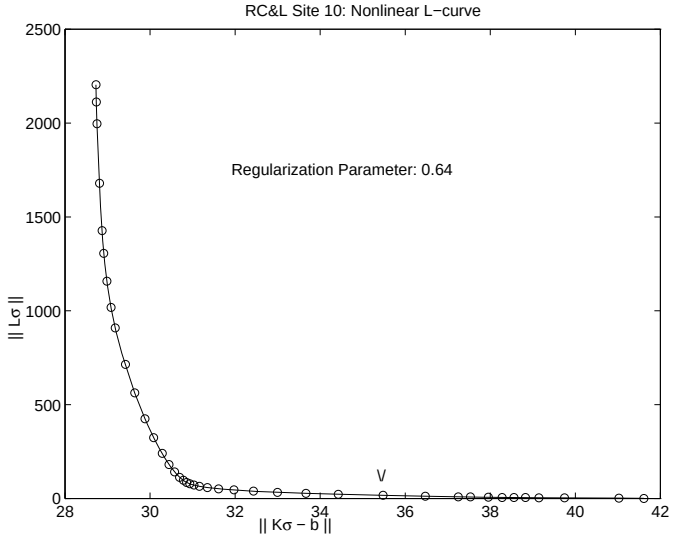

Figure D.10: Site 10: Nonlinear L-curve.

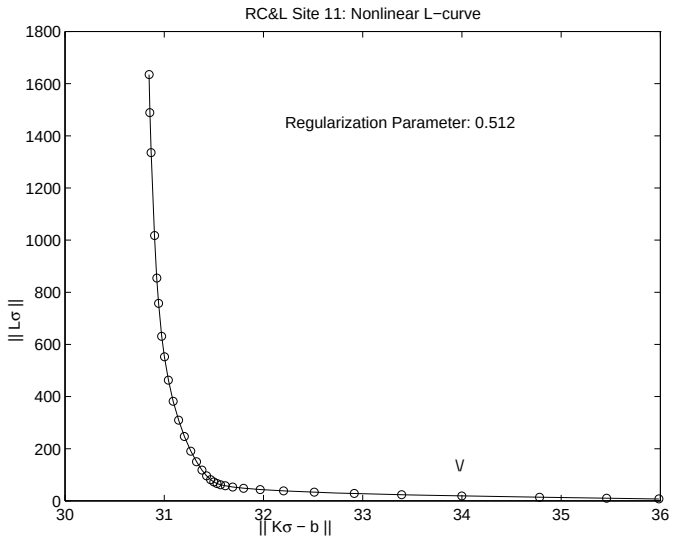

Figure D.11: Site 11: Nonlinear L-curve.

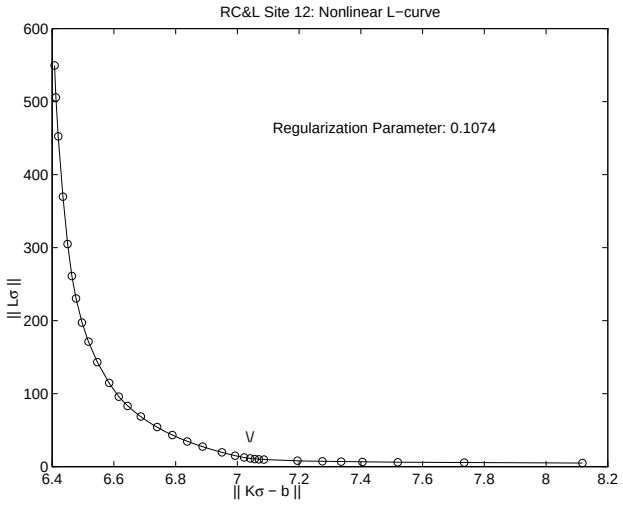

Figure D.12: Site 12: Nonlinear L-curve.

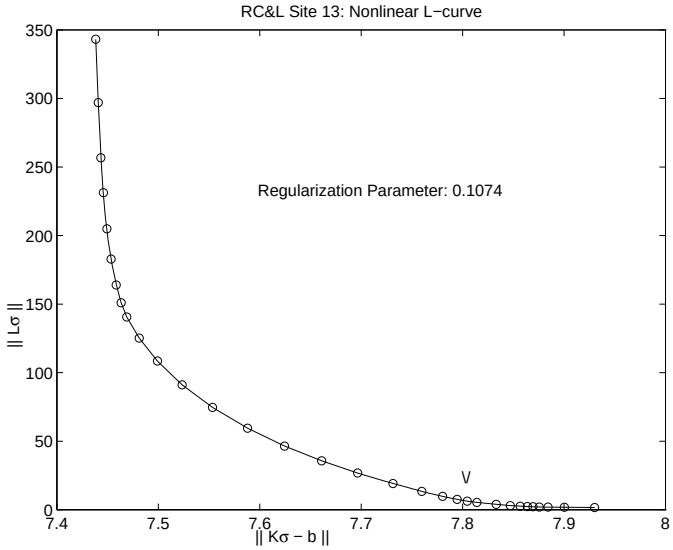

Figure D.13: Site 13: Nonlinear L-curve.

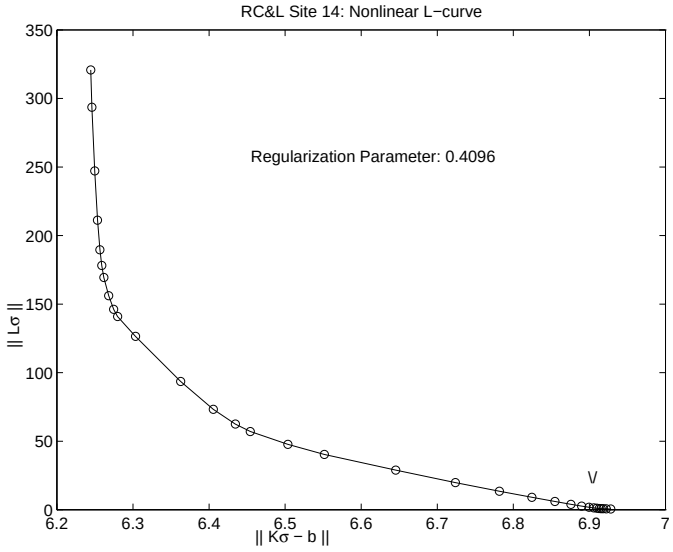

Figure D.14: Site 14: Nonlinear L-curve.

References

[Alumbaugh and Newman, 1997] Alumbaugh, D. L. and Newman, G. A. (1997). Three-dimensional massively parallel electromagnetic inversion II. analysis of a crosswell electromagnetic experiment. *Geophysics*, 128:355–363.

[Alumbaugh et al., 1996] Alumbaugh, D. L., Newman, G. A., Prevost, L., and Shadid, J. N. (1996). Three-dimensional wideband electromagnetic modeling on massively parallel computers. *Radio Science*, 31:1–23.

[Anderson, 1979] Anderson, W. L. (1979). Computer program numerical integration of related Hankel transforms of orders 0 and 1 by adaptive digital filtering. *Geophysics*, 44(7):1287–1305.

[Borchers et al., 1997] Borchers, B., Uram, T., and Hendrickx, J. M. (1997). Tikhonov regularization of electrical conductivity depth profiles in field soils. *Soil Science Society of America Journal*, 61(4):1004–1009.

[Engl, 1993] Engl, H. W. (1993). Regularization methods for the stable solution of inverse problems. *Surveys on Mathematics for Industry*, 3:71–143.

[Forsythe et al., 1977] Forsythe, G. E., Malcom, M. A., and Moler, C. B. (1977). *Computer methods for mathematical computations.* Prentice-Hall, Englewood Cliffs, New Jersey.

[Groetsch, 1993] Groetsch, C. W. (1993). *Inverse Problems in the Mathematical Sciences*. Vieweg, Bruanschweig; Wiesbaden.

[Hanke and Hansen, 1993] Hanke, M. and Hansen, P. C. (1993). Regularization methods for large-scale problems. *Surveys on Mathematics for Industry*, 3:253–315.

[Hansen, 1992] Hansen, P. C. (1992). Numerical tools for analysis and solution of Fredholm integral equations of the first kind. *Inverse Problems*, 8:849–872.

[Lawson and Hanson, 1974] Lawson, C. L. and Hanson, R. J. (1974). *Solving Least Squares Problems*. Prentice-Hall, Englewood Cliffs, New Jersey.

[McNeill, 1980] McNeill, J. D. (1980). Electromagnetic terrain conductivity measurement at low induction numbers. Technical Report TN-6, Geonics Limited.

[McNeill, 1985] McNeill, J. D. (1985). EM34–3 measurements at two inter-coil spacings to reduce sensitivity to near–surface material. Technical Report TN-19, Geonics Limited.

[McNeill, 1990] McNeill, J. D. (1990). Use of electromagnetic methods for groundwater studies. In Ward, S. H., editor, *Geotechnical and Environmental Geophysics*, volume 1, pages 191–218. Society of Exploration Geophysicists.

[Nabighian, 1987] Nabighian, M. N., editor (1987). *Electromagnetic Methods in Applied Geophysics*, volume 1. Society of Exploration Geophysicists, Tulsa.

[Newman and Alumbaugh, 1997] Newman, G. A. and Alumbaugh, D. L. (1997). Three-dimensional massively parallel electromagnetic inversion - I. Theory. *Geophysics*, 128:345–354.

[Parker, 1994] Parker (1994). *Geophysical Inverse Theory.* Princeton University Press.

[Rhoades et al., 1990] Rhoades, J., Corwin, D., and Lesch, S. (1990). Effect of soil ECa - depth profile pattern on electromagnetic induction measurements. Technical Report 125, US Dept. of Agriculture - Agricultural Research Service US - Salinity Laboratory.

[Rhoades, 1993] Rhoades, J. D. (1993). Electrical conductivity methods for measuring and mapping soil salinity. In Sparks, D. L., editor, *Advances In Agronomy, volume 46*, pages 201–251. Academic Press.

[Twomey, 1997] Twomey, S. (1997). *Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements.* Dover Publications, Inc., Mineola, New York.

[Wahba, 1990] Wahba, G. (1990). *Spline Models for Observational Data.* SIAM, Philadelpqhia.

[Wait, 1982] Wait, J. R. (1982). *Geo-Electromagnetism.* Academic Press, New York.

[Wing, 1991] Wing, G. M. (1991). *A Primer on Integeral Equations of the First Kind.* SIAM, Philadelphia.

[Zauderer, 1983] Zauderer, E. (1983). *Partial Differential Equations of Applied Mathematics*. John Wiley & Sons, Inc., New York.

This thesis is accepted on behalf of the faculty of the Institute by the following committee:

Advisor

Date