

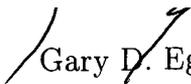
AN ABSTRACT OF THE DISSERTATION OF

Weerachai Siripunvaraporn for the degree of Doctor of Philosophy in Geophysics
presented on July 15, 1999.

Title: An Efficient Data-Subspace Two-Dimensional Magnetotelluric Inversion
and its Application to High Resolution Profile across the
San Andreas Faults at Parkfield, California.

Redacted for privacy

Abstract approved: _____


Gary D. Egbert

This dissertation is divided into two different aspects of the Magnetotelluric (MT) method: the development of numerical algorithms for interpreting and modeling of MT data, and the interpretation of high resolution MT data across the San Andreas Fault (SAF). In the first part, it involves two major developments: a three-dimensional MT forward modeling program on a parallel computer, and an efficient inversion scheme for two-dimensional MT data.

Modeling geologically realistic problems in 3-D requires such large grids that is not feasible on workstations. To allow for large problems, we developed an algorithm for 3-D forward modeling to run on a massively parallel distributed memory computer. The development involves a parallel implementation of existing serial algorithms, and the development of preconditioners to speed up the convergence of the iterative solver.

Generally, most reliable and efficient 2-D inversion methods require extensive computational time and memory. To overcome these difficulties, we present a new and much more efficient inversion algorithm, the reduced basis Occam's inver-

sion (REBOCC). The algorithm is based on the fact that MT data are smooth and redundant. Therefore, the solution can be constructed from subsets of the data, without significant loss of detail, and still fit all data adequately. Numerical experiments on synthetic and real data sets show that REBOCC is stable and relatively faster than most other 2-D inversion algorithms.

Fluid which strongly influences electrical resistivities has been related to earthquake mechanisms at the SAF. Thus, MT which has an advantage over other methods for mapping the fluid distribution in the fault zone, was collected across the SAF at Parkfield, California, and were interpreted using REBOCC inversion, along with hypothesis and case studies. Three main features are required by the data: resistive Salinian granite to the west, a large volume of low resistivity to the east, and high conductivity ($3 \Omega - m$) in the fault zone. These features are consistent with velocity models. Experiments with constrained inversions show that low resistivity ($3 \Omega - m$, 250-m wide) can extend to 10-km depth. If interpreted to result from fluid, this would require a porosity around 8-18 %, consistent with other geophysical observations.

© Copyright by Weerachai Siripunvaraporn

July 15, 1999

All rights reserved

An Efficient Data-Subspace Two-Dimensional Magnetotelluric Inversion
and its Application to High Resolution Profile across the
San Andreas Faults at Parkfield, California

by

Weerachai Siripunvaraporn

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Completed July 15, 1999
Commencement June 2000

Doctor of Philosophy dissertation of Weerachai Siripunvaraporn presented on
July 15, 1999

APPROVED:

Redacted for privacy

Major Professor, representing Geophysics ✓ ✓

Redacted for privacy

Dean of College of Oceanic & Atmospheric Sciences

Redacted for privacy

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Redacted for privacy

Weerachai Siripunvaraporn, Author

ACKNOWLEDGMENT

First and foremost thanks go to my advisor Dr. Gary Egbert, who provided me with support and guidance throughout my long journey towards the Ph.D, with his consistent encouragement and patience, especially in regards to my English. This is a good opportunity for me to sincerely thank him for everything he has done for me. To me, he is not just an advisor, he is more than that.

I like to thank my other committee members Dr. Anne Trehu, Dr. Shaul Levi, Dr. John Chen, and graduate representative Dr. Cyrus Field for their helpful comments and suggestions. I also wish to thank other geophysics faculty members at OSU, Dr. John Nabelek and Dr. Bob Lillie who used to be a member of graduate committee, but cannot attend my final oral exam due to his sabbatical out of state. Another thank is to Dr. Marcus Eisel who did a postdoc here for two years, and now working in Germany. We have had a good memory throughout his staying.

A special thank goes to all of my geophysics friends who had graduated ahead of me, Axel, Beate, Bernd, Christof, Ganyuan, Guibiao, Jochen, Kelly, Kurt, Maren, Rugang, Sean, and Xiao-Qing, and also to Cindy and Sierra (from UBC) who consistently help me correcting English, both pronunciation and written. Also, I would like to thank all of my Thai friends at OSU. Some of them are still around and some of them are already in Thailand.

Most of all, I wish to thank Manaviga Nualmanee (Ying), my heart and soul, for supporting and helping me through my tough time. Without my mom, my dad and all of my six brothers and four sisters in Thailand, I would not be able to come this far. My mom provides me not only good home cooking meals but also love, care, understanding and of course financial support. My dad taught me many good things, not only by his words but also his actions. My brothers and sisters fully

support me, and help me in many different aspects; academics, financial, personal problems, and many little things. I would like to dedicate this degree to all of these people.

Finally, I would like to thank the Development of Sciences and Technology (DPST) of Thailand who provides most of my financial support throughout the time I stayed in the US.

TABLE OF CONTENTS

	<u>Page</u>
1 INTRODUCTION	1
1.1 Introduction to the Magnetotelluric (MT) Method	1
1.1.1 Source of Signal	2
1.1.2 Maxwell's Equations	3
1.1.3 Data Acquisition and Processing	5
1.1.4 One-Dimensional MT	7
1.1.5 Two-Dimensional MT	9
1.1.6 Static Distortions	11
1.2 Electrical Properties of Rocks	11
1.2.1 Conduction Mechanisms	14
1.2.2 Resistivity Ranges of Geological Structure	19
1.3 Magnetotelluric Studies - Chapters of this Thesis	19
1.4 References	21
2 THREE-DIMENSIONAL MAGNETOTELLURIC FORWARD MODEL- ING ON A MASSIVELY PARALLEL COMPUTER.....	24
2.1 Introduction	24
2.2 Staggered Finite Difference Grid and Linear System of Equations	25
2.3 Boundary Conditions	30
2.4 Iterative Relaxation Methods	31
2.5 Divergence Correction	33
2.6 Parallel Implementations	35
2.6.1 Partitioning Vectors and Matrices	35
2.6.2 Parallel Computations	36
2.6.3 Iterative Methods and Parallel Preconditioners	37

TABLE OF CONTENTS (Continued)

	<u>Page</u>
2.7 Numerical Experiments and Discussions	39
2.7.1 Different Preconditioners	41
2.7.2 Without Divergence Correction	43
2.7.3 With Divergence Correction	44
2.8 Conclusion	49
2.9 References	51
3 REBOCC: AN EFFICIENT DATA-SUBSPACE INVERSION FOR TWO-DIMENSIONAL MAGNETOTELLURIC DATA	53
3.1 Introduction	53
3.2 Overview of Inversion Methods	54
3.3 The Data Space Occam Method (DASOCC)	62
3.4 Reduced Data Space Occam Approaches (REBOCC)	63
3.4.1 Representer Subsets for REBOCC	67
3.4.2 Comparison of Computational Resource Requirements	68
3.5 REBOCC : Algorithm Details	70
3.5.1 Forward Modeling	70
3.5.2 Model Covariance C_m	73
3.5.3 Line Search for λ	76
3.5.4 Interpolation Matrix	78
3.5.5 Static Shift Correction	79
3.6 Numerical Experiments	80
3.6.1 Synthetic Data	81
3.6.2 Results of REBOCC Inversion	81
3.6.3 Comparison with Other Inversions	87
3.6.4 Example with Field Data	89

TABLE OF CONTENTS (Continued)

	<u>Page</u>
3.7 Discussions and Conclusions	90
3.8 Acknowledgments	93
3.9 References	93
4 HIGH RESOLUTION MT PROFILES ACROSS SAN ANDREAS FAULT AT PARKFIELD, CALIFORNIA	97
4.1 Introduction	97
4.1.1 Tectonic Development of the SAF and Seismicity	98
4.1.2 Geologic Setting of SAF at Parkfield	101
4.1.3 Past Geophysical Experiments at Parkfield	103
4.2 MT Data across the San Andreas Fault at Parkfield	105
4.2.1 Data Acquisition and Data Processing	106
4.2.2 Qualitative Interpretation of MM1	109
4.2.3 Comparisons of Middle Mountain Profiles	115
4.2.4 Qualitative Interpretation of the Cholame Valley Profile	118
4.3 Dimensionality and Distortion Analysis	121
4.3.1 Strike Estimation for Middle Mountain Profiles	123
4.3.2 Strike Estimation for the Cholame Valley Profiles	127
4.4 Inversion Results	128
4.4.1 Single Mode Inversions (MM1)	130
4.4.2 Joint Inversion (MM1): Best Fitting Models	134
4.4.3 Joint Inversion (MM1): Preferred Models	136
4.4.4 Joint Inversion Results for MM2 and MM3	139
4.4.5 Inversion Results for the Cholame Valley	143
4.5 Case Studies on 2-D Synthetic Models	146
4.6 Hypothesis and Sensitivity Tests	153

TABLE OF CONTENTS (Continued)

	<u>Page</u>
4.6.1 Hypothesis Tests on MM1 profile	153
4.6.2 Hypothesis Tests on CV profile	163
4.7 Discussion	163
4.7.1 Sources of Low Resistivity	168
4.7.2 Conductive Fault Zones at Middle Mountain	173
4.7.3 Low Resistivity in the Franciscan, East of the SAF	176
4.8 Conclusions	179
4.9 References	180
5 CONCLUSION	186
BIBLIOGRAPHY	188

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
1.1 Power spectrum of the magnetic field.	3
1.2 Skin depth comparisons for a) different frequencies, and b) different conductivities. (Adapted from Dobrin and Savit, 1988).	5
1.3 Hypothetical 1-D MT apparent resistivity and phase for a simplified continental region, with a conductive upper crust, a resistive lower crust and conductive mantle (After Mackie et al., 1988).	8
1.4 Effect of conductivity contrast to the apparent resistivities and phases for (blue) TM and (red) TE modes.	12
1.5 Plots of the apparent resistivity and phase of the real field data collected at Parkfield, California.	13
1.6 Generalized plots of resistivity as a function of inverse temperature for different rocks.	16
1.7 Plots of conductivity versus salinity in log ₁₀ Molality, at different pressure and temperature condition (redrawn from Merzer and Klemperer, 1997).	16
1.8 Typical ranges of resistivities of earth materials (redrawn from Palacky, 1987).	18
2.1 Positions of magnetic fields \mathbf{H} and electric fields \mathbf{E} on a staggered grid.	26
2.2 Sparse structure of matrix \mathbf{A}_{nn} for a $5 \times 5 \times 5$ model, where the dot indicate the non-zero elements.	30
2.3 Mapping the magnetic field \mathbf{H}_x vectors to the processors for the $3 \times 3 \times 3$ model.	36
2.4 The test 3-D model.	40
2.5 A plot of log 10 normalized residual versus relaxation iteration from solving the 3-D Poisson's equation by using different parallel preconditioners.	41

LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
2.6 (a) The log10 normalized residual (\mathbf{r}) vs. iteration for the model with the coarse grid, at period of 1 (blue) and 100 seconds (green) of \mathbf{J}_x polarization, and without divergence correction; (b) log10 of the divergence of the magnetic fields ($\nabla \cdot \mathbf{H}$); (c) log10 of the normalized residual of the surface magnetic fields after each iteration (\mathbf{H}_{sit}) and the converged surface magnetic fields (\mathbf{H}_c); and (d) log10 of the relative change of the magnetic fields after each iteration (\mathbf{H}_{it}) comparing with the magnetic fields of the previous iteration ($\mathbf{H}_{\text{it}-1}$).	43
2.7 Similar plot to Figure 2.6 but divergence correction is applied every 50 iterations for 100 second and every 100 iterations for 1 second. . .	46
2.8 Similar plot to Figure 2.7 but for \mathbf{J}_y polarization.	46
2.9 Similar plot to Figure 2.7 but for finer discretization of $60 \times 60 \times 49$ model.	49
2.10 Magnitude of the surface magnetic fields at 100 second, relative to a source field value of 1.0.	50
2.11 Magnitude of the surface magnetic fields at 1 second, relative to a source field value of 1.0.	50
3.1 Plots of representers $\mathbf{C}_m \mathbf{J}^T$ for TM mode apparent resistivity at one station for nearby periods, for the synthetic data discussed later. . . .	64
3.2 Examples of subset of representers.	67
3.3 Schematic interpolation example.	79
3.4 Synthetic model and data responses.	80
3.5 Memory requirements for OCCAM, DASOCC and REBOCC for different representer subsets calculated from Table 3.1 for TM and TM+TE modes.	82
3.6 RMS vs. cpu time required for inversion with different representer subsets.	83
3.7 Models produced by the 6th-stripe REBOCC inversion from various mode combinations.	84

LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
3.8	Result from 6th-stripe REBOCC inversion using all data. 86
3.9	Models obtained from (a) NLCG and (b) RRI inversions of TM, TE and TM+TE mode. 86
3.10	Plots of RMS versus cpu time of RRI, 6th-stripe REBOCC, and NLCG for TM, and TM+TE inversions. 87
3.11	Results from an 8th-stripe REBOCC inversion on a real data set from a high density profile across the SAF. 91
4.1	Geological map of Central California (after Irwin, 1990). 100
4.2	Geological cross-section across central California (adapted from Mooney and Weaver, 1989) 103
4.3	Map of San Andreas fault in the vicinity of Parkfield. 106
4.4	MT configurations. 107
4.5a	The MT responses (apparent resistivities, phases and the real and imaginary part of the tipper) of MM1 from 0.01 second to around 700 second. 110
4.5b	Continued. 111
4.6	Pseudo-section plots of the (distorted) apparent resistivity of the dense profiles (no remote sites), MM1 (top row), MM2 (middle), and MM3 (bottom). 112
4.7	Same as Figure 4.6, but for phase. 113
4.8	Plot of phase differences between profiles. 117
4.9a	Same as 4.5a, but for Cholame Valley profile (no remote sites were plotted). 119
4.9b	Continued. 120
4.10	Pseudo-section plots of the responses of the dense Cholame Valley profile (no remote stations). 121

LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
4.11 The misfit plots from distortion analysis (using wide period band) of the series of rotational angles (-45° to 45°) against stations along the profiles from west to east (no remote sites) of MM1, MM2, MM3 and CV profiles.	124
4.12 A plot of the residual (in % for apparent resistivity, and \circ for phases) between the actual response estimates and the best fitting distorted 2-D responses assuming a strike angle of 40°	126
4.13 Resistivity cross-section from single mode inversions.	132
4.14 The same plots as in Fig 4.13, but from joint inversion of TM and TE modes with tipper.	134
4.15 The same plots as in Fig 4.14, but inverting with increasing the data variance on the area the estimated strike is not consistent with the SAF strike.	137
4.16 The plot of the conductance which is the conductivity vertically integrated over the depth of 10 km, from (blue) the best fitting model (Figure 4.14a), and from (red) the preferred model (Figure 4.15a).	138
4.17 Comparison of preferred models from three profiles, MM1, MM2, MM3, from north to south in the Middle Mountain area.	141
4.18 An alternative model for MM3, referred to as MM3-B, whereas the model in Figure 4.17c is MM3-A.	142
4.19 Comparison of conductances from three profiles, MM1 (red), MM2 (green), MM3 (blues) from north to south at the Middle Mountain area.	143
4.20 The same plots as in Fig 4.14, but for CV profile.	144
4.21 The conductance plots for CV (pink) and MM1 (red) profiles.	145
4.22 Case studies	149
4.23 Similar to Figure 4.22, but show the results from single mode inversions, TM (middle) or TE (right).	151
4.24 Results from inversion based on different vertically sharp boundaries.	155

LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
4.25 Results from inversion based on different vertically sharp boundaries on the case study models showing in Figure 4.22.	156
4.26 Results after inversion based on fixed fault zone with different resistivities	159
4.27 Plot of observed responses (blue) and the calculated responses from various models.	160
4.28 Results after inversion based on fixed fault zone with different shapes of fault zones ($3 \Omega - m$)	162
4.29 Results from inversion based on vertically sharp boundary from 2 to 10 km depths at two different locations as prior models	164
4.30 Results from inversion based on fixed fault structure with $1 \Omega - m$ and 244-m-wide	164
4.31 The MM1 preferred model.	165
4.32 Same as Figure 4.31, but for CV profile.	165
4.33 Plots of experimentally determined conductivity of NaCl (Quist and Marshall, 1968) and KCl (Hwang et al., 1970) solutions at different molality, temperature and pressure (redrawn from Merzer and Klemperer, 1997).	170
4.34 Resistivity of 3.8 % saline (KCl) with depth in the crust at different conditions; lithostatic pressure and no CO ₂ , lithostatic pressure with CO ₂ and hydrostatic pressure with CO ₂	171
4.35 a) Plot of steady state crustal temperature profile at Parkfield-Cholame region (redrawn from Sass et al, 1997). b) Fluid resistivity at Varian Well, estimated at lithostatic pressure (No CO ₂) and temperature gradient of a). Water is predominantly NaCl at 0.37 M. . .	172
4.36 Plots of porosity required for different values of saline water conductivity to achieve bulk conductivities of 0.1, 0.3, 1 and 3 S/m.	173
4.37 Resistivity model of MM1 profile with subsurface faults (Sims, 1990) plotted on top.	178

LIST OF TABLES

<u>Table</u>		<u>Page</u>
1.1	Constant parameters a and m in Archie's law for different lithologies (Keller, 1987).	17
3.1	Top: Comparison of number of operations (multiplications and additions) needed to form cross-product (X-product) matrices, solve system of equations (SEs). For REBOCC additional operations are required to compute (3.18). Bottom: Comparison of real words required to store sensitivity (Sens.) matrices, cross-product matrices and the interpolation (Interp.) matrix (REBOCC only).	69
4.1	Lists of fault parameters (depths, widths and resistivities) used as a fixed structure in the inversion tests.	158
4.2	Lists of fault parameters (depths, widths and resistivities) used as a fixed structure in the inversion tests.	161

AN EFFICIENT DATA-SUBSPACE TWO-DIMENSIONAL MAGNETOTELLURIC INVERSION AND ITS APPLICATION TO HIGH RESOLUTION PROFILE ACROSS THE SAN ANDREAS FAULTS AT PARKFIELD, CALIFORNIA

1. INTRODUCTION

1.1. Introduction to the Magnetotelluric (MT) Method

The magnetotelluric (MT) method is a natural-source technique utilizing the time-varying electromagnetic (EM) fields as a source to determine the conductivity distribution of the Earth. The MT method is generally used for reconnaissance of unexplored area and as a supplement to seismic surveys.

The external time varying plane-wave magnetic field is vertically incident on the Earth's surface and penetrates into the ground (which is very conductive compared to the air) resulting in an induced electric field and secondary magnetic fields. Because the penetration depth of EM energy depends on frequency, the depth of investigation based on MT data is much more definitive than that based on gravity or magnetic data (Vozoff, 1972). MT data can probe the near surface to as deep as thousand kilometers (Egbert and Booker, 1992) depending on frequency of waves measured and also on the purpose of the experiment.

In MT, the total magnetic and electric fields are measured by an observer at the surface of the Earth. The relationship between these fields give us a physical

response that depends on a volumetric average of the Earth's conductivity. MT data can thus be inverted for conductivity structure which can then be related to other physical properties, such as porosity of the rocks. This has an impact on both economical aspects, e.g., identifying zones of mineralization (Orange, 1989), and more fundamental studies, e.g. detecting fluid in the deep crust or determining the structure of the upper mantle (see Jones, 1992).

The basic physics for the MT method is simple and has been around for a century. MT experiments were first carried out in the early 1950s in the Soviet Union by Tikhonov (Tikhonov, 1950) and in France by Cagniard (Cagniard, 1953). Since then substantial progress has been made in many aspects of this method including the theory (see Vozoff, 1986), and the development of the tools for interpretation, e.g., data processing (Egbert and Booker, 1986; Egbert, 1997), and 1-D, 2-D and 3-D inversions and forward modelings, (see Mackie et al., 1994; Rodi and Mackie, 1999; Siripunvaraporn and Egbert, 1999) as well as its application to exploration (e.g. Orange, 1989) and to fundamental studies (e.g. EMSLAB, 1989; Jones, 1992).

1.1.1. Source of Signal

The frequency range that is commonly used for the MT method is from around 10^{-4} Hz to about 10^4 Hz. The source of the EM signals for frequencies below 1 Hz is due to the current flow in the ionized layers surrounding the Earth. The current flow is powered by the solar activity and the relative motions of the earth, sun and moon (Vozoff, 1972). The energy of the ionospheric source increases significantly as frequency decreases (Fig 1.1). For the frequencies above 1 Hz, most of the energy comes from worldwide electrical thunderstorm activity within the atmosphere. The energy of the natural source is very weak in the frequency range

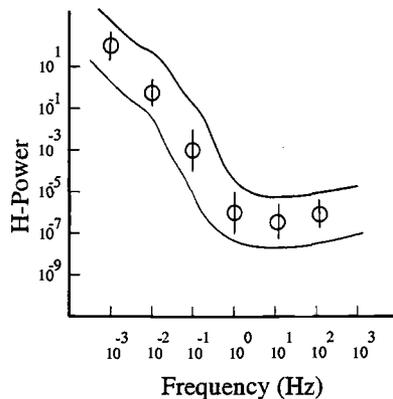


FIGURE 1.1. Power spectrum of the magnetic field. The energy is high for the low frequency and very low around 1 to 10 Hz.

from 0.1 to 10 Hz, the so-called “dead” band. With low signal in this band, cultural noise can dominate the signal resulting in low signal to noise (S/N) ratio.

1.1.2. Maxwell's Equations

In MT, the fundamental equations for describing the behavior of the EM waves are given by the quasi-static approximation to Maxwell's equations (assuming a time dependence $e^{-i\omega t}$ and no displacement currents):

$$\nabla \times \mathbf{E} = i\omega\mu\mathbf{H}, \quad (1.1)$$

$$\nabla \times \mathbf{H} = \sigma\mathbf{E}. \quad (1.2)$$

where \mathbf{E} and \mathbf{H} are the electric and magnetic fields respectively, ω is the angular frequency, σ and ρ are the conductivity (unit of Siemens-m) and resistivity (Ohm-m) respectively and μ is the permeability approximated to the permeability of the air, μ_0 . Either \mathbf{E} or \mathbf{H} may be eliminated to yield the second order differential forms,

$$\nabla \times \nabla \times \mathbf{E} = i\omega\mu\sigma\mathbf{E}, \quad (1.3)$$

$$\nabla \times \rho\nabla \times \mathbf{H} = i\omega\mu\mathbf{H}. \quad (1.4)$$

For a 1-D Earth, the conductivity (or resistivity) will be assumed to not vary in either of the horizontal directions (x and y), and will only vary with the depth (z). From Maxwell's equation (1.3), the electric field in the 1-D case can be written as

$$\frac{\partial^2 E_x}{\partial z^2} = -i\omega\mu\sigma E_x \quad (1.5)$$

A similar equation can be obtained from equation (1.4) for the magnetic field.

In the case of a half space of constant conductivity, σ , the solution for equation (1.5) as a function of depth is

$$E_x(z) = E_0 e^{(-1+i)z/\delta}. \quad (1.6)$$

The skin depth (δ) is defined as the depth (in meters) at which the amplitude decays by a factor of $1/e$ from its original value,

$$\delta(T) = \sqrt{\frac{2}{\omega\mu\sigma}} \approx 500\sqrt{\rho T} \quad (1.7)$$

where T is period in second (the inverse of frequency, f). From equation (1.7), the skin depth depends on two values, the conductivity and the frequency. In the case of a homogeneous medium (Figure 1.2a), the electric fields of the lower frequencies decay more slowly with depth than that of the higher frequencies, which result in deep penetration. For a given frequency (Figure 1.2b), the lower the conductivity (or the higher the resistivity), the greater the skin depth. The physical meaning of the skin depth in MT is quite important, especially for designing an experiment. Spies (1989) estimated the "maximum" depth of investigation to be about $1.5\delta(T)$. If the goal is to study deep structure in a conductive region, low frequency data must be used to achieve large penetration depths.

Simple algebra from the 1-D solution (1.6) gives us an expression of the conductivity (or resistivity) in term of the induced fields,

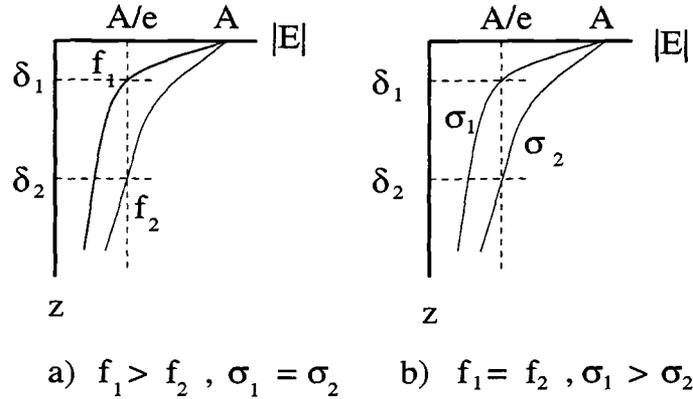


FIGURE 1.2. Skin depth comparisons for a) different frequencies, and b) different conductivities. (Adapted from Dobrin and Savit, 1988).

$$\sigma^{-1} = \rho = \frac{i}{\omega\mu} \left(\frac{E_x}{H_y} \right)^2. \quad (1.8)$$

In practice, where the conductivity is not homogeneous, the right hand side of (1.8) still reflects a volume average of the conductivities in the skin depth below the measurement point at the surface.

1.1.3. Data Acquisition and Processing

As described above, the magnetic and electric fields are measured at the surface of the Earth. In the case of the magnetic field, all three components, H_x , H_y and H_z are measured, while for the electric field only horizontal components E_x and E_y are measured. Note that the vertical magnetic field H_z is assumed to be generated via the induction process, and the vertical electric field is theoretically zero at the surface of the Earth. The fields are measured in the time domain and then transformed into the frequency domain. The electric and magnetic fields in the frequency domain are related via the complex impedance tensor \mathbf{Z} ,

$$\mathbf{E}(\omega) = \mathbf{Z}(\omega)\mathbf{H}(\omega), \quad (1.9)$$

$$\text{or} \quad \begin{bmatrix} E_x(\omega) \\ E_y(\omega) \end{bmatrix} = \begin{bmatrix} Z_{xx}(\omega) & Z_{xy}(\omega) \\ Z_{yx}(\omega) & Z_{yy}(\omega) \end{bmatrix} \begin{bmatrix} H_x(\omega) \\ H_y(\omega) \end{bmatrix}.$$

A similar linear relationship can be found between the vertical and horizontal magnetic field:

$$H_z(\omega) = AH_x + BH_y, \quad (1.10)$$

where A and B are complex transfer functions. These relationships follow from the assumption that the sources are spatially uniform (plane waves; e.g. Egbert and Booker, 1989).

The linear relationships of (1.9) and (1.10), normally, do not hold exactly, because of measurement errors, and/or because the plane wave source field assumption is not exactly valid. A simple technique to estimate the impedance tensor can yield results that are very biased or wildly oscillatory. The estimation of the impedance tensor from noisy data has progressed substantially in the last few decades. The remote reference technique (Gambel et al., 1979) in which the horizontal magnetic fields at the remote site are simultaneously recorded and then used to correlate with the EM fields at the local site, has dramatically reduced bias due to local noise. Another development that has improved the quality of the impedance tensor is use of improved statistical techniques, e.g. robust data adaptive weighting schemes (see Jones et al., 1989 for comparison of different schemes). Egbert (1997) has developed a practical robust processing scheme which is based on multivariate statistical methods. To get rid of coherent noise (e.g. noise from the electric train BART in the San Francisco area), more than two stations were recorded simultaneously at different locations, and then data from all channels (not only magnetic fields) are used in the multivariate statistical method to improve the signal-to-noise ratios.

1.1.4. One-Dimensional MT

The 1-D assumption that the resistivity varies only with depth may be valid for high frequency data over a sedimentary basin, or at very long periods when the EM fields are sampling the continental upper mantle (Jones, 1992). In this case, the diagonal terms of the impedance tensor (1.9) are both zero, and the off-diagonal terms are equal but opposite in sign. The apparent resistivity and phase are then defined in terms of the (essential equal) off-diagonal terms as

$$\begin{aligned}\rho_a(\omega) &= \frac{1}{\omega\mu} |Z_{xy}(\omega)|^2, \\ \phi(\omega) &= \tan^{-1} \frac{\text{Im}[Z_{xy}(\omega)]}{\text{Re}[Z_{xy}(\omega)]}.\end{aligned}\tag{1.11}$$

The apparent resistivity ρ_a must be clearly distinguished from the resistivity ρ . The resistivity ρ is a property of the medium, whereas ρ_a represents the scaled total impedance of the layers and zones of varying conductivity in the Earth measured at the surface. For a uniform earth, ρ_a is the true resistivity (see equation (1.8)) of the half space, and the phase is 45 degrees.

The apparent resistivity and phase for a 1-D earth (and almost for the 2-D case) are not independent functions, but are related to each other via a Hilbert-transform type relationship (Weidelt, 1972). The first-order approximation for this Hilbert-transform is given by

$$\frac{\partial \log \rho_a}{\partial \log \omega} \cong \frac{\phi}{45} - 1.\tag{1.12}$$

This relationship is very helpful in controlling the quality of the data, because the phase curve can be predicted from the apparent resistivity. However, the reverse is not completely true. From the phase curve, only the shape of the apparent resistivity can be predicted, because there is an unknown scaling factor involved in the apparent resistivity.

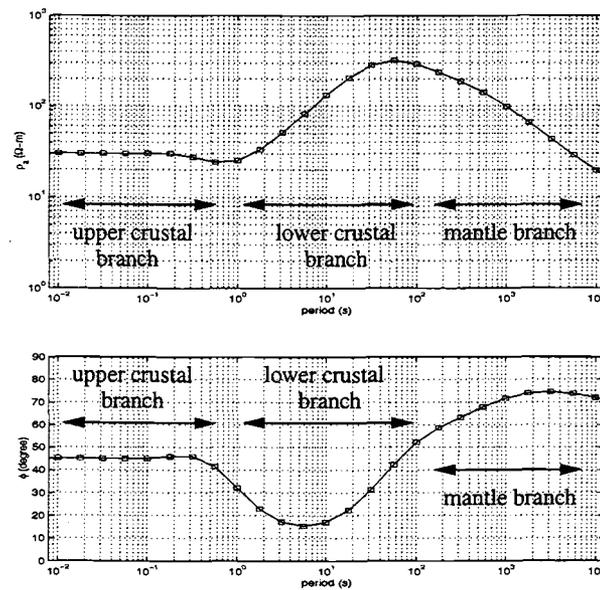


FIGURE 1.3. Hypothetical 1-D MT apparent resistivity and phase for a simplified continental region, with a conductive upper crust, a resistive lower crust and conductive mantle (After Mackie et al., 1988).

For a half-space model, the apparent resistivity is the true resistivity value, and the phase is 45 degrees. For a layered 1-D Earth (e.g. in the hypothetical 1-D MT model, Figure 1.3), long period responses correspond to greater depth, and short period responses correspond to shallow depth. At sufficiently short periods, the skin depth is small enough that practically no energy is able to penetrate to the second resistive unit, therefore, the apparent resistivity approaches the actual resistivity of the upper crust, and the phase is 45 degrees. In the region of high electrical conductivity (mantle branch), the induced electric field is reduced, the apparent resistivities are therefore low and the phases are high (near 90 degrees). In the intervening region of low conductivity (the lower crustal branch), the apparent resistivities are high and the phases are low (close to 0 degrees). Note that the phases

and apparent resistivities in this example (Figure 1.3) do obey the Hilbert-transform relationship described in (1.12).

1.1.5. Two-Dimensional MT

In a 2-D Earth, the coordinate axes are rotated until one of them is along strike, e.g. y is along the strike and x is perpendicular to the strike. As in the 1-D, the diagonal terms (Z_{xx} and Z_{yy}) of the impedance tensor (1.9) are still zero. The non-diagonal terms (Z_{xy} and Z_{yx}) are not zero, and in general are now unequal. Maxwell's equations, (1.3) and (1.4), separate completely into two modes: the TE-mode (transverse electric) describes the case for which the electric field (E_x) is parallel to the strike, and the magnetic field (H_y and H_z) is perpendicular to the strike. The TM-mode (transverse magnetic) describes the case for which the magnetic field (H_x) is parallel to the strike, and the electric field (E_y and E_z) is perpendicular.

The apparent resistivities and phases for the TE and TM modes respectively are

$$\rho_{xy}(\omega) = \frac{1}{\omega\mu} |Z_{xy}(\omega)|^2, \quad (1.13)$$

$$\phi_{xy}(\omega) = \tan^{-1} \frac{\text{Im}[Z_{xy}(\omega)]}{\text{Re}[Z_{xy}(\omega)]},$$

and

$$\rho_{yx}(\omega) = \frac{1}{\omega\mu} |Z_{yx}(\omega)|^2, \quad (1.14)$$

$$\phi_{yx}(\omega) = \tan^{-1} \frac{\text{Im}[Z_{yx}(\omega)]}{\text{Re}[Z_{yx}(\omega)]}.$$

These two modes exhibit quite different behavior. To understand the behavior of the responses of both modes, we consider a classic example, a fault model

(Figure 1.4a) with a conductor of $10 \Omega\text{-m}$ on one side, and a resistor of $1000 \Omega\text{-m}$ on the other side. The MT stations are measured along the x-axis at the surface.

Figure 1.4b displays the apparent resistivities and phases versus period (blue for TM and red for TE) for stations on both sides of the fault (3 km from the fault). At station A, on the conductive side, at short periods where the fields have not sensed the conductivity discontinuity, both TM and TE apparent resistivities and phases are at $10 \Omega\text{-m}$ and 45 degrees respectively. At longer periods, the responses of both modes are spread in opposite directions. The apparent resistivity of the TM mode starts to dip below its actual value (phase increases), while that of the TE mode rises above its actual value (phase decreases). On the resistive side (station B), the modes separate at shorter periods. The TM mode, which is less sensitive to the conductor shows almost constant apparent resistivities and phases. This is not the case for the TE mode which is sensitive to the conductor.

This can be explained because in the TM mode the currents (flowing across the fault) are tightly gathered close to the surface at great distances from the interface on the conductive side, but are broadly distributed near the surface on the resistive side. Approaching closer to the fault from the conductive side (at station A), the current density near the surface becomes less (since the current density must be continuous), causing the apparent resistivity to dip. For TE mode the currents (flowing parallel to the fault) are able to adjust smoothly across the fault, causing the apparent resistivity to increase.

Figure 1.4c displays the apparent resistivities and phases along the x-axis at the surface from left to right across the fault at a period of 1 second. At large distances from the fault, apparent resistivity is asymptotic to the true resistivity value, and phase approaches 45° . From left to right closer to the fault, in the TE mode, the apparent resistivity and phase vary smoothly, while in the TM mode, both

apparent resistivity and phase are discontinuous across the fault. The discontinuous TM response can be explained by the fact that current density ($\mathbf{J} = \sigma\mathbf{E}$) must be continuous across the contact. Since the conductivity is discontinuous, the electric fields must be discontinuous (leading to a discontinuous responses across the fault).

1.1.6. Static Distortions

Near surface inhomogeneities can cause a shift in the log apparent resistivity (Figure 1.5) as can clearly be seen in the TM mode data collected at the Parkfield area (Unsworth, 1998). Each station is less than 100 meters away. The apparent resistivities plotted on a log-log scale appear to shift by a constant factor away from each other, while the phases are unaffected. These shifts by the local near surface inhomogeneities are due to electric charges at boundaries between conductivity contrasts which effect the electric field. These effects have been termed “static shifts”, which could be misleading. A more general term is “static distortions”. It is necessary to correct the MT responses affected by the static distortions prior to interpretation, to prevent an erroneous model (Jones, 1992). These static distortions can be corrected manually, with a prior information such as from drilling or these parameters can be solved for in an inversion algorithms (Smith and Booker, 1991; deGroot-Hedlin and Constable, 1993).

1.2. Electrical Properties of Rocks

The electrical properties of the rocks relevant to interpretation of electromagnetic geophysical data include electrochemical activity caused by electrolytes in the ground, the dielectric constant (i.e., the capacity of the rock material to capacitatively store electric charge), and resistivity (the reciprocal of conductivity)

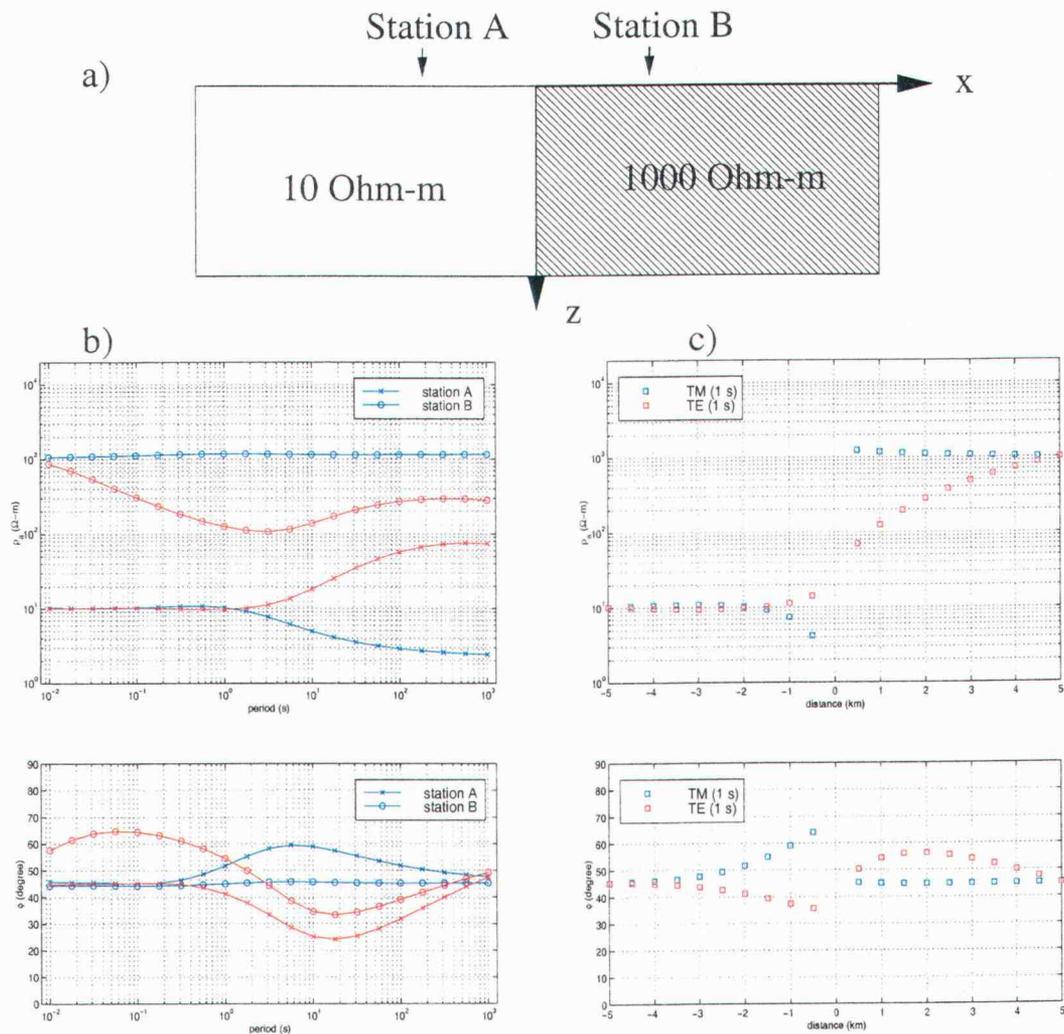


FIGURE 1.4. Effect of conductivity contrast to the apparent resistivities and phases for (blue) TM and (red) TE modes. a) Two-dimensional fault model consists of 10 Ω -m on the west side and 1000 Ω -m on the east side. b) Apparent resistivity and phase curves versus periods for the model displayed in a) at 3 km from fault. c) Apparent resistivity and phase curves versus distances from the fault for the model displayed in a) at the period of 1 second.

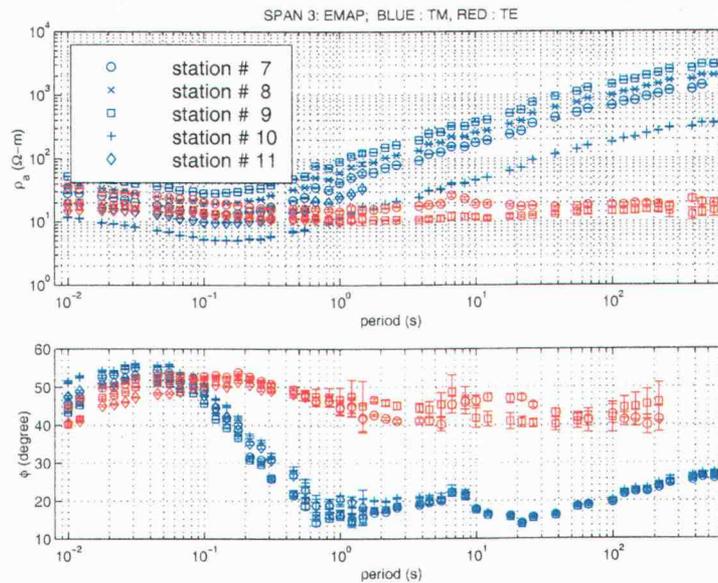


FIGURE 1.5. Plots of the apparent resistivity and phase of the real field data collected at Parkfield, California. Note that the spacing between stations is really closed, less than 100 m. A shift in the log-log apparent resistivities can be observed from stations to stations, while unaffected to the phases.

controlling the amount of the current that passes through the rock when a specified potential difference is applied. In MT applications, the resistivity is measured via the relationship between the electric and the magnetic fields measured at the surface of the Earth. At the frequency ranges used in MT, the dielectric constant does not appear in the quasi-static approximation to Maxwell's equations, and is thus not relevant. Resistivity is expressed in $\Omega - m$ (ohm meters), and its reciprocal, the conductivity, in S/m (Siemens per meter). The resistivity of rocks depend on temperature, pressure and other environmental factors (e.g. fluid present in the rock).

1.2.1. Conduction Mechanisms

The mechanisms of conduction differ significantly from one material to another. Four major conduction mechanisms are metallic conduction, electronic semiconductor, solid electrolytic conduction, and aqueous electrolytic conduction. This section is summarized from Keller (1987) and Zhdanov and Keller (1994).

Metallic conduction dominates in metallic materials such as gold and copper, which have extremely high conductivities, exceeding 10^5 S/m. The conductivity of a metal depends greatly on the number of the free electrons, and also the scattering time of the electrons under an applied electric field. In the case of higher temperatures, the probability that the free electrons will collide with each other increases, which results in a decrease of conductivity with temperature in a metal.

In a semiconductor relatively few electrons are free to move through the material, in comparison with the number of free electrons in a metal. In contrast to the free movement of the electrons, electronic semiconductor requires a significant amount of energy to drive the electrons out of the home atoms to move through the crystal lattice. This energy is commonly provided as heat, and so conductivity is found to increase with temperature for this mechanism. This type of material is called a semiconductor, and has a conductivity greater than a resistor (10^{-8} S/m) and less than a conductor (10^5 S/m). Some economically important semiconductors are mineral sulfides or oxides. Carbons in the form of graphite can behave as a metallic conductor along one crystal plane, but can behave like a semiconductor in the third direction. In some semiconductors, the energy required for activation of conduction electrons is so large that essentially no conduction occurs. These materials are termed as insulators; silicates, and carbonates are examples. The con-

ductivity of both metals and semiconductors are unlikely to change with frequency in the frequency ranges used in geophysics (Keller, 1987).

Most rock forming minerals are solid electrolytes. Electrolytic conduction can take place in ionic bonded crystals. When an external electric field is applied to an ionic bounded crystal structure, the force exerted on each ion by that field is quite small compared to the strength of the binding forces. For the ideal crystal, there will be no conduction by ion movement, or the solid electrolysis, taking place. Generally, crystals are imperfect, with ions of the wrong valence substituted in the lattice or missing from the lattice. Higher temperature increases the movement of ions in random directions. When an electric field is applied, the ions jump statistically in one direction more than the other direction resulting in a net flow of current.

Generally, extensive studies of rock and mineral conductivities have shown that at high temperature the conductivity in a dry rock is a function of temperature, regardless of composition (Figure 1.6). In molten rocks (such as in and around volcanos, and possibly at depth in the crust and upper mantle), the conduction mechanism is basically similar to the mechanism of the solid electrolytes.

In the majority of rocks, water in pore spaces is the most important factor in controlling electrical conductivity. Rock conductivity with fluid present depends mainly on the amount of fluid in the pores, the salinity of the water, the temperature, and the way the water is distributed through the rock. In this case, the conduction is in the form of aqueous electrolytes, with ions moving in the fluid with the application of an electric field. Figure 1.7 shows that increasing salt concentration (NaCl and KCl) in the water increases the conductivity. Also, at a given salinity, elevated temperature increases the conductivity.

The amount of water present in rocks controls the conductivity of the rocks. An empirical Archie's law given as

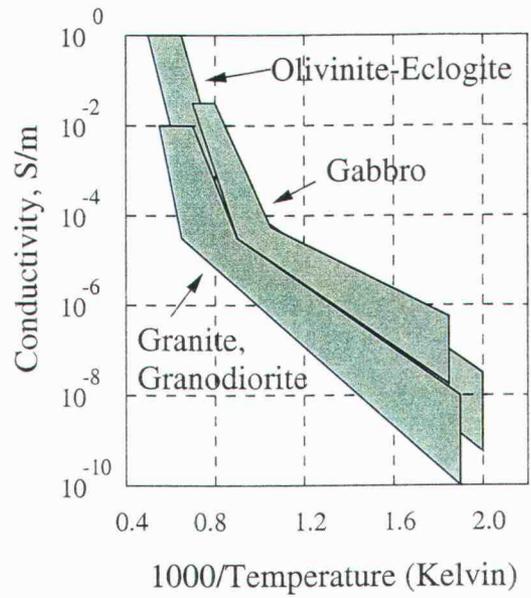


FIGURE 1.6. Generalized plots of resistivity as a function of inverse temperature for different rocks. At low temperatures the rocks with more silicic composition tend to have lower conductivity than rocks of more basic compositions. At higher temperature, there is little conductivity variation caused by compositional difference. (redrawn from Zhdanov and Keller, 1994).

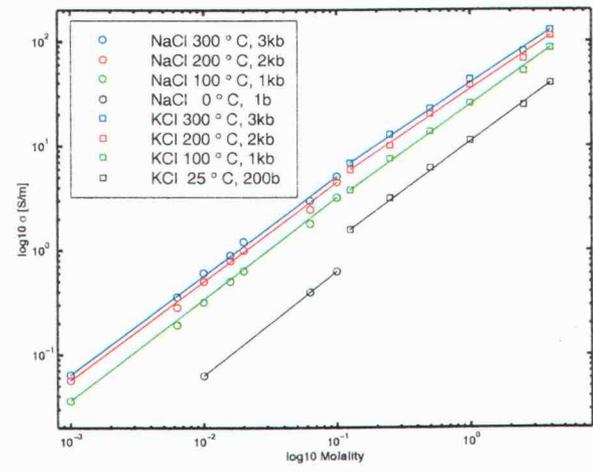


FIGURE 1.7. Plots of conductivity versus salinity in log10 Molality, at different pressure and temperature condition (redrawn from Merzer and Klemperer, 1997).

$$\sigma_B = a\sigma_w W^m, \quad (1.15)$$

is normally used to relate electrical resistivity to porosity, where σ_B is the bulk conductivity of the rock, σ_w is the conductivity of the electrolyte in the pore structure, W is the volume fraction of electrolyte-filled pore space or the porosity, and a and m are empirically determined parameters.

Description of rock	a	m
Weakly cemented detrital rocks, such as sand, sandstone, and some limestones, with a porosity range from 25 to 45 %, usually Tertiary in age	0.88	1.37
Moderately well cemented sedimentary rocks, including sandstones and limestones, with a porosity range from 18 to 35 %, usually Mesozoic in age	0.62	1.72
Well-cemented sedimentary rocks with a porosity range from 5 % to 25 %, usually Paleozoic in age	0.62	1.95
Highly porous volcanic rocks, such as tuff, aa, and pahoehoe, with porosity in the range 20 % to 80 %	3.5	1.44
Rocks with less than 4 % porosity, including dense igneous rocks and metamorphosed sedimentary rocks	1.4	1.58

TABLE 1.1. Constant parameters a and m in Archie's law for different lithologies (Keller, 1987).

Pore geometry controls the parameters a and m , and can be categorized in three different styles; the intergranular space (space left over when grains are deposited) in clastic detrital rocks, the vesicular porosity in rocks such as extruded

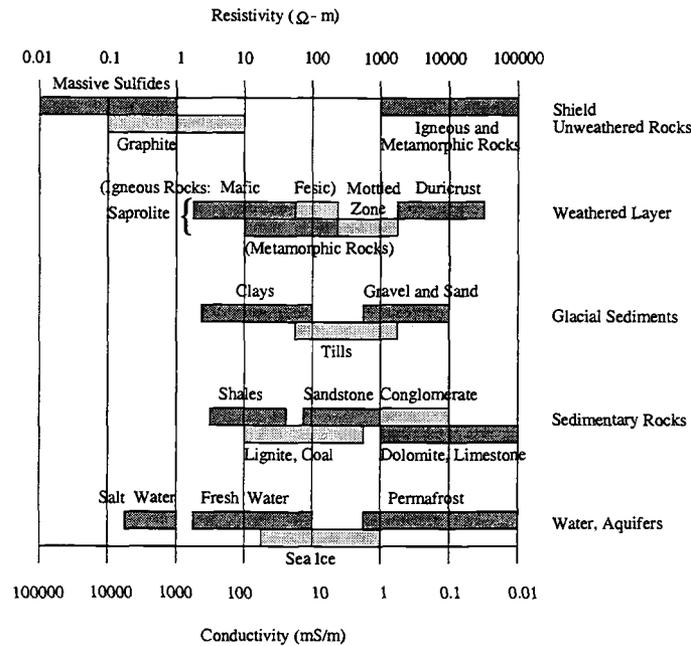


FIGURE 1.8. Typical ranges of resistivities of earth materials (redrawn from Palacky, 1987).

lavas and fracture, joint and microcrack pores in crystalline rocks. Vesicular porosity occurs where large, poorly connected equant pores form as expanding gas is frozen in a solidifying lava flow. For a given porosity and water content, in these three types, fracture porosity will result in the highest rock conductivity, while the vuggy porosities will yield lowest conductivity. Various rock lithologies with different parameters of a and m are given in Table 1.1. The parameter a can vary from slightly less than unity in clastic detrital rocks to values as large as 4.5 in vesicular rocks. The parameter m varies from values slightly greater than unity in dense, fractured rock to values as great as 2.3 in tightly cemented clastic rocks.

1.2.2. Resistivity Ranges of Geological Structure

This section is summarized from Palacky (1987). Figure 1.8 summarizes the typical ranges of resistivities of some common earth materials. Resistivities can be found outside of these typical ranges at some locations. The resistivity of sulfide deposits depends on many factors such as their formation, geologic history and weathering. In most cases, volcanic-associated massive sulfide deposits are normally very conductive (less than $1 \Omega - m$). Igneous and metamorphic rocks in the shield areas where sulfide deposits commonly occur are very resistive (over $1000 \Omega - m$). In areas where the weathered layer has not been removed, saprolite is the most conductive component, with resistivities between 2 and $200 \Omega - m$, while the mottled zone and duricrust are less conductive. The resistivity of the glacial sediments covering most of the previously glaciated areas, such as Canada and Scandinavia, vary from conductive (clay) to poorly conductive (sand, gravel and tills). The resistivity of the sedimentary rocks depends on various factors, the clay content, porosity, dissolved mineral content, and water saturation of rocks. The values given in Figure 1.8 are purposely to suggest that in most area dolmites, limestones, and conglomerates are more resistive than shales.

1.3. Magnetotelluric Studies - Chapters of this Thesis

The body of this dissertation is comprised of three main chapters. Chapters 2 and 3 are the developments of numerical algorithms: 3-D forward modeling to run on a parallel computer and 2-D inversion technique for using in interpreting MT data. The last chapter is about the analysis and interpretation of MT data collected across the San Andreas Fault at Parkfield, California.

The main object in Chapter 2 is to develop a 3-D forward modeling algorithm that works in a parallel environment. Most of the existing algorithms are mainly developed to work on a workstation or personal computer. Modeling geologically realistic problems in 3-D requires huge computation resources, which is not feasible to perform on a workstation. To allow for solving very large problems, we develop an algorithm for the 3-D forward modeling to run on a massively parallel distributed memory computer, the Thinking Machines Co. CM-500e with memory upto 8 Gbytes available at the College of Oceanic and Atmospheric Sciences. We first describe theory and algorithms available for serial computations. We then explain how we implement these serial algorithms to work on a parallel machine. The main problem here is the lack of parallel preconditioners. One of our main goals is to find the appropriate parallel preconditioner that improves the convergent rate.

In Chapter 3, the primary objective is to present a new 2-D MT inversion that is efficient, fast, and requires fewer computation resources, compared to other available algorithms. The new technique is called the reduced basis Occam's inversion (REBOCC). It is based on the fact that MT data is smooth and redundant. Therefore, the solution can be constructed from subset of the data, without significantly loss of detail, and still fit all data adequately. Numerical experiments on synthetic and real data sets show that REBOCC is relatively faster than other 2-D inversion algorithms, such as NLCG (Rodi and Mackie, 1999), and Occam (deGroot-Hedlin and Constable, 1990), but slower than RRI (Smith and Booker, 1991). However, we could not always get RRI to converge to a reasonable solution. The basic idea behind REBOCC should be more broadly applicable, in particular to 3-D MT inversion.

In the last chapter, the main interest is to apply REBOCC inversion to interpret the MT data set across the San Andreas Fault at Parkfield in term of the resistivity structure. This area is special because of the moderate "characteristic"

earthquakes occurring at an average of 22 years. Several pieces of evidence indicate that the SAF is weak, and that high pressure fluids may play an important role in controlling the mechanism of earthquakes. Electrical resistivities of rocks are strongly influenced by fluids in pore spaces and fractures, thus MT have an advantage over other methods for mapping the fluid distribution in the fault zones. We first start with the data acquisitions and qualitative interpretations of the data, the dimensionality analysis, and the results from REBOCC inversion. Many hypothesis tests and case studies are performed in order to constrain other geophysical results and also propose some new idea.

1.4. References

- [1] EMSLAB, 1989, Special Issue of Journal of Geophysical Research devoted to the EMSLAB experiment (14 papers)., *J. Geophys. Res.*, **94**, 14093-14283.
- [2] Cagniard, L., 1953, Basic Theory of the Magnetotelluric Method of Geophysical Prospecting, *Geophysics*, **18**, 605-635.
- [3] Constable, C. S., Parker, R. L., and Constable, C.G., 1987, Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics*, **52**, 289-300.
- [4] deGroot-Hedlin, C., and Constable, S., 1990, Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data, *Geophysics*, **55**, 1613-1624.
- [5] deGroot-Hedlin, C., and Constable, S., 1993, Occam's Inversion and the North American Central Plains Electrical Anomaly, *J. Geomag. Geoelectr.*, **45**, 985-999.
- [6] Dobrin, M.B. and C.H. Savit, 1988, Introduction to Geophysical Prospecting 4th, McGraw-Hill Inc.
- [7] Egbert, G. D., and J. R. Booker, 1986, Robust estimation of geomagnetic transfer functions, *Geophys. J. R. astr. Soc.*, **87**, 173-194.
- [8] Egbert, G. D., and J. R. Booker, 1989, Multivariate analysis of geomagnetic array data, 1, The response space, *J. Geophys. Res.*, **94**, 14227-14247.

- [9] Egbert, G.D. and J.R. Booker, 1992, Very Long Period Magnetotellurics at Tucson Observatory: Implications for Mantle Conductivity, *J. Geophys. Res.*, **97**, 15,099-15,112.
- [10] Egbert, G. D., 1997, Robust multiple station magnetotelluric data processing, *Geophys. J. Int.*, **130**, 475-496.
- [11] Keller G. V., 1987, Rock and Mineral Properties, in *Electromagnetic Methods in Applied Geophysics—Theory*, vol. 1, by M. N. Nabighian, Society of Exploration Geophysicists, Tulsa, Okla, 13-52.
- [12] Jones, A. G., 1992, Electrical conductivity of the continental lower crust, *in* Fountain, D. M., Arculus, R. J., and Kay, R. W., Eds., *Continental Lower Crust*: Elsevier Science Publ. Co., Inc., 81-143.
- [13] Jones, A. G., A. D. Chave, G. Egbert, D. Auld, and K. Bahr, 1989, A Comparison of Techniques for Magnetotelluric Response Function Estimation, *J. Geophys. Res.*, **94**, 14201-14213.
- [14] Gamble, T. D., W. M. Goubau and J. Clarke, 1979, Magnetotellurics with a remote magnetic reference, *Geophysics*, **44**, 53-68.
- [15] Mackie R. L., B. R. Bennett, and T. R. Madden, 1988, Long-period magnetotelluric measurements near the central California coast: a land-locked view of the conductivity structure under the Pacific Ocean, *Geophys. J.*, **95**, 181-194.
- [16] Mackie, R.L., J.T. Smith and T.R. Madden, 1994, Three-dimensional electromagnetic modeling using finite difference equations: The magnetotelluric example, *Radio Science*, **29**, 923-935.
- [17] Merzer M. and S. L. Klemperer, 1997, Modeling Low-frequency Magnetic-field Precursors to the Loma Prieta Earthquake with a Precursory Increase in Fault-zone Conductivity, *Pure and Appl. Geophys.*, **150**, 217-248.
- [18] Palacky G. J., 1987, Resistivity characteristics of geologic targets, in *Electromagnetic Methods in Applied Geophysics—Theory*, vol. 1, by M. N. Nabighian, Society of Exploration Geophysicists, Tulsa, Okla, 53-129.
- [19] Rodi, W. L., and Mackie, R. L., 1999, Nonlinear Conjugate Gradients Algorithm for 2-D Magnetotelluric Inversion, *Geophysics*, submitted.
- [20] Orange, A. S., 1989, Magnetotelluric Exploration for Hydrocarbons, *Proc. IEEE*, **77**, 287-317.
- [21] Siripunvaraporn W. and G. Egbert, 1999, REBOCC: An Efficient Data-Subspace Inversion for Two-Dimensional Magnetotelluric Data, *Geophysics*, submitted.

- [22] Smith J. T, and Booker, J. R., 1988, Magnetotelluric inversion for minimum structure, *Geophysics*, **53**, 1565-1576.
- [23] Smith J. T, and Booker, J. R., 1991, Rapid Inversion of Two- and Three-Dimensional Magnetotelluric Data, *J. Geophys. Res.*, **96**, 3905-3922.
- [24] Spies B. R., 1989, Depth of investigation in electromagnetic sounding methods, *Geophysics*, **54**, 872-888.
- [25] Tikhonov, A.V., 1950, Determination of Electrical Characteristics of the Deep Strata of the Earth's Crust, *Dokl. Akad. Nauk*, **73**, 295.
- [26] Unsworth M. J., P. Bedrosian, J. R. Booker, G. D. Egbert, M. Eisel, W. Siripunvaraporn, 1998, Imaging Fault Zone Structure with Electromagnetic Exploration techniques, *EOS, Trans. Am. Geophys. Union*, **79**, 45, Fall Meet. Suppl., F227.
- [27] Vozoff, K., 1972, The Magnetotelluric Method in the Exploration of Sedimentary Basins, *Geophysics*, **37**, 98-141.
- [28] Vozoff, K., ed., 1986, Magnetotelluric Methods. Soc. Explor. Geophys. Reprint Ser. No. 5, Publ. by Soc. Expl. Geophys., Tulsa, Okla., ISBN 0-931830-36-2.
- [29] Weidelt, P., 1972, The inverse problem of geomagnetic induction, *Z. Geophys.*, **38**, 257-289.
- [30] Zhdanov M. S. and G. V. Keller, 1994, *The Geoelectrical Methods in Geophysical Exploration*, Elsevier.

2. THREE-DIMENSIONAL MAGNETOTELLURIC FORWARD MODELING ON A MASSIVELY PARALLEL COMPUTER

2.1. Introduction

In the past decade, the development of 3-D Magnetotelluric (MT) modeling algorithms has seen a tremendous amount of progress, both on the integral equation approach and on differential equation methods. The recent developments by Mackie et al. (1993, 1994) and Smith (1996a, 1996b) makes it possible to solve a small problem accurately and efficiently on a desktop workstation.

Mackie and Madden (1993) solved the second order system of equations in the magnetic fields (\mathbf{H}) derived from the integral forms of Maxwell's equations, on a staggered finite difference grid (SFD) by using an iterative relaxation method. They found that 3-D MT relaxation solutions are difficult because the system is so ill-conditioned. Smith (1996a, 1996b) developed a similar algorithm for solving a second order system in terms of the electric fields (\mathbf{E}). He also introduced a procedure designed to reduce the divergence of the magnetic fields. This procedure greatly helps improving the accuracy of the solution and also the speed of the convergent rate. Mackie et al. (1994), later, adapted this procedure to work in their system, which significantly helped solving the problems and obtained an accuracy solutions. Both methods of Smith (1996b) and Mackie et al. (1994) have been shown to be robust, efficient and accurate in a reasonable computational time for small 3-D grids.

However, modeling of many realistic geological problems requires grids which are so large that solution on a workstation is not feasible. Here, to allow for a significantly larger grids, we have implemented and developed an iterative solution algorithm, based on Mackie et al. (1994) and Smith (1996a,1996b), that is appro-

priate for the architecture of a massively parallel distributed memory computer, the Thinking Machines Co. CM-5 which is available at the COAS.

We begin with a review of the basic concepts of the staggered finite difference (SFD) grid used in the 3-D MT problems to obtain an accurate solution, and then derive the linear system of equations corresponding to Maxwell's equations on the SFD. Then, the iterative solution method and the divergence correction are introduced along with an implementation for parallel computation. Finally, we discuss tests of our solver using a 3-D test model previously used by Mackie et al. (1994), including a comparison to the fully converged solution.

2.2. Staggered Finite Difference Grid and Linear System of Equations

In MT, where the displacement currents can be neglected, the differential form of Maxwell's equation (assuming a time dependence $e^{-i\omega t}$) is :

$$\nabla \times \mathbf{E} = i\omega\mu\mathbf{H}, \quad (2.1)$$

$$\nabla \times \mathbf{H} = \sigma\mathbf{E}, \quad (2.2)$$

where \mathbf{E} and \mathbf{H} are the electric and magnetic fields respectively, ω is the angular frequency, σ and ρ are the conductivity and resistivity respectively, μ is the permeability approximated to the permeability of the air, μ_0 , and $i = \sqrt{-1}$ (when not an index). Either \mathbf{E} or \mathbf{H} may be eliminated to yield the second order differential forms,

$$\nabla \times \nabla \times \mathbf{E} = i\omega\mu\sigma\mathbf{E}, \quad (2.3)$$

$$\nabla \times \rho\nabla \times \mathbf{H} = i\omega\mu\mathbf{H}. \quad (2.4)$$

To model the fields numerically, we divide the Earth (and the overlying air layers) into rectangular blocks labeling them as (i, j, k) . Each block has a resistivity

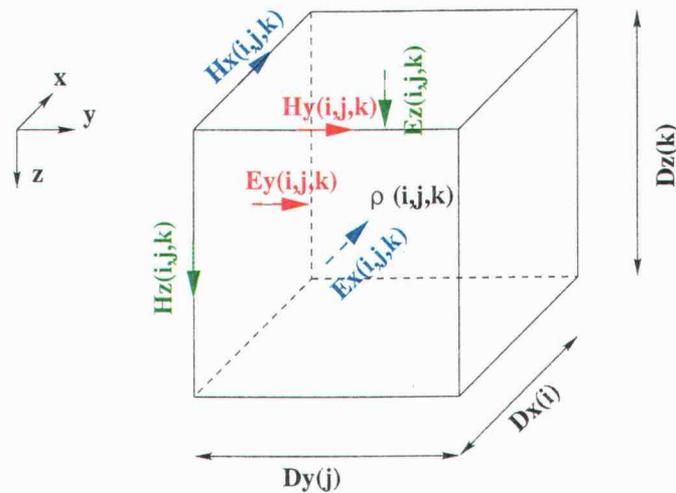


FIGURE 2.1. Positions of magnetic fields \mathbf{H} and electric fields \mathbf{E} on a staggered grid.

value of $\rho(i, j, k)$ where i is the index running from 1 to N_x (number of blocks in the x -direction) and similarly for j and k in y - and z -directions respectively. Using staggered-grid finite difference method, the components of \mathbf{H} are defined on the edges of each block and the components of \mathbf{E} are defined normal to the the block faces, as shown in Figure 2.1, instead of at the corners of the block as for the standard finite difference method. The advantage of sampling the fields on a staggered grid have been discussed previously by Smith (1996a, 1996b) and Mackie et. al. (1993, 1994). In particular, the SFD grid holds the conservation of the currents while the standard FD is not (Smith, 1996a).

To compute the fields, we can either solve the coupled first order equations (2.1) and (2.2) together, or one of the second order systems (2.3) or (2.4). Previous work by Mackie et al. (1993) showed that a direct solution of (2.1) and (2.2) together proves to be accurate but computationally inefficient. Smith (1996a) shows that computing the fields by using SFD from the second order differential equations

either on \mathbf{E} or on \mathbf{H} gives the same result as from a direct solution of the coupled first order system.

Here, we consider solving for the magnetic fields \mathbf{H} from the second order differential equation (2.4). Because the amplitude of the magnetic source field value can be given 1.0 at the top of air layer, the magnetic field values are then automatically computed relative to a normalized source field value. Therefore, the boundary conditions for \mathbf{H} are easy to construct especially at the top.

Writing (2.4) explicitly in the three components \hat{x} , \hat{y} and \hat{z} :

$$\begin{aligned} \frac{\partial}{\partial y} \rho_z \frac{\partial H_x}{\partial y} + \frac{\partial}{\partial z} \rho_y \frac{\partial H_x}{\partial z} - \frac{\partial}{\partial y} \rho_z \frac{\partial H_y}{\partial x} - \frac{\partial}{\partial z} \rho_y \frac{\partial H_z}{\partial x} + i\omega\mu H_x &= 0. \\ \frac{\partial}{\partial x} \rho_z \frac{\partial H_y}{\partial x} + \frac{\partial}{\partial z} \rho_x \frac{\partial H_y}{\partial z} - \frac{\partial}{\partial x} \rho_z \frac{\partial H_x}{\partial y} - \frac{\partial}{\partial z} \rho_x \frac{\partial H_z}{\partial y} + i\omega\mu H_y &= 0. \\ \frac{\partial}{\partial x} \rho_y \frac{\partial H_z}{\partial x} + \frac{\partial}{\partial y} \rho_x \frac{\partial H_z}{\partial y} - \frac{\partial}{\partial x} \rho_y \frac{\partial H_x}{\partial z} - \frac{\partial}{\partial y} \rho_x \frac{\partial H_y}{\partial z} + i\omega\mu H_z &= 0. \end{aligned} \quad (2.5)$$

where ρ_x , ρ_y and ρ_z are the resistivities in the \hat{x} , \hat{y} and \hat{z} directions respectively.

Applying the SFD on the \hat{x} component of equation (2.5), we got

$$\begin{aligned} & \left[i\omega\mu - \frac{\bar{\rho}_z(i, j, k)}{D_y(j)\Delta_y(j)} - \frac{\bar{\rho}_z(i, j-1, k)}{D_y(j-1)\Delta_y(j)} \right. \\ & \quad \left. - \frac{\bar{\rho}_y(i, j, k)}{D_z(k)\Delta_z(k)} - \frac{\bar{\rho}_y(i, j, k-1)}{D_z(k-1)\Delta_z(k)} \right] H_x(i, j, k) \\ & + \frac{\bar{\rho}_z(i, j, k)}{D_y(j)\Delta_y(j)} H_x(i, j+1, k) + \frac{\hat{\rho}_z(i, j-1, k)}{D_y(j-1)\Delta_y(j)} H_x(i, j-1, k) \\ & + \frac{\bar{\rho}_y(i, j, k)}{D_z(k)\Delta_z(k)} H_x(i, j, k+1) + \frac{\bar{\rho}_y(i, j, k-1)}{D_z(k-1)\Delta_z(k)} H_x(i, j, k-1) \\ & + \frac{\bar{\rho}_z(i, j, k)}{D_x(i)\Delta_y(j)} H_y(i+1, j, k) - \frac{\bar{\rho}_z(i, j, k)}{D_x(i)\Delta_y(j)} H_y(i, j, k) \\ & + \frac{\bar{\rho}_z(i, j-1, k)}{D_x(i)\Delta_y(j)} H_y(i+1, j-1, k) - \frac{\bar{\rho}_z(i, j-1, k)}{D_x(i)\Delta_y(j)} H_y(i, j-1, k) \\ & + \frac{\bar{\rho}_y(i, j, k)}{D_x(i)\Delta_z(k)} H_z(i+1, j, k) - \frac{\bar{\rho}_y(i, j, k)}{D_x(i)\Delta_z(k)} H_z(i, j, k) \\ & + \frac{\bar{\rho}_y(i, j, k-1)}{D_x(i)\Delta_z(k)} H_z(i+1, j, k-1) - \frac{\bar{\rho}_y(i, j, k-1)}{D_x(i)\Delta_z(k)} H_z(i, j, k-1) = 0, \end{aligned} \quad (2.6)$$

where

$$\bar{\rho}_z(i, j, k) = \frac{\rho(i, j, k)D_z(k) + \rho(i, j, k-1)D_z(k-1)}{D_z(k) + D_z(k-1)},$$

$$\bar{\rho}_y(i, j, k) = \frac{\rho(i, j, k)D_y(j) + \rho(i, j-1, k)D_y(j-1)}{D_y(j) + D_y(j-1)},$$

$$\Delta_y(j) = \frac{D_y(j) + D_y(j-1)}{2}, \quad \text{and} \quad \Delta_z(k) = \frac{D_z(k) + D_z(k-1)}{2}$$

Whenever we encounter discontinuities of the resistivity model, an average value $\bar{\rho}$ is used which is defined so that the current, $\mathbf{J} = \sigma\mathbf{E}$, is continuous across the discontinuities. Similar equations can be defined on the SFD grid for the \hat{y} and \hat{z} components of equation (2.5).

The set of SFD equations for \mathbf{H}_x , \mathbf{H}_y and \mathbf{H}_z can be grouped together, and expressed as the linear system of equation,

$$\tilde{\mathbf{A}}\mathbf{x} = \begin{bmatrix} \tilde{\mathbf{A}}_{nn} & \tilde{\mathbf{A}}_{nb} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{x}_n \\ \mathbf{x}_b \end{bmatrix} = \mathbf{b} = \begin{bmatrix} \mathbf{0} \\ \mathbf{x}_b \end{bmatrix} \quad (2.7)$$

or, in another way,

$$\tilde{\mathbf{A}}_{nn}\mathbf{x}_n = -\tilde{\mathbf{A}}_{nb}\mathbf{x}_b, \quad (2.8)$$

where \mathbf{x}_b is the magnetic fields (all three components) on the boundaries of the model domain, \mathbf{x}_n is the unknown magnetic fields inside the model (all three components), and \mathbf{I} is the identity matrix. The matrix $\tilde{\mathbf{A}}_{nb}$ is the coefficient matrix which couples the known boundary fields, \mathbf{x}_b , with the unknown fields on adjacent nodes. The matrix $\tilde{\mathbf{A}}_{nn}$ is the coefficient matrix of the unknown interior fields, \mathbf{x}_n , coupling among themselves inside the model, and can be written as

$$\tilde{\mathbf{A}}_{nn} = \begin{bmatrix} \tilde{\mathbf{A}}_{xx} & \tilde{\mathbf{A}}_{xy} & \tilde{\mathbf{A}}_{xz} \\ \tilde{\mathbf{A}}_{yx} & \tilde{\mathbf{A}}_{yy} & \tilde{\mathbf{A}}_{yz} \\ \tilde{\mathbf{A}}_{zx} & \tilde{\mathbf{A}}_{zy} & \tilde{\mathbf{A}}_{zz} \end{bmatrix}, \quad (2.9)$$

where $\tilde{\mathbf{A}}_{xx}$ gives the coefficients of the \mathbf{H}_x coupling with \mathbf{H}_x , $\tilde{\mathbf{A}}_{xy}$ is the coefficients of the \mathbf{H}_x coupling with \mathbf{H}_y , and similarly for other terms.

Matrix $\tilde{\mathbf{A}}_{nn}$ is not symmetric as written, if the grid spacing is not uniform. To make $\tilde{\mathbf{A}}_{nn}$ symmetric, we can scale the equations with appropriate volume elements (Smith, 1996a). For example, for $H_x(i, j, k)$, equation (2.6) is multiplied by $D_x(i)\Delta_y(j)\Delta_z(k)$ to become

$$\begin{aligned}
& \left[i\omega\mu D_x(i)\Delta_y(j)\Delta_z(k) \right. \\
& \quad - \frac{\bar{\rho}_z(i, j, k)D_x(i)\Delta_z(k)}{D_y(j)} - \frac{\bar{\rho}_z(i, j-1, k)D_x(i)\Delta_z(k)}{D_y(j-1)} \\
& \quad \left. - \frac{\bar{\rho}_y(i, j, k)D_x(i)\Delta_y(j)}{D_z(k)} - \frac{\bar{\rho}_y(i, j, k-1)D_x(i)\Delta_y(j)}{D_z(k-1)} \right] H_x(i, j, k) \\
& + \frac{\bar{\rho}_z(i, j, k)D_x(i)\Delta_y(j)}{D_y(j)} H_x(i, j+1, k) + \frac{\bar{\rho}_z(i, j-1, k)D_x(i)\Delta_y(j)}{D_y(j-1)} H_x(i, j-1, k) \\
& + \frac{\bar{\rho}_y(i, j, k)D_x(i)\Delta_z(k)}{D_z(k)} H_x(i, j, k+1) + \frac{\bar{\rho}_y(i, j, k-1)D_x(i)\Delta_z(k)}{D_z(k-1)} H_x(i, j, k-1) \\
& \quad + \bar{\rho}_z(i, j, k)\Delta_z(k) \left[H_y(i+1, j, k) - H_y(i, j, k) \right] \\
& + \bar{\rho}_z(i, j-1, k)\Delta_z(k) \left[H_y(i+1, j-1, k) - H_y(i, j-1, k) \right] \\
& \quad + \bar{\rho}_y(i, j, k)\Delta_y(j) \left[H_z(i+1, j, k) - H_z(i, j, k) \right] \\
& + \bar{\rho}_y(i, j, k-1)\Delta_y(j) \left[H_z(i+1, j, k-1) - H_z(i, j, k-1) \right] = 0. \tag{2.10}
\end{aligned}$$

With this scaling, we can see the coupling coefficients between fields are symmetric, e.g., the coefficients between $H_x(i, j, k)$ to $H_x(i, j-1, k)$ and $H_x(i, j-1, k)$ to $H_x(i, j, k)$ are the same. After scaling, equation (2.8) becomes

$$\mathbf{A}_{nn}\mathbf{x}_n = -\mathbf{A}_{nb}\mathbf{x}_b. \tag{2.11}$$

Matrix \mathbf{A}_{nn} is now symmetric but complex (only on the diagonal). Because of the huge contrast of resistivity between the air (very resistive) and the Earth, the matrix \mathbf{A}_{nn} is stiff and ill-conditioned. Matrix \mathbf{A}_{nn} is also sparse, therefore, we store only

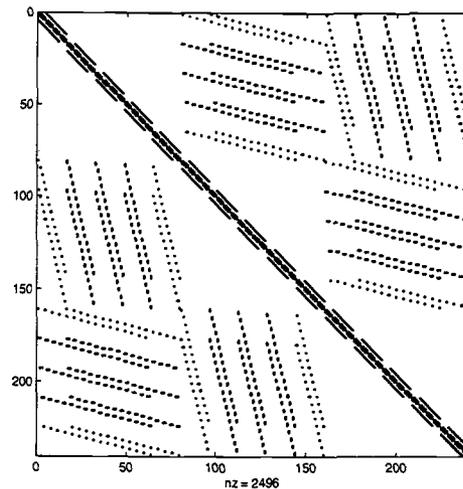


FIGURE 2.2. Sparse structure of matrix \mathbf{A}_{nn} for a $5 \times 5 \times 5$ model, where the dot indicate the non-zero elements.

the non-zero elements. A schematic example of the sparse structure of matrix \mathbf{A}_{nn} is shown in Figure 2.2 for a $5 \times 5 \times 5$ model. A 5-stripped sparse structure is shown for each of the submatrices \mathbf{A}_{xx} , \mathbf{A}_{yy} and \mathbf{A}_{zz} .

2.3. Boundary Conditions

To solve (2.11), boundary fields \mathbf{x}_b are required. These could be computed by solving a larger-scale 3-D problem, or by using the fields calculated from 2-D models of the vertical boundary planes. Mackie et al. (1994) divided the problem into two polarizations depending on the direction of the uniform external sources.

Considering the case that the current source flows in the x -direction (\mathbf{H}_y mode; \mathbf{J}_x polarization), if the strike is in x -direction, then TE mode is used on the boundary yz -planes to compute for \mathbf{H}_y and \mathbf{H}_z . If the strike is in the y -direction, then TM mode is used on the extended xz -planes of every y slices to solve for \mathbf{H}_y . Then, only \mathbf{H}_y located on the boundary yz -planes are used as the boundary fields.

If there is no extension of the xz -planes, \mathbf{H}_y at the boundary yz -planes are identical to the 1-D boundary fields, which should be avoided. Similarly, for the case that current source flows in the y -direction (\mathbf{H}_x mode; \mathbf{J}_y polarization), if the strike is in y -direction, then TE mode is used on the xz -planes to compute for \mathbf{H}_x and \mathbf{H}_z . If the strike is in the x -direction, then TM mode is used on the extended yz -planes of every x slices to solve for \mathbf{H}_x . Then, only \mathbf{H}_x located on the boundary xz -planes are used as the boundary fields. Here, we have used an approach similar to that of Mackie et al. (1994) for our boundary conditions.

Smith (1996a) used a more complicated method by mixing boundary fields computed from the 2-D calculations. This allowed their boundary fields to vary from their 2-D values. Mackie et al. (1994) has stated that the approximation made in the boundary conditions will not have much effect on solution for the fields inside. The main difficulty is primary from the ill-conditioning of the system of equation.

2.4. Iterative Relaxation Methods

There are several ways of solving equation (2.11). A direct method would be to invert \mathbf{A}_{nn} by decomposing it into upper and lower triangular matrices, then using back and forward substitution to get the solution. This requires a lot of computational time and storage for decomposing \mathbf{A}_{nn} . Thus, it is not computationally practical to use a direct solution approach for the 3-D MT modeling.

Iterative methods have been developed to solve the linear system of equation of the form, $\mathbf{Ax} = \mathbf{b}$. The advantages of using iterative methods over the direct method are that they are faster and do not require much memory. Examples of iterative methods are conjugate gradients (CG), minimum residual (MINRES) and bi-conjugate gradients (BiCG). Most of these algorithms are designed for real sys-

tems of equations. However, it is still possible that these algorithms would work reasonably well in our 3-D MT forward problem, where the coefficient matrix is complex (but only in the diagonal term) and symmetric. This is because the diagonal complex term is considerably smaller than the real term (Mackie and Madden, 1993).

The CG method is an extremely effective method especially when the coefficient matrix is real, symmetric and positive definite (Barret et al., 1994). It requires less storage and fewer operations in each iteration than other methods. The BiCG method (Jacobs, 1986), used by Smith (1996b), is particularly well suited to the general system (not necessarily symmetric). BiCG generates two CG-like sequences of vectors, one for \mathbf{A} and another for \mathbf{A}^H where H denotes the conjugate transpose. Therefore, its storage and computational requirement are double that of the CG method. However, in our case, the symmetric property of equation (2.11) helps reduce computation and storage of BiCG to the same as those for the CG method but for complex matrices, with all conjugate transposes H replaced with simple transpose T (Smith, 1996b).

The MINRES method (Fletcher, 1975), used by Mackie et al. (1994), is another alternative to CG useful when the coefficient matrix is symmetric but possibly indefinite. MINRES requires more operations and storages in each iteration than the CG does. However, MINRES which minimizes $\mathbf{r}^T \mathbf{r}$ where $\mathbf{r} = \frac{\|\mathbf{b} - \mathbf{A}\mathbf{x}\|}{\|\mathbf{b}\|}$ is the residual, usually does not show the erratic convergence behavior that can be seen with CG or BiCG.

The success of the relaxation methods depends greatly upon the conditioning of the system. A large condition number corresponds to a wide spread of eigenvalues of the system (Mackie and Madden, 1993). Because of the high resistivity of the air layers, the coefficient matrix \mathbf{A}_{nn} is stiff and ill-conditioned resulting in a

slow convergence rate of all of the relaxation methods. Dramatic improvements in convergence can be achieved by preconditioning the system of equation, to reduce the spread of eigenvalues and lower the condition number.

Preconditioning the system of equation is to solve the transformed system of equation $\mathbf{M}^{-1}\mathbf{Ax} = \mathbf{M}^{-1}\mathbf{b}$, where \mathbf{M} is a preconditioner. The transformed system of equation has the same solution as the original system $\mathbf{Ax} = \mathbf{b}$, but the condition of the transformed system is better. Normally, \mathbf{M} is the approximates of \mathbf{A} in some ways.

The simplest preconditioner is the diagonal term of the coefficient matrix. A better preconditioner can be obtained by using an incomplete decomposition (e.g., Smith, 1996b) or using the diagonal subblocks of \mathbf{A}_{nn} (e.g., Mackie et al., 1994). Both preconditioners require backward and forward substitution which work fairly well in a sequential computation. However, in our case where we want to solve the system of equation in the parallel machine, the backward and forward substitution become inefficient.

2.5. Divergence Correction

Mackie and Madden (1993) found that relaxation methods alone cannot yield an accurate result especially at long periods even though a small residual is obtained. Theoretically, this is because, as $\omega \rightarrow 0$, the ratio $\mu\omega/\rho$ becomes very small (Mackie et al., 1994), the right hand side of (2.4) which contains information about the conductivity structure (Smith, 1996b) is disappeared. This is particularly true in the air layers. The result is a very large nonzero divergence of the magnetic fields. This can be interpreted that the magnetic monopoles were added to the solution

where resistivity contrasts are large, such as at the interface of the air and the Earth's surface.

Taking the divergence of (2.2) yields the requirement that $\nabla \cdot \mathbf{H} = \mathbf{0}$, or in other words, magnetic monopoles do not exist (no sources and no sinks). Therefore, to obtain an accurate result, the condition $\nabla \cdot \mathbf{H} = \mathbf{0}$ must be enforced explicitly.

Smith (1996b) introduced a simple divergence correction scheme (to the electric current $\sigma\mathbf{E}$ in his case) which is extremely effective and requires small amount of computational time. He applied it alternately with a series of BiCG relaxation iterations, and found that it not only helps in improving the accuracy of the solution, but also speeds up the convergence rate of the iterative solver. Mackie et al. (1994) applied a similar divergence correction scheme (to the magnetic field), again alternated with the MINRES relaxation iterations.

The divergence correction works as follows. At a given iteration, the divergence of the magnetic fields is calculated as $\psi = \nabla \cdot \mathbf{H}$, and a potential ϕ is then found by solving Poisson's equation

$$\nabla \cdot \nabla\phi = \psi, \quad (2.12)$$

with Neumann boundary conditions. Once ϕ is obtained, the magnetic fields required to correct the magnetic divergence are computed as,

$$\mathbf{H}_{\text{corrected}} = \mathbf{H}_{\text{old}} - \nabla\phi. \quad (2.13)$$

Note that if \mathbf{H}_{old} satisfies (2.4), $\mathbf{H}_{\text{corrected}}$ does also since $\nabla \times (\mathbf{H} - \nabla\phi) = \nabla \times \mathbf{H}$.

Poisson's equation (2.12) can be written as a linear system of equation, similar to (2.11), and can be solved the same ways via any iterative methods. However, in this case, the coefficient matrix is real, symmetric and positive definite. Most importantly, it is also not so ill-conditioned as \mathbf{A}_{nn} (since there are no variable resistivity values used to form the coefficient matrix).

Furthermore, an accuracy of the solution to the Poisson's equation (2.12) is not that important (Mackie et al., 1994), as long as the divergence of the magnetic fields is reduced. This implies that the relaxation method for solving (2.12) can be terminated early, leading to a small computational load for each divergence correction step (Mackie et al., 1994; Smith, 1996b).

2.6. Parallel Implementations

With the iterative approach and the divergence correction, problems of modest size can be solved on a typical workstation. To allow for a larger size of model so that geological realistic problems can be solved, we have adapted this scheme to run on a massively parallel computer. The parallel machine that we are using is the CM-5 and CM-500e developed by Thinking Machine Corporation. The CM machines available in the COAS have processor nodes from 32 nodes (with 4 Gigabytes of RAM) to 128 nodes (with 16 Gigabytes of RAM).

2.6.1. Partitioning Vectors and Matrices

In a parallel system, all of the vectors are divided among processors such that they are load balanced. For example, the vector storing \mathbf{H}_x of a $3 \times 3 \times 3$ model (i.e., with size of $3 \times 4 \times 4$) is partitioned among the twelve physical processors as shown in Figure 2.3. Similarly, the other vectors are also distributed for each processor in the same way, which makes any mathematical operations among vectors efficient, compared to the performance of a single processor.

Because the coefficient matrix \mathbf{A}_{nn} is sparse (see Figure 2.2), there is no need to store the non-zero elements. We therefore decompose \mathbf{A}_{nn} into many vectors.

P0 Hx(1,1,1) Hx(1,1,2) Hx(1,1,3) Hx(1,1,4)	P1 Hx(1,2,1) Hx(1,2,2) Hx(1,2,3) Hx(1,2,4)	P2 Hx(1,3,1) Hx(1,3,2) Hx(1,3,3) Hx(1,3,4)	P3 Hx(1,4,1) Hx(1,4,2) Hx(1,4,3) Hx(1,4,4)
P4 Hx(2,1,1) Hx(2,1,2) Hx(2,1,3) Hx(2,1,4)	P5 Hx(2,2,1) Hx(2,2,2) Hx(2,2,3) Hx(2,2,4)	P6 Hx(2,3,1) Hx(2,3,2) Hx(2,3,3) Hx(2,3,4)	P7 Hx(2,4,1) Hx(2,4,2) Hx(2,4,3) Hx(2,4,4)
P8 Hx(3,1,1) Hx(3,1,2) Hx(3,1,3) Hx(3,1,4)	P9 Hx(3,2,1) Hx(3,2,2) Hx(3,2,3) Hx(3,2,4)	P10 Hx(3,3,1) Hx(3,3,2) Hx(3,3,3) Hx(3,3,4)	P11 Hx(3,4,1) Hx(3,4,2) Hx(3,4,3) Hx(3,4,4)

FIGURE 2.3. Mapping the magnetic field \mathbf{H}_x vectors to the processors for the $3 \times 3 \times 3$ model. Each processor contains four elements of \mathbf{H}_x .

Each vector represents the coupling coefficient among fields (see equation (2.6)).

Again, these vectors will be mapped into each processor.

2.6.2. Parallel Computations

Solving by any iterative schemes require these simple operations: scalar-vector multiplication, vector addition, vector inner products, and matrix-vector multiplication. These operations can be computed quickly using a parallel machine, especially when all vectors are distributed uniformly among all processor nodes.

1. Scalar-vector multiplication and vector addition. Suppose using p processors, then n/p multiplications and n/p additions are required for scalar-vector multiplication and vector addition respectively for parallel system comparing to n multiplications and n additions respectively required for a single processor, where n is the size of the vector.

2. Vector inner products. Corresponding elements of each vectors are multiplied locally (n/p multiplications), summed up ($n/p - 1$ additions) and the result kept locally in each processor. Then all the partial sums in each processor are added up, using data communication among nodes. The time for data communication depends upon the architecture of the parallel machine. On a machine like the CM-5, communication time for such operations can be ignored (Gupta et al., 1995) in comparison with the communication time incurred in other parts of the algorithm such as matrix-vector multiplication.

3. Matrix-vector multiplication. This operation requires a significant communication time among processor nodes. For example, to multiply \mathbf{A}_{xx} with \mathbf{H}_x at node (2,2,3) located on P5 processor (Figure 2.3), the values of \mathbf{H}_x from node (2,2,2), (2,2,4), (2,1,3) and (2,3,3) are required from (2.10). \mathbf{H}_x at node (2,1,3) and (2,3,3) are located at different processors, P4 and P6 respectively. Therefore, P4 need to communicate with P5 and P6 to obtain \mathbf{H}_x at those locations. However, even with the addition of the communication time, the total execution time for matrix-vector multiplication in parallel system is still a lot less than that in the sequential system.

2.6.3. Iterative Methods and Parallel Preconditioners

Preconditioners used to speed up the convergence rate of the relaxation methods are not so easily implemented in parallel. For example, backward and forward substitution used by an incomplete decomposition require sequential computations. Using the diagonal part of the coefficient matrix as a preconditioner requires only vector multiplication, and therefore works efficiently in parallel. Although a signifi-

cant improvement over no preconditioner, However, the convergence rate using this simple preconditioner is still slow.

The diagaonal scaling can be applied once before the relaxation started. The system of equation $\mathbf{Ax} = \mathbf{b}$ is modified in this manner

$$[\mathbf{D}^{-\frac{1}{2}}\mathbf{A}\mathbf{D}^{-\frac{1}{2}}]\mathbf{D}^{\frac{1}{2}}\mathbf{x} = \mathbf{D}^{-\frac{1}{2}}\mathbf{b} \quad (2.14)$$

where \mathbf{D} is the diagonal matrix consisting of the diagonal elements of \mathbf{A} . By this scaling the coefficient matrix $\mathbf{D}^{-\frac{1}{2}}\mathbf{A}\mathbf{D}^{-\frac{1}{2}}$ has unity on its main diagonal, but still remains symmetric and complex. Because the system of equation is scaled, therefore after the relaxation method converging to a target level, the solution $(\mathbf{D}^{\frac{1}{2}}\mathbf{x})$ must be scaled back to the true values of \mathbf{x} .

The Symmetric Successive Over Relaxation (SSOR) preconditioner introduced by Evans (1968) can be implemented to work in the parallel environment. After the diagonal scaling, the symmetric coefficient matrix $\mathbf{D}^{-\frac{1}{2}}\mathbf{A}\mathbf{D}^{-\frac{1}{2}}$ can be decomposed as

$$\mathbf{D}^{-\frac{1}{2}}\mathbf{A}\mathbf{D}^{-\frac{1}{2}} = \mathbf{I} + \mathbf{L} + \mathbf{L}^T, \quad (2.15)$$

where \mathbf{L} is the lower-triangular matrix consisting of the lower diagonals of $\mathbf{D}^{-\frac{1}{2}}\mathbf{A}\mathbf{D}^{-\frac{1}{2}}$, and \mathbf{I} is the identity matrix.

The SSOR preconditioner \mathbf{M} (Seager, 1986) is then defined by

$$\mathbf{M} = (\mathbf{I} + \mathbf{L})(\mathbf{I} + \mathbf{L}^T). \quad (2.16)$$

One advantage of this preconditioner is that no additional storage is required. However, an SSOR-preconditioner still requires backward and forward substitutions making it not appropriate for a parallel machine architecture. However, the inverse of \mathbf{M} can be approximated to first order as $(\mathbf{I} + \mathbf{L})^{-1} \approx \mathbf{I} - \mathbf{L}$, or to second order as $(\mathbf{I} + \mathbf{L})^{-1} \approx \mathbf{I} - \mathbf{L} + \mathbf{L}^2$. Thus, \mathbf{M}^{-1} can be written as,

$$\mathbf{M}^{-1} \approx (\mathbf{I} - \mathbf{L})(\mathbf{I} - \mathbf{L}^T), \quad (2.17)$$

as the first-order inverse approximation of SSOR (ISSOR-1), or

$$\mathbf{M}^{-1} \approx (\mathbf{I} - \mathbf{L} + \mathbf{L}^2)(\mathbf{I} - \mathbf{L}^T + (\mathbf{L}^2)^T), \quad (2.18)$$

as the second-order approximation (ISSOR-2). With these approximations, only matrix-vector multiplications and diagonal scalings are involved in the preconditioner.

Here, relaxation methods are used to solve several systems of equations; Poisson's equation (2.12), (2.11) and systems of equations similar to (2.11) but for 1-D and 2-D problem. Due to the lower cpu time required per iteration, we prefer to use the BiCG method (modified for symmetric system) to solve these system of equations, along with preconditioner.

2.7. Numerical Experiments and Discussions

To test our iterative scheme, we have applied our parallel 3-D forward modeling program to the conductivity model of Figure 2.4. The model consists of two adjacent rectangular blocks (20 km in width, 40 km in length and 10 km in depth each), one block is conductive (1 Ω -m; on the west) and the other is resistive (100 Ω -m; on the east). These two blocks are imbedded in host layer of 10 Ω -m and 10 km depth which is underlain by a 100 Ω -m with 20 km of thickness and a half-space of resistivity of 0.1 Ω -m.

This model was also used by Mackie et al. (1994), so that our results can be compared to this previous work. Here, we discretize the model on two different grids: a coarse discretization of $28 \times 21 \times 17$ (including 7 airs) (same as of Mackie et al., 1994), and a finer discretization of $60 \times 60 \times 49$ (including 10 airs). The air layers

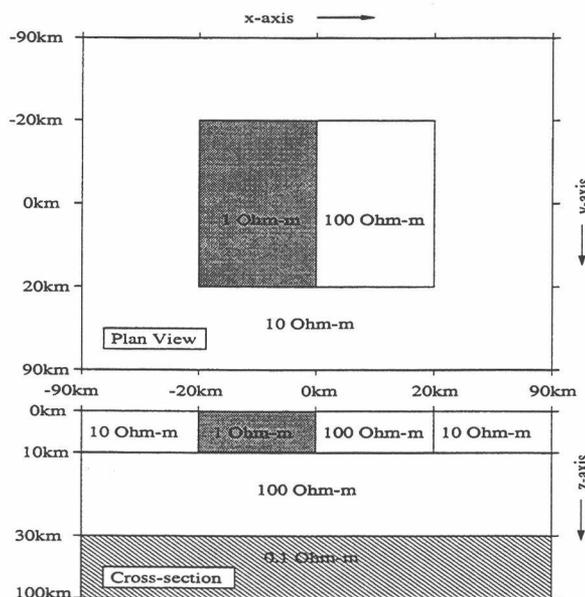


FIGURE 2.4. The test 3-D model. The model is discretized into the original discretization, $28 \times 21 \times 17$ (including 7 airs), and a finer discretization, $60 \times 60 \times 49$ (including 10 airs).

at the top of the model have a resistivity of $10^{10} \Omega\text{-m}$ to provide a large resistivity contrast with the Earth.

As discussed earlier, we solve for two polarizations: the \mathbf{H}_x mode, where the external source current \mathbf{J}_y flows in the y -direction, and the \mathbf{H}_y mode, where the source current \mathbf{J}_x flows in the x -direction. The boundary values are computed using a 2-D TM on the extended yz - and xz -planes at the boundaries for \mathbf{H}_x and \mathbf{H}_y mode respectively. The starting fields for the relaxation method are also calculated this way. Because we solve the 2-D problems on a parallel machine, all 2-D vertical slices in the xz -plane with extension (or yz -planes) are simultaneously solved for \mathbf{H}_x , and then projected onto the 3-D grid. Two periods (at 100 second and 1 second) are used to demonstrate that the program works for a range of periods.

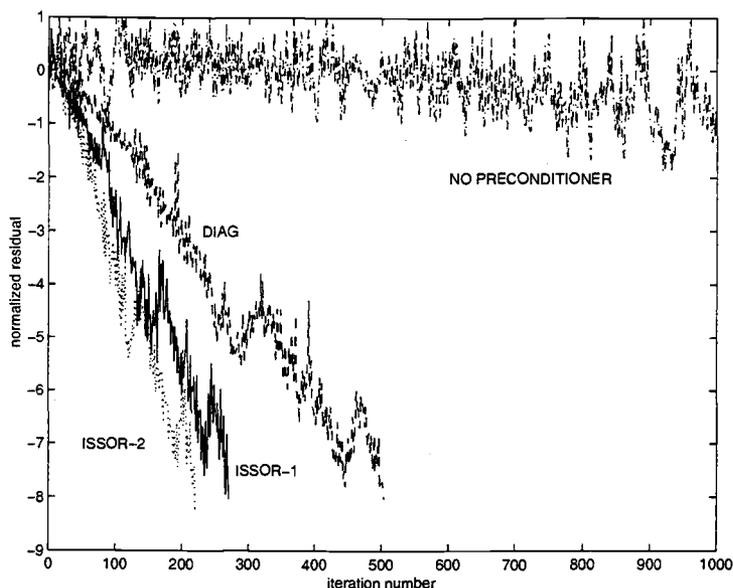


FIGURE 2.5. A plot of \log_{10} normalized residual versus relaxation iteration from solving the 3-D Poisson's equation by using different parallel preconditioners. DIAG is the diagonal preconditioner. ISSOR-1 is the first order inverse approximation of the SSOR. ISSOR-2 is the second order inverse approximation of the SSOR.

2.7.1. Different Preconditioners

The main relaxation method is intervened after numbers of iterations to correct the divergence of the magnetic fields, by solving the 3-D Poisson's equation. Different preconditioners are tested to compare the efficiency of the preconditioners in term of convergence rate. The normalized residual defined as $\|\mathbf{b} - \mathbf{Ax}\|/\|\mathbf{b}\|$ for the linear system $\mathbf{Ax} = \mathbf{b}$, versus relaxation iteration is plotted in Figure 2.5 for the different preconditioner types. Here, we set the desired normalized residual at 10^{-8} . As discussed earlier that the solution of the Poisson's equation is not importantly accurate, as long as it can reduce the divergence of the magnetic fields. Therefore, the desired normalized residual can be a lot lower than this value.

Without the preconditioner, the normalized residual is gradually decreased approximately a decade in about 1000 iterations. The convergence rate is significantly improved with the preconditioner as simple as the diagonal preconditioner, and more improvement achieved with the ISSOR-1 and ISSOR-2. Increasing the order of approximation in SSOR case will increase the amount of computation and communication time in each iteration, this increasing time is tradeoff by the reduction of the number of total iteration number. However, in this case, the reduction of the number of iteration of ISSOR-2 from ISSOR-1 is not so great, resulting in no gain of using ISSOR-2.

Therefore, we apply ISSOR-1 as the inverse of preconditioner to solve the Poisson's equation and the 2-D linear system of equation. Both of these systems of equation are very similar. The coefficient matrices of these systems of equation are sparse in diagonal format (5-diagonals for 2-D forward problem, and 7-diagonals for 3-D Poisson's problem) similar to the matrix \mathbf{A}_{xx} or other diagonal subblocks matrices of the 3-D forward problem. This diagonal format makes it efficient to use SSOR (2.16) as preconditioner, and also its approximate inverse (2.17).

However, in 3-D forward problem where the non-diagonal blocks (such as \mathbf{A}_{xy}) are not in diagonal format (see Figure 2.2), approximate inverse of SSOR becomes less efficient, comparing to other cases (2-D forward problem or 3-D Poisson's equation), or in the other word, the error from truncating the series expansion is larger. The convergence rate of ISSOR will not be as obvious as in those cases (Figure 2.5) resulting in less gain by using ISSOR. In addition, in 3-D case, the relaxation stops periodically to correct the divergence of the magnetic fields, which significantly helps to improve the solution in each correction (Smith 1996b; Mackie 1994). Therefore, to avoid extra computations by ISSOR, we apply only diagonal scaling to the 3-D case.

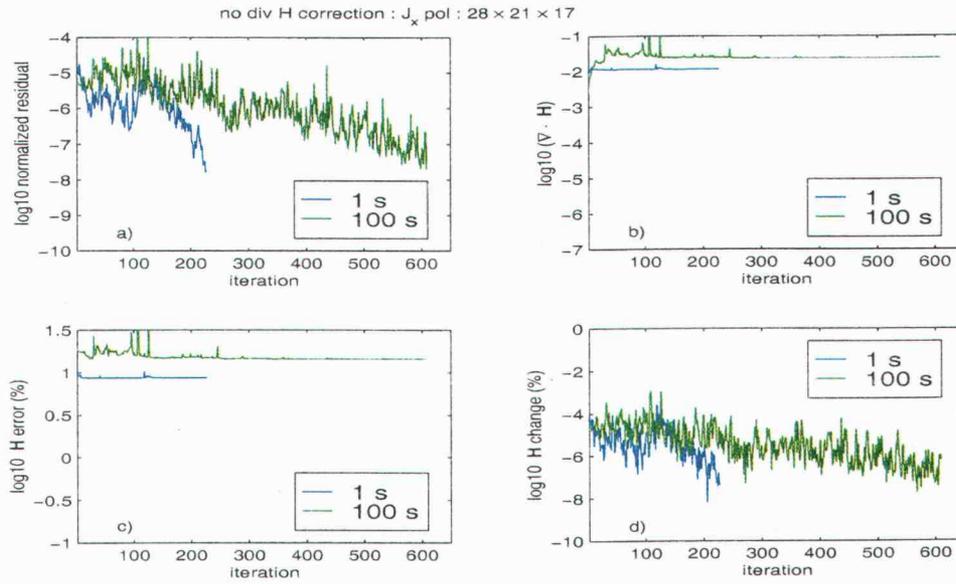


FIGURE 2.6. (a) The \log_{10} normalized residual (\mathbf{r}) vs. iteration for the model with the coarse grid, at period of 1 (blue) and 100 seconds (green) of \mathbf{J}_x polarization, and without divergence correction; (b) \log_{10} of the divergence of the magnetic fields ($\nabla \cdot \mathbf{H}$); (c) \log_{10} of the normalized residual of the surface magnetic fields after each iteration (\mathbf{H}_{sit}) and the converged surface magnetic fields (\mathbf{H}_c); and (d) \log_{10} of the relative change of the magnetic fields after each iteration (\mathbf{H}_{it}) comparing with the magnetic fields of the previous iteration ($\mathbf{H}_{\text{it}-1}$).

2.7.2. Without Divergence Correction

Next, we ran the programs on the model of Figure 2.4 with the coarse grid and without the divergence correction. After each iteration, we compute the normalized residual, defined as $\mathbf{r} = \frac{\|\mathbf{b} - \mathbf{A}\mathbf{x}\|}{\|\mathbf{b}\|}$ for the system of equation $\mathbf{A}\mathbf{x} = \mathbf{b}$. We also compute the divergence of the magnetic fields ($\nabla \cdot \mathbf{H}$), and the relative change of the magnetic fields from the previous iteration, defined as $\mathbf{H}_{\text{change}} = \frac{\|\mathbf{H}_{\text{it}} - \mathbf{H}_{\text{it}-1}\|}{\|\mathbf{H}_{\text{it}-1}\|} * 100$ (%). The magnetic fields on the surface after each iteration \mathbf{H}_{sit} are also compared directly to the fully converged magnetic fields \mathbf{H}_c at the surface computed from Mackie's code, $\mathbf{H}_{\text{error}} = \frac{\|\mathbf{H}_{\text{sit}} - \mathbf{H}_c\|}{\|\mathbf{H}_c\|} * 100$ (%).

Figure 2.6 show plots of r , $\nabla \cdot \mathbf{H}$, $\mathbf{H}_{\text{error}}$ and $\mathbf{H}_{\text{change}}$ for periods of 1 and 100 seconds for the \mathbf{J}_x polarization. Result for the \mathbf{J}_y polarization are similar. We terminated the run, once the normalized residuals are below 10^{-8} (about 230 iterations for 1 second and 600 iterations for 100 second; Figure 2.6a). Generally, long period equation is stiffer and more ill-conditioned, and thus requires more iterations than the short period equation.

In a 2-D MT forward problem where there is no divergence of the magnetic fields, this level of normalized residual indicates that we have obtained a solution which is of accuracy comparable to results from direct method. However, in the 3-D case the divergence of the magnetic fields builds up after the first iteration and remains constantly high (Figure 2.6b). The magnetic fields are changed only beneath the surface, and hardly change in the air resulting in relatively small normalized change from iterations to iterations (Figure 2.6d). However, the solutions are still far from the fully converged solutions (Figure 2.6c).

Therefore, we cannot use the normalized residual alone to judge when to terminate the iterative scheme. The divergence of the magnetic fields must also be taken into account. As expected, the magnetic divergence from the long period (100 seconds) is higher than that of the short period (1 second), (Figure 2.6b)), because the ratio of the $\mu\omega/\rho$ is smaller for long periods. Therefore, applying the divergence correction will benefit the long period more than the short period solution (Smith, 1996b).

2.7.3. With Divergence Correction

To correct for the divergence of the magnetic fields and obtain an accurate solutions, the divergence correction must be included in the iterative scheme. From

experiments on the coarse grid, we conclude that there are no exact rules of when to apply the divergence correction.

The strategy of Mackie et al. (1994) is to make divergence corrections every d relaxation steps, with d initially four, and then increased by two after each update. However, they make sure that $\mathbf{r}^h \mathbf{r}$ is decreased before applying the divergence correction.

Similar to Mackie et al. (1994), we apply divergence correction every d relaxation steps. However, because Mackie et al. (1994) used a more sophisticated preconditioner the normalized residual is generally reduced more in each iteration than we achieve with our simple diagonal preconditioner. We thus use larger values of d than Mackie et al. (1994).

Here, we set the goal to obtain a solution with normalized residual $\mathbf{r} < 10^{-8}$ and $\nabla \cdot \mathbf{H} < 10^{-5}$. By experimenting with several values of d , we found that applying the divergence correction too frequently (small d) sometimes does more harm than good. We suggest about between every 50th iteration and 100th iteration in our scheme. However, the optimal value of d needs further study. The relaxation method alone seeks the solution fields within the Earth (cannot penetrate the stiffness of the resistive air; small residual but high magnetic divergence). The divergence correction help clearing out the magnetic monopoles existed at the boundary of the high contrast of resistivity (such as beneath surface) and breaking through the stiffness of the air layers. If the divergence correction is applied frequently, the relaxation does not have a chance to build up solutions beneath the surface, and the divergence correction will not improve the solution, resulting in a high number of iteration to obtain a reasonable solution.

Figure 2.7 and 2.8 shows plots similar to Figure 2.6, except that in this case we applied the divergence correction every 50 iterations for the 100 second case and

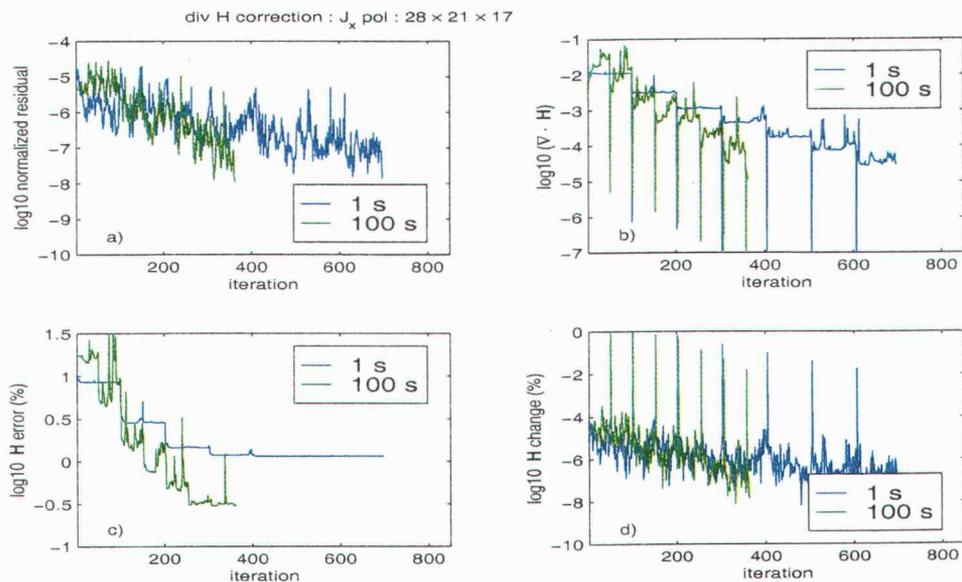


FIGURE 2.7. Similar plot to Figure 2.6 but divergence correction is applied every 50 iterations for 100 second and every 100 iterations for 1 second.

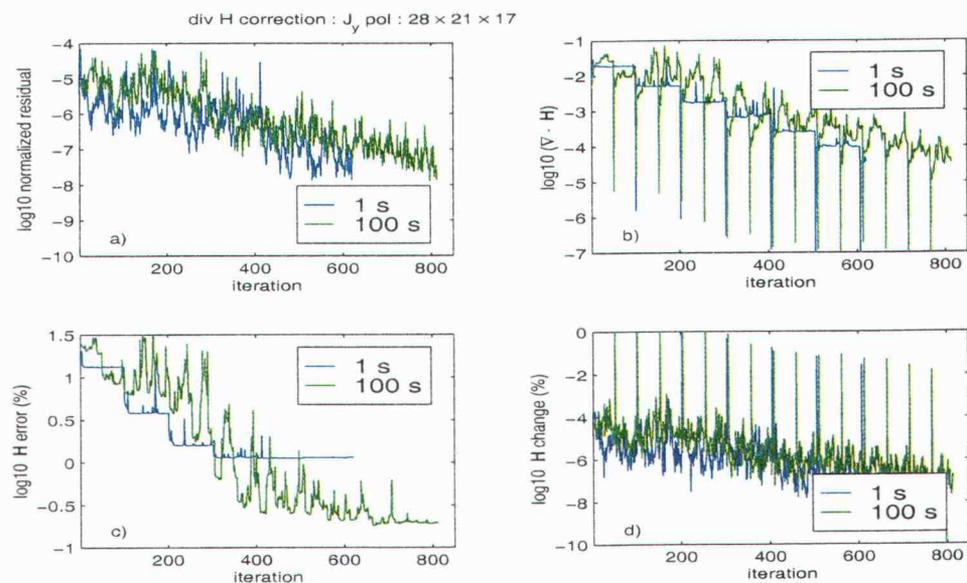


FIGURE 2.8. Similar plot to Figure 2.7 but for J_y polarization.

every 100 iterations for the 1 second case. In both cases, at the final iterations, we obtained solutions that have normalized residual and the divergence of the magnetic fields less than the desired values.

With the divergence correction, the number of iterations for the 100 second case drops dramatically from about 600 to below 400 for the \mathbf{J}_x polarization (Figure 2.7a), but not for the \mathbf{J}_y polarization (Figure 2.8a), where the number of iteration is increased to around 800 iterations. Also, the numbers of iterations are significantly increased for both \mathbf{J}_x and \mathbf{J}_y of the 1 second case, (Figure 2.7a and 2.8a).

The magnetic divergence is sharply reduced by the divergence correction, but then significantly increases again to a level that is a little bit lower than before divergence correction (Figure 2.7b and 2.8b) after restarting the relaxation scheme. Also, after each correction, the magnetic fields are dramatically changed from relaxation step before the correction as indicated with the spike in Figure 2.7d and 2.8d. This change is mostly in the air and at the surface. At the early stages, the change of the magnetic fields after the divergence correction is often more than 1 %. It gradually becomes smaller at later iterations. Note that during the relaxation steps, the magnetic fields are hardly changed from iteration to iteration.

After each divergence correction, the magnetic fields at the surface becomes closer to the converged magnetic fields (Figure 2.7c and 2.8c). For the 100 second case, the magnetic fields are within 1 % of the converged values, after about 300 iterations for the \mathbf{J}_y polarization, and after about 200 iterations for the \mathbf{J}_x polarization. Similarly, for the 1 second case, after 300 iterations of both polarizations the magnetic fields are within 2 % of the fully converged magnetic fields, and remain constant after that. Including roundoff and discretization error, errors less

than about 2 % of the magnetic fields is probably accurate enough even though the normalized residual is not at the desired level.

Here, we propose another condition to terminate the program, while keeping the desired normalized residual at the same level as before. We terminate when the change of the magnetic fields after the divergence correction drops below a specified minimum level. We still obtain a reasonable solution with this strategy. Figure 2.7d and 2.8d show that after about 300 iterations for both polarizations and both periods H_{change} is less than 0.1 % after the divergence correction. Even though the normalized residuals are still above the desired level, the solutions are already within 2 % accuracy, as long as the divergences of the magnetic fields are small.

The divergence correction helps the fields to diffuse into the highly resistive air, whereas the relaxation method alone cannot break the resistive air layer. This is why at the early stage after the divergence correction, the magnetic fields are significantly changed. Once the solutions get close to the real solution, the change of the magnetic fields gets smaller and smaller.

Next we consider solution of the same model, but with a more finely discretized grid ($60 \times 60 \times 49$) including 10 air layers. As the size of the grid gets bigger, it is become more difficult to obtain solutions with the desired normalized residual. Again, we choose to stop once the magnetic fields change very little after the divergence correction. This way the number of iterations required is about 1000 iterations for the 100 second case and about 500 iterations for the 1 second case. At this point the normalized residual is reduced to roughly 10^{-6} to 10^{-7} (Figure 2.9a). However, the accuracies of the magnetic fields compared to fully converged solutions from Mackie et al. (1994) are about 2 % for 1 second and less than 1 % for 100 second (Figure 2.9c).

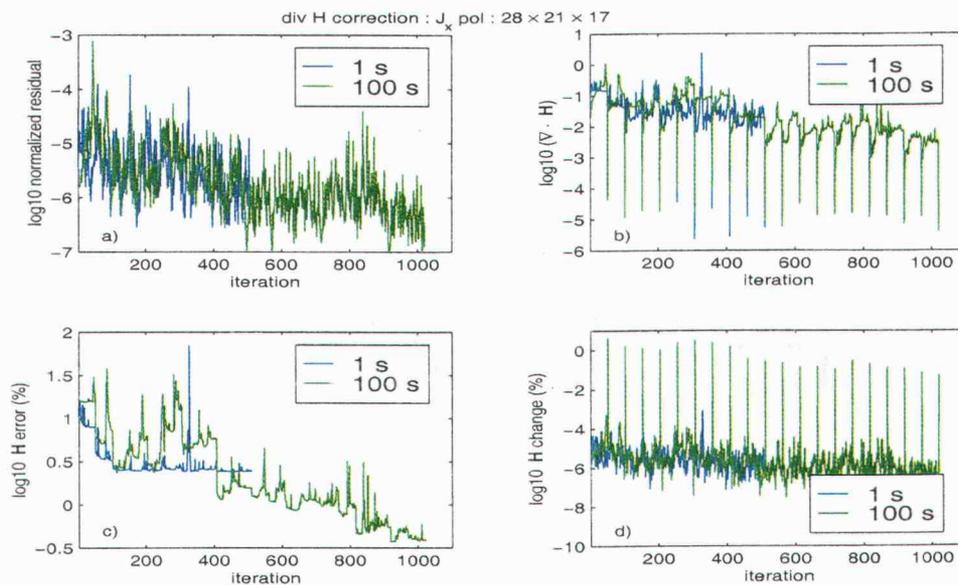


FIGURE 2.9. Similar plot to Figure 2.7 but for finer discretization of $60 \times 60 \times 49$ model.

The surface magnetic fields interpolated to the center of the rectangular grid for 100 second and 1 second are shown in Figure 2.10 and Figure 2.11 respectively.

2.8. Conclusion

An iterative 3-D MT modeling program with divergence correction has been implemented and developed on a massively parallel computer so that large problems can be solved. An accurate result can be achieved within a few hundred of iterations. The magnetic fields from the second order Maxwell's equation are sampling by using the staggered finite difference (SFD) grid which are more accurate than the standard finite difference grid, and then forming the linear system of equation which can then be solved via the iterative relaxation methods. Here, the precondition conjugate gradient (PCG) is used to solve the system of equation, along with the divergence

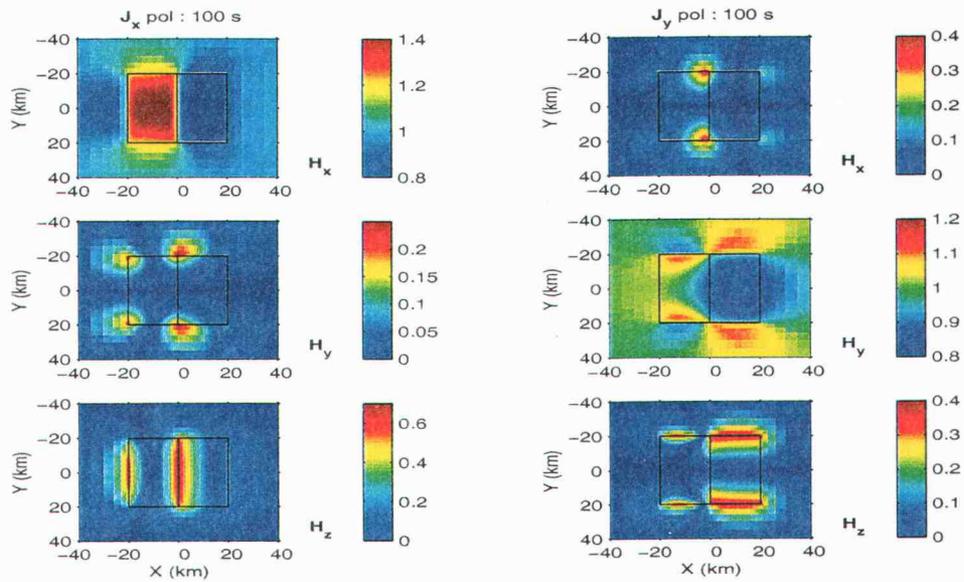


FIGURE 2.10. Magnitude of the surface magnetic fields at 100 second, relative to a source field value of 1.0.

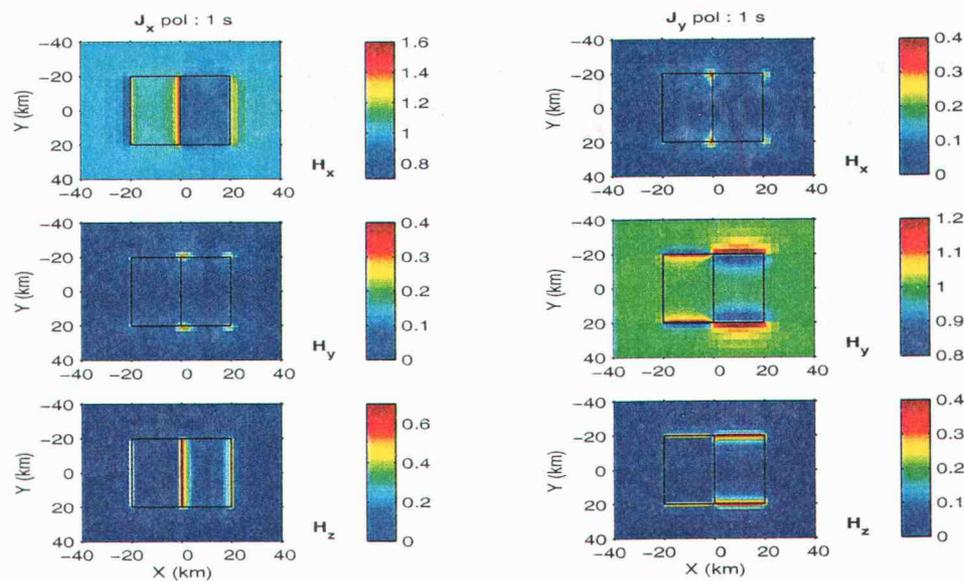


FIGURE 2.11. Magnitude of the surface magnetic fields at 1 second, relative to a source field value of 1.0.

correction which solve the Poisson's equation. With the divergence correction, the convergence rate is greatly improved.

The most challenge in development the 3-D MT forward modeling to run on a parallel machine is the designing of the efficient preconditioner. Different preconditioners are then developed from the most simple case, the diagonal preconditioner, to the more complicated cases, the inverse approximation of the Symmetric Successive Over Relaxation (ISSOR) both first and second orders. We found that the ISSOR-1 is greatly improving the convergent rate when solving the Poisson's equation in the divergence correction process. Since we intervene the main relaxation process for the divergence correction, both ISSOR-1 and diagonal preconditioners do not make much different in term of convergent rate. ISSOR-1 preconditioner requires higher computational costs than diagonal preconditioner, and therefore will not be used.

We also found one criteria that we can use to terminate the run is the change of the magnetic fields after the divergence correction. At the early iterations, the change of magnetic fields after the divergence correction is very significant. This change becomes minimal at later iteration when the calculated fields get closer to the field solutions. The iteration can then be stopped even though the normalized residual is still higher than the desired residual. This help reducing the computational time.

2.9. References

- [1] Barret, R., M. Berry, T.F. Chan, J. Demmel, J.M. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine and H. van der Vorst, 1994, Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, *SIAM*, Philadelphia.
- [2] Fletcher, R., 1975, Conjugate gradient methods for indefinite systems in Numerical Analysis by A. Dold and B. Eckmann, Springer, New York.

- [3] Gupta, A., V. Kumar, and A. Sameh, 1995, Performance and Scalability of Preconditioned Conjugate Gradient Methods on Parallel Computers, *IEEE Trans. Parallel and Distrib. System.*, **6**, 455–469.
- [4] Jacobs, D. A. H., 1986, A generalization of the conjugate-gradient method to solve complex systems, *IMA J. Numer. Anal.*, **6**, 447-452.
- [5] Mackie, R.L., T.R. Madden and P.E. Wannamake, 1993, Three-dimensional magnetotelluric modeling using difference equations: Theory and comparisons to integral equation solutions, *Geophysics*, **58**, 215–226.
- [6] Mackie, R.L and T.R. Madden, 1993, Conjugate direction relaxation solutions for 3-D magnetotelluric modeling, *Geophysics*, **58**, 1052-1057.
- [7] Mackie, R.L., J.T. Smith and T.R. Madden, 1994, Three-dimensional electromagnetic modeling using finite difference equations: The magnetotelluric example, *Radio Science*, **29**, 923–935.
- [8] Smith, J.T., 1996, Conservative modeling of 3-D electromagnetic fields, Part I: Properties and error analysis, *Geophysics*, **61**, 1308–1318.
- [9] Smith, J.T., 1996, Conservative modeling of 3-D electromagnetic fields, Part II: Biconjugate gradient solution and an accelerator, *Geophysics*, **61**, 1319–1324.

3. REBOCC: AN EFFICIENT DATA-SUBSPACE INVERSION FOR TWO-DIMENSIONAL MAGNETOTELLURIC DATA

3.1. Introduction

The Magnetotelluric (MT) method for imaging crustal and upper mantle electrical conductivity has found increasing use in both geophysical exploration application (e.g. Orange, 1989; Vozoff, 1972) and in fundamental studies of large scale tectonics (e.g. Jones, 1992; Wannamaker, 1994; Chen et al., 1996; Unsworth et al., 1999). Initial applications of MT were based on local 1-D interpretations, for which theories (e.g. Weidelt, 1972; Parker, 1980) and inversion methods (e.g. Jupp and Vozoff, 1975; Constable et al., 1987; and Smith and Booker, 1988) are well developed. It is now clear that two or even three dimensional interpretation is essential for most real MT data sets. Over the past decade very substantial progress has been made on the development of 2-D inversion methods. These have included straightforward extensions of linearized search methods developed previously for 1-D regularized inversion (deGroot-Hedlin and Constable, 1990; Uchida, 1993) efficient approximate methods (Smith and Booker, 1991; Farquharson and Oldenburg, 1996), the subspace method (Oldenburg et al., 1993) and methods based on direct iterative minimization of a regularized penalty functional (Rodi and Mackie, 1999). These program are freely available, and are widely used for interpretation of modern MT surveys. However the available inversion algorithms are not without flaws. The fastest (RRI) intrinsically limit the model space search and can fail to converge (without substantial user intervention); the most general and flexible (e.g. OCCAM) run very slowly and require substantial memory.

We begin our discussion with a review of 2-D MT inversion methods including a data space variant on OCCAM. We then introduce our inversion algorithm which is

stable, flexible and general, and much faster than most previously proposed schemes. We demonstrate the stability and efficiency of our approach by inverting synthetic MT data, comparing our results to those obtained with other methods, including RRI (Smith and Booker, 1991) and NLCG (Rodi and Mackie, 1999). Finally we briefly consider an example application to field data from a dense profile across the San Andreas fault near Parkfield, CA (Unsworth et al., 1997).

3.2. Overview of Inversion Methods

We begin with a broad overview of previously developed 2-D MT inversion methods. To be explicit we consider the Earth as discretized into a series of M constant resistivity blocks, $\mathbf{m} = [m_1, m_2, \dots, m_M]$. There are N observed data $\mathbf{d} = [d_1, d_2, \dots, d_N]$, with estimated uncertainties $\mathbf{e} = [e_1, e_2, \dots, e_N]$. The fit of the theoretical model responses $\mathbf{F}[\mathbf{m}]$ to the observational data can be expressed as

$$X_d^2 = (\mathbf{d} - \mathbf{F}[\mathbf{m}])^T \mathbf{C}_d^{-1} (\mathbf{d} - \mathbf{F}[\mathbf{m}]), \quad (3.1)$$

where the superscript T represents matrix transpose and \mathbf{C}_d is the data covariance matrix, which in practice is diagonal.

Because of the non-uniqueness of the inverse problem, an infinite number of models can produce the same misfit. Most modern MT inversion schemes resolve this non-uniqueness by seeking models that have minimum possible structure (in some sense) for a given level of misfit (Parker, 1994). This makes the inversion stable, with resulting models less likely to contain spurious features.

To quantify “model structure”, we consider a model norm of the general form

$$X_m^2 = (\mathbf{m} - \mathbf{m}_0)^T \mathbf{C}_m^{-1} (\mathbf{m} - \mathbf{m}_0), \quad (3.2)$$

where \mathbf{m}_0 is a base (or prior) model, and \mathbf{C}_m is a model covariance matrix which characterizes the expected magnitude and smoothness of resistivity variations relative to \mathbf{m}_0 . Other approaches to minimum structure inversion are similar, though in some cases \mathbf{C}_m^{-1} is replaced by a model roughness operator. The minimum structure inverse problem is to minimize X_m^2 subject to $X_d^2 = X_*^2$, where X_*^2 is the desired level of misfit.

To solve this minimization problem, a Lagrange multiplier λ^{-1} can be introduced, resulting in an unconstrained functional $U(\mathbf{m}, \lambda)$

$$U(\mathbf{m}, \lambda) = (\mathbf{m} - \mathbf{m}_0)^T \mathbf{C}_m^{-1} (\mathbf{m} - \mathbf{m}_0) + \lambda^{-1} \{ (\mathbf{d} - \mathbf{F}[\mathbf{m}])^T \mathbf{C}_d^{-1} (\mathbf{d} - \mathbf{F}[\mathbf{m}]) - X_*^2 \}, \quad (3.3)$$

for which we seek stationary points (with respect to both \mathbf{m} and λ). Alternatively, we may consider the penalty functional $W_\lambda(\mathbf{m})$

$$W_\lambda(\mathbf{m}) = (\mathbf{m} - \mathbf{m}_0)^T \mathbf{C}_m^{-1} (\mathbf{m} - \mathbf{m}_0) + \lambda^{-1} \{ (\mathbf{d} - \mathbf{F}[\mathbf{m}])^T \mathbf{C}_d^{-1} (\mathbf{d} - \mathbf{F}[\mathbf{m}]) \}. \quad (3.4)$$

In (3.4), λ acts to “trade-off” between minimizing the norm of data misfit and the norm of the model (Parker, 1994). When λ is large, the data misfit is de-emphasized, leading to a smoother model. In contrast, as $\lambda \rightarrow 0$, the inverse problem becomes closer to the ill conditioned least-square inversion problem, resulting in an erratic model (see Parker, 1980).

Note that both U and W have the same stationary points with respect to variation of the model, i.e. $\partial U / \partial \mathbf{m} = \partial W / \partial \mathbf{m}$, where λ is fixed. Parker (1994) uses this to show that stationary points of (3.3) can be found by minimizing (3.4) for a series of λ values, and then choosing λ so that the misfit satisfies the constraint $X_d^2 = X_*^2$.

For linear $\mathbf{F}[\mathbf{m}]$ this is straightforward, since in this case (for fixed λ) $\partial U / \partial \mathbf{m} = 0$ is a linear system of equations which may be solved for \mathbf{m} . Because

$\mathbf{F}[\mathbf{m}]$ is non-linear, iterative solution methods are required. We briefly consider some of the approaches which have been taken by previous workers, and then outline our approach.

Rodi and Mackie (1999) provide a good review of several approaches, including a straightforward Gauss-Newton (GN) method. This approach is based on linearizing $\mathbf{F}[\mathbf{m}]$ with a Taylor series expansion,

$$\mathbf{F}[\mathbf{m}_{k+1}] = \mathbf{F}[\mathbf{m}_k + \delta\mathbf{m}] = \mathbf{F}[\mathbf{m}_k] + \mathbf{J}_k(\mathbf{m}_{k+1} - \mathbf{m}_k), \quad (3.5)$$

where k denotes iteration number, and $\mathbf{J}_k = \left. \frac{\partial \mathbf{F}}{\partial \mathbf{m}} \right|_{\mathbf{m}_k}$ is the $N \times M$ sensitivity matrix calculated at \mathbf{m}_k . Calculation of \mathbf{J}_k , which describes the perturbations in the data due to changes in the model, is described in detail by Rodi and Mackie (1999), Mackie and Madden (1993) and Rodi (1978). Substituting (3.5) in (3.4), we obtain

$$\begin{aligned} \widetilde{W} = & (\mathbf{m}_{k+1} - \mathbf{m}_0)^T \mathbf{C}_m^{-1} (\mathbf{m}_{k+1} - \mathbf{m}_0) + \\ & \lambda^{-1} \{ (\hat{\mathbf{d}}_k - \mathbf{J}_k(\mathbf{m}_{k+1} - \mathbf{m}_0))^T \mathbf{C}_d^{-1} (\hat{\mathbf{d}}_k - \mathbf{J}_k(\mathbf{m}_{k+1} - \mathbf{m}_0)) \}, \end{aligned} \quad (3.6)$$

where $\hat{\mathbf{d}}_k = \mathbf{d} - \mathbf{F}[\mathbf{m}_k] + \mathbf{J}_k(\mathbf{m}_k - \mathbf{m}_0)$. This \widetilde{W} is then quadratic in \mathbf{m}_{k+1} and thus can be minimized exactly (for fixed λ). For numerical stability (Marquardt, 1963), an extra damping parameter ϵ is generally required to control step size for each iteration. The system of equation to be solved for each iteration then becomes

$$\mathbf{m}_{k+1} - \mathbf{m}_k = [\lambda \mathbf{C}_m^{-1} + \mathbf{\Gamma}_k^m + \epsilon_k \mathbf{I}]^{-1} [\mathbf{J}_k^T \mathbf{C}_d^{-1} (\mathbf{d}_k - \mathbf{F}[\mathbf{m}_k]) - \lambda \mathbf{C}_m^{-1} (\mathbf{m}_k - \mathbf{m}_0)], \quad (3.7)$$

where the ‘‘model space cross-product’’ matrix $\mathbf{\Gamma}_k^m = \mathbf{J}_k^T \mathbf{C}_d^{-1} \mathbf{J}_k$ is an $M \times M$ positive semi-definite symmetric matrix, and \mathbf{I} is the identity matrix. Note that with the GN approach λ is fixed. Therefore, the algorithm will converge to a stationary point of (3.4), not (3.3). To achieve the stationary point of (3.3) (with respect to both λ and \mathbf{m}), the process would have to be repeated with different values of λ until the constraint $X_d^2 = X_*^2$ was satisfied.

The OCCAM approach, first proposed by Constable et al. (1987) (see also Parker, 1994; deGroot-Hedlin and Constable, 1993; Uchida, 1993), is also based on linearizing $\mathbf{F}[\mathbf{m}]$ and then solving for the stationary points of (3.6). Differentiating (3.6) with respect to \mathbf{m} and setting the result to zero leads to an iterative sequence of approximate solutions

$$\mathbf{m}_{k+1}(\lambda) = [\lambda \mathbf{C}_m^{-1} + \mathbf{\Gamma}_k^m]^{-1} \mathbf{J}_k^T \mathbf{C}_d^{-1} \hat{\mathbf{d}}_k + \mathbf{m}_0. \quad (3.8)$$

The unique feature of the OCCAM approach is that the parameter λ is used in each iteration both as a step length control and a smoothing parameter. That is, (3.8) is solved for a series of trial values of λ and the misfit $X_d^2(\mathbf{m}_{k+1}(\lambda))$ for each λ is evaluated by solving the 2-D forward problem. As for the linear problem, λ should be chosen so that the condition $X_d^2 = X_*^2$ is met. Usually, in the early iterations, the true misfit X_d^2 is higher than the desired X_*^2 for all possible λ . The OCCAM process thus chooses the model with the minimum misfit as the basis for the next iteration. The process is then repeated until the misfit reaches the desired level. Parker (1994) called this process of bringing the misfit down to the target level Phase I. Once the misfit reaches the desired level, Phase II begins by keeping the misfit at the desired level, but varying λ to search for the model with smallest norm. Since the problem is non-linear, the desired misfit may never be reached. However, in practice, improvement of the misfit from iteration to iteration can be expected, until a minimum is achieved.

Both GN and OCCAM share similar computational steps. For each iteration, \mathbf{J}_k must be calculated, and an $M \times M$ system of equations, (3.7) for GN and (3.8) for OCCAM, must be solved. These methods are thus very time-consuming (e.g. Smith and Booker, 1991; Rodi and Mackie, 1999). Furthermore, these methods require much memory to store the sensitivity and cross-product matrices. These

computational inefficiencies are the result of strongly dependence on the model space dimension M .

Several approaches have been proposed to avoid the heavy computational burden of the direct linearized search schemes. One approach is to use approximate sensitivities to eliminate calculation of the full sensitivity matrix. A good example in this category is the Rapid Relaxation Inversion (RRI) introduced by Smith and Booker (1991). RRI turns the 2-D inverse problem into a series of 1-D inverse problems, by computing the approximate sensitivity of data at each site to variations of resistivity directly below the site. The model is updated by solving a series of 1-D inverse problems and interpolating horizontally to form the 2-D resistivity model. Fit to the data is tested with a full 2-D forward calculation, and step length is adjusted if necessary. The process is repeated until the misfit condition is met. Note that this approach eliminates both the 2-D sensitivity calculation, and the need to solve a large $M \times M$ system of equations. Generally, RRI requires many iterations but overall is very fast. RRI can handle very large 2-D data sets, and has been applied to interpretation of many MT data sets. Oldenburg and Ellis (1991) suggested a very similar approach based on using a series of 1-D inversions as an approximate inverse mapping (AIM) to map the data back to model space, followed by full calculation of the 2-D forward problem to assess model fit.

The efficiency of approximate inversion schemes comes at a price. Because of the incomplete search of the model space, schemes based on 1-D inversions can be insensitive to features that are not directly beneath the locations of measurement (Farquharson and Oldenburg, 1996). For example, if the data is dominated by a significant feature outside of the profile, the inversion may have difficulty finding any models which actually fit the data (e.g. Unsworth et al., 1999), or the inversion may insert a geologically unreasonable feature beneath the profile. In addition, inversion

of vertical magnetic transfer functions is difficult with this approach, since vertical magnetic fields are sensitive to nearby structures, rather than features directly beneath the sites.

Farquharson and Oldenburg (1996) proposed a somewhat different approximate approach based on using 2-D sensitivities of a homogeneous or layered half-space, instead of the exact sensitivities. This scheme eliminates the need for calculation of \mathbf{J}_k but still requires inversion of large ($M \times M$) matrices. However, in many cases these simple approximations to the sensitivities are good enough to allow convergence of the inversion to an acceptable level of misfit. The effectiveness of this scheme depends on many factors: the complexity of structure of the true model, the closeness of the structure to the data sites and the magnitude of the resistivity contrasts.

Another way to reduce the computational burden in 2-D MT inversion is provided by the subspace approach of Oldenburg et al. (1993): the inverse solution is sought in a low dimensional subspace of the original M -dimensional model space. The success of this approach depends greatly on the choice of subspace basis vectors. Unfortunately, the proper choice is often not obvious, and a bad selection can lead to a poor solution. A combination of both approximate sensitivities and the subspace approach has been used by Oldenburg and Ellis (1993). Current implementation of OCCAM also allows for a simple sort of subspace approach, since the resistivity can be parameterized on a grid coarser than that used for numerical computations.

Rather than approximate the sensitivities, or impose prior restrictions on the model space, Mackie and Madden (1993) considered an approach based on a conjugate gradient (CG) relaxation solution of the linearized normal equations derived from (3.6). With the relaxation method, the actual computation of the sensitivity matrix can be avoided by using the fact that evaluating the gradient of the linearized

penalty functional requires only one forward solution (per period) with a distributed set of sources either in the volume or on the surface. CG thus significantly reduces the computational requirements (both cpu time and memory) for each iteration, making attempts at even 3-D MT inversion feasible (Mackie and Madden, 1993). The model is updated for each iteration and the CG relaxation solution process is repeated until the stationary point of (3.4) is reached. Note however that since CG is a descent method which does not directly use any information about curvature of the penalty functional, many iterations may be required compared to GN or OCCAM. Also, as for the GN approach the entire process must be repeated for different values of λ to find a true minimum structure model which achieves a specified misfit X_*^2 .

Rodi and Mackie (1999) considered a variant on this CG approach. These authors applied non-linear conjugate gradients (NLCG) directly to minimization of (3.4), with λ fixed. Similar to the CG method (Mackie and Madden, 1993), NLCG requires only a few forward solutions (per period) in each line minimization step. To improve the convergence rate, a simple preconditioner is used. Again, to obtain a norm minimizing solution with minimum structure at the desired misfit, one needs to minimize (3.4) for various values of λ . Rodi and Mackie (1999) show that the two descent methods (CG and NLCG) are comparable, and both are much more efficient than the GN method in terms of cpu time and memory requirements.

Here, we develop a variant on the OCCAM approach which is significantly more efficient than previously proposed methods. We begin by transforming the inverse problem from the model space into the data space, by expressing the solution as a linear combination of rows of the sensitivity matrix smoothed by the model covariance. This transformation reduces the size of the system of equations to be solved from $M \times M$ to $N \times N$. Since the number of model parameters M is often

much larger than the number of data N , a significant decrease in both cpu time and memory can be achieved with this approach. More importantly, the data space formulation leads naturally to a simple approximation which can result in very significant computational savings in most cases.

Generally, MT data are smooth (in period, and for closely spaced sites, in space) and “redundant”. Therefore, in the data space approach, there is no need to use all of the sensitivities as basis functions. A subset is typically sufficient to construct the model without significantly loss of detail. With this approximation it is unnecessary to compute all sensitivities, and the size of the system of equations that must be solved can be significantly reduced. We call this approach the **REduced Basis OCCam’s (REBOCC) Inversion**. Note that even though we construct the solution from subset of the smoothed sensitivities, the goal of the inversion remains to find the norm minimizing model subject to fitting all of the data well enough. As we shall discuss in more detail below, in the data space the choice of basis functions is very natural and is dictated by what features can be resolved by the available data. This is in contrast to the choice of a model subspace (Oldenburg et al., 1993), where the choice of subspace is rather arbitrary.

With careful implementation of forward modeling and sensitivity calculations, REBOCC runs in a fraction of the time required by methods such as GN or OCCAM and is also considerably faster than NLGG. In addition, memory requirements are significantly reduced so that large data sets can be inverted with REBOCC on a standard workstation. The basic idea behind REBOCC generalizes readily to the 3-D case.

3.3. The Data Space Occam Method (DASOCC)

Parker (1994) shows that the minimizer of (3.6) for iteration k can be expressed as a linear combination of rows of the smoothed sensitivity matrix $\mathbf{C}_m \mathbf{J}_k^T$,

$$\mathbf{m}_{k+1} - \mathbf{m}_0 = \mathbf{C}_m \mathbf{J}_k^T \boldsymbol{\beta}_{k+1} = \sum_{j=1}^N \Upsilon_{kj} \beta_{k+1,j} \quad (3.9)$$

where $\boldsymbol{\beta}_{k+1}$ is an unknown expansion coefficient vector of the basis functions Υ_{kj} ; $j = 1, \dots, N$. The basis vectors Υ_{kj} are sometimes referred to as the “representers” of the linearized data functionals for iteration k (e.g., Parker, 1994). Substituting (3.9) into (3.6), we obtain

$$\widetilde{W} = \boldsymbol{\beta}_{k+1}^T \Gamma_k^n \boldsymbol{\beta}_{k+1} + \lambda^{-1} \{ (\hat{\mathbf{d}}_k - \Gamma_k^n \boldsymbol{\beta}_{k+1})^T \mathbf{C}_d^{-1} (\hat{\mathbf{d}}_k - \Gamma_k^n \boldsymbol{\beta}_{k+1}) \}. \quad (3.10)$$

Here $\Gamma_k^n = \mathbf{J}_k \mathbf{C}_m \mathbf{J}_k^T$ is the $N \times N$ “data space cross-product” matrix, which is again symmetric and positive semi-definite. Differentiating (3.10) with respect to $\boldsymbol{\beta}$ and rearranging, the unknown expansion coefficients can be obtained as

$$\boldsymbol{\beta}_{k+1} = (\lambda \mathbf{C}_d + \Gamma_k^n)^{-1} \hat{\mathbf{d}}_k. \quad (3.11)$$

The inverse problem thus becomes a search for the N real expansion coefficients $\boldsymbol{\beta}_{k+1}$, instead of the M dimensional model, \mathbf{m}_{k+1} . Exactly as for the standard Occam, we can solve for $\boldsymbol{\beta}_{k+1}$, update the model, and then check the misfit for various values of λ . We again choose λ to achieve the minimum misfit if this exceeds the desired level X_*^2 (Phase I), and use this model as the basis for the next iteration. Once the desired misfit is achieved, Phase II begins to wipe out unnecessary features, while keeping the misfit at the desired level.

We emphasize here that we have only transformed the inverse problem solution method from the model space to the data space. Solutions obtained in both spaces will be identical if we choose all parameters (i.e., λ and \mathbf{C}_m) the same. For

brevity we refer to this variant on Occam as the **Data Space OCCam (DASOCC)** inversion. Note that a data space approach was also used by Smith and Booker (1988) in their treatment of the 1-D MT inverse problem.

The data space formulation offers several advantages. The most obvious is the reduction in the dimension of the system of equations which must be solved ($N \times N$ in the data space instead of $M \times M$ in the model space). Generally, $N \ll M$. This will be particularly true for the 3-D case.

Also, calculation in the model space requires \mathbf{C}_m^{-1} . Since it is not practical to specify a full $M \times M$ model covariance matrix \mathbf{C}_m and then compute the inverse, \mathbf{C}_m^{-1} is replaced by the first derivative roughness penalty in deGroot-Hedlin and Constable (1990). In the data space approach \mathbf{C}_m is required, not its inverse. This offers some advantages since the model covariance can be readily used to include prior information such as an ocean, or faults which should be fixed in the model. We discuss these issues in more detail in the model covariance section.

3.4. Reduced Data Space Occam Approaches (REBOCC)

The data space formulation clearly shows that the solution is a linear combination of natural basis functions or representers. Each representer corresponds to a single data element (at a particular period, station, response and mode). Just as the MT data should be smooth and redundant, the representers vary slowly with period and site location for a given response and mode.

Figure 3.1 shows a plot of the representers of the apparent resistivity (TM mode) for three nearby periods. The plot reveals a smooth and, in fact, insignificant change from one period to another. These basis functions are thus highly redundant, so that an excellent approximation to the solution can be found in a subspace of

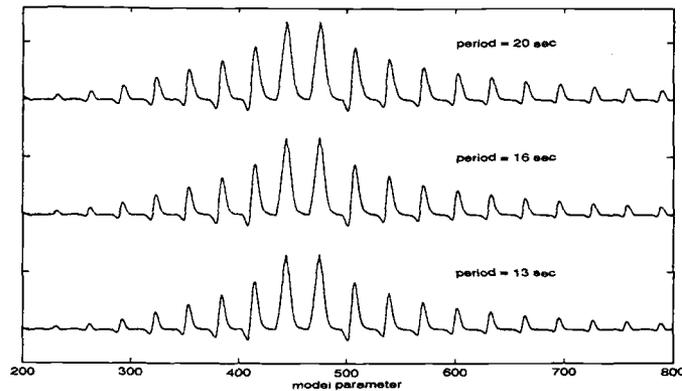


FIGURE 3.1. Plots of representers $\mathbf{C}_m \mathbf{J}^T$ for TM mode apparent resistivity at one station for nearby periods, for the synthetic data discussed later. The representers change very slowly as a function of period.

much lower dimension (Parker and Shure, 1982; see also Parker, 1994). This simple but critical concept of data redundancy can significantly speed up the inversion while also substantially decreasing memory requirements.

Prior to solving the inverse problem, we will select a subset of L (out of N) data for which representers will be calculated at each iteration. As a simple example, we could choose all data for every other period (so $L = N/2$). Note that, L can typically be considerably smaller than this, as we shall show later. For iteration $k + 1$, we seek solutions of the form,

$$\mathbf{m}_{k+1} = \mathbf{C}_m \mathbf{G}_k^T \boldsymbol{\alpha}_{k+1} + \mathbf{m}_0, \quad (3.12)$$

where $\boldsymbol{\alpha}_{k+1}$ is the L -dimensional unknown coefficient vector for the reduced basis, and \mathbf{G}_k is the $L \times M$ subset sensitivity matrix.

To fit all of the data and to derive equations for $\boldsymbol{\alpha}_{k+1}$ analogous to (3.10) and (3.11), we require the linearized relationship between $\delta \mathbf{m}$ and $\hat{\mathbf{d}}_k$. In fact, we do not strictly have this relationship, unless we calculate all of the sensitivities. However, the data varies smoothly, and so a data value would be well approximated

by interpolation of “nearby” data (e.g., adjacent frequencies from the same site). In the same way, sensitivities vary smoothly with frequency and/or site location (Fig. 3.1) and can be interpolated from nearby sensitivities. We thus express the approximation to the full sensitivity matrix \mathbf{J}_k in terms of the subset sensitivity matrix \mathbf{G}_k using an interpolation matrix, \mathbf{B} of size $N \times L$, that is ,

$$\mathbf{J}_k \approx \mathbf{B}\mathbf{G}_k. \quad (3.13)$$

The interpolation matrix does not need to be very sophisticated. Recall that Farquharson and Oldenburg (1996) have had success in 2-D MT inversion using a very simple sensitivity matrix generated from either a homogeneous or layered half-spaced, while RRI uses only 1-D approximate sensitivities. By comparison, the approximate sensitivity matrix of (3.13) generated by even a crude interpolation of a sparse subset of representers will actually be quite close to the exact sensitivity matrix \mathbf{J}_k . Further details of the interpolation are given in the interpolation matrix section.

Substituting (3.12) and (3.13) into (3.6), we find

$$\widetilde{W} = \boldsymbol{\alpha}_{k+1}^T \boldsymbol{\Gamma}_k^1 \boldsymbol{\alpha}_{k+1} + \lambda^{-1} \{ (\hat{\mathbf{d}}_k - \mathbf{B}\boldsymbol{\Gamma}_k^1 \boldsymbol{\alpha}_{k+1})^T \mathbf{C}_d^{-1} (\hat{\mathbf{d}}_k - \mathbf{B}\boldsymbol{\Gamma}_k^1 \boldsymbol{\alpha}_{k+1}) \}, \quad (3.14)$$

where $\boldsymbol{\Gamma}_k^1 = \mathbf{G}_k \mathbf{C}_m \mathbf{G}_k^T$ is the $L \times L$ “data subspace cross-product” matrix. Following Egbert et al. (1994), we decompose $\mathbf{C}_d^{-\frac{1}{2}} \mathbf{B}$ into the $N \times N$ orthonormal matrix \mathbf{Q} where $\mathbf{Q}^T = \mathbf{Q}^{-1}$, and the $N \times L$ matrix \mathbf{R}

$$\mathbf{C}_d^{-\frac{1}{2}} \mathbf{B} = \mathbf{Q}\mathbf{R} = \bar{\mathbf{Q}}\bar{\mathbf{R}}, \quad (3.15)$$

$$\text{where} \quad \mathbf{Q} = [\bar{\mathbf{Q}} \mid \bar{\mathbf{Q}}_0] \quad \text{and} \quad \mathbf{R} = \begin{bmatrix} \bar{\mathbf{R}} \\ \mathbf{0} \end{bmatrix}.$$

Matrices $\bar{\mathbf{Q}}$ and $\bar{\mathbf{Q}}_0$ have dimensions $N \times L$ and $N \times N - L$ respectively. Matrix $\bar{\mathbf{R}}$ is the square $L \times L$ upper triangular matrix, and $\mathbf{0}$ is the $N - L \times L$ zero matrix, i.e., all elements are zeros.

Equation (3.14) becomes

$$\widetilde{\mathbf{W}} = \bar{\boldsymbol{\alpha}}_{k+1}^T \bar{\boldsymbol{\Gamma}}_k^1 \bar{\boldsymbol{\alpha}}_{k+1} + \lambda^{-1} \{ (\mathbf{C}_d^{-\frac{1}{2}} \hat{\mathbf{d}}_k - \mathbf{Q} \bar{\boldsymbol{\Gamma}}_k^1 \bar{\boldsymbol{\alpha}}_{k+1})^T (\mathbf{C}_d^{-\frac{1}{2}} \hat{\mathbf{d}}_k - \mathbf{Q} \bar{\boldsymbol{\Gamma}}_k^1 \bar{\boldsymbol{\alpha}}_{k+1}) \} \quad (3.16)$$

$$\text{where} \quad \bar{\boldsymbol{\alpha}}_{k+1} = (\bar{\mathbf{R}}^{-1})^T \boldsymbol{\alpha}_{k+1} \quad (3.17)$$

$$\text{and} \quad \bar{\boldsymbol{\Gamma}}_k^1 = \bar{\mathbf{R}} \boldsymbol{\Gamma}_k^1 \bar{\mathbf{R}}^T. \quad (3.18)$$

Inserting $\mathbf{Q}\mathbf{Q}^T = \mathbf{I}$ in between $\bar{\boldsymbol{\Gamma}}_k^1$ and $\bar{\boldsymbol{\alpha}}_{k+1}$ on the right side of (3.16) and rearranging, this can be rewritten as

$$\widetilde{\mathbf{W}} = \bar{\boldsymbol{\alpha}}_{k+1}^T \bar{\boldsymbol{\Gamma}}_k^1 \bar{\boldsymbol{\alpha}}_{k+1} + \lambda^{-1} \{ X_{min}^2 + (\bar{\mathbf{d}}_k - \bar{\boldsymbol{\Gamma}}_k^1 \bar{\boldsymbol{\alpha}}_{k+1})^T (\bar{\mathbf{d}}_k - \bar{\boldsymbol{\Gamma}}_k^1 \bar{\boldsymbol{\alpha}}_{k+1}) \}, \quad (3.19)$$

where $\bar{\mathbf{d}}_k = \bar{\mathbf{Q}}^T \mathbf{C}_d^{-\frac{1}{2}} \hat{\mathbf{d}}_k$, and $X_{min}^2 = \|\bar{\mathbf{Q}}_0 \mathbf{C}_d^{-\frac{1}{2}} \hat{\mathbf{d}}_k\|^2 = \|\mathbf{C}_d^{-\frac{1}{2}} \hat{\mathbf{d}}_k\|^2 - \|\bar{\mathbf{Q}}^T \mathbf{C}_d^{-\frac{1}{2}} \hat{\mathbf{d}}_k\|^2$ is the approximate minimum achievable total square misfit for the selected basis. If we use all representers (i.e. $\mathbf{B} = \mathbf{I}$), then $X_{min}^2 = 0$. This corresponds to the fact that for a linear problem we can fit the data exactly if we use all representers. This will not be true for the non-linear MT problem. Thus X_{min}^2 only provides a very rough estimate of the magnitude of data misfit that might be achieved with the chosen reduced basis. In general, X_{min}^2 is high in the early iterations, and decreases to a constant in the later iterations.

Differentiating (3.19) with respect to $\bar{\boldsymbol{\alpha}}$ and setting the result to zero, the unknown expansion coefficients can be obtained in a form similar to (3.11),

$$\bar{\boldsymbol{\alpha}}_{k+1} = (\lambda \mathbf{I} + \bar{\boldsymbol{\Gamma}}_k^1)^{-1} \bar{\mathbf{d}}_k. \quad (3.20)$$

Again, just as in the model and the data space methods, after solving (3.20), we update the model using (3.12) and (3.17), then solve the forward problem to evaluate

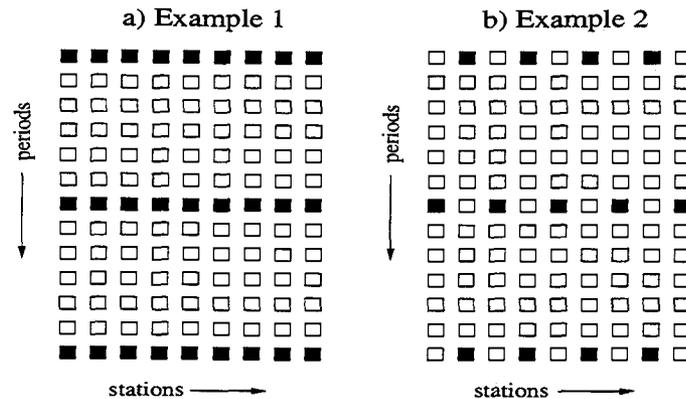


FIGURE 3.2. Examples of subset of representers. a) First example: “*pth*-stripe” pattern where every *pth* period is selected for all sites (here $p = 6$). Representers are calculated for data corresponding to the filled squares; the open circles represent the remaining data. b) Second example: “*pth:sth*-checker” (with $p = 4$ and $s = 2$) pattern where the selected data form a checker pattern.

X_d^2 . The procedure is repeated to find the appropriate λ . The outer loop of the iterative minimization of (3.4) proceeds exactly as for OCCAM or for DASOCC.

3.4.1. Representer Subsets for REBOCC

The success of the data subspace approach depends to some extent on the selection of data points which determine the representers used in the model expansion. Clearly it is necessary to select the basis to uniformly cover the full data set, so that the simple interpolation scheme used here is effective. Beyond this basic criteria, the optimal choice of data subsets is an issue that needs further study. Here we offer some simple schemes based on our experience so far.

Two classes of data subsets which generally seem to work well are shown in Figure 3.2. In the first example, every *pth* period is chosen for all sites (“*pth*-stripe” pattern). This pattern is safe to apply in almost all cases because for physically consistency the data must be a smooth function of period (e.g. Weidelt, 1972).

Selecting at least two periods per decade is our recommendation for this stripe pattern. Test runs with additional periods should be made (if feasible, given available computer resources) to verify that the data space has been adequately resolved.

This pattern offers a significant computation advantage: calculations for the sensitivity matrix need only be done for a reduced set of periods. For the 2-D problem, these sensitivities can be calculated by direct factorization of the coefficient matrix (see forward modeling section). With the *pth*-stripe pattern this factorization need only be done for a small subset of periods. This idea is clearly directly transferable to the 3-D case.

In the second example every *pth* period and every *sth* station are selected in a staggered pattern to build the basis ("*pth:sth*-checker" pattern). This pattern is probably reasonable only in the case where the site spacing is small (e.g., continuous EM profiling). As noted above, because of the way we compute sensitivities, it is generally most efficient to use as many sites as possible (i.e. small *s*). However, this pattern can be used to reduce storage requirements for very large data sets.

3.4.2. Comparison of Computational Resource Requirements

In this section we summarize the computer resources required by each method (OCCAM, DASOCC and REBOCC). In all three methods, most of the major cost lies in calculating the sensitivity matrices, computing the cross-product matrices, and, then (for each λ) solving equations (3.8), (3.11) or (3.20) respectively, followed by forward modeling. Similarly, the memory required by each method is dominated by storage of two matrices: the sensitivity matrices and the cross-product matrices. For REBOCC, extra memory is required to store the $N \times L$ interpolation matrix \mathbf{B} , and an additional calculation is required to factor \mathbf{B} (but this only be done once at

Number of Operations (Multiplication and Additions)			
Methods	X-product Matrix*	Solve SEs**	Compute (3.18)*
OCCAM	NM^2	$M^3/6$	-
DASOCC	MN^2	$N^3/6$	-
REBOCC	ML^2	$L^3/6$	L^3
*once every iteration, **for every trial value of λ			
Storage Requirement (words)			
Methods	X-product Matrix	Sens. Matrix	Interp. Matrix
OCCAM	$M(M+1)/2$	NM	-
DASOCC	$N(N+1)/2$	NM	-
REBOCC	$L(L+1)/2$	LM	NL

TABLE 3.1. Top: Comparison of number of operations (multiplications and additions) needed to form cross-product (X-product) matrices, solve system of equations (SEs). For REBOCC additional operations are required to compute (3.18). Bottom: Comparison of real words required to store sensitivity (Sens.) matrices, cross-product matrices and the interpolation (Interp.) matrix (REBOCC only).

the start of the inversion), and to compute the inner products of (3.18) (for every iteration).

Table 3.1 summarizes the number of operations (multiplication and additions) needed, along with the minimum storage requirements for each method. As L approaches N , the extra calculations required with the REBOCC method become significant. However, our experience shows that L need only be 10 - 30 % of N to ensure convergence of REBOCC. Thus REBOCC can reduce memory requirements by at least 60 % and cpu time by more than 80 % compared to DASOCC.

3.5. REBOCC : Algorithm Details

The overall stability and efficiency of the REBOCC scheme depends on many details including forward modeling, the model covariance, the 1-D line search for the Lagrange multiplier λ , the interpolation matrix, and static shift corrections. In this section we describe our implementations of these parts of the inversion algorithms.

3.5.1. Forward Modeling

Forward modeling is the heart of the inversion and thus must be reliable, fast and accurate. It is used in two parts of the inversion: to compute the sensitivity matrix, and to compute responses for calculating the misfit.

Assuming a time dependence of $\exp(-i\omega t)$, the second order Maxwell's equations to be solved are

$$\nabla \times \nabla \times \mathbf{E} = i\omega\mu\sigma\mathbf{E} \quad (3.21)$$

for the transverse electric (TE) mode in which the electric currents flow parallel to the strike of the structure, and

$$\nabla \times \rho\nabla \times \mathbf{H} = i\omega\mu\mathbf{H} \quad (3.22)$$

for the transverse magnetic (TM) mode in which the electric currents flow perpendicular to the strike. Here \mathbf{E} and \mathbf{H} are the electric and magnetic fields, σ is the conductivity (inverse of the resistivity ρ), μ is magnetic permeability for free space and ω is the angular frequency.

We use a finite difference (FD) method with the TE and TM mode differential equations discretized as in Smith and Booker (1991). As usual, the discrete form of the differential equations can be expressed as $\mathbf{Ax} = \mathbf{b}$, where \mathbf{b} contains the terms

associated with the known boundary values and the source fields, and \mathbf{x} represents the unknown fields (\mathbf{E} for TE and \mathbf{H} for TM). The matrix \mathbf{A} is sparse (5-nonzero diagonals) and symmetric, but not Hermitian since the diagonal elements are not real. Boundary conditions for the top, right and left sides of the model domain are as in Smith and Booker (1991). At the bottom of the model domain, we extend the last layer at least one skin depth and set the field at the bottom of this extra layer to zero. The accuracy of the solution is controlled by the quality of the mesh. The reader is referred to Rodi (1976), deGroot-Hedlin and Constable (1990) and Smith and Booker (1991) for discussion of these issues.

Sensitivities for MT data (e.g., the apparent resistivity or phase) can be readily calculated in terms of sensitivities of the electric and magnetic field components at the surface (Rodi, 1976; Mackie and Madden, 1993). The surface electric and magnetic field components in turn can always be expressed in the general form $\mathbf{a}^T \mathbf{x}$ (Rodi, 1976) where \mathbf{a}^T may depend upon \mathbf{m} , and \mathbf{x} is the discrete electric (TE) or magnetic (TM) field solution. These sensitivities may be calculated from

$$\frac{\partial(\mathbf{a}^T \mathbf{x})}{\partial m_j} = \frac{\partial \mathbf{a}^T}{\partial m_j} \mathbf{x} - \mathbf{a}^T \mathbf{A}^{-1} \left[\frac{\partial \mathbf{A}}{\partial m_j} \mathbf{x} \right] \quad j = 1, \dots, M. \quad (3.23)$$

The second term of the right hand side in (3.23) can be computed by solving the same system of equations required for the forward problem, but with a different right hand side, $(\partial \mathbf{A} / \partial m_j) \mathbf{x}$.

Rodi (1976) shows that there are two ways to calculate the field component sensitivities of (3.23): one requires solving the forward problem M times per period, once for each m_j . The other uses the reciprocity property of the forward problem and requires N_s (number of stations) forward solutions per period. Good reviews of the sensitivity calculations can be found in Rodi (1976), Rodi and Mackie (1999)

and Mackie and Madden (1993). Since $N_s \ll M$ the second approach is generally much more efficient. We use this approach for REBOCC.

The sparse system of linear equations $\mathbf{Ax} = \mathbf{b}$ can be solved in two general ways: by a direct method using the LU decomposition or with an iterative method (Press et al., 1992), such as preconditioned conjugate gradients (PCG). Both approaches have advantages. With a direct method, after \mathbf{A} is decomposed into lower (\mathbf{L}) and upper (\mathbf{U}) triangular matrices, solution (by forward and back substitution) is extremely fast. The direct approach is thus best for constructing the sensitivity matrix where the same system of equations must be solved for multiple right hand sides. To solve for a single right hand side (e.g., when calculating the misfit during search for λ) an iterative method which takes advantage of sparseness is more efficient. We use an iterative approach in these circumstances.

PCG is an extremely effective solution method for positive definite symmetric systems (Barrett et al., 1994). PCG requires less storage and fewer operations per iteration than other iterative methods such as the Biconjugate gradient method. In our case, the coefficient matrix \mathbf{A} is symmetric but the diagonal is not real. However, since the imaginary parts of the diagonal terms are considerably smaller than the real parts, it is still possible that this algorithm will work reasonably well (Mackie and Madden, 1993). To use the classical PCG on the complex symmetric system (non-Hermitian), the conjugate transpose is not applied when computing the inner product, i.e. we use $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$ (Barrett et al., 1994). To speed up convergence, we use the incomplete LU decomposition level 3, ILU(3), (Kershaw D.S., 1978; Smith and Booker, 1991) as a preconditioner.

3.5.2. Model Covariance \mathbf{C}_m

In the penalty functional (3.4), the inverse of the model covariance is required to evaluate the norm $\mathbf{m}^T \mathbf{C}_m^{-1} \mathbf{m}$. Inversion of a non-diagonal $M \times M$ model covariance would not be computationally practical. Thus, for model space inversion approaches it is conventional to formulate the model norm in terms of a “roughness penalty”, instead of the inverse of a covariance matrix (for example, deGroot-Hedlin and Constable (1990) used the first derivative roughness penalty, in equation (1)). The relationship of some common roughness penalties to a generalized sort of model covariance is discussed by Wahba (1990).

With a data space approach, inversion of \mathbf{C}_m is not required (see (3.9)). This gives us a great deal of freedom to define a model covariance which can include prior information of various sorts. However, it is not practical to even form an arbitrary full (non-diagonal) $M \times M$ model covariance matrix for typical 2-D problems. We thus consider a fairly general class of model covariance functions which allow reasonable flexibility for including prior information, and at the same time allow for efficient computation of the matrix product $\mathbf{C}_m \mathbf{J}^T$ needed in (3.9) and (3.12) without actually computing \mathbf{C}_m explicitly.

Initially we consider covariances of the general form

$$\mathcal{C}_M(\mathbf{r}, \mathbf{r}') = \eta^2(\mathbf{r}) \varrho(\mathbf{r} - \mathbf{r}') \quad (3.24)$$

where $\eta^2(\mathbf{r})$ is a prior model variance at \mathbf{r} in the model domain and $\varrho(\mathbf{r} - \mathbf{r}') = \exp(-[(\mathbf{r} - \mathbf{r}')/r_e]^2)$ is the model correlation (with length scale r_e). For our initial discussion we assume $\eta^2(\mathbf{r})$ is a constant. \mathbf{C}_m is the discrete representation of the covariance function $\mathcal{C}_M(\mathbf{r}, \mathbf{r}')$. Multiplication of a vector \mathbf{u} in the discrete model parameter space (e.g. \mathbf{u} could be a column of \mathbf{J}^T) by \mathbf{C}_m is a discrete representation of the integral

$$\int \mathcal{C}_{\mathcal{M}}(\mathbf{r}, \mathbf{r}') \mathcal{U}(\mathbf{r}') d\mathbf{r}' \quad (3.25)$$

in which $\mathcal{U}(\mathbf{r}')$ is smoothed by convolution with $\mathcal{C}_{\mathcal{M}}$, and \mathbf{u} is discrete representation of \mathcal{U} .

Egbert et al. (1994) show that this integral can be computed (up to a scalar factor) by introducing a “pseudo-time” t and solving the diffusion equation

$$\frac{\partial \mathcal{U}}{\partial t} = \gamma \nabla^2 \mathcal{U} \quad (3.26)$$

with initial condition $\mathbf{u}(\mathbf{r}')$. Here ∇^2 is the 2-D Laplacian operator and γ is a diffusion parameter. If the pseudo diffusion time τ is chosen so that $r_e = \sqrt{4\gamma\tau}$ the integral (3.25) is given by $\mathbf{u}(\mathbf{r}, t = \tau)$. The matrix product $\mathbf{C}_m \mathbf{u}$ can thus be computed by explicit time stepping of (3.26) on the model grid from $t = 0$ to τ . In general, allowing for a spatially varying variance $\eta^2(\mathbf{r})$ the scheme can be expressed as

$$\mathbf{C}_m = \mathbf{\Sigma} \mathbf{D}_0^{-\frac{1}{2}} \mathbf{D}^\tau \mathbf{D}_0^{-\frac{1}{2}} \mathbf{\Sigma}, \quad (3.27)$$

where $\mathbf{\Sigma}$ is the discrete model space variance matrix, \mathbf{D} is the sparse diffusion operator matrix and \mathbf{D}_0 is a diagonal normalization matrix. The normalization factor of each block is calculated from solving (3.26), by replacing \mathcal{U} with the delta function as an initial condition. Therefore, this normalization matrix requires solving the diffusion equation M times, but only once at the beginning of the process. Note that each step (i.e., multiplication by \mathbf{D}) represents a local smoothing of the field which requires minimal storage or computation time.

Egbert et al (1994) applied this approach to an oceanographic inverse problem. Here, we consider a slightly different approach which is more efficient and better suited to the MT inverse problem. Stability considerations limit the size of

γ (see Press et al., 1992), and force τ to be a very large number if r_e is large. This makes a direct application of (3.27) impractical.

For REBOCC we have developed a more efficient approach based on a simple application of operator splitting methods (Press et al., 1992). With this approach the 2-D diffusion equation is replaced by a series of 1-D problems alternating between vertical and horizontal directions.

$$\mathbf{C}_m = \Sigma \mathbf{D}_0^{-\frac{1}{2}} [\mathbf{D}_H^{\frac{1}{2}} \mathbf{D}_V \mathbf{D}_H^{\frac{1}{2}}]^\tau \mathbf{D}_0^{-\frac{1}{2}} \Sigma, \quad (3.28)$$

Solutions of each 1-D diffusion equation (i.e. multiplication by \mathbf{D}_H and \mathbf{D}_V) is rapid with a fully implicit method which is trivially implemented for a 1-D problem. This scheme is stable for arbitrarily long “time steps”. Since we do not need an accurate solution of the diffusion equation to define a reasonable model correlation function, we use only a small number of pseudo-time steps, τ .

With this approach, it is not necessary to use a constant correlation length scale (or variances) throughout the model. Different length scales can be used in vertical and horizontal directions. The model covariance approach is thus very flexible, and allows rather general statistical specification of prior information. For example, known faults can be allowed for by letting the smoothing length scale go to zero at the known fault location, or part of the resistivity structure can be frozen by letting variances go to zero. However, the choice of proper correlation length scale is important, but difficult to justify rigorously. Making r_e too large can result in difficulties in finding any models that fit the data. Choosing r_e too small can also result in problems (e.g., confusion of static shifts with deep lateral structure). Inevitably some experimentation with correlation length scales will be required.

As a default strategy we make the vertical correlation length scale of each layer proportional to the depth of that layer, with the horizontal correlation length

scale set to the maximum of the depth and the gap between stations. On the edges of the model domain, the length scale is set equal to the distance from sides of the model to the edge stations. This choice of correlation length scales coincides with the loss of the resolving power of the data at depth, and near the boundaries of the model.

The model covariance of (3.28) allows only for perturbations around the assumed background model, which may in fact be grossly wrong. We can make some allowance for this in the model covariance by adding a constant matrix \mathbf{K} (all elements are κ) so that the covariance matrix \mathbf{C}_m becomes

$$\mathbf{C}_m = \Sigma(\mathbf{K} + \mathbf{D}_0^{-\frac{1}{2}}[\mathbf{D}_{\mathbf{H}}^{\frac{1}{2}}\mathbf{D}_{\mathbf{V}}\mathbf{D}_{\mathbf{H}}^{\frac{1}{2}}]^{\tau}\mathbf{D}_0^{-\frac{1}{2}})\Sigma.$$

The addition of \mathbf{K} to the model covariance corresponds to allowing for uncertainties in the level of a constant background resistivity. In our experience $\kappa = 1$ usually works well (with $\eta^2(\mathbf{r}) = 1$), however, the limits of this choice need to be better understood.

Finally, note that the data subspace cross-product matrix can be written

$$\begin{aligned} \Gamma^1 &= \mathbf{J}\mathbf{C}_m\mathbf{J}^T \\ &= \mathbf{J}\Sigma\mathbf{K}\Sigma\mathbf{J}^T + ([\mathbf{D}_{\mathbf{H}}^{\frac{1}{2}}\mathbf{D}_{\mathbf{V}}\mathbf{D}_{\mathbf{H}}^{\frac{1}{2}}]^{\frac{\tau}{2}}\mathbf{D}_0^{-\frac{1}{2}}\Sigma\mathbf{J})^T([\mathbf{D}_{\mathbf{H}}^{\frac{1}{2}}\mathbf{D}_{\mathbf{V}}\mathbf{D}_{\mathbf{H}}^{\frac{1}{2}}]^{\frac{\tau}{2}}\mathbf{D}_0^{-\frac{1}{2}}\Sigma\mathbf{J}). \end{aligned} \quad (3.29)$$

Computation of the cross-product matrix is thus accomplished most efficiently by doing half the smoothing of the sensitivities to compute $([\mathbf{D}_{\mathbf{H}}^{\frac{1}{2}}\mathbf{D}_{\mathbf{V}}\mathbf{D}_{\mathbf{H}}^{\frac{1}{2}}]^{\frac{\tau}{2}}\mathbf{D}_0^{-\frac{1}{2}}\Sigma\mathbf{J})$ followed, by computation of the cross products.

3.5.3. Line Search for λ

At every iteration, we search for the λ that gives the model (defined by (3.12) and (3.20)) with the minimum misfit (Phase I), or at the desired misfit (Phase

II). For each trial value of λ , the system of equation (3.20) must be solved, the model updated, and the misfit computed by solving the forward problem. Therefore, minimizing the number of λ s tried can help us to further reduce computational costs.

We use a relatively simple search method which takes advantage of several facts: (a) the misfit is a smooth function of λ , (b) the range of $\log_{10} \lambda$ is generally within the interval $[0,6]$ (for our default model variance $\eta^2(\mathbf{r}) = 1$), and (c) the optimal choice of λ changes little between iterations.

For the first iteration, we start with three different equally spaced λ s covering one decade of λ . Misfits computed for these initial λ s will tell us whether we should go left or right, or stop if a minimum has already been bracketed. Once a minimum is bracketed, we use a parabolic interpolation (Press et al., 1992) on these three points to estimate the minimum of the misfit. This estimated misfit might be significantly lower than the misfits from the three λ s. If not, we just keep the best λ so far and move on to the next iteration. Usually, this scheme requires about 3 to 8 trials values of λ , about half of the 8 to 12 values of λ per iteration reported for the 2-D Occam's inversion by deGroot-Hedlin and Constable, (1990).

With the prior knowledge of the previous iteration, the previous three bracketing points can be used as a starting point for the next iteration, and the process is repeated. Generally, the optimal λ does not change much between iterations, so only a small number of trial values of λ (but at least 3) are required.

Once the desired misfit is achieved within the range of trial λ s bracketing the minimum, parabolic interpolation is used to locate λ providing the desired level of misfit. If two or more values of λ bracketing the minimum have the same (desired) misfit, we choose the larger λ , which usually corresponds to a smaller model norm.

In general, the desired misfit may never be reached, so that the smoothing process is not performed. The model with the minimum misfit (higher than the de-

sired misfit) might then contain some unnecessary features inserted by the inversion to make the fit better. Therefore, we would recommend an additional run with a higher target misfit to find the minimum norm model corresponding to a larger (and this time achievable) misfit.

3.5.4. Interpolation Matrix

As described above, we require an approximation to the full sensitivity matrix \mathbf{J}_k to use all of the data in the approximate penalty functional (3.19). This approximation can be made via the subset sensitivity matrix \mathbf{G}_k and an $N \times L$ interpolation matrix \mathbf{B} .

Here, we describe our interpolation scheme which is fairly simple, while still allowing for a general pattern of data subsets. Each type of data (e.g. TM phase) is tabulated into a 2-dimensional array ordered by period and station location (see Figure 3.3). To interpolate the sensitivity to a site/period for which the sensitivity is not computed (j_s, j_p) (filled circle) we first divide the area surrounding it into eight parts, A_1, \dots, A_8 , as indicated in Figure 3.3.

In each area, we search for the nearest calculated sensitivity, indexed by (i_s, i_p). The interpolated sensitivity is then a weighted average of the nearest calculated sensitivities. The weighting coefficients for the eight nearest elements are given as

$$\Pi(i_s, j_s; i_p, j_p) = \frac{\Pi(i_s, j_s)\Pi(i_p, j_p)}{\sum \Pi(i_s, j_s)\Pi(i_p, j_p)}$$

where $\Pi(i_s, j_s) = \exp[-\frac{P_{ij}^s}{\Delta_{ij}^s}]$, $\Pi(i_p, j_p) = \exp[-\frac{P_{ij}^p}{\Delta_{ij}^p}]$, and Δ_{ij}^s and Δ_{ij}^p are the distance between stations and log10 period difference respectively, and the summation is over the eight subareas of Figure 3.3.

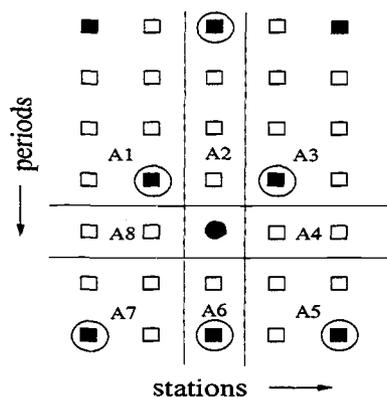


FIGURE 3.3. Schematic interpolation example. Filled squares correspond to calculated sensitivities, and open circles the remaining data. To interpolate the calculated sensitivities to the filled circle, the data distribution is divided into eight areas, A_1, \dots, A_8 , as indicated. The nearest calculated sensitivities (filled squares inside the circle) are then chosen from each area to form the interpolated sensitivity at the target point. Sub-areas where there are no calculated sensitivities, e.g. A_4 and A_8 , are omitted from the weighted average.

This simple interpolation scheme can fail when the station spacing is large. However, in such a case, it would probably be best to avoid interpolation between stations, and use a *pth*-stripe pattern.

3.5.5. Static Shift Correction

Shallow local inhomogeneities can distort the regional electric field, and cause a frequency-independent shift in the log apparent resistivity while leaving the phase unaffected. For REBOCC, static shifts are incorporated as additional model parameters, which are automatically estimated by the inversion (if requested by the user). For each iteration static shifts for each site are estimated using the median residual (observed minus (undistorted) calculated) log₁₀ apparent resistivity. Then, the misfits are recalculated using the (distorted) calculated responses. More sophisticated

ways can be used to obtain the static shift factor (e.g., Wu et. al., 1993; deGroot-Hedlin and Constable, 1993; Ogawa and Uchida, 1996), but tests with synthetic distorted data indicate that this simple scheme is effective.

3.6. Numerical Experiments

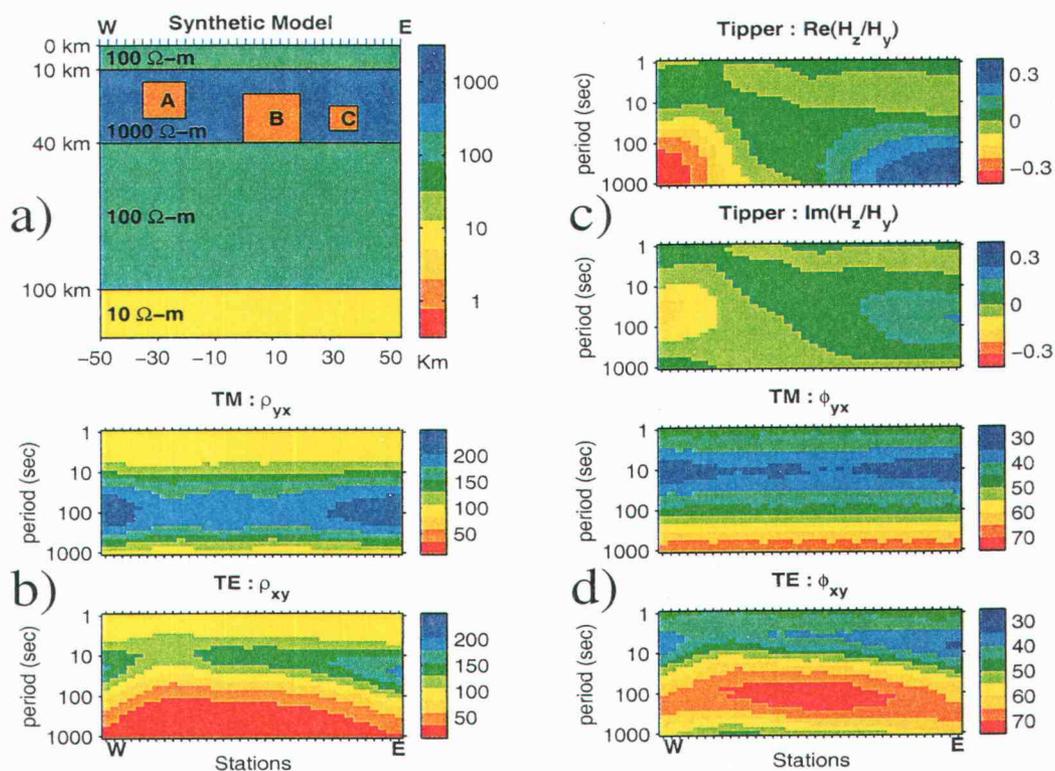


FIGURE 3.4. Synthetic model and data responses. (a) model used to generate synthetic data responses. Inside the $1000 \Omega - m$ resistive layer, there are three $1 \Omega - m$ conductors, A, B and C. Conductor A is $15 \text{ km} \times 15 \text{ km}$, buried at 15 km depth. Conductor B is $20 \text{ km} \times 20 \text{ km}$, buried at 20 km depth. Conductor C is $10 \text{ km} \times 10 \text{ km}$, buried at 25 km depth. A and B are separated by 20 km, while B and C are separated by only 10 km. Apparent resistivities and phases from TM and TE mode and Tippers generated from the model are shown in (b) - (d).

3.6.1. Synthetic Data

To test REBOCC, we generated synthetic data from a 2-D model adapted from the COPROD2 inversion results of Wu et. al. (1993). The model (Figure 3.4) consists of four layers, with three $1 \Omega - m$ rectangular conductors embedded (see Figure 3.4 for details). The grid used to form the synthetic data was well discretized at 170×183 (with an additional 10 air layers for the TE mode) to ensure accuracy of the solution. Model responses, including apparent resistivity and phases for transverse magnetic (TM) and transverse electric (TE) modes, and also vertical magnetic field transfer functions (which we will refer to as tipper, or TP), were generated using the finite element forward modeling program of Wannamaker et al. (1986). Data for 36 stations spaced at 3 km intervals, and 31 periods increasing logarithmically from 1 s to 1000 s (about 10 periods per decade) were used for the inversion test. Two percent Gaussian noise was added to the data prior to the inversion. Note that the three conductors are not clearly evident in either TE or TM model responses (Figure 3.4).

3.6.2. Results of REBOCC Inversion

Our first experiment is to invert the synthetic data using different subsets of calculated representers for TM and combined TM+TE mode data sets. We consider one checker pattern and three stripe patterns in this comparison. All inversions are run on a Sun UltraSparc I workstation. The desired RMS is set to 1 (i.e., 2 % misfit).

Due to limitations of computer resources, the model grid used for inverting the data is necessarily coarser. Different forward modeling programs and different grids will help to reveal any systematic errors occurring in our forward program.

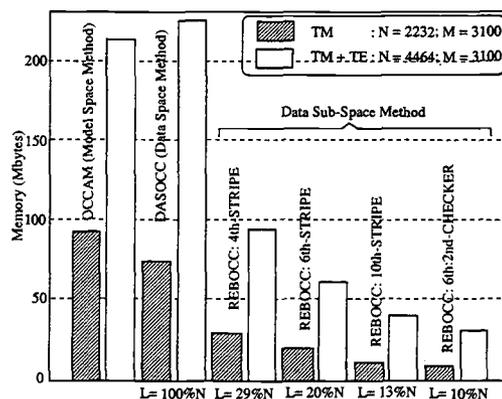


FIGURE 3.5. Memory requirements for OCCAM, DASOCC and REBOCC for different representer subsets calculated from Table 3.1 for TM and TM+TE modes. Note that double precision is used.

The model grid used for the inversion was thus discretized at 100×31 , plus 10 air layers for the TE mode. A $100 \Omega - m$ half space was used as a starting model for all inversions.

The correlation length scale r_e for all runs was set as described in the model covariance section. With a station spacing of 3 km this leads to larger horizontal length scales from the surface to 3 km depth, and a 1:1 ratio below.

Memory requirements for storing the cross-product and sensitivity matrices for all three methods (OCCAM, DASOCC and REBOCC) and the interpolation matrix (for REBOCC only) are calculated directly from Table 3.1, and are shown in Figure 3.5.

For TM+TE inversion memory requirements of the data space (DASOCC) approach exceed those of the model space (OCCAM) approach, because N is actually larger than M for this joint inversion case. However, using the reduced basis method, memory requirements can be significantly reduced (Figure 3.5). For example, only 12 MBytes of memory are needed for a TM mode inversion using the

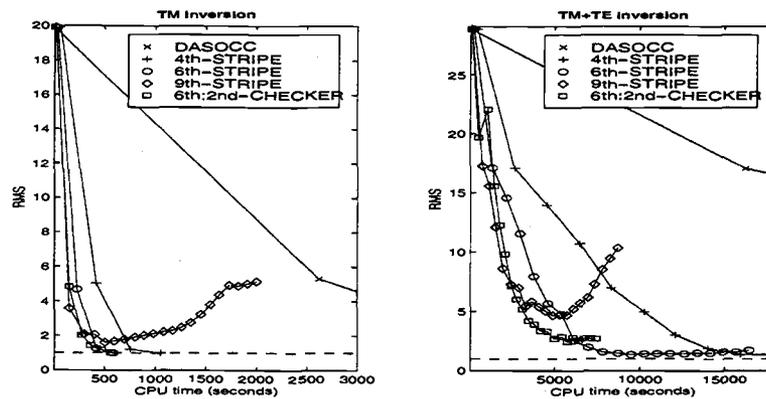


FIGURE 3.6. RMS vs. cpu time required for inversion with different representer subsets. Iteration numbers are indicated by the plotted symbols.

6th:2nd-checker pattern of calculated representers. These computations would be practical on even a small PC.

Although memory requirements for OCCAM are comparable to those of DASOCC, joint TM and TE inversion with OCCAM is not practical. Cpu time on a Sun UltraSparc I workstation for the TM only inversion ($N = 2232$ and $M = 3100$) required about 22 hours per iteration, significantly longer than required by data space methods (DASOCC and REBOCC). The long run times result from larger matrices in the model space, and from direct calculation of the sensitivity matrix elements without the reciprocity theorem. Because of the long run times required, we do not consider OCCAM further in our comparison.

The convergence rates of DASOCC and REBOCC are plotted in Figure 3.6. For both the TM and TM+TE inversions, the full basis inversion (DASOCC) requires much longer per iteration than any of the REBOCC inversions (by a factor of 6 - 35). Both the 9th-stripe and the 6th:2nd-checker data subset patterns have comparable numbers of representers (L), and require about the same amount of cpu time per iteration. However, the 6th:2nd-checker converges while the 9th-stripe does

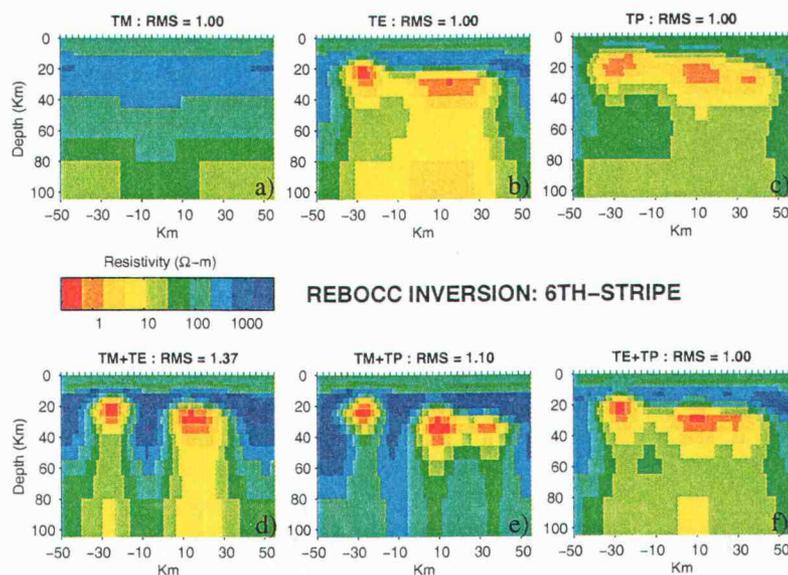


FIGURE 3.7. Models produced by the 6th-stripe REBOCC inversion from various mode combinations. The upper panels are the models from single mode inversion: (a) TM alone; (b) TE alone; (c) Tipper alone. The lower panels are models from joint inversions: (d) TM and TE; (e) TM and Tipper; (f) TE and Tipper.

not. Clearly representers for more than one period per decade is required. For the TM mode case, all of the inversions converge (except the 9th-stripe) to the desired misfit (5 iterations for the 6th:2nd-checker subset; 3 iterations for the rest). Models resulting from the different inversions are indistinguishable from one another. Only the models from the 6th-stripe subset are shown in Figure 3.7. For TM+TE inversion, none of the inversions are able to find a model with an RMS of 1. Except for the 9th-stripe, all converge to some minimum level, which tends to get larger as L gets smaller. The minimum misfit is around 1.15 RMS for the full basis, and 2.5 RMS for the 6th:2nd-checker subset.

We used a 6th-stripe representer subset for additional single and joint mode inversions, using the same synthetic data with the same starting model. Figure 3.7 and 3.8 show the results of these runs. The TM mode inversion reveals the layered

host resistivity structure with little lateral variation. In contrast, the TE inversion reveals two conductors beneath the resistive layer. The second conductor on the east side of the model from $y = 0$ to 40 km corresponds to a combination of conductors B and C in the synthetic model. The conductive features extend continuously to the deeper part of the model, and the layered host resistivity structure is not well-resolved. Interestingly, the tipper inversion displays the boundaries of the conductors more accurately than either single mode MT inversion. Even the small third conductor (C) can be distinguished. However, the layered host resistivity structure is very poorly resolved with the tipper inversion.

The joint inversion of TM and TE modes shows that the MT data is fit adequately with only two conductors inside the resistive layer (although the desired level of misfit is not quite achieved). Using tipper and the TE mode, the inversion finds all three conductors (although the last one is not well-separated from the middle one). However, the upper resistive layer is not well-defined. Similarly, with tipper and the TM mode, all three conductors inside the resistive layer are recovered. All of these joint inversions show that the tipper is needed for the inversions to resolve the smallest conductor. The model from tipper and TM mode inversion seems to be closest to the true model of Figure 3.4.

Finally, we inverted all data, TM, TE and tipper using the REBOCC inversion with a 6th-stripe pattern ($N = 6696$, $M = 3100$, and $L = 1296$). The inversion requires about 108 MBytes of memory and approximately 11 hours of run time (on a Sun UltraSparc 1) to obtain a model with a RMS of 1.0 at the 18th iteration. One can stop the inversion after the 8th iteration (about 5 hrs) to obtain a model with an RMS of about 2 (or 4 % error), which is in fact indistinguishable from the fully converged run. The model and the fit are shown in Figure 3.8.

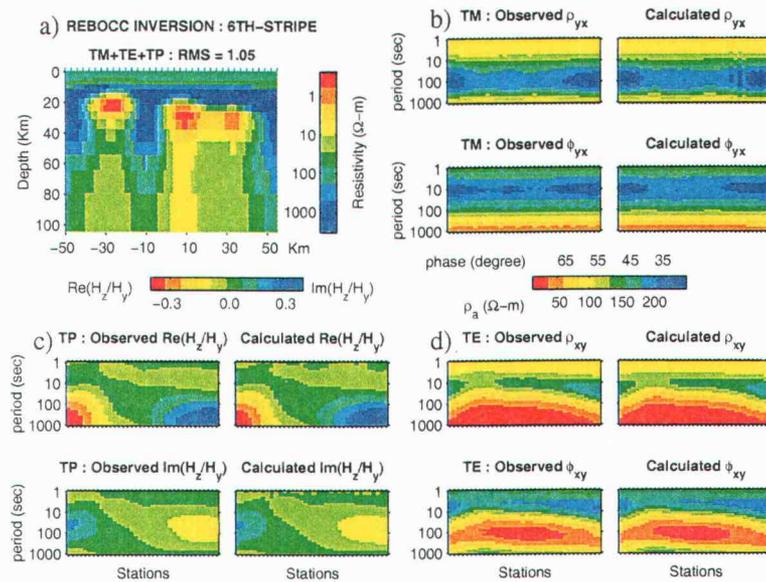


FIGURE 3.8. Result from 6th-stripe REBOCC inversion using all data. (a) Model from three mode inversion (TM, TE and tipper) with the 6th-stripe REBOCC (b)-(d) observed and predicted pseudo-sections for all data.

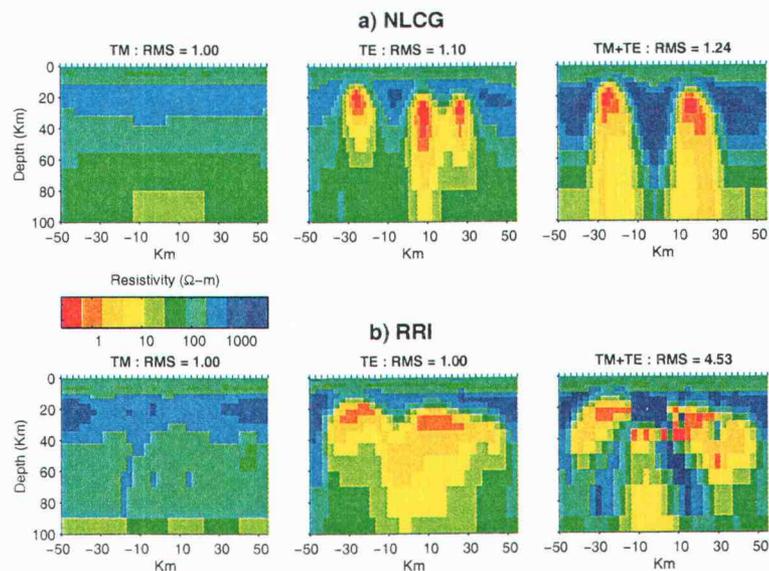


FIGURE 3.9. Models obtained from (a) NLCG and (b) RRI inversions of TM, TE and TM+TE mode. RRI fails to converge to a reasonable model for the TM+TE mode. The final result fits poorly and is very rough.

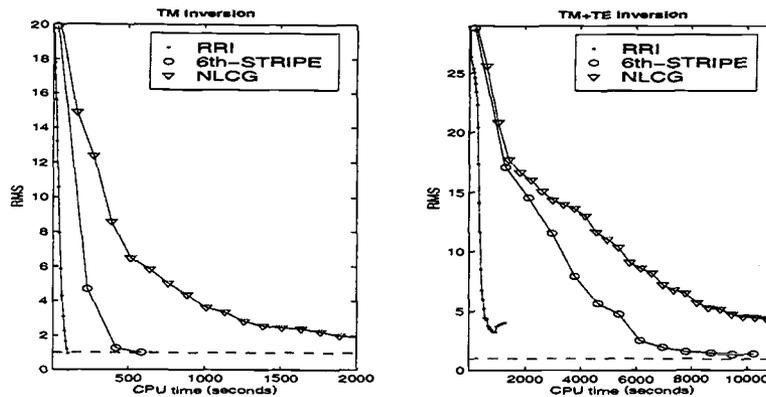


FIGURE 3.10. Plots of RMS versus cpu time of RRI, 6th-stripe REBOCC, and NLCG for TM, and TM+TE inversions. Iteration numbers are indicated by the symbols.

3.6.3. Comparison with Other Inversions

Using the same data set, we also compared REBOCC with several other inversion programs: NLCG of Rodi and Mackie (1999) and RRI of Smith and Booker (1991). The same starting model ($100 \Omega - m$ half space) and the same model mesh ($M = 100 \times 31 = 3100$) were used for all inversions. For timing comparisons, we run the inversions on the same machine, a Sun UltraSparc I. We use the 6th-stripe subset for the REBOCC inversion. For other inversion programs, we generally used the default parameter values provided with the programs.

Figure 3.9 shows the models produced by NLCG and RRI for TM, TE and TM+TE data sets (Note that NLCG and RRI do not presently allow for inversion of tippers). These plots should be compared with the REBOCC results of Figure 3.7 (a), (b) and (d). In Figure 3.10 we plot RMS misfit against cpu time (and iteration number) of RRI, REBOCC (6th-stripe) and NLCG algorithms for TM inversion and joint inversion of TM and TE modes. The plot of RMS misfit and cpu time for the TE inversion is very similar to that of the TM case, but cpu times are greater.

RRI which is incredibly fast per iteration requires 12 iterations to converge to the solution at the desired level of misfit for TM mode (27 iterations for TE mode). However, for the joint inversion of TM and TE, RRI fails to converge to a reasonable model. The minimum misfit achieved is quite large (4.53), and the final model is very rough. The small scale features (D+ like) between $y = -15$ km and 0 km (between conductor A and B in the synthetic model; Figure 3.4) are inserted in the early iterations and, once there, are very difficult to get rid of in the subsequent iterations (Smith and Booker, 1991). This confusion of the inversion is perhaps the result of the 1-D sensitivity approximation used by RRI. Beneath the sites from $y = -15$ to 0 km, the TM data (which is not sensitive to the isolated conductors) suggests a resistive structure. On the other hand, the TE data which is highly sensitive to the nearby conductors, tries to place conductive features beneath these sites.

NLCG converges to the desired misfit in 31 iterations (about 4000 seconds) for the TM mode, and about 27500 seconds (97 iterations to reach an RMS of 1.1) for the TE mode. The same misfits of both single mode inversions can be accomplished much more quickly with REBOCC about 600 seconds (3 iterations) for the TM mode (Figure 3.10) and about 1600 seconds (5 iterations) for the TE mode.

For joint inversion, NLCG requires about 51000 seconds (120 iterations) to reach an RMS of 1.4 while REBOCC only need 9500 seconds (in 11 iterations) to reach a similar level of misfit. However, in this example NLCG can reduce the RMS to a lower minimum (1.2 RMS after 165 iterations) than the the 6th-stripe REBOCC (1.37 RMS after 11 iterations). To reduce the RMS further, more basis function may be required for REBOCC.

Note that the convergence times plotted in Figure 3.10 for REBOCC reflect only Phase I iterations where the goal is to bring the misfit down to the desired

level (1.0 RMS). Since for the joint (TM+TE) inversion the desired level cannot be reached, one should restart the inversion so that Phase II can be completed with the desired misfit level set to a higher value (e.g., 1.5 RMS starting with the model from the 9th iteration of Phase I). Phase II is necessary in order to obtain the minimum structure model (Parker, 1994). At the same time, for NLCG, to obtain a model with minimum structure, several trial values of λ would be required.

It is difficult to compare the models produced by NLCG, RRI and REBOCC (Figure 3.9 and Figure 3.7) because of the difference of the smoothing parameters (model covariance). However, all of these inversions (except the joint mode inversion for RRI) reveal the main features and produce responses that fit the data adequately.

3.6.4. Example with Field Data

Finally we briefly consider application of REBOCC to a real data set, a high density MT profile across the San Andreas fault (SAF) near Parkfield, CA (Unsworth et al., 1997). This is a large data set with 55 TM mode responses, 37 TE mode responses and 15 tippers each at 41 periods. We discretized the model at 200×74 blocks (plus 10 air layers for the TE mode), and used a $10 \Omega - m$ half space as a starting model. The static shift distortion parameters are set free so that the program can automatically adjust the values. An 8th-stripe subset ($L = .15N$) is used for joint inversion of TM, TE and vertical magnetic transfer functions. The results of inversion are shown on Figure 3.11. After the 4th iteration (about 12 hrs of cpu time on Sun UltraSparc I), the inversion finds a model with a misfit of about 13 % (an RMS of about 2.7). A slightly better misfit can be obtained after a few more iterations. However, the norm of the model increases significantly when the data is fit better. Dimensionality analysis of the full impedance tensor suggests

that the misfit of the model of Figure 3.11 is reasonable, given the degree of 3-D complications in this data set (Siripunvaraporn et al., 1998). Thus, we prefer the model with the misfit at the 4th iteration. These cpu times are large, but this is a very large data set ($N = 8774$) and big model ($M = 14800$).

3.7. Discussions and Conclusions

The REBOCC inversion has been shown to be effective in practice. By using a relatively small subset of the representers (i.e., smoothed sensitivities) computational requirements (both memory and cpu time) can be substantially reduced. The p th-stripe subset requires sensitivity calculations for only every p th frequency. Our results show that even with p up to about 8 the inversion usually converges to a reasonable model, which does not depend in any significant way upon p .

In our comparisons with other inversions, we found that REBOCC was significantly faster than both OCCAM and NLCG. None of these inversions (including REBOCC) are competitive with RRI in terms of speed. However, we could not get RRI to converge for joint mode (TM+TE) inversion of our test data set. Possibly, a more experienced user of this program might be able to vary or adjust some parameters to make the inversion work successfully. Similarly for NLCG, the convergence rate might vary with different values of λ . Perhaps a reasonable model could be obtained more rapidly with a different choice of λ . However, to find a true minimum structure solution (as REBOCC does) NLCG would have to be run with several trial values of λ (some of which might converge more slowly).

CG and NLCG are descent methods which make no use of the second derivative (Hessian) of the penalty functional. GN and OCCAM essentially calculate the full Hessian, while REBOCC make a very good (but not perfect) approximation.

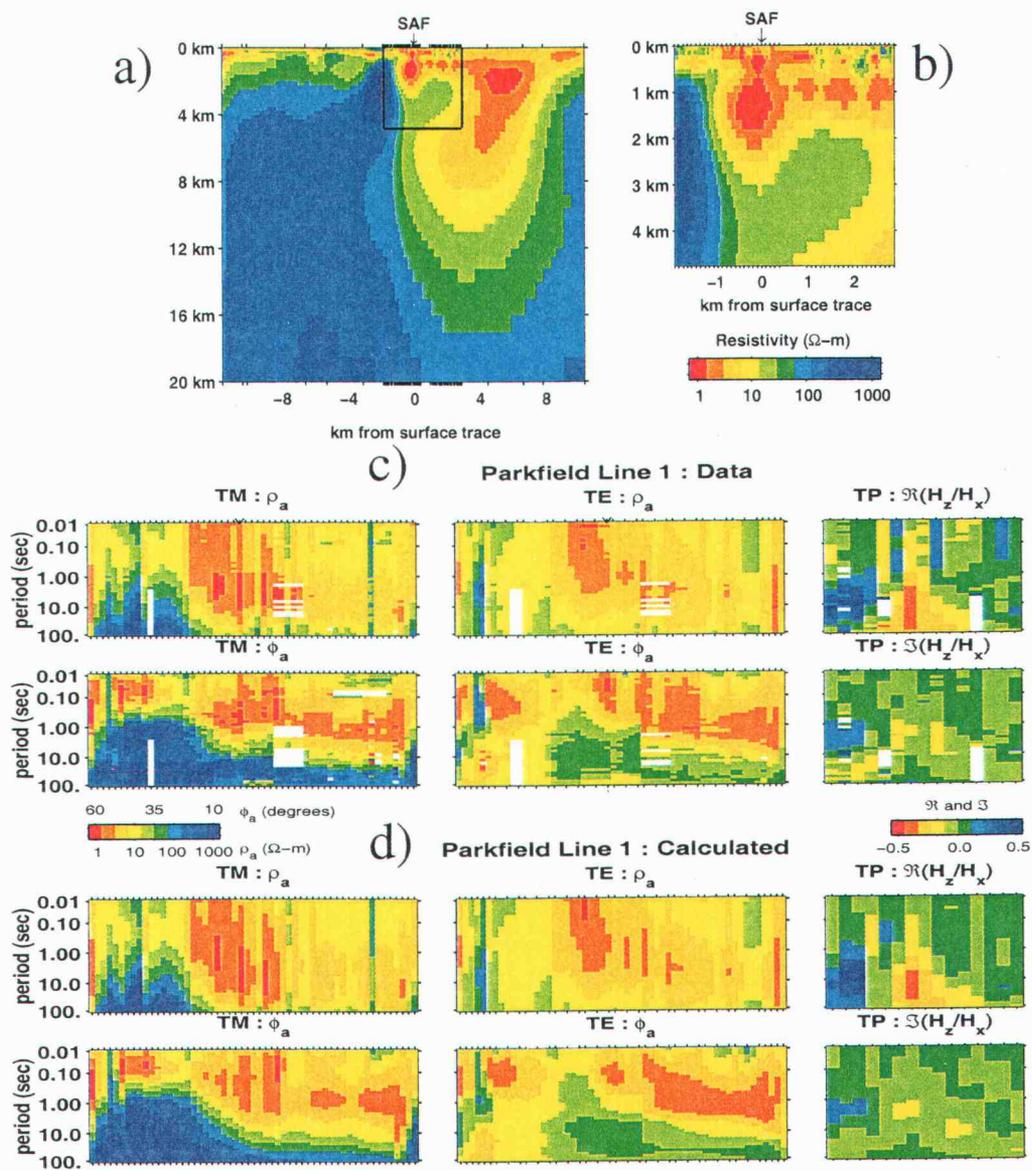


FIGURE 3.11. Results from an 8th-stripe REBOCC inversion on a real data set from a high density profile across the SAF. (a) gives the inverse model covering the whole profile; (b) is a zoom of the rectangular region near the fault. Data sites are indicated by tick marks along the surface. (c) measured data for TM, TE and tippers; (d) corresponding calculated responses.

Our results indicate that this approximate calculation is worth the effort. This is particularly true for the 2-D case considered here, where direct LU factorization of the differential equation coefficient matrix is feasible, since in this case many sensitivities can be computed quickly once the factorization is complete.

The basic idea behind the REBOCC algorithm could also be applied to the 3-D inversion problem. We thus do not agree with the statement made by Rodi and Mackie (1999) that any sort of inversion based on a sensitivity calculation will not be practical for realistic 3-D EM problems, even allowing for improvement of the computer hardware. A straightforward extension of REBOCC to 3-D inversion is readily apparent, although for calculating the sensitivity matrix the LU decomposition would probably be impractical (except on a supercomputer) and have to be replaced by a relaxation method. The possibility of adapting the reduced basis method to other inversion approaches also deserves consideration.

For example, one could try to solve (3.20) using conjugate gradients. As in the model space conjugate gradient approach used by Mackie and Madden (1993) the matrix multiplications required can be reduced to solving the forward problem twice per inner loop iteration. Using the reduced basis idea, forward calculations could be made for only a subset of frequencies. Sensitivity and cross-product matrices would not have to be constructed or stored. However, a full conjugate gradient solution would be required for each λ , so a true OCCAM type approach would probably not be practical. A similar data space conjugate gradient scheme has been applied to oceanographic inversion problems by Egbert (1997) and Bennett et al. (1996).

We believe that REBOCC is a significant advance in practical methods for solution of 2-D MT inverse problems, and that the basic ideas of our implementation should be highly relevant to the 3-D problem in some form.

3.8. Acknowledgments

This research has been partially supported by the Royal Thai Embassy via the Development and Promotion of the Science and Technology (DPST) scholarship to WS, and by grants from the US DOE (DE-FG03-96ER1495) and NSF (9614411-EAR).

3.9. References

- [1] Barret, R., Berry, M., Chan, T. F., Demmel, J., Donato, J., Dongarra, J., Eijkhout, V., Pozo, R., Romine, C., and Van der Vorst, H., 1994, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*, Soc. Ind. Appl. Math.
- [2] Bennett, A. F., Chua, B. S., and Leslie, L. M., 1996, Generalized inversion of a global numerical weather prediction model, *Meteorology and Atmospheric Physics*, **60**, 165-178.
- [3] Chen, L., Booker, J. R., Jones, A. G., Wu, N., Unsworth, M., Wei, W., and Tan, H., 1996, Electrically conductive crust in southern Tibet from INDEPTH magnetotelluric surveying, *Science*, **274**, 1694-1696.
- [4] Constable, C. S., Parker, R. L., and Constable, C.G., 1987, Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics*, **52**, 289-300.
- [5] deGroot-Hedlin, C., and Constable, S., 1990, Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data, *Geophysics*, **55**, 1613-1624.
- [6] deGroot-Hedlin, C., and Constable, S., 1993, Occam's Inversion and the North American Central Plains Electrical Anomaly, *J. Geomag. Geoelectr.*, **45**, 985-999.
- [7] Egbert, G. D., Bennett, A. F., and Foreman, M. G., 1994, TOPEX/POSEIDON tides estimated using a global inverse model, *J. Geophys. Res.*, **99**, 24821-24852.
- [8] Egbert, G. D, 1997, Tidal data inversion: interpolation and inference, *Prog. Oceanog.*, **40**, 53-80.

- [9] Farquharson, C. G., and Oldenburg, D. W., 1996, Approximate sensitivities for the electromagnetic inverse problem, *Geophys. J. Int.*, **126**, 235-252.
- [10] Jones, A. G., 1992, Electrical conductivity of the continental lower crust, *in* Fountain, D. M., Arculus, R. J., and Kay, R. W., Eds., *Continental Lower Crust*: Elsevier Science Publ. Co., Inc., 81-143.
- [11] Jupp, D. L. B., and Vozoff, K., 1975, Stable iterative methods for the inversion of geophysical data, *Geophys. J. R. astr. Soc.*, **42**, 957-976.
- [12] Kershaw, D.S., 1978, The Incomplete Cholesky-Conjugate Gradient Method for the Iterative Solution of Systems of Linear Equations, *J. Comput. Phys.*, **26**, 43-65.
- [13] Mackie, R. L., and Madden, T. R., 1993, Three-dimensional magnetotelluric inversion using conjugate gradients, *Geophys. J. Int.*, **115**, 215-229.
- [14] Marquardt, D. W., 1963, An Algorithm for least-squares estimation of nonlinear parameters, *J. Soc. Indust. Appl. Math*, **11**, 431-441.
- [15] Oldenburg, D. W., and Ellis, R. G., 1991, Inversion of geophysical data using an approximate inverse mapping, *Geophys. J. Int.*, **105**, 325-353.
- [16] Oldenburg, D. W., McGillivray, P. R., and Ellis, R. G., 1993, Generalized subspace methods for large scale inverse problems, *Geophys. J. Int.*, **114**, 12-20.
- [17] Oldenburg, D. W., and Ellis, R. G., 1993, Efficient inversion of magnetotelluric data in two dimensions, *Phys. Earth Planet. Inter.*, **81**, 177-200.
- [18] Ogawa, Y. and Uchida, T., 1996, A two-dimensional magnetotelluric inversion assuming Gaussian static shift, *Geophysics*, **126**, 69-76.
- [19] Orange, A. S., 1989, Magnetotelluric Exploration for Hydrocarbons, *Proc. IEEE*, **77**, 287-317.
- [20] Parker, R. L., 1980, The inverse problem of electromagnetic induction: existence and construction of solutions based upon incomplete data, *J. Geophys. Res.*, **85**, 4421-4425.
- [21] Parker, R. L., and Shure, L., 1982, Efficient Modeling of the Earth's Magnetic Field with Harmonics Splines, *Geophys. Res. Lett.*, **9**, 812-815.
- [22] Parker, R. L., 1994, *Geophysical Inverse Theory*: Princeton Univ. Press.
- [23] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., 1992, *NUMERICAL RECIPES in FORTRAN: The Art of Scientific Computing - 2nd edition*: Cambridge Univ. Press.

- [24] Rodi, W. L., 1976, A technique for improving the accuracy of finite element solutions for magnetotelluric data, *Geophys. J. R. astr. Soc.*, **44**, 483-506.
- [25] Rodi, W. L., and Mackie, R. L., 1999, Nonlinear Conjugate Gradients Algorithm for 2-D Magnetotelluric Inversion, *Geophysics*, submitted.
- [26] Siripunvaraporn, W., Egbert, G. D., Eisel, M., and Unsworth, M., 1998, A High Resolution EM Survey of the San Andreas Fault (SAF): Local Conductivity Structure in a Regional Context, *EOS, Trans. Am. Geophys. Union*, **79**, 45, Fall Meet. Suppl., F227.
- [27] Smith, J. T, and Booker, J. R., 1988, Magnetotelluric inversion for minimum structure, *Geophysics*, **53**, 1565-1576.
- [28] Smith, J. T, and Booker, J. R., 1991, Rapid Inversion of Two- and Three-Dimensional Magnetotelluric Data, *J. Geophys. Res.*, **96**, 3905-3922.
- [29] Uchida, T., 1993, Smooth 2-D Inversion for magnetotelluric data based on statistical criterion ABIC, *J. Geomag. Geoelectr.*, **45**, 841-858.
- [30] Unsworth, M. J., Bedrosian, P., Egbert, G. D., and Eisel, M., 1997, 3-D Electrical Resistivity Structure of the San Andreas Fault at Parkfield, California, *EOS, Trans. Am. Geophys. Union*, **78**, 46, Fall Meet. Suppl., F696.
- [31] Unsworth, M., Egbert, G., and Booker, J., 1998, High resolution electromagnetic imaging of the San Andreas fault in Central California, *J. Geophys. Res.*, **104**, 1131-1150.
- [32] Wannamaker, P. E., Stodt, J. A., and Rijo, L., 1986, Two-dimensional topographic responses in magnetotelluric modeled using finite elements, *Geophysics*, **51**, 2131-2144.
- [33] Wannamaker, P. E., Booker, J. R., Filloux, J. H., Jones, A. G., Jiracek, G. R., Chave, A. D., Tarits, P., Waff, H. S., Egbert, G. D., Young, C. T., Stodt, J. A., Martinez, M. G., Lwaw, L. K., Yukutake, T., Segawa, J. S., White, A., and Green, A. W., Jr, 1994, Magnetotelluric Observations Across the Juan de Fuca Subduction System in the EMSLAB Project, *J. Geophys. Res.*, **94**, 14111-14125.
- [34] Wahba, G., 1990, Spline Methods for Observational Data, *Soc. Ind. Appl. Math.*
- [35] Vozoff, K., 1972, The Magnetotelluric Method in the Exploration of Sedimentary Basins, *Geophysics*, **37**, 98-141.
- [36] Weidelt, P., 1972, The inverse problem of geomagnetic induction, *Z. Geophys.*, **38**, 257-289.

- [37] Wu, N., Booker, J. R., and Smith, J. T., 1993, Rapid Two-Dimensional Inversion of COPROD2 Data, *J. Geomag. Geoelectr.*, **45**, 1073-1087.

4. HIGH RESOLUTION MT PROFILES ACROSS SAN ANDREAS FAULT AT PARKFIELD, CALIFORNIA

4.1. Introduction

Several pieces of evidence indicate that the San Andreas fault (SAF) is weak, and that high pressure fluids may play an important part in controlling earthquake processes along the SAF (Johnson and McEvilly, 1995). The maximum principal stresses along the SAF in central California are observed to be nearly normal to the strike (Zoback et al., 1987). Examples that support normal compression are the uplift of the Coast Ranges, and the numerous active reverse faults and folds that trend nearly parallel to the SAF. In spite of the normal compression, most earthquakes along the SAF have strike-slip, instead of thrust or normal fault plane solutions. One possible explanation for this paradox is the presence of high fluid pressure in the fault zone, which would relieve the effective normal stress across the fault, and allow the fault to slide at low shear stress. This also helps explain the lack of heat flow anomaly (Brune et al., 1969; Lachenbruch, 1980), and is also consistent with other geophysical data, including low seismic velocities in the fault zone (e.g., Trehu and Wheeler, 1987), along with low density and low resistivities which imply a highly deformed and overpressured fault zone (Zoback et al., 1987).

The main problem with this explanation of fault weakness is that the maximum the pore pressure can reach is the minimum horizontal principal stress (or else hydraulic fracturing would occur). This is not high enough to explain the angle of maximum horizontal compressive stress, given the relatively high coefficient of friction of rocks in the fault (Zoback et al., 1987). As a result, many theoretical models (Sleep and Blanpied, 1992; Byerlee, 1990 and 1993; Johnson and McEvilly, 1995) have been proposed to explain the dynamic role of fluids in the earthquake

process. To support and constrain these theoretical models, detailed studies of fault related fluids *in situ* up to 10 km depth are necessary.

Eberhart-Phillips et al. (1995) provides a good review of how surface-based geophysical methods such as seismic and electromagnetic methods, can be used to constrain the presence and distribution of fluids along the fault zone and in the surrounding crust. Different geophysical methods provide estimates of variations of different physical properties, which are generally related. Seismic methods (e.g., reflection and refraction methods) are used to detail the velocity structure, which can be linked to rock type, porosity, etc. On the other hand, electromagnetic (EM) methods offer a unique view of the geological structure at great depth in terms of the electrical resistivity (or conductivity). Electrical resistivities of rocks are strongly influenced by fluids in pore spaces and fractures, and thus have an advantage over the other methods for mapping the fluid distribution in the fault zones. However, it is still very difficult to distinguish the actual fluid features from other anomalous structures in the fault zone. Here, we use the magnetotelluric (MT) method to investigate the fault zone, and the conductivity of the crust near Parkfield, California.

4.1.1. Tectonic Development of the SAF and Seismicity

Based on the patterns of the magnetic anomalies in the oceanic crust, the San Andreas fault developed no earlier than 30 Ma (late Oligocene) when the ridge between the Pacific plate and the Farallon plate began subducting beneath the North American plate (McKenzie and Morgan, 1969; Atwater, 1970), forming two triple junctions. As the Farallon plate continued to subduct beneath the North American plate, the triple junctions migrated along the right-lateral transform fault to the northwest and southeast. The transform fault developed during this period

of time is not the current SAF trace, but a fault now located mostly to the west and at the edge of the continent (Irwin, 1990). At the present time, these migrated triple junctions are separated by 2500 km, with the Mendocino triple junction in the north, and the Rivera triple junction at the mouth of Gulf of California. The modern San Andreas fault did not apparently form until the opening of the Gulf of California, about 4 Ma. Since then, Baja California has moved about 260 km away from the Mexican mainland (Larson et al., 1968).

In central California, the modern SAF section is fairly simple and straight. It is oriented at about N40°W, between the latitudes of San Jose and Bakersfield in California (Figure 4.1). Farther to the north and south, the San Andreas fault is very complex. It spreads out into several branches, or changes abruptly, such as the Big Bend, south of Bakersfield. In central California, the 200-km-long section of the SAF is recognized as the creeping section (red segment of Figure 4.1) because of its close correlation with microseismic activities. It separates the south end of the 1906 break near San Juan Bautista from the north end of the 1857 break near Cholame Valley where the locked segment begins (blue segment of Figure 4.1). Microseismicity in the locked segment is considerably less compared to the seismicity of the creeping section.

The most notable segment in the central section is the 30-km-long Parkfield segment defined by the 5°W bend on the north end and a 1-km right-stepping offset at the southern end in the Cholame Valley (Hill et al., 1990). This segment is the transition between the south end of the creeping segment and the north end of the of the locked segment. Microseismicity in the Parkfield area is very active. A sequence of five moderate earthquakes ($M \approx 6$) has occurred on this segment at nearly 22 year interval since 1881 (Bakun and McEvilly, 1984). The last rupture on this segment was during 1966. Because of its characteristic 22-year interval, the

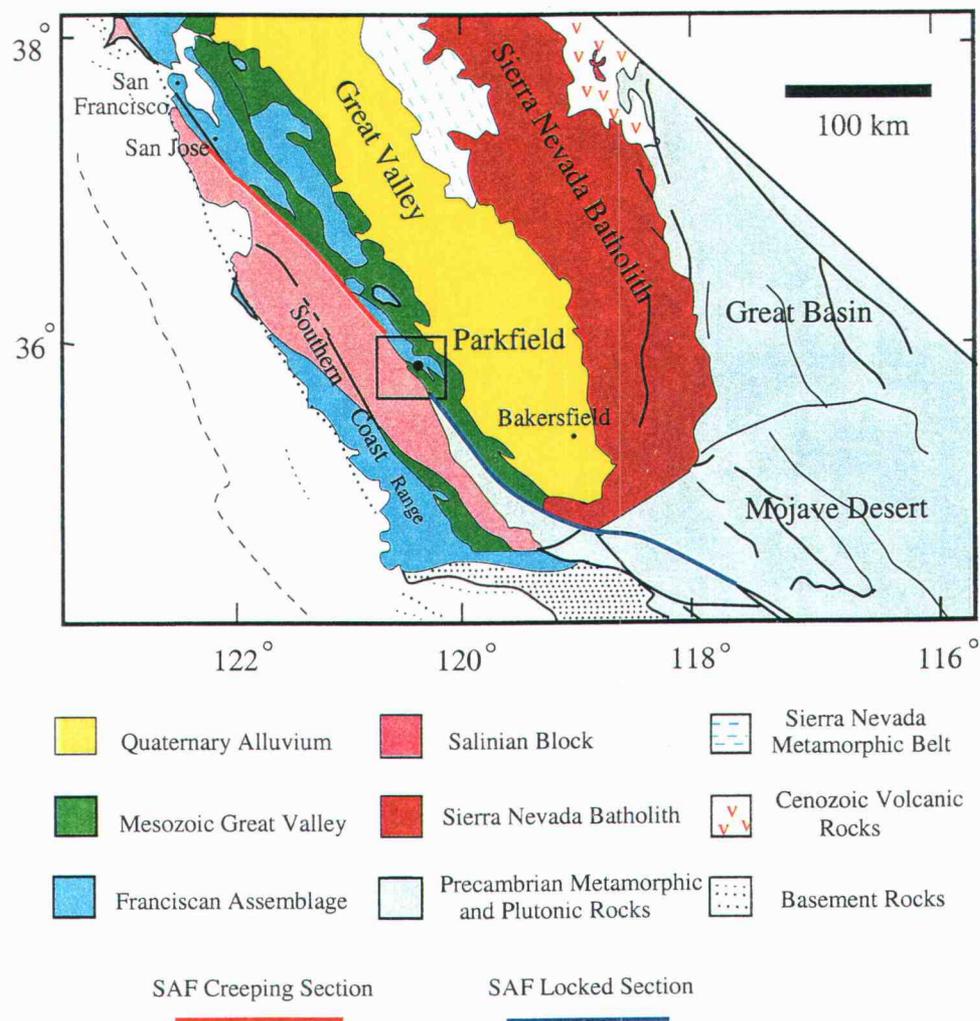


FIGURE 4.1. Geological map of Central California (after Irwin, 1990). The rectangular region is the area of study.

next rupture was expected to be between 1988 ± 10 . Because of its accessibility, this expectation leads to intensive monitoring experiments in this area to study the fault behavior during the 1980s and 1990s, in the hope of observing precursor phenomena before the next "characteristic" earthquake. Up to now, the expected earthquake of magnitude 6 has not yet occurred on this segment. This area is also the subject of intensive geological and geophysical studies, including seismic reflection and refraction (e.g. McBride and Brown, 1986; Trehu and Wheeler, 1987), earthquake tomography (e.g. Eberhart-Philips and Michael, 1993), heat flow (Sass et al., 1997), and magnetotelluric (e.g. Unsworth et al., 1999) experiments. The goal of this effort has been to understand the fault behavior in terms of the physical parameters derived from those geophysical experiments.

4.1.2. Geologic Setting of SAF at Parkfield

The geological structure in coastal central California (Figure 4.1) is composed of two distinct terranes, with the Salinian block to the southwest of the SAF and the Franciscan assemblage and Great Valley to the northeast. The Salinian block is composed of granitic and metamorphic rocks, overlain by a thin layer of sedimentary rocks. Many geologists believe that the Salinian block may be a displaced part of the Sierra Nevada (Page, 1981) because of the similarities in composition and age of the basement rocks. However, paleomagnetic data indicate that the Salinian block may have been transported 2500 km northward since Cretaceous time from its origin in Mexico (Champion et al., 1984).

The Franciscan assemblage is very complex. It is composed mainly of a melange of metasedimentary and volcanic rocks deposited as an accretionary prism associated with subduction processes (Dickinson, 1970). Despite the age and com-

positional similarities to the Great Valley sequence, they are largely different in depositional histories (Mooney and Weaver, 1989). Whereas the Franciscan is associated with subduction, the Great Valley sequence shows regular bedding in normal stratigraphic sequence, as a forearc in the subduction process.

At Parkfield, the stable Salinian block consists of Cretaceous granitic rocks and granodiorite overlain by Cenozoic marine and non-marine strata and local volcanic rocks. It is bounded on the southwest side by the Franciscan assemblage that traveled together with the Salinian block (Page, 1981), which may thus be different from the Franciscan found to the northeast. The Franciscan complex is much more heterogeneous and deformed than the Salinian block. It contains numerous folds, thrust faults, and strike-slip faults. Some of the faults are parallel to the SAF. It consists largely of dismembered sequences of graywackes, shales, and lesser amounts of mafic volcanic rock, thin-bedded chert, and limestone (Irwin, 1990).

Geologic mapping of the fault zone (Sims, 1988; 1990) indicates a variation of the fault character from place to place. At Middle Mountain, a clear fault trace can be observed through an uplift centered on the fault. About 20 km south of the Middle Mountain, at the Cholame Valley, the fault is disrupted, and separated by a 1-km-wide offset. North of this offset the SAF lies on the east side of the valley (the Parkfield segment), and south of the offset the fault lies on the west side of the valley (the Cholame segment). This right-stepping offset is believed to be the cause of the pull-apart basin that formed the Cholame Valley (Lindh and Boore, 1981) although definitive evidence for this model is still lacking (Sims, 1988). Earthquake studies shows that the surface rupture resulting from the moderate earthquake in the Parkfield segment does not continue across the Cholame Valley, but instead terminates at or just beyond the right stepover (Bakun and McEvilly, 1984). Thus, this offset may represent a crucial barrier to the propagation of the

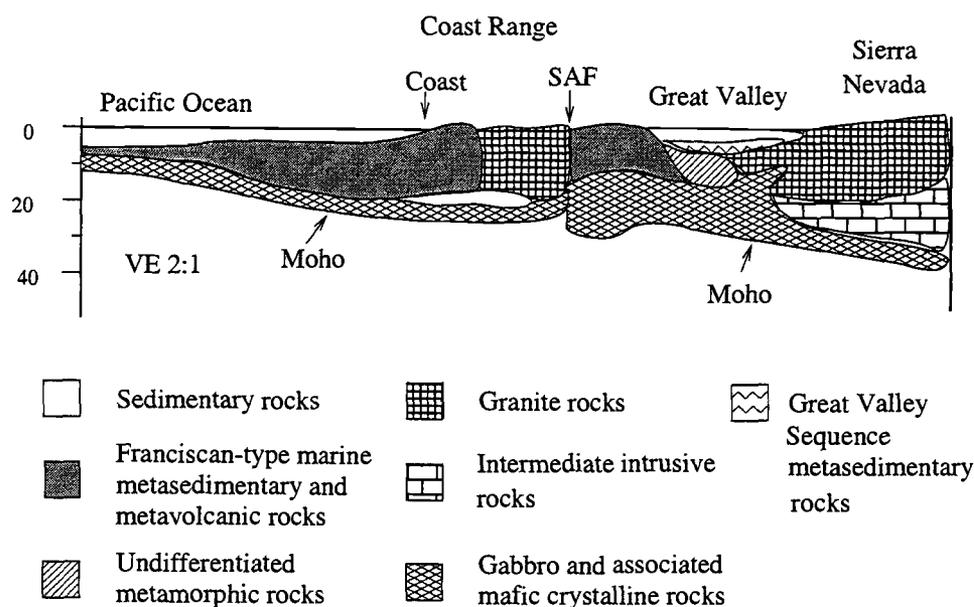


FIGURE 4.2. Geological cross-section across central California (adapted from Mooney and Weaver, 1989)

rupture associated with moderate and large earthquakes, between two segments of the SAF (Shedlock et al., 1990).

4.1.3. Past Geophysical Experiments at Parkfield

Because of the easy access to this area, many geophysical studies have been conducted to study fault characteristics and the geological structure across the fault zone. There are also plenty of data and many interpretations of the surrounding area. All of these can be used to help constrain the results from this area. The geological cross-section inferred from many geophysical experiments is given in Figure 4.2 which is adapted from Mooney and Weaver (1989).

McBride and Brown (1986) reinterpreted the 22-km-long COCORP reflection profile at Parkfield, and concluded that the Salinian block consists of a 2-km

thick layer of young sediment thinning toward the fault where the basement rock arches upward. The base of the Salinian crust is probably marked as the decay of the reflection density and amplitude at about 8-10 km depth. They speculate that this boundary marks the Franciscan or other material subducted under the Salinian basement. To the east, the basement beneath the Franciscan assemblage is located at about 13 km depth, while the Moho is at about 29 km. This is deeper than the estimated Moho at 22 km depth beneath the Salinian block. The shallow portion of the fault zone is imaged in the reflection data as a 3 km wide and 5-6 km deep disrupted zone which could be a flower structure, developed by fault-normal compression. At great depth, the fault is marked as a near-vertical structure truncating crustal boundaries as deep as 31 km.

From the reflection and refraction data across the SAF south of Cholame Valley, about 20 km south of the town of Parkfield, Trehu and Wheeler (1987) found that there is an east-dipping low velocity zone (LVZ) at around 15-22 km depth beneath the Salinian block (about 25-55 km west of SAF trace). They interpreted this LVZ as subducted sedimentary rock estimated to be about 23 Ma (assuming a subduction rate of 5 cm/yr), before the transition from subduction to strike-slip at about 20 Ma (Atwater, 1970). Beneath the fault zone, they observed another LVZ extending vertically to at least mid-crustal depths.

High-resolution reflection data across the SAF in Cholame Valley by Shedlock et al. (1990) found that the shallow structure at the south end of the Cholame Valley is composed of a zone of incoherent energy, narrow at the surface but widening with depth. This zone is located in between the active traces of the SAF fault. The offset between traces of the main fault in the Cholame Valley is characterized by the presence of many small faults and discontinuous reflections consistent with distributed shear. However, the actual fault zone is not well imaged.

Eberhart-Phillips and Michael (1993) have inverted P-wave first arrival times from local earthquakes and shots to obtain a 3-D velocity image in a 60- by 80-km region containing the Parkfield segment of the SAF. They found that the SAF is well characterized by a sharp across-fault velocity gradient, with about a 5-20 % lateral change in P-wave velocity over a 4 km width. The variations in the velocity along the SAF are also modeled. A large volume of low velocity is detected northeast of Middle Mountain, in the Franciscan and Great Valley. This low velocity zone is consistent with low resistivity material observed with MT data (W.D. Stanley, described in Eberhart-Phillips and Michael, 1993), and is interpreted to be overpressured fluid trapped beneath a low permeability serpentinite body. This low velocity zone is absent in the Gold Hill. Eberhart-Phillips and Michael (1993) also relocated all seismicity from 1969 to 1991 using the inverted 3-D velocity model. These studies show that the SAF is planar and vertical at seismogenic depths (3-14 km).

4.2. MT Data across the San Andreas Fault at Parkfield

During the fall of 1994 and 1997, University of Washington and Oregon State University conducted four high resolution MT profiles to image the electrical resistivity structure of the San Andreas fault (SAF) near Parkfield, CA. Three profiles were near Middle Mountain, and one in the Cholame Valley. Each profile runs approximately perpendicular to the SAF (Figure 4.3). MM2 which is between MM1 and MM3, was carried out earlier in 1994, and the interpretation of this line has been given by Unsworth et al. (1997) and Unsworth et al. (1999). MM1 and MM3 were both conducted in 1997 to further constrain resistivity structure of the SAF near Parkfield. Preliminary results for the new lines are given by Unsworth et

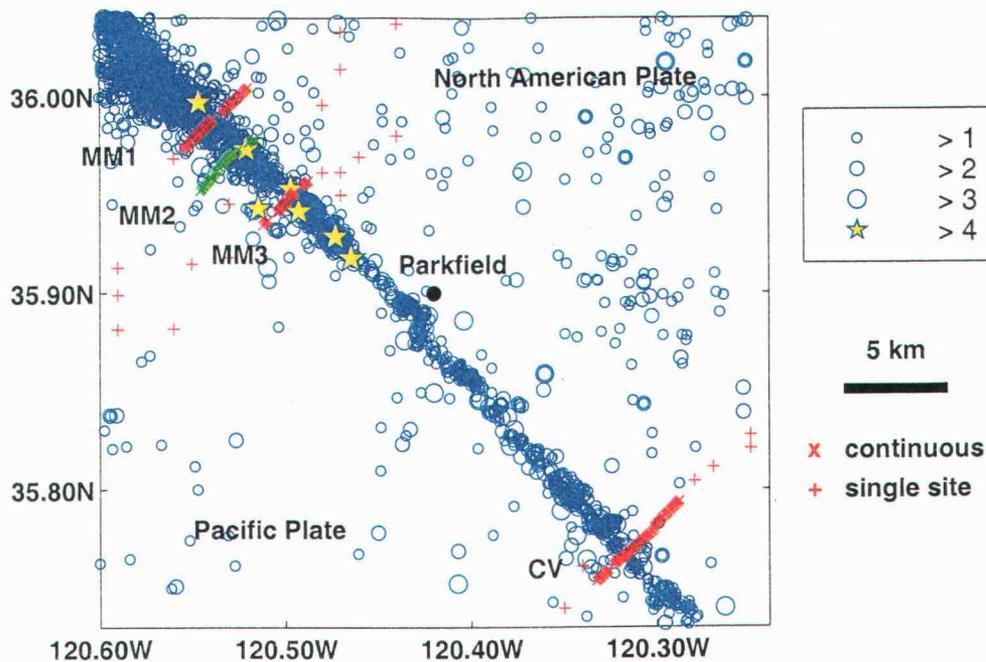


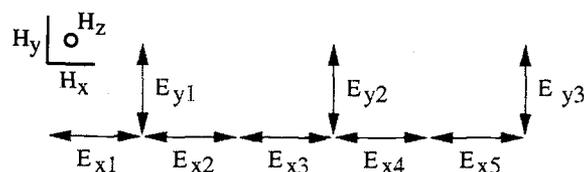
FIGURE 4.3. Map of San Andreas fault in the vicinity of Parkfield. Locations of the high resolution magnetotelluric profiles and single site magnetotelluric are plotted on top of the seismicity (from 1967-1998; Northern California Earthquake Data Center).

al. (1998) and Siripunvaraporn et al. (1998). Another profile, CV, was also carried out in 1997 in the Cholame Valley, about 20 km south of MM3 (Figure 4.3).

4.2.1. Data Acquisition and Data Processing

The data acquisition can be divided into two sections. The dense array profiles (MM1, MM2 and MM3 from north to south in Middle Mountain and CV in the Cholame Valley, Figure 4.3) use electric dipoles deployed continuously across the SAF. The Electromagnetic Instruments (EMI) MT-24 system with 11 channels, three magnetic field components (two horizontal and one vertical) and eight electric field components as shown in Figure 4.4a, were used to collect the data. Each

a) High resolution MT survey (11 channels)



b) Conventional MT single station (5 channels)

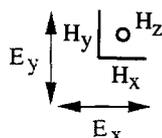


FIGURE 4.4. MT configurations. a) High resolution MT profile. It consists of three magnetic components and eight electric components. The arrow heads indicate the location of the electrodes. The length from one electrode to another is around 100 meters or less. Each span can cover the area roughly around 500 meters. The magnetometers are buried at the beginning or at the end of each span. Two consecutive spans use the same magnetometers. b) Conventional MT single station. It consists of three magnetic components and two electric components. This scheme is used as a remote site.

deployment covering about 500 meters was occupied for almost one day. The same system configuration was laid end to end continuously across the SAF from west to east over a distance of 4-5 km.

The remote sites (marked with “+” on Figure 4.3) were deployed sparsely on both ends of the dense profiles. A conventional MT single station configuration was used (Figure 4.4b), also using the EMI MT-24 system with 5 channels. The remote sites were occupied simultaneously with the profiles and are very useful for eliminating bias in the impedance due to local noise (e.g. Gamble et al., 1979). The data at the remote sites are also used to constrain the structure outside the dense profile. We recorded all data in three bands with sampling rates of 960, 60 and 3.125 Hz.

The dense profiles are not aligned on the same (vertical) plane with the remote sites. We therefore projected the remote sites to align with the dense profiles before we performed the dimensionality analysis and the 2-D inversion, presented in the next sections. This results in the same remote sites outside of the dense profiles for MM1, MM2 and MM3.

The time series data were processed using robust remote reference and multiple-station schemes (Egbert, 1997) to estimate the MT responses (apparent resistivity, phase, and the vertical magnetic field transfer function or tipper) in the period range between 0.01 to around 700 seconds. The multiple-station technique requires simultaneous operation of at least two sites (including the dense profile). The data quality in general was very good, except at long periods for some stations.

The remote reference technique introduced by Gamble et al. (1979), along with the robust scheme of Egbert and Booker (1986) has dramatically reduced the bias from local incoherent noise. However, in Northern and Central California, the DC electric commuter train BART in the San Francisco Bay Areas produces electromagnetic signals covering a very large area (e.g. Egbert et al., 1999). This becomes coherent noise in the MT measurements which can cause the apparent resistivities to steeply increase, and the phases to be negative at periods around 1 to 10 second for the stations in the Hollister area, around 50 to 100 km south of San Francisco (Siripunvaraporn et al., 1993). When MT data are contaminated by coherent noise, the robust remote reference technique becomes ineffective. To remove this coherent noise, Egbert (1997) used a robust multiple-station scheme based on multivariate statistical methods to enhance the signal to noise ratio, and to diagnose possible biases due to coherent noise.

In the Parkfield area, about 200 km south of San Francisco, the effect from BART is subtle compared to the stations in the Hollister area. There is no steep

increase in apparent resistivity, and no negative phase. Because the measurement area is far from the source, the coherent electromagnetic signal generated by BART behaves locally like a plane-wave source, and may actually make the signal to noise ratio higher. Note however that for periods around 10 second, Eisel and Egbert (1999) show that there are subtle biases in impedances at Parkfield due to BART.

4.2.2. Qualitative Interpretation of MM1

The data of MM1 plotted on a log-log scale for apparent resistivity, and phase and tipper are shown in Figure 4.5a and 4.5b. The MT responses are divided into 12 “spans” from west to east (for plotting). The remote sites on both ends of the profile, projected to align with the dense profile, were grouped together as spans 1 and 2 (on the west with 3 stations each) and span 12 (on the east with 4 stations). In the dense profile, two consecutive spans were used with the same magnetometer location, because magnetic fields change more slowly over a distance. The magnetometers are buried at the middle of the two spans, resulting in tipper data (i.e. vertical magnetic transfer functions) at spans 4, 6, 7, 9 and 11. The TM and TE responses are represented with the blue and red symbols respectively, while the real and imaginary parts of the tipper are green and cyan respectively.

The MM1 profile consists of 45 stations (plus 6 remote stations on the west and 4 on the east) for the TM mode, 27 stations (plus 10 remotes) for the TE mode, and 5 stations (plus 10 remotes) for the tipper. All of these parameters are estimated at periods between 0.01-666 seconds, with two responses each (apparent resistivity and phase for TM and TE, real and imaginary for tipper). Pseudo-section plots of the apparent resistivity and phase of TM and TE modes with period on the y-axis and stations on the x-axis (no remote stations) are given in the top rows

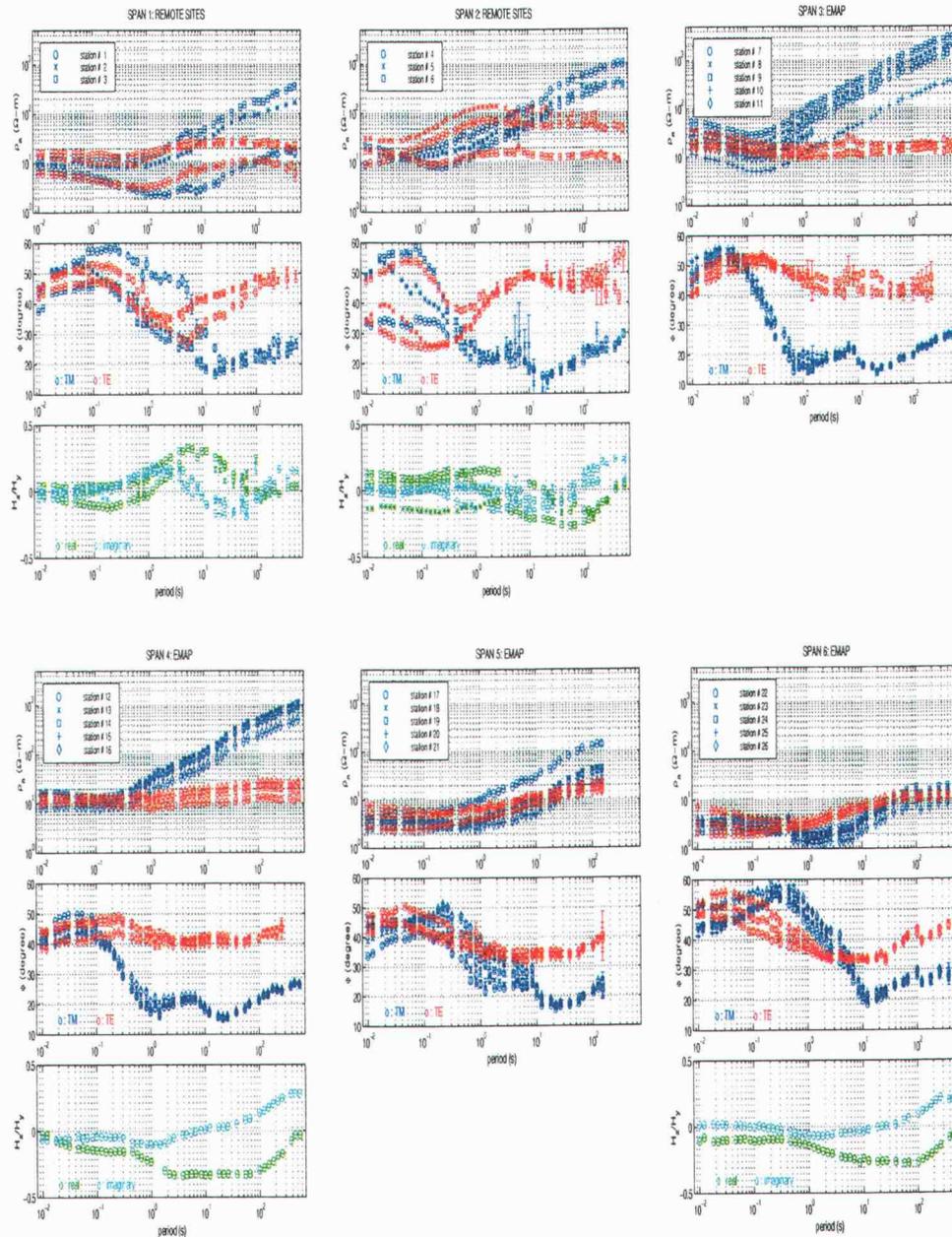


FIGURE 4.5a. The MT responses (apparent resistivities, phases and the real and imaginary part of the tipper) of MM1 from 0.01 second to around 700 second. Spans 1 and 2, and 12 are the remote sites in the west and east respectively of the dense profile. The SAF surface trace is in between 6 and 7. TM is blue, and TE is in red. Real part of tipper is green, imaginary part is cyan.

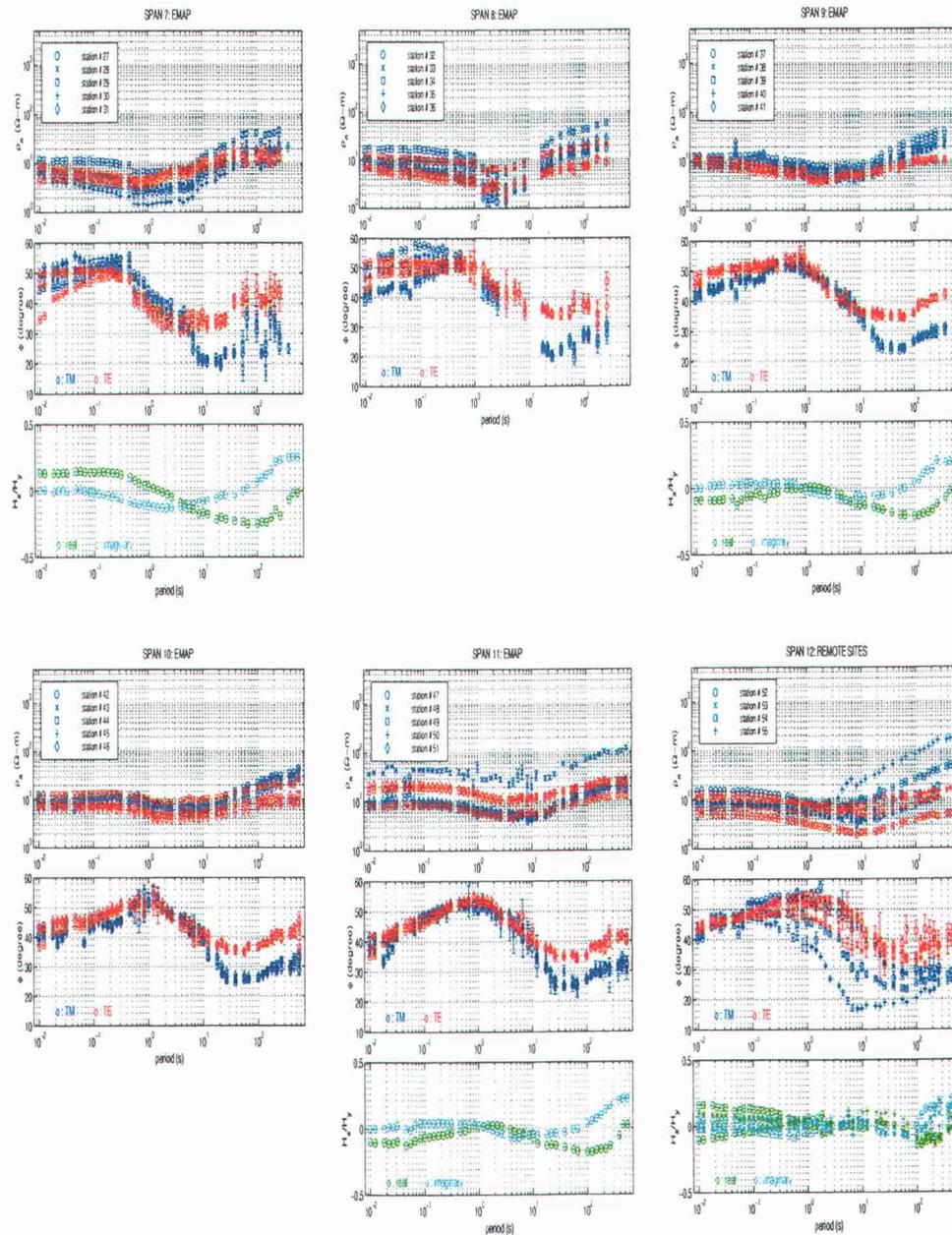


FIGURE 4.5b. Continued.

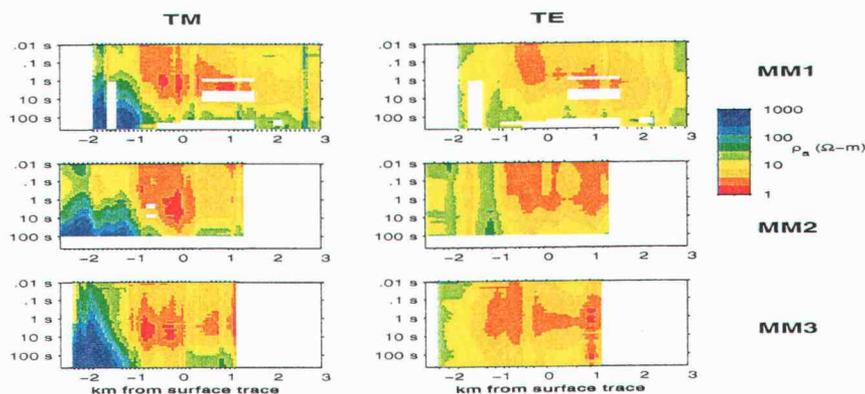


FIGURE 4.6. Pseudo-section plots of the (distorted) apparent resistivity of the dense profiles (no remote sites), MM1 (top row), MM2 (middle), and MM3 (bottom). TM mode is on the left column, and TE mode is on the right. White patches on the plots are the bad data segments. X-axis is the distance from the SAF surface trace in km. Y-axis is the period from short periods on top, corresponding to the shallow penetration depths at those periods, to longer periods with greater penetration depths.

of Figure 4.6 and Figure 4.7 respectively, with TM in the left column and TE in the right column. In MT, long period data are sensitive to deeper structure. As a result, the frequency dependence gives some idea of resistivity variation with depths. This dense profile covers about two (three) km west (east) of the surface fault trace, which is marked at zero distance.

The static distortions caused by the local near surface inhomogeneities or by topographic effects (which might be expected to be severe in this area), will shift log apparent resistivities from station to station by constant values, while the phases are unaffected. The effect is clearly visible at span 3 (also at other spans) of MM1 in Figure 4.5a. Before interpreting the data, the shifts of the apparent resistivities must be corrected, to avoid misinterpretation in terms of deeper structure. The

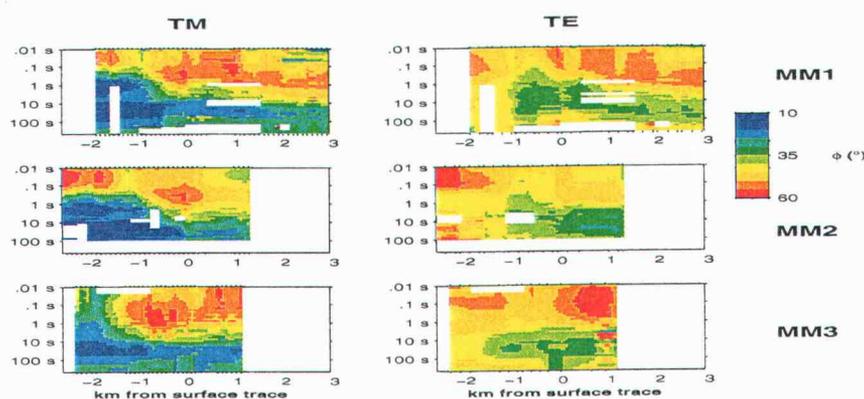


FIGURE 4.7. Same as Figure 4.6, but for phase.

static distortion can be corrected either manually (if there is any prior information to guide the correction), or automatically via the available inversion programs.

For the TM mode, the currents flow across the fault or, in other words, from one side of the profile to the other. The apparent resistivities of the TM mode start flat at short periods on both west and east sides of the SAF. At longer periods (> 1 sec), the apparent resistivities on the west rise steeply, resulting in low phases (e.g., span 3 and 4 of MM1 in Figure 4.5a). This steep increase of the TM apparent resistivity is not seen (and phases are higher) closer to the fault and to the east of the SAF (span 5, 6, 7, 8, 9, 10 and 11 in Figure 4.5a and 4.5b). Note that the surface fault trace is located between spans 6 and 7. The pseudo-section plots give us a clear picture of distinct resistivity changes across the fault. These distinctive responses on the west and east of the SAF tell us that different materials must be present, with generally higher resistivity on the west and lower resistivity on the east. Comparing the resistivity structures with other geological studies, we can qualitatively interpret that the more resistive material is the Salinian granite on the west, and the conductive material is the Franciscan assemblage on the east.

In the TE mode, the currents flow parallel to the fault. The apparent resistivities stay relatively flat from short to long periods in the west, but rise slightly to higher values at long period in the east. The phases are also flat at almost 45° for all periods in the west (consistent with the flat apparent resistivity), but change to higher values at short periods, and lower values at long periods at stations near the fault and to the east. The pseudo-section plots (Figures 4.6 and 4.7) again show clear differences between west and east, especially in the phase. Qualitative interpretation from the TE data suggests to us that an almost 1-D resistivity structure dominates on the west, while resistivity structure on the east is more complicated.

The splitting of the data responses between the TM and the TE modes can be clearly seen throughout the profile, from around 0.1 second on the west (e.g., phase of span 3 of Figure 4.5a) to a bit longer period at around 10 second on the east (e.g., phase of span 10 of Figure 4.5b). One possible explanation for the splitting is the effect of the Pacific Ocean, about 60 km to the west of the measurement area. Because of the high conductance of the ocean, most of the currents are concentrated in the ocean (with phases near zero), even to very long periods. Long period MT measurements made near the ocean will be biased by the extra ocean currents trapped in the upper crust (Ranganayaki and Madden, 1980; Mackie, Bennett and Madden, 1988). Thus, in the TM mode, the currents flowing from the ocean to the continent carry the low ocean phases and result in low phases, and the increase in apparent resistivity with period measured on land. In the TE mode, the currents which flow in the ocean parallel to the strike do not affect land data in the same way. In addition to the ocean, there might be similar effects due to other structure, e.g., the east-west resistivity contrast across the SAF or the conductive Great Valley.

When current flows parallel to the resistivity variations, it will concentrate in the more conductive material. The lateral variations in electric current density lead to vertical magnetic fields which can be detected at the stations along the profiles. The tipper, which is the ratio of the vertical magnetic field over the horizontal magnetic field, is collected roughly every kilometer (not as densely as the TM and TE data). It is actually another part of the TE mode. When currents flow parallel to the strike through the heterogeneous medium, they produce vertical magnetic fields which can be measured at the surface (but not directly above the source). The tipper is thus sensitive to structures that are not directly beneath the measurement location.

The real part of the tipper at spans 4, 6 and 7, west of the SAF trace (Figure 4.5a and Figure 4.5b) is negative at longer periods, but this anomaly is reduced at spans 9 and 11 on the east of the SAF. This is simply explained by a concentration of currents on the east due to higher conductivity in the area. The imaginary part of the tipper is small, except at periods longer than 100 seconds where it becomes positive. The long period behavior is probably the result of structures far away from the measurement sites, such as the Pacific Ocean, or the conductive Great Valley on the east side.

4.2.3. Comparisons of Middle Mountain Profiles

We have qualitatively interpreted the northern profile MM1. In this section, we will describe the similarities and differences of the data from the three Middle Mountain profiles, to get an idea about possible 3-D effects. MM2 and MM3 are located about 2 km and 6 km southeast of the MM1 profile. The pseudo-section plots for these profiles from north to south are given as the top, middle and bottom

rows for MM1, MM2 and MM3, respectively, in Figure 4.6 for (distorted) apparent resistivities and Figure 4.7 for phases. The x-axis is plotted as distance relative to the surface fault traces. Note that the remote sites are not included in these plots.

The MM2 profile, conducted earlier in 1994, southeast of MM1, consists of 34 stations for the TM mode, and 21 stations for the TE mode with periods from 0.01-100 seconds and two responses as well. There is no tipper for the MM2 profile. MM2 covers about three km on the west of the surface trace, and about one km on the east (with a 500 m gap). The MM3 profile consists of 31 stations for the TM mode, 17 stations for the TE mode, and 8 stations for the tipper, with periods from 0.01-700 seconds and two responses.

A direct comparison of the apparent resistivity plots might not be reasonable due to the static distortions from local inhomogeneities. However, the phase is unaffected by the distortions, and therefore more reasonable to use for comparison. Figure 4.8 shows the difference of the phases between profiles. The phases of each profile are interpolated to the same grid locations. Only the area under the overlapping stations between the two profiles should be considered, rather than the non-overlapping area. The yellow color represents the region where the two profiles show similarity, while red and blue represent the regions where they are different.

We clearly see that the MM1 and MM2 profiles, separated by 2 km, are not much different in phase compared to the differences of these with the MM3 profile. The TE phase differences are less significant than the differences of the TM phase. For the TM phases, the two biggest differences between the two northern profiles and MM3 occur around 0.01-0.1 seconds (at -2 km) and around 0.1-10 seconds (at -1 km). These differences are evident in Figure 4.7, where MM1 and MM2 have high phases at short periods (0.01-0.1 seconds) at around -2 km, while MM3 has lower phases. Profile MM3 shows a very broad high phase ($> 50^\circ$) zone starting

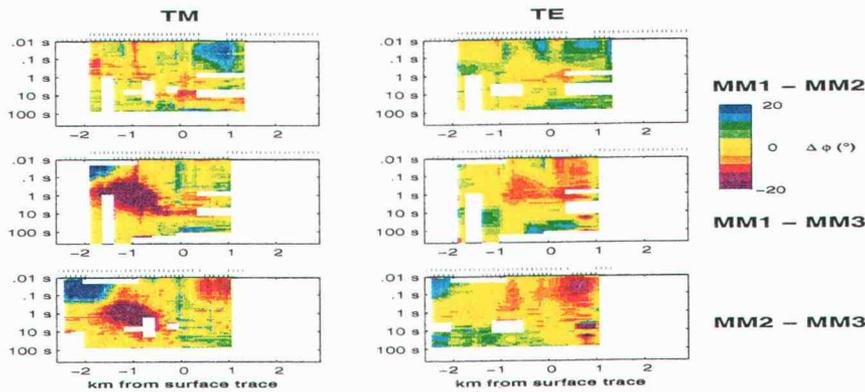


FIGURE 4.8. Plot of phase differences between profiles. The TM mode is on left column, and TE mode is on right column. Station locations are on top of the plots. The yellow color tone is where two profiles share similarity, and blue or pink tones are where they are different.

from around -1.2 km, while for both MM1 and MM2, the high phase zone appears after -1 km to the east. This leads to a big difference of the phase at around -1 km and between 0.1-10 seconds. Qualitatively, this suggests that zone of low resistivity is probably located further to the west from the surface trace for profile MM3.

Due to the uncertainties of mapping the surface fault trace, we tried shifting the MM3 profile within a distance of 1 km to either west or east of the surface fault trace. This does not help to reduce the phase differences between MM3 and the northern profiles. Another region where differences are large is near the eastern end of MM3 (0.01 to 1 seconds). Unfortunately, there is inadequate station coverage to the east. These local differences in phase, within 6 km separation from north to south, have demonstrated to us that this area has some 3-D complications, which is not surprising since the geological structure around this area is so complex.

4.2.4. Qualitative Interpretation of the Cholame Valley Profile

The Cholame Valley (CV) profile is located mostly in the valley where there is a right stepover of about 1 km in the SAF surface trace, from the northwest Parkfield segment to the southeast Cholame segment. The profile covers almost six kilometers and consists of 58 stations for the TM mode, 35 stations for the TE mode, and 6 stations for the tipper, with periods from 0.01-700 seconds and two responses. An additional 7 remote stations were used to extend the dense profile on both sides of the fault. Note that these remotes are not the same as those used for the Middle Mountain profiles.

The data for the CV profile plotted on log-log scale for apparent resistivity, phase and tipper are shown in Figure 4.9a and 4.9b. The surface trace of the SAF Parkfield segment is located around station 26 of span 5, and the surface trace of the Cholame segment is around station 18 of span 4. The pseudo-section plots are given in Figure 4.10. The Cholame segment surface trace is about 0.8 km west of the Parkfield segment surface trace, which is marked at zero distance in Figure 4.10.

Unlike the Middle Mountain profiles, the TM and TE mode responses are relatively constant across the SAF. Apparent resistivities are flat at short periods before rising to higher values, with a corresponding decrease to lower phases. Phases in the west are lower than in the east for the TM mode. Qualitatively, in a large scale, the resistivity structure on both sides of the fault zone consist of a near surface conductive layer overlaying more resistive units. The resistive unit on the west is probably the Salinian granite, similar to at Middle Mountain. To the east, the data indicate that the conductive zone does not extend as deeply as at Middle Mountain. Around -0.5 km to 1 km, in the middle of the profile, the data are obviously distinct

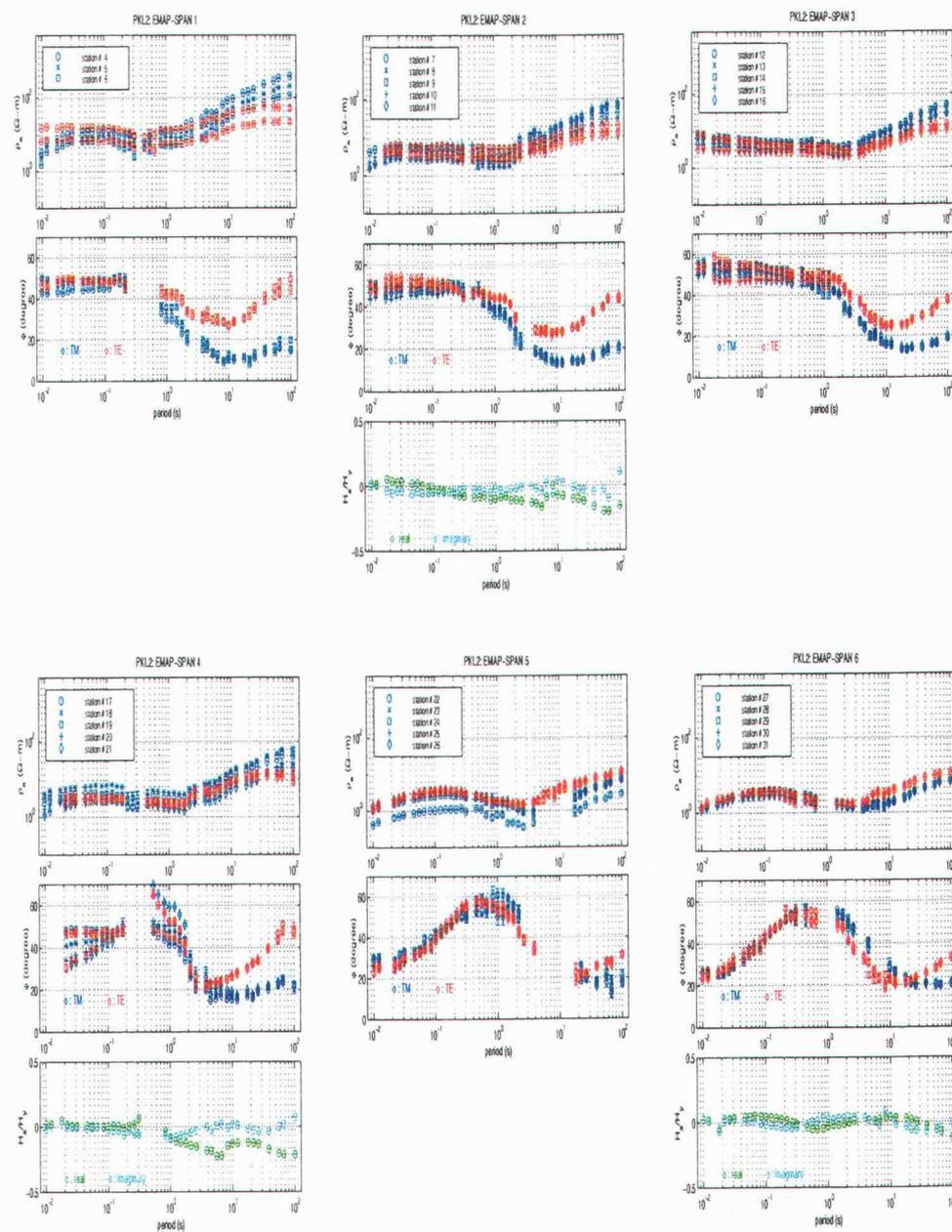


FIGURE 4.9a. Same as 4.5a, but for Cholame Valley profile (no remote sites were plotted). The Parkfield SAF segment is located around station 26 of span5, and the Cholame SAF segment is around station 18 of span 4.

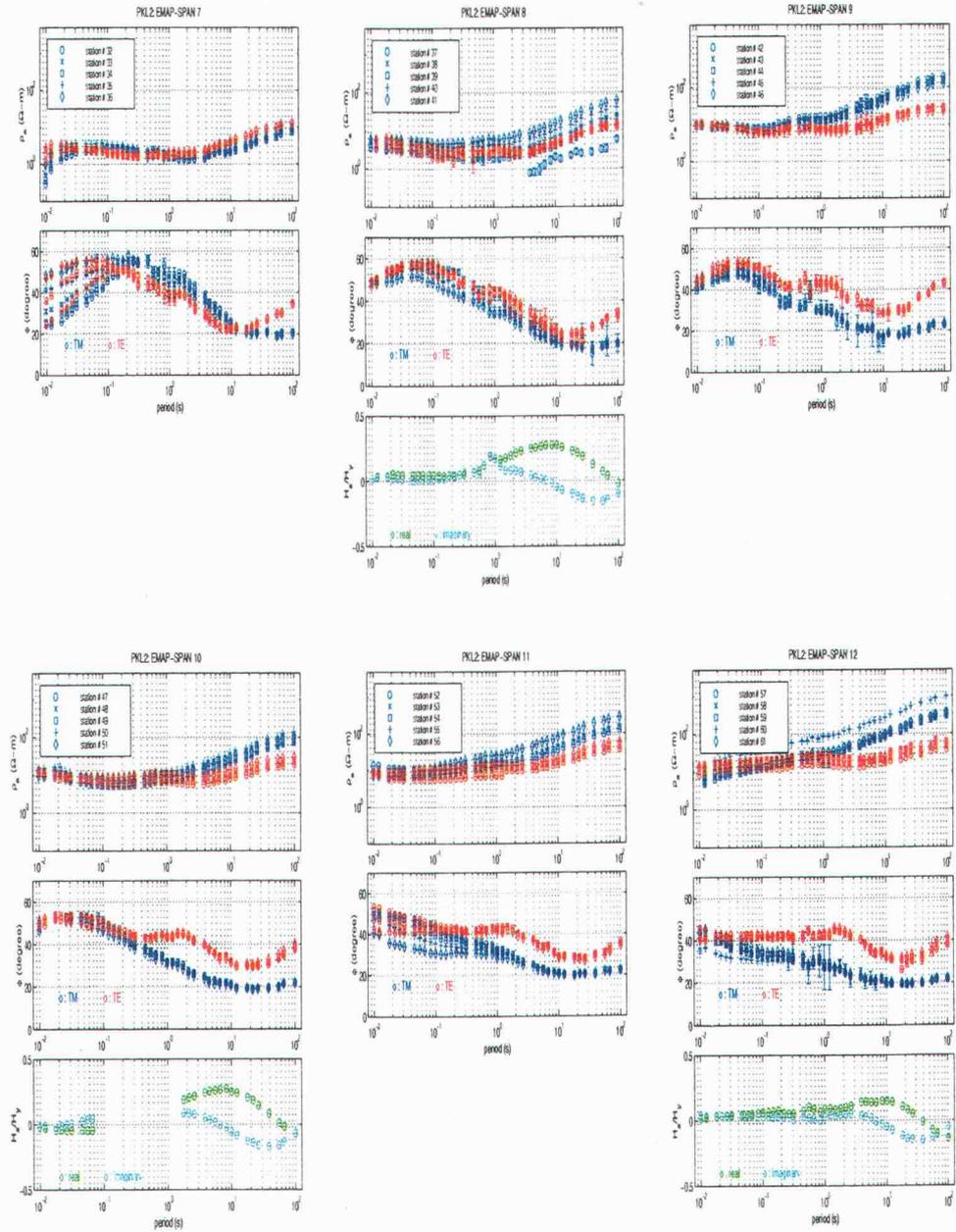


FIGURE 4.9b. Continued.

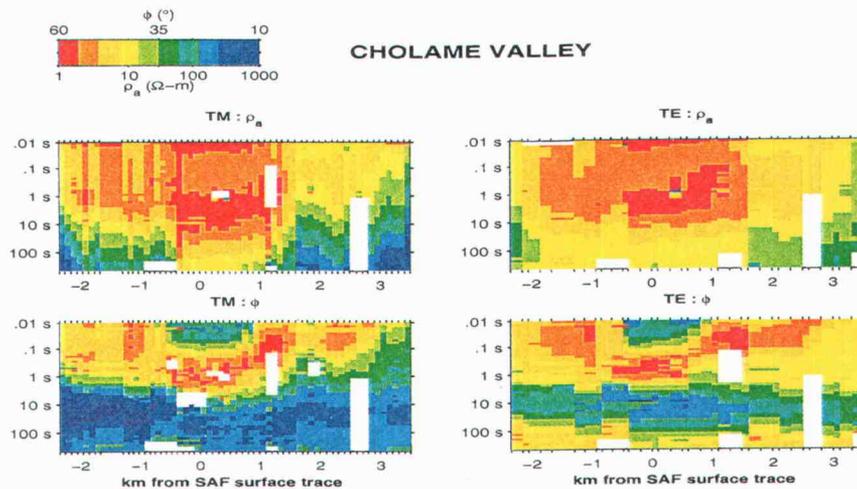


FIGURE 4.10. Pseudo-section plots of the responses of the dense Cholame Valley profile (no remote stations). X-axis is the distance from the SAF surface trace (Parkfield segment) in km. Y-axis is the period from short (top) to long (bottom).

from nearby. The central sites show consistently high phases and low resistivities at around 0.1-10 seconds, which probably indicates a shallow low resistive zone.

Interestingly, negative tipper anomalies (real part) appear at longer periods for the spans west of the Cholame segment fault trace (spans 2 and 4; Figure 4.9a), while positive tipper anomalies appear on the east (span 8, 9 and 12), and the tipper at span 6 (about the fault trace location of Parkfield segment) is almost null. This indicates that there must be a highly conductive structure centered beneath span 6, resulting in positive anomalies to the east, and negative to the west.

4.3. Dimensionality and Distortion Analysis

In MT, the relationship of the electric fields \mathbf{E} to the magnetic fields \mathbf{H} for an arbitrary (possibly 3-D) resistivity structure is defined via the impedance tensor \mathbf{Z} ,

$$\mathbf{E} = \mathbf{ZH}, \quad (4.1)$$

or

$$\begin{bmatrix} E_x \\ E_y \end{bmatrix} = \begin{bmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{bmatrix} \begin{bmatrix} H_x \\ H_y \end{bmatrix}.$$

For 2-D resistivity variations the MT impedance tensor should have the diagonal elements (Z_{xx} and Z_{yy}) of the tensor equal to zero when expressed in a coordinate system consistent with the geoelectric strike.

Practically, even when the large scale structure is 2-D, MT data is usually distorted by local inhomogeneity in the near surface, which can cause static (i.e., frequency independent) shifts (e.g., span 3 in Figure 4.5a) of the apparent resistivities, or a rotation of electric fields. At low enough frequencies the relationship between the undistorted impedance tensor \mathbf{Z}_u and the distorted impedance tensor \mathbf{Z}_d is,

$$\mathbf{Z}_d = \mathbf{DZ}_u, \quad (4.2)$$

where \mathbf{D} is the frequency-independent real distortion matrix (Larsen, 1977). Because of the distortion, the 2-D MT condition is often violated in all coordinate system. We therefore must check if there is a coordinate system where the estimated impedances are consistent with a 2-D interpretation, while allowing for these near surface distortion effects. This analysis can help us to understand the dimensionality of the data.

A classical method of find the correct strike (θ) for a 2-D conductivity compute \mathbf{Z} for many values of θ , and then find the maxima of the mean square modulus of off-diagonal terms in (4.1) which corresponds to a minima of the mean square modulus of the diagonal terms (Vozoff, 1972). However, this procedure will not in general recover the correct geoelectric strike in the presence of static distortion. Smith (1995), following earlier work of Zhang and Pederson (1986), Bahr (1987),

and Groom and Bailey (1989), discusses a decomposition technique that recovers the orientation of the regional strike over a frequency range. They start by replacing the simple (4.2) with

$$\mathbf{Z}_d = \mathbf{U}_\theta \mathbf{D} \mathbf{Z}_u \mathbf{U}_\theta^T, \quad (4.3)$$

where \mathbf{U}_θ is a rotation matrix. To estimate θ and \mathbf{D} the problem becomes minimizing $\chi^2 = \|\mathbf{C}_d^{-1}(\hat{\mathbf{Z}} - \mathbf{Z}_d)\|$, for a sequence of tentative angles, stepping from -45° to 45° relative to the geographical coordinates, where \mathbf{C}_d is the data covariance matrix. The magnitude of residual is then used to evaluate the 2-D assumption. The strike angle (or ranges of strike angles) which minimizes the misfit is selected as the geoelectric strike.

4.3.1. Strike Estimation for Middle Mountain Profiles

Eisel et al. (1999) implemented a variant of the decomposition scheme of Smith (1995) to estimate the strikes for these data sets. For each profile, the distortion analysis is conducted for each station for a series of rotational angles, over a wide period band from 0.1 to 700 seconds. Figure 4.11 show the misfits of all profiles in dB, defined as $10 * \log_{10}[\frac{\chi^2}{\text{mean}(\chi^2)}]$, where the mean is over all rotational angles.

For the Middle Mountain profiles, a clear strike angle at about 40° is observed along these three profiles on the west of the surface trace. This strike angle indeed corresponds well with the strike of the SAF. Closer to the fault and on the east, the strike angles are variable from site to site, from profile to profile. The resistive Salinian granite, with a conductive sedimentary layer on top, dominates the west side of the SAF, while on the east the conductive Franciscan assemblage is the major geological unit. This results in a greater skin depth on the west than on the east.

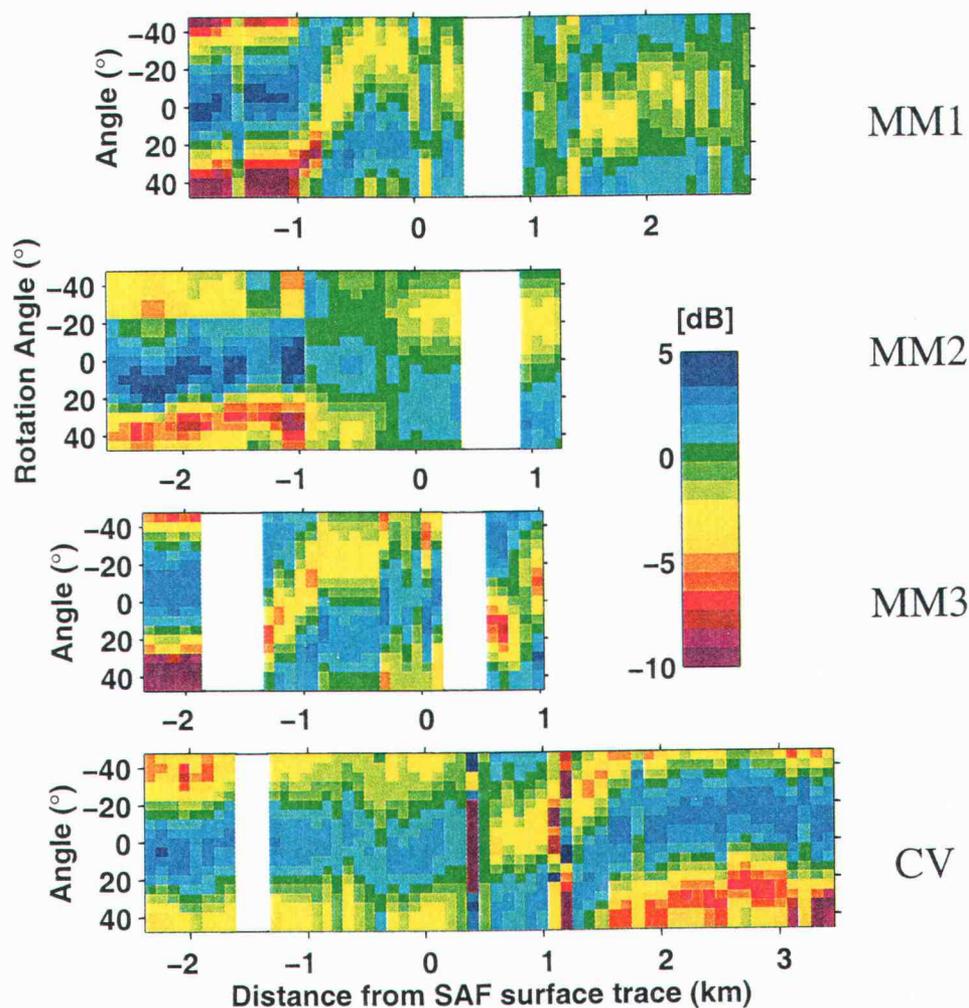


FIGURE 4.11. The misfit plots from distortion analysis (using wide period band) of the series of rotational angles (-45° to 45°) against stations along the profiles from west to east (no remote sites) of MM1, MM2, MM3 and CV profiles. Pink color indicates the best fit from the distortion model, or, basically, the preferred strike angle, while blue color indicates worst fit to the distortion model.

Due to the greater skin depth the contact between the resistive and conductive crust affects the data over a larger distance. On the conductive eastern side, the effect of this major conductivity contrast disappears over a much shorter distance. Instead, it senses the local heterogeneous Franciscan assemblage consisting of many fragments of sedimentary (conductive) and volcanic (resistive) rocks, as well as active folds, thrust faults and strike-slip faults. Thus, it is not surprising to see more variable (and less well-defined) strike angles on the east.

The dimensionality analysis for MM2 is consistent with the results obtained by Unsworth et al. (1999), who concluded that the data on this profile is mostly consistent with the distorted 2-D model. However, there are fewer stations on this profile east of the fault trace (where 3-D complications tend to be greatest) than on the west side.

The inconsistency of the estimated strike angles on these profiles indicate that these data sets at Middle Mountain are not exactly 2-D. To permit valid 2-D modeling, no matter what the strike angles of each station are, it is important that all data should be rotated into one consistent reference frame, with one regional strike for the whole profile. Since the data were collected along profiles perpendicular to the SAF, and since this is the strongly preferred direction on the west side, all data are interpreted in a coordinate system consistent with the SAF strike, or basically in the measurement coordinates.

To see how well the distorted model can fit the data overall, we use the distortion matrix estimated in Figure 4.11 and the undistorted impedance tensors estimated assuming a strike angle of 40° for the whole profile to produce the apparent resistivity and phase responses of both TM and TE modes. A plot of the difference between the actual response estimates and the best fitting distorted 2-D responses are shown in Figure 4.12 for MM1 profile. The TE phase (ϕ_{yx}) fits the

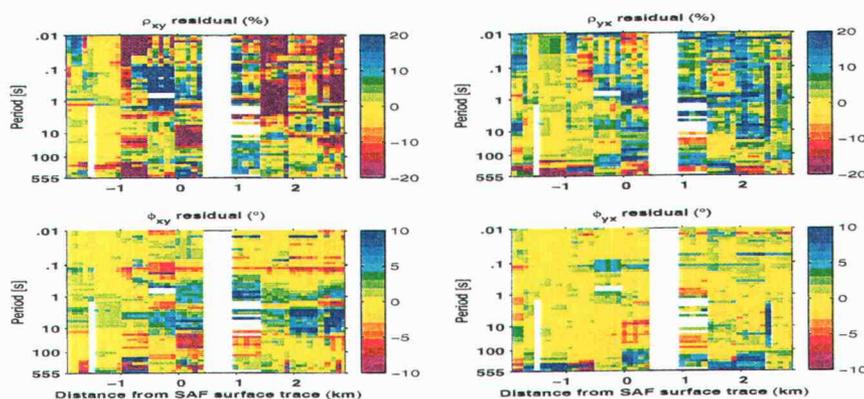


FIGURE 4.12. A plot of the residual (in % for apparent resistivity, and $^{\circ}$ for phases) between the actual response estimates and the best fitting distorted 2-D responses assuming a strike angle of 40° . XY is TM mode, and YX is TE mode. Yellow color indicates the area of the best fit, while either blue tone or red tone indicate the regions of poor fit.

distorted 2-D best in this coordinated system, while the other responses do not fit as well. Fits are also clearly better west of -1 km. The worst fit region is on the east and near the fault, consistent with the result of Figure 4.11. If we chose a different regional strike the pattern of misfit would be different. For example, if a strike angle of around $0-20^{\circ}$ were used, the area of good and poor fit might be reversed. Note that since there is no clear strike in the east, the overall fit here would still probably be worse.

The residual plot (Figure 4.12) suggests a “lower bound” on how well we can really fit the data in one fixed coordinate system, to allow for 3-D effects. Therefore, in the area east of the fault where 3-D complications appear to be more severe, we should not expect to fit the data perfectly, especially in the area where the best fitting strike is not consistent with the coordinate system used for interpretation. Note that the residuals plotted in Figure 4.12 are computed from the simple distorted model of (4.3), which is not necessarily geologically interpretable, or even physically

realizable. A physically realizable 2-D model might result in significantly larger residuals.

4.3.2. Strike Estimation for the Cholame Valley Profiles

At Cholame Valley, The result of the distortion analysis is somewhat different. Figure 4.11 shows that the best fitting strike angles for CV vary rapidly in the fault zone region (0 to 1 km), while a pretty clear strike of around 40° is found outside of the two surface fault traces.

On a very large scale, the geology in the Cholame Valley is not obviously different from Middle Mountain. It is still dominated by Salinian granite on the west and Franciscan melange on the east. It is therefore not surprising that the preferred strike angle on the west side of CV is the same as that of MM1. However, the story is different on the east. As discussed earlier, the CV data suggest to us that the conductive zone east of the SAF at the Cholame Valley might not be as deep or as conductive as seen at Middle Mountain. This would results in a greater penetration depth for EM field, and therefore sensitive to distant structures. Between 0 and 1 km, the variable strike might be explained by the many small faults located in between the traces of the main fault, as observed in the high-resolution reflection profile (Shedlock et al., 1990).

In summary, we find that the data from our four profiles are not exactly 2-D. The data sets from Middle Mountain (MM1, MM2 and MM3) show a best fitting and strongly preferred geoelectric strike angle consistent with the SAF strike on the west, with variable and more poorly defined strike angles east of the SAF, and near the fault zone. The data from the Cholame Valley (CV) show the same clear strike angle (consistent with the strike of the SAF) both west and east, but not in the fault

zone area (at 0 to 1 km distance). Since these MT data sets are not exactly 2-D, the question arises as to whether we can interpret the data sets with 2-D resistivity models. As pointed out earlier, in an area with 3-D complications, we should not expect to have 2-D models that fit the data perfectly. In the next section, we use what we have learned from the dimensionality analysis (and the residual plots of Figure 4.12) to search for 2-D models that fit the data reasonably.

4.4. Inversion Results

Because of the non-uniqueness of the inverse problem, an infinite number of models can produce the same misfit. Most modern MT inversion schemes (OCCAM of Constable et al., 1987; RRI of Smith and Booker, 1991; and NLCG of Rodi and Mackie, 1999) resolve this non-uniqueness by seeking models that have minimum possible structure subject to a desired level of misfit. This makes the inversion stable, with resulting models less likely to contain spurious features. Here, we use the REBOCC inversion (Siripunvaraporn and Egbert, 1999; Chapter 3) which is a variant of OCCAM inversion to convert the data responses (apparent resistivity, phase, and real and imaginary part of the vertical magnetic transfer function) into smooth electrical resistivity models. Generally, most reliable and efficient inversion methods require extensive computational time and significant memory. REBOCC overcomes these difficulties by using the fact that MT data is smooth and redundant. The solution can be constructed from subsets of natural basis function, or representers, without significantly loss of detail, and still fit all of the data well enough. This results in significantly less computational time and lower memory requirements for inverting large data sets, compared to other inversion methods (Siripunvaraporn and Egbert, 1999; Chapter 3).

Static distortions of the apparent resistivities are incorporated as additional model parameters, which are automatically estimated by REBOCC. The “6th-stripe” (see Siripunvaraporn and Egbert, 1999; Chapter 3) pattern was used in the REBOCC inversion, (i.e. only 5 out of 25 periods) for inversion of the MM1, MM3 and CV data, while for MM2 we used “4th-stripe” pattern. The model covariance controls the smoothness of modeled resistivity variations relative to the base model. In most of the inversions discussed here, the REBOCC default was used, that is in the near surface the horizontal smoothing length scale (equal to the station spacing) is larger than the vertical smoothing parameter (equal to the depth). At great depth, smoothing length scales in both directions are the same, resulting in a 1:1 smoothing ratio. This choice of correlation length scales coincides with the loss of resolving power of the data at depth, and outside of the profiles.

A vertical grid discretization of 63 layers (plus 10 air layers for the TE mode) is used for all profiles, while a horizontal grid discretization of 210 columns are used for MM1, 198 columns for both MM2 and MM3, and 224 columns for CV. All models are extended both west and east of the profile ends to allow for the effect of more distant features, such as the Pacific ocean about 60 km to the west. Different starting models for the inversion were tried, including one with a fixed ocean. However, the results depended little on the starting models. Therefore, most of our inversion presented in this section used a 10 Ω -m half space as the starting model.

Even though our data sets are not exactly 2-D due to the different geoelectric strikes at various station locations, for several reasons we feel that a 2-D interpretation in the measurement coordinate system is possible, if proper caution is used to avoid overinterpretation. One of the main reasons is that the large scale geology shows a clear 2-D variation across the SAF. Also from the dimensionality analysis,

we knew which parts of the data are consistent with the regional strike (i.e. the SAF strike), and which portions are not.

Because of the 3-D complications, before inverting the data we assume a noise floor of 2 %. An additional increase in error bars is used for data at a particular station and period, if this data seems to be biased or inconsistent with neighboring data. We learned from the dimensionality analysis that if one regional strike is used, the fit to the data where the deviation of the best fitting strike from the SAF strike occurs will likely be poor (Figure 4.12). To avoid overfitting these areas of 3-D complication, these data should have less influence on the inverse solution than data which is more consistent with the assumed regional strike. To accomplish this, error bars for data east of and near the fault for MM1, MM2 and MM3 (and only near the fault for CV) are significantly increased.

We start this section with the results from single mode inversions to explore what structures are really required by the data. Then results of the joint inversions (using data from both modes) are presented, with comparison of the best fitting model and the preferred model. We focus on inversion of data from the MM1 profile, because of its quality and quantity. Then, preferred models for the MM2 and MM3 profiles will be given. The same process is then repeated for the data from the CV profile.

4.4.1. Single Mode Inversions (MM1)

In a case study of a conductive fault in a resistive background, Eberhart-Phillips et al. (1995) found that TM data is not effective at resolving the conductivity and depth extent of a vertical fault, though the boundaries of a vertical feature are well-resolved. This is because of the electrostatic effects due to the currents flow-

ing perpendicular to the fault structure. In the TE mode the currents flow parallel to the fault strike. This mode provides a stronger constraint on fault conductivity and depth extent, due to current gathering in the conductive fault zone. However, in the TE mode lateral resolution is poorer. For example, a wide conductive zone can be replaced with several narrow conductive zones to fit the TE data. This problem might be solved with TM data if the measurement locations are closely spaced allowing resolution of the the individual conductive zones.

The study of Eberhart-Phillips et al. (1995) demonstrates that the two modes are sensitive to different sorts of resistivity variations. A joint interpretation of all modes can help resolve ambiguities which might not be resolved by either mode. Note that the case study of Eberhart-Phillips et al. (1995) was limited to a conductive fault zone in a resistive environment. A different environment, as at Parkfield where the fault is located between resistive and conductive sides, might be somewhat different. The electromagnetic fields attenuate rapidly in the conductive structure, which probably makes it harder to resolve the fault zone conductivity and shape.

We start by inverting single mode data, TM, TE and also tipper in order to see which features are required by each mode separately. Figures 4.13 give the models (left column), the calculated responses, and the observed responses (right columns) from the three cases. The fit to the observed responses is good for all modes, with a normalized root mean square (RMS) misfit of about 3.0 (to within 3. standard deviations) for the TM mode, 2.3 RMS for the TE mode, and 5.2 RMS for the tipper. There is some difficulty in fitting the observed responses at high frequencies (short periods) which probably reflects near surface 3-D structure, and also at long periods, which are sensitive to deeper and more distant structure.

The TM mode model (Figure 4.13a) consists of a moderate resistive area west of the SAF, covered with a thin conductive surface layer, and a conductive

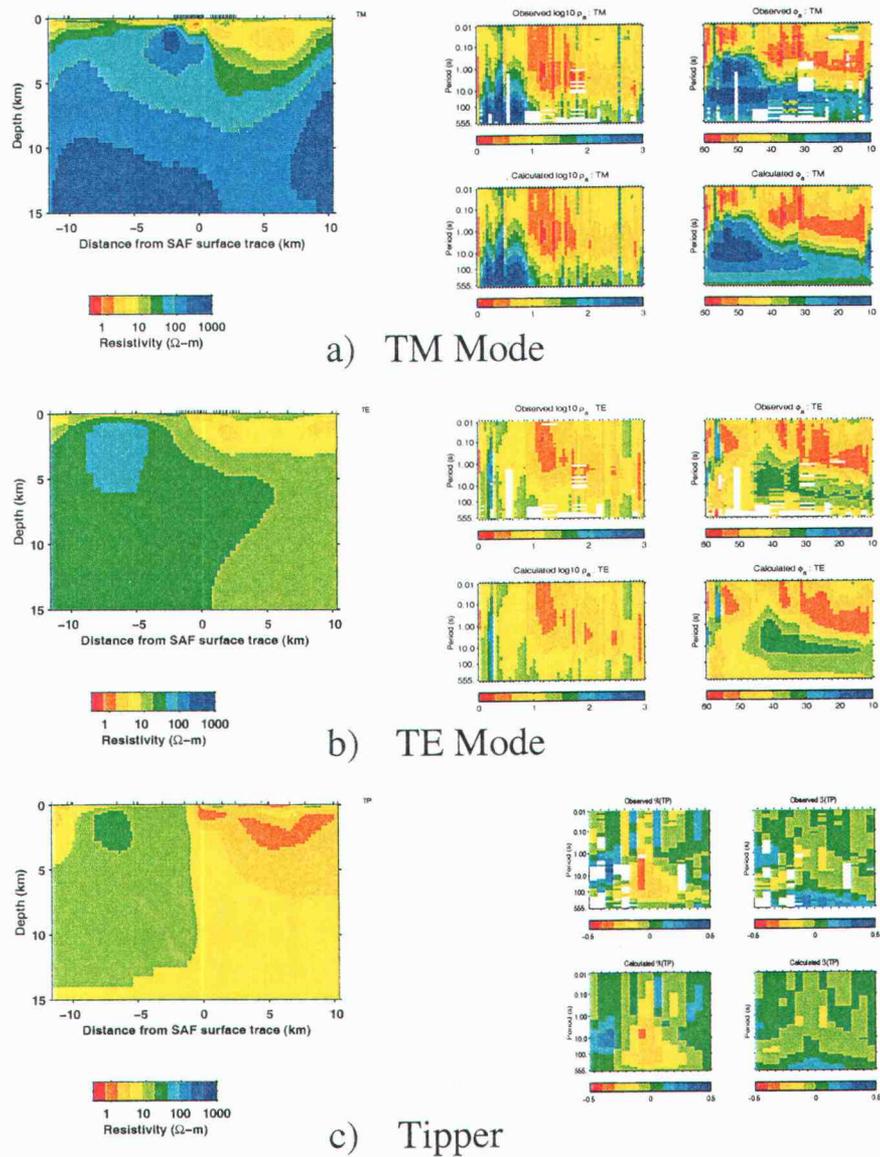


FIGURE 4.13. Resistivity cross-section from single mode inversions. a) Left: Resistivity cross-section from inversion of the TM mode data. The stations are shown on top of the model. Right: The pseudo-section plot of the observed and calculated ρ_a and ϕ . The stations are laid out linearly (not a function of distance) from west to east. Therefore, the first six stations and the last four stations are the remote sites. The overall RMS for this model is at 3.00. b) Same as a) but from the TE mode inversion. The overall RMS is at 2.33. c) Same as b) but from the tipper inversion. The responses are the real and the imaginary part, instead of the ρ_a and ϕ as of TM and TE. The overall RMS is at 5.15.

zone (about $1-10 \Omega - m$) in the upper few kilometers east side of the fault. A conductor of approximately 1-km width is detected from the surface to a depth of about 1 km directly beneath the surface fault trace (at zero distance). Below 5-km depth, the crust is resistive on both sides of the fault.

The single mode inversions of the TE (Figure 4.13b) and tipper (Figure 4.13c) data also suggest a resistivity contrast across the fault, with more resistive crust to the west, and more conductive crust to the east, especially in the upper few kilometers. This resistivity contrast is expected as discussed earlier in our qualitative interpretation of the data. The TE model does not clearly separate the fault zone conductor from the conductive crust on the eastern side of the SAF. Note that the absolute resistivity values from these three models differ significantly, particularly at great depth. This could be the result of static distortions, or 3-D effects which could wrongly imply low resistivity at depth if interpreted with 2-D modeling (Wannamaker et al, 1984).

The tipper is most sensitive to resistivity variations of nearby lateral variations in resistivity (not directly beneath the measurement location), and should not be expected to constrain the absolute resistivity value. However, it can contribute significantly to estimate the location and size of anomalously conductive features. The lateral resolution of the tipper in our profile is reduced due to the smaller number of stations compared to either the TM or the TE resistivities and phases. The fit to the tipper is significantly improved if we include the ocean in the the starting model. The model from the tipper inversion suggests that the resistive unit on the west is at least 12 km deep. Below that depth no resistivity difference between east and west is required. However, the long period tipper data is generally biased due to the 3-D structures far away from the profile, such as from the irregularity of the ocean-continent boundaries, and probably should not be taken too seriously.

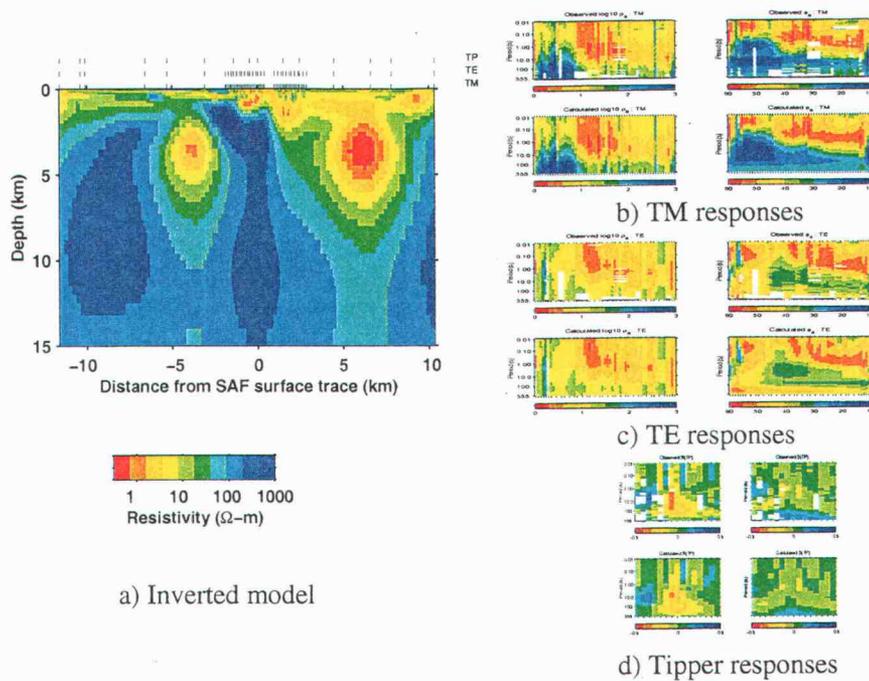


FIGURE 4.14. The same plots as in Fig 4.13, but from joint inversion of TM and TE modes with tipper. The overall RMS is at 3.73. The station locations of each mode are shown on top of the model. Note an extra conductor appears on the west side of the profile.

4.4.2. Joint Inversion (MM1): Best Fitting Models

The result of joint inversion of the TM, the TE apparent resistivities and phases and the tipper is given in Figure 4.14 along with the observed and calculated responses. All data were fit quite well, with an overall RMS of 3.7. Similar to the single mode inversions, the difficulties in fitting the data are greatest at the shortest (local near surface structure) and longest periods (deeper and more distant structure). Interestingly, the model from the joint inversion (Figure 4.14a) is quite different from the models generated from any of the single mode inversions.

The most striking difference is located west of the SAF where there is a huge conductor (about $2\text{-}3 \Omega\text{-m}$ or $0.3\text{-}0.5 \text{ S/m}$) embedded in the resistive structure

about 5 km west of the surface fault trace and 2-5 km depth. To the east, the single mode models all have conductive crust. However, in this case, the conductivity in this area is extremely high (about 3 S/m or $0.3 \Omega - m$, which is as conductive as the ocean) at depths of around 3-5 km. Neither the extra conductor on the west nor the extreme conductivity east of the SAF are features required by the single mode inversions. So, what are these? Are these features consistent with the results from other geophysical experiments or, are they just artifacts generated from the inversion? Note that Unsworth et al. (1997; 1999) interpreted the data from profile MM2. Their models are limited to the area of the dense profile, and shallow depths, outside of the area where these questionable structures occur.

The extra conductor 5-km west of the SAF is not consistent with the prevalent geological interpretation of the area west of the SAF as Salinian granite (Sims, 1990). The interpretation of the COCORP seismic reflection profile in this region by McBride and Brown (1986) did not reveal any structures with geometry similar to this conductive feature. Seismic tomography inversions show uniform high velocities (around 5-6 km/s) in this area (Eberhart-Phillips and Michael, 1993). Also the resistivity model based on widely spaced MT sites by W.D. Stanley shows a uniform resistivity of more than $300 \Omega - m$ in this area (Eberhart-Phillips and Michael, 1993). Because there is no other geophysical or geological evidence (e.g., gravity and magnetic studies of Griscom and Jachens, 1990) to support the anomalous structure, we believe that this extra conductive feature is geologically unrealistic, and probably result from overfitting the data. The extremely conductivity east of the SAF is also unrealistic, and probably has a similar cause.

From this evidence, we therefore conclude that this model generated from the joint inversion of TM, TE and tipper is geologically questionable. However, this model fits the full data set best, so we refer to the model of Figure 4.14 as the

best fitting model, which we abbreviate as BF. By allowing a larger RMS misfit to the data a more reasonable model can be obtained. At the second iteration of the inversion where the RMS misfit is at about 5.1, (Model BF requires 4 iterations) the unrealistic features discussed above are minimal. This model will later be referred to as model BF2. The conductivity of the extra conductor to the west is significantly decreased to around 0.1-0.03 S/m ($10-30 \Omega - m$), while to the east the conductivity is also sharply reduced to around 1 S/m ($1 \Omega - m$).

We have shown clear evidence for 3-D effects in the MT data. Fitting this data with 2-D models might result in erroneous and uninterpretable models, such as model BF. The dimensionality analysis from the previous section suggests that there is some variation in strike angles across the profile. It also shows that the data fits a distorted 2-D model better in the area where the estimated strikes are consistent with the geologically inferred regional strike of the SAF, and more poorly in the other regions. We therefore consider adjusting the fit to the data as suggested by the dimensionality analysis.

4.4.3. Joint Inversion (MM1): Preferred Models

In the inverse problem, the solution does not depend only on the data but also on the assumed data error variances. Data with large error variances will have less influence on the solution than data with smaller error variances. We thus increase data variances in the area where the fit to the distorted 2-D model is poor.

This approach is then applied to the joint inversion using all of the data. The result (Figure 4.15), which will be referred as the preferred model, is dramatically changed from the best fitting model of Figure 4.14. The overall RMS for this model is about 5.6 (calculated with the same data variances as for the best fitting model),

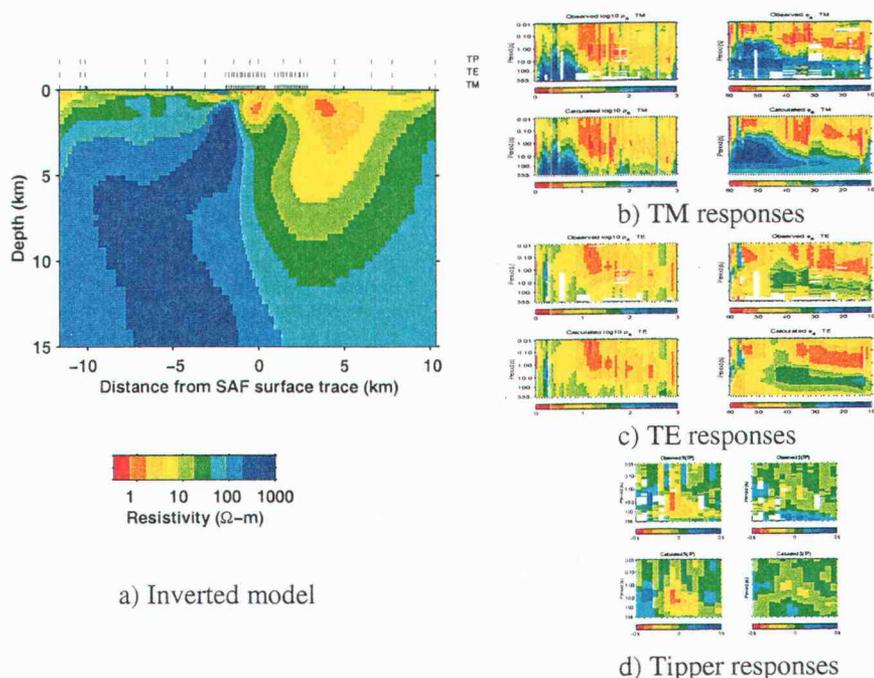


FIGURE 4.15. The same plots as in Fig 4.14, but inverting with increasing the data variance on the area the estimated strike is not consistent with the SAF strike. The overall RMS is at 5.65, relative to the level before adjusting the variance.

which is about the same level as of model BF2. The difference is that this model fits the data better west of the SAF, and more poorly to the east and in the fault zone area, while the other model fits the data more uniformly across the profile.

Comparing the preferred model to the best fitting model, the extra conductor west of SAF surface trace disappears, and the conductivity east of the SAF is significantly reduced. The plot of conductance (i.e., the vertically integrated conductivity) from both models is shown in Figure 4.16, along with the conductance of the model BF2. The crustal conductance of the preferred model is flat across the Salinian block west of the surface trace. A very small peak in conductance is still seen with model BF2, and a very large peak is seen for the best fitting model. To the east, the conductance of the preferred model is significantly reduced by about 50

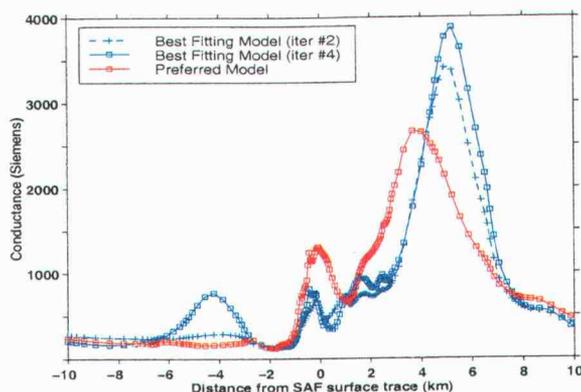


FIGURE 4.16. The plot of the conductance which is the conductivity vertically integrated over the depth of 10 km, from (blue) the best fitting model (Figure 4.14a), and from (red) the preferred model (Figure 4.15a).

%, and shifted a bit to the west, while the conductance in the fault zone is increased by a factor of two relative to the best fitting models. One interesting point from the conductance plot is the sharp increase of conductance beginning about 1 km west of the surface trace.

The preferred model consists of three main resistivity structures: the resistive basement west of the SAF, the deep zone of high conductivity to the east, and a separate zone of high conductivity directly beneath the surface trace. We interpret the resistive structure west of the SAF as the Salinian granite, similar to earlier MT interpretations of Unsworth et al. (1997;1998;1999), and W. D. Stanley published in Eberhart-Phillips and Michael (1993). In the preferred model (Figure 4.15), the resistivity increases from about $100 \Omega - m$ at 2-3 km depth to more than $300 \Omega - m$. The resistive structure is overlain by a thin (1-2 km) conductive sedimentary layer of about $7-30 \Omega - m$. The conductive sedimentary rock appears thicker at the western end of the profile, and thinner toward the fault zone, where more resistive rock comes closer to the surface ($y = -2$ km).

Another prominent feature in Figure 4.15 is the large volume of low resistivity material on the east side of SAF, which geologically corresponds to the Franciscan melange. This conductive zone appears to extend to almost 12 km depth, in contrast to the results from the single mode inversions where it is limited to the upper 5 km. Resistivity increases with depth from less than $10 \Omega - m$ above 6 km depth to about $100 \Omega - m$ at around 12 km depth. Conductivity is highest (0.25-1. S/m; $1-4 \Omega - m$), about 3-6 km east of the surface trace (outside of the dense profile) at about 3 km depth.

Beneath the surface trace of the SAF is a 1 km wide and 2 km deep fault zone conductor. This feature is not symmetric around the surface trace, but extends more to the west. Its average conductivity is around $3 \Omega - m$. Beneath the fault zone is a sharp conductivity gradient across the fault, nearly vertical to 8 km depth before curving to the east. We believe that the bending of the conductivity contrast is probably the effect of the smoothing of the inversion, along with lack of the long period data (> 1000 second) to constrain the structure. Below roughly 12-km depth, lateral variations of resistivity are small.

4.4.4. Joint Inversion Results for MM2 and MM3

A similar approach of increasing data variances to correspond with the results from the dimensionality analysis (Figure 4.11) was applied to invert for resistivity on the other profiles at Middle Mountain, MM2 and MM3. Results are presented in Figure 4.17 from north to south, along with the preferred MM1 model. The separation between the northern (MM1) and southern (MM3) profile is 6 km. Note that all profiles share the same remote sites, projected to align with the dense profiles. However, it is not obvious that the structure beneath the remote sites for

these three profiles has to be the same. The data from the dense profiles also have an impact on the structure beneath the remote sites.

Mapped surface fault traces (Simm, 1990) vary within one km between profiles (relative to the western-most remote site). All of these models (Figure 4.17) display the same three prominent features; the resistive structure west of the SAF, the conductive crust to the east, and the fault zone conductor, with a strong resistivity contrast directly beneath the surface fault trace. However, there are some variations from north to south.

Comparing the MM2 model (Figure 4.17b) with the MM1 model (Figure 4.17a), both fault zone conductors are located a bit to the west of the surface trace, and have about the same width and depth. However, for MM2, the separation of the fault zone conductor from the conductive crust to the east is less clear. The conductive zone in the east also appears to extend a bit deeper, with the western edge staying vertical as deep as 10 km. These small variations between MM1 and MM2 might be the result of inadequate station coverage east of the surface trace on line MM2, or be due to the lack of long period data (> 100 second). At the same time, MM2 extends about 700 meters (one full span) further west than MM1. This gives a clearer view of the near surface structures west of the SAF.

For the MM3 profile, about 4 km south of MM2, the fault zone appears to be more conductive, and is located entirely west of the surface trace. The resistive crust in the near surface just west of the fault core at MM1 and MM2 seems to be shifted further to the west. Note that for MM3, there is one less 5-dipole measurement span west of the SAF due to the difficulty of the topography. Similar to MM2, this profile lacks station coverage east of the fault. This could be a factor in some of the variations between profiles.

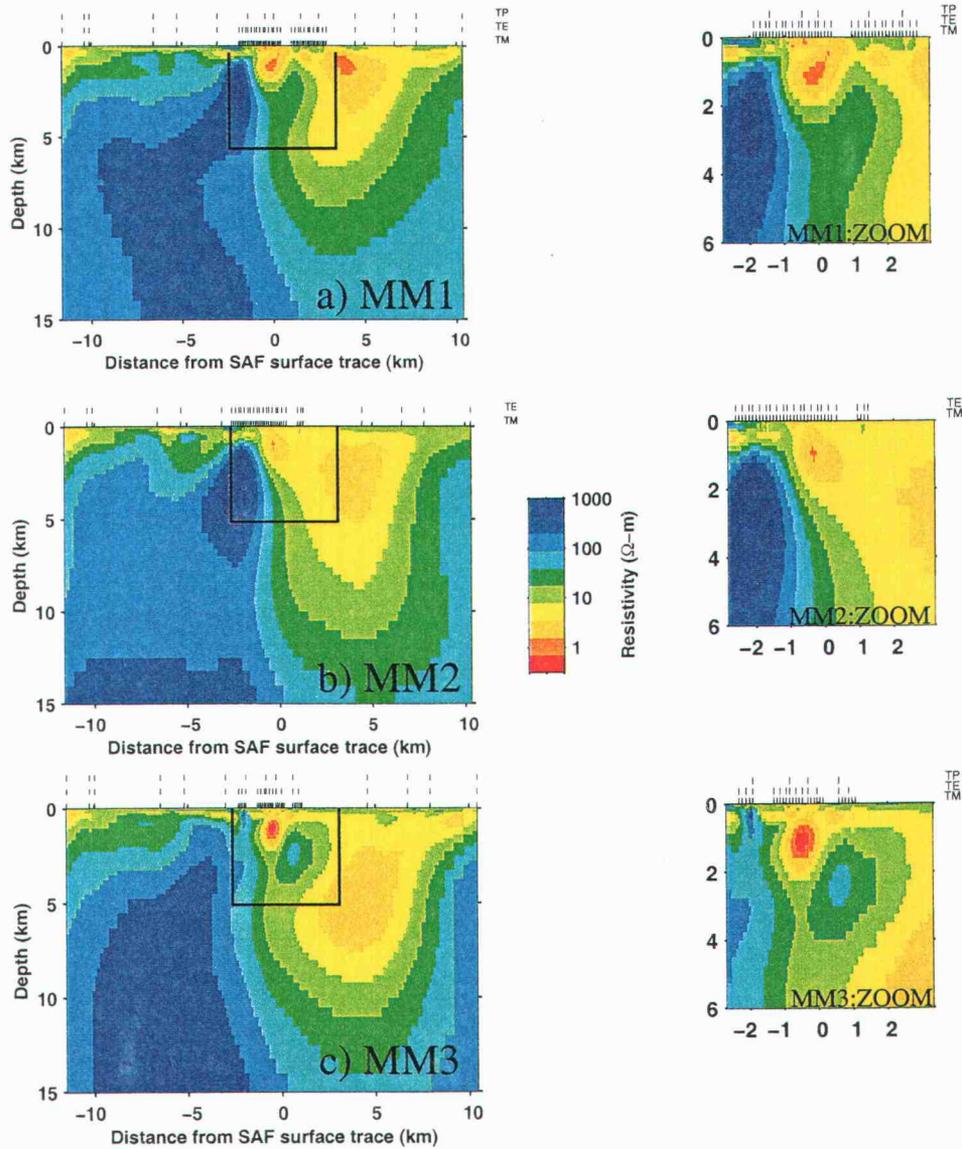


FIGURE 4.17. Comparison of preferred models from three profiles, MM1, MM2, MM3, from north to south in the Middle Mountain area. The pictures on the right column are the zoom parts in the near fault zone area of the models on the left, in the square block. All of these models are inverted from increasing data variance to correspond with the results from the dimensionality analysis (Figure 4.11).

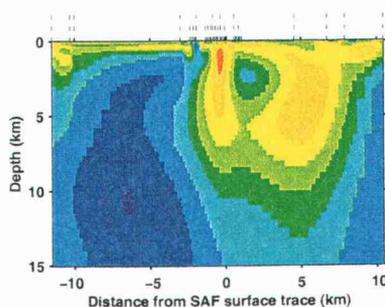


FIGURE 4.18. An alternative model for MM3, referred to as MM3-B, whereas the model in Figure 4.17c is MM3-A.

Interestingly, we found that the very high conductivity in the fault core in the near surface of Figure 4.17c can be distributed as less conductive material extending to greater depth (5-6 km), Figure 4.18. This model is referred to as MM3-B, whereas the preferred model is MM3-A. MM3-B is obtained from the inversion in the same way as the preferred model, except that the data at remote sites 4 and 5 are omitted, and the error floor is increased from 2 to 3 %. With the same error bars (and omitting remote station 4 and 5 in the calculation of misfit for the preferred model), the fit of MM3-B is actually slightly better, mainly due to the TE mode. These two models are a good example of the non-uniqueness of solution to the MT inverse problem.

Figure 4.19 is a plot of conductances for all models, including MM3-B. The conductances of both MM1 and MM2 are almost identical for the whole profile, except just east of the fault (at $y = 1$ km) where there is a difference of around 800 Siemens due to the resistive area between the fault zone conductor in MM1 and the conductive crust to the east (this feature could well be an artifact of the inversion). To the east, the conductances of both models are about the same, peaking at around 3700 Siemens 4 km east of the SAF. The conductance of MM3-A

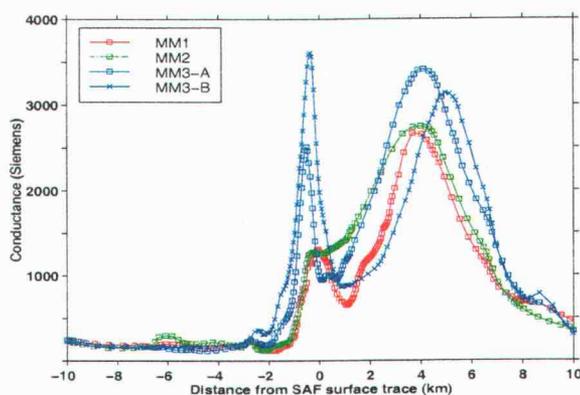


FIGURE 4.19. Comparison of conductances from three profiles, MM1 (red), MM2 (green), MM3 (blues) from north to south at the Middle Mountain area.

and MM3-B are more significantly different from MM1 and MM2, particularly near the fault trace and east of the SAF. The conductance of the fault zone is about 2-3 times higher for MM3. The peaks of conductance at the fault are also shifted slightly to the west (less than 1 km) from the peaks of MM1 and MM2. East of the SAF, conductance of MM3-A is also somewhat larger than for the other two profiles. Some of these differences might be the result of static shifts which have not been corrected adequately.

4.4.5. Inversion Results for the Cholame Valley

Similar methods were used for inverting the Cholame Valley data. The dimensionality analysis (Figure 4.11d) suggests that the data error variance in the area close to the fault zone should be increased for the model search. The resistivity cross-section from joint inversion of TM, TE and tipper is shown in Figure 4.20a, along with the calculated and observed responses.

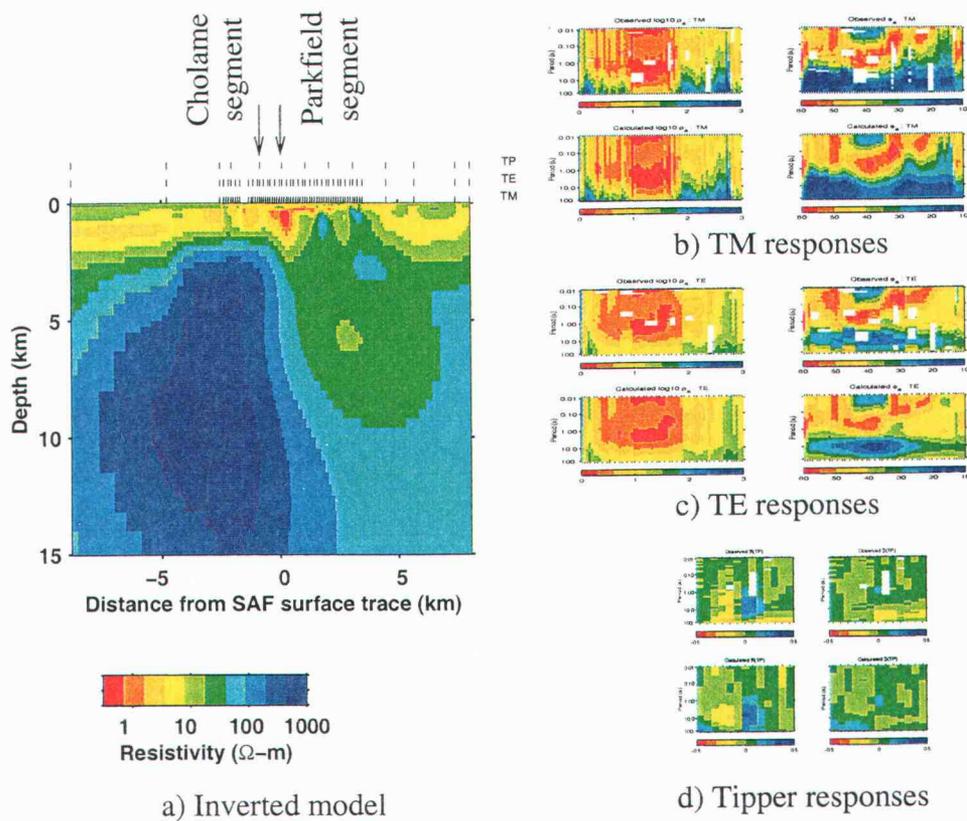


FIGURE 4.20. The same plots as in Fig 4.14, but for CV profile.

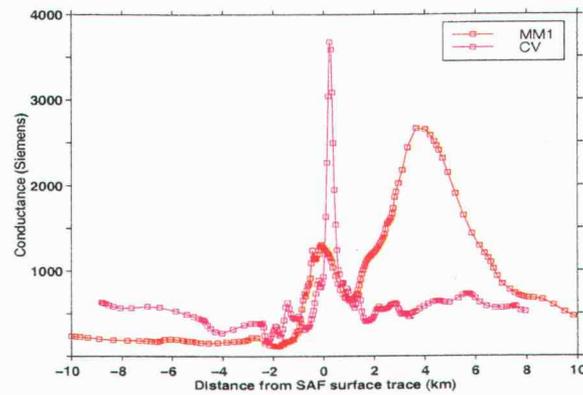


FIGURE 4.21. The conductance plots for CV (pink) and MM1 (red) profiles.

Our profile in the Cholame Valley is deployed across two segments of the SAF where a 1 km right-stepping offset is observed (Hill et al., 1990). However, the location of the Parkfield segment surface trace is not so clear across our profile (Simms, 1988). We therefore linearly project the well defined position of the surface fault trace of the Parkfield segment from further north to define the 0 km point on our profile. The joint inversion result (Figure 4.20) shows that the highly conductive fault core is located entirely to the east of this reference point. However, we need to account for the uncertainties of locating the fault trace. At the SAF Cholame segment, about 0.8 km west of the SAF Parkfield segment, the fault zone is significantly less conductive, and less well-defined. Beneath the fault zones, the crust is resistive ($> 100\Omega - m$) to the west and the less resistive ($< 100\Omega - m$) to the east. However, because of the smooth inversion, the sharpness of the resistivity contrast is not well-resolved (but it clearly shows a near-vertical shape).

The major differences between the Middle Mountain profiles and the Cholame Valley profile are as follows. 1) the fault zone is more conductive. The conductance in the fault zone (Figure 4.21) at the CV profile is about three times higher than

at MM1. 2) the Franciscan east of SAF is much less conductive (but still clearly distinct from the resistive Salinian rocks west of the SAF). The conductance plot (Figure 4.21) clearly displays the absence of the highly conductive crust east of the fault zones for the CV profile. 3) the conductive layer on top of the resistive unit west of the SAF is thicker, extending to about 2-3 km depth.

We summarize here that with the guide of the dimensionality analysis, the inversion shows that at Middle Mountain and Cholame Valley three main features are required, with some variations along the fault from north to south. The resistive Salinian granite dominates the west side of the fault. It is overlain with conductive sedimentary rocks that are thicker in the Cholame Valley. At Middle Mountain, the fault zone conductor is located slightly west of the surface trace and extends to a depth of 2-3 km for the northern two profiles, and probably either deeper or with higher conductivity at MM3. The conductance in the fault zone is estimated to be around 1300 S beneath MM1 and MM2, and 2-3 times higher than that at MM3 and CV. At the Cholame Valley, the fault zone conductor appears east of the Parkfield segment. Beneath the fault zone conductors, there is a near-vertical contact between resistive rocks on the west and more conductive Franciscan sediment extending to 8 km depth, and possibly deeper. The conductivity in the Franciscan sediment is much higher near the Middle Mountain profiles.

4.5. Case Studies on 2-D Synthetic Models

The inversions are generally designed to search for a simple and smooth resistivity model, by minimizing a norm (e.g. $(\mathbf{m} - \mathbf{m}_0)^T \mathbf{C}_m^{-1} (\mathbf{m} - \mathbf{m}_0)$ is used in REBOCC) subject to fitting the data adequately. The smooth model might not necessarily be the most geologically realistic, and sometimes might be misleading. In

the case where the data are generated from a model with a known sharp boundary, a smooth model that adequately fits the data will not generally resolve the sharp boundaries. Instead, the inversion will produce a gradual smooth transition across the boundaries, making it difficult to locate the exact position of the boundaries.

In this section, we further investigate the resolution of the MT data, in order to help distinguish artifacts from the inversion from real structures, and also to understand the characteristics of the inversion scheme. The 2-D MT forward model is solved for many periods on known models to generate the data, Gaussian noise is added to the data, and the REBOCC inversion is then used to reconstruct the solutions, which are directly compared to the known models. Our case studies are similar to those described by Eberhart-Phillips et al (1995) who studied a fault zone of about 5 km width and 10 km depth with resistivity of $10 \Omega - m$, in a $1000 \Omega - m$ host overlain by the $30 \Omega - m$ surface layer of 1 km thickness. The results we obtained from the inversion are more complicated than this simple model.

Our known 2-D model was developed based on the regional scale geological cross-section of Mooney and Weaver (1989; Figure 4.2). Resistivity values are assigned to each major feature as follows: The ocean resistivity is set to $0.33 \Omega - m$. The resistivity beneath the ocean floor increases with depth, similar to the model obtained by Kurtz et al. (1986) from the Vancouver Island area. The resistivity structure beneath the Salinian block is extracted from the preferred model, Figure 4.15a, and extrapolated far to the west. For the Franciscan assemblage on the east, resistivity values similar to those of the preferred model are used, except that the large volume of low resistivity is replaced with a layered structure with resistivity increasing with depth. The Great Valley resistivity structure is assumed to be similar to the Franciscan, except the first 3-km is assumed to be $3 \Omega - m$. The Sierra-Nevada batholith is set at $300 \Omega - m$. At 17-60 km depth, the whole model

is assigned a resistivity of $300 \Omega - m$, which decreases to $10 \Omega - m$ below 100 km depth.

Data sets for five different fault zone structures (the other parts of the model are the same) are generated. For the first case the fault zone is not conductive, Figure 4.22a1, with only the contrast between resistive crust on the west and conductive crust to the east. In the second (Figure 4.22b1) and third cases (Figure 4.22c1), we added conductive fault zones of $1 \Omega - m$ to depths of 2 km and 10 km respectively, and additional $10 \Omega - m$ zones to depths of 2.5 km and 13 km, respectively. The fourth and fifth models are like the second and third but another conductive block ($1-3 \Omega - m$) was added at 1-3 km depth 3 km east of the fault. The purpose of the first model is to make sure that the inversion algorithm can resolve the resistivity gradient, without adding artifacts (such as a fault-like structure) to the model. For the second and third cases, we would like to know how well the inversion can recover the fault zone geometry and conductivity of possible fault zones. Similarly for the fourth and fifth models, we consider how a conductive block east of the fault (similar to the models from MM1, MM2 and MM3) will affect the geometry and conductivity of the fault. Note that different inversion schemes might produce different results. Therefore, the same inversion scheme used for our analysis above (REBOCC) will be used for these test models.

The model covariance used in the inversion controls the smoothness of the model. For REBOCC, the smoothing length scales of the model covariance in both the horizontal and vertical directions are defined as the distances where the model correlation drops by a factor e . By default, the horizontal length scale is automatically set to 500 m (the station spacing) from the surface down to 500 m depth, while the vertical length scales vary with depth. This makes the horizontal length

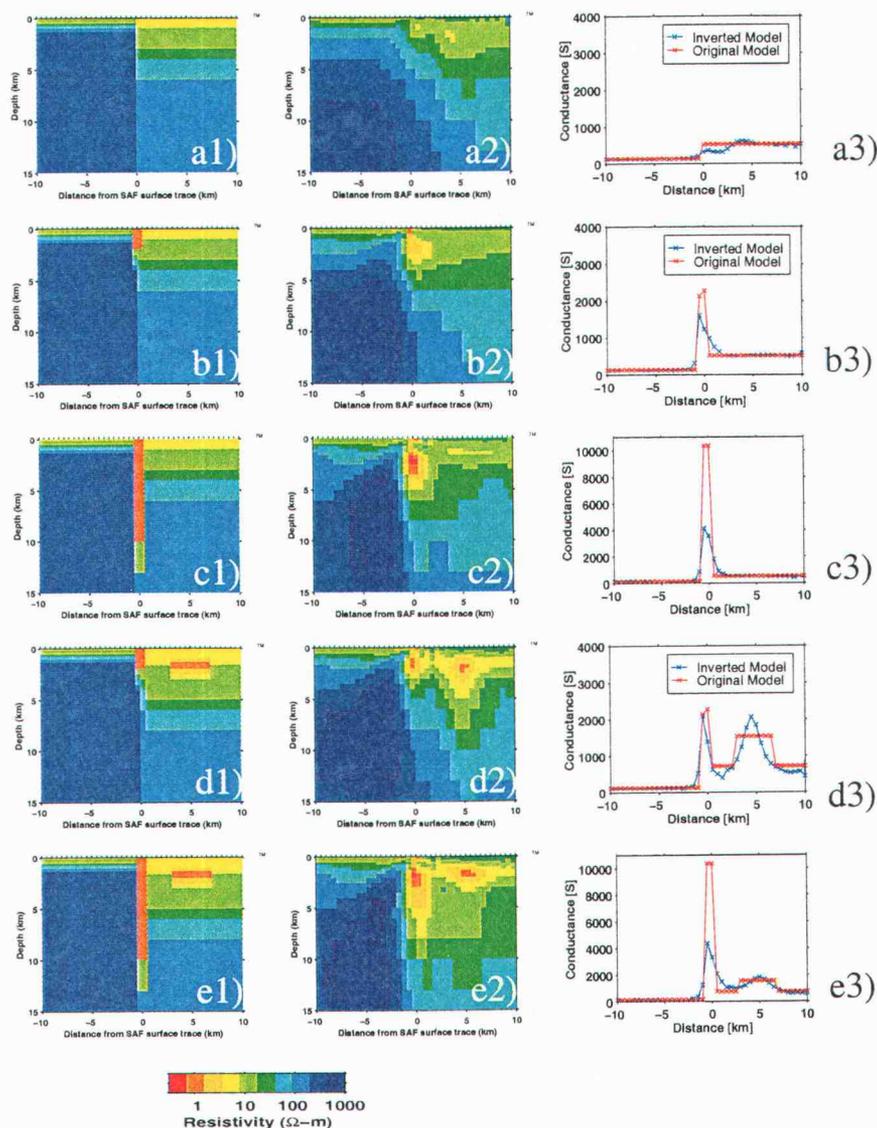


FIGURE 4.22. Case studies: (a1) Original model used to generate the data. In this case, there is no conductive fault zone, only the resistivity contrast from west to east. (b1) Same as (a1) but two-km fault zone of $1 \Omega - m$ is added in between the resistive crust on the west, and conductive crust on the east. (c1) Same as (b1) but the fault zone extends deeper to 10 km. (d1) and e1) are the same models as (b1) and (c1), but were added one extra conductive block to the east of the fault. (a2), (b2), (c2), (d2) and (e2) are the final inverted models using the data generated from (a1), (b1), (c1), (d1) and (e1) respectively. (a3), (b3), (c3), (d3) and (e3) are the conductance plots; red is from the original model and blue is the inverted model. Note a different scaling for (c3) and (e3).

scale larger than the vertical length scale at shallow depths. Below 500 m, both horizontal and vertical length scale vary with depth in the same manner.

The data (TM and TE modes and tipper) from these five base models are generated with the 2-D forward modeling program at 25 periods from 0.01-700 seconds, and 41 stations where stations are equally spaced 500 m apart covering 10 km on both side of the fault. Results from finer spaces between sites are not much different. About 10 % Gaussian noise were added to the data prior to the inversion to account for the 3-D effects. No static shifts were added to the data. Grid discretization for all models is the same, at 190×33 (plus 10 air layers for the TE mode). The results from the joint inversions using all data are shown in Figure 4.22(a2), (b2), (c2), (d2) and (e2) respectively, and the conductances for each model (in blue) are plotted in Figure 4.22(a3), (b3), (c3), (d3) and (e3) respectively, along with the conductances of the original models (in red).

Because of the smoothness criterion, the resistivity contrast of Figure 4.22(a2) is not sharp. The resistivities decrease slowly across the fault zone from west to east. This causes a small drop in the conductance east of the fault (Figure 4.22a3) The results from the single mode inversions (Figure 4.23a) show that TM data resolve the increase in resistivity with depth better than the TE data. The tipper inversion alone starting from the half-space model does not prove to be very useful in any test models. However, including tippers with TM and TE data helps constrain the model.

A comparison of Figure 4.22(a2) to (b2) and (c2) gives us an idea that the REBOCC inversion can indeed distinguish between models with and without a conductive fault zone, giving us some confidence that the highly conductive fault zones seen in all profiles are not artifacts generated from the inversion. However, the resistivity values and the fault zone geometries are not always as well-resolved as

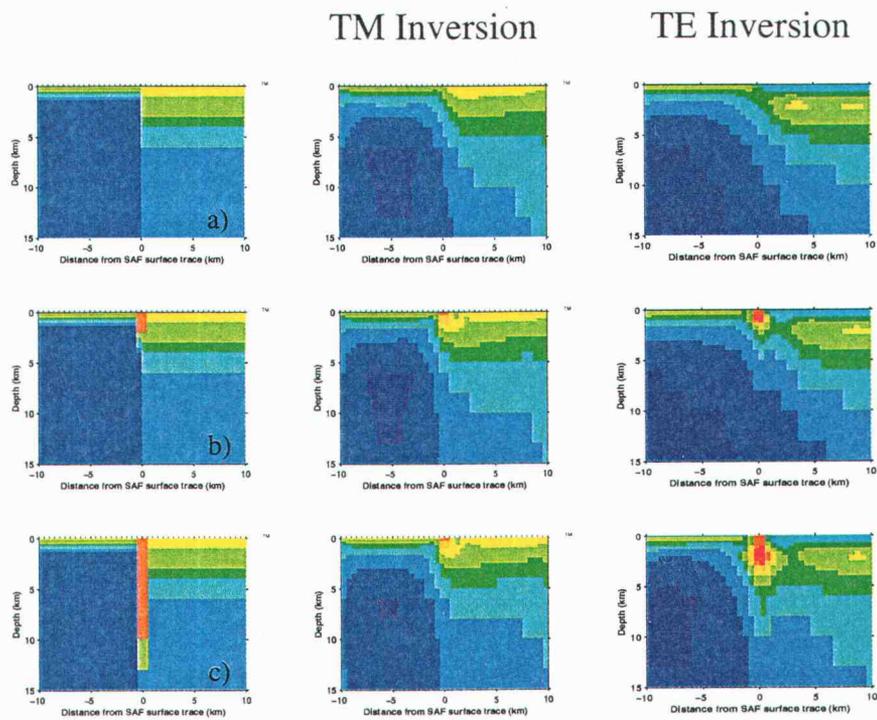


FIGURE 4.23. Similar to Figure 4.22, but show the results from single mode inversions, TM (middle) or TE (right). Original models are on the left column. Model A, B and C are the same as (a1), (b1) and (c1) of Figure 4.22.

desired. In particular for the 10-km-deep fault zone (Figure 4.22c2), the inversion restricts the fault zone to the upper 5 km. An extension of low resistivity can still be seen continuing to greater depth. The TM mode inversions (Figure 4.23b and c) do not show any indication of the conductive fault zone, except for some high conductivity in the near surface at the fault trace. In contrast, the TE mode inversions distinguish the fault zone from the surrounding rocks, but make the fault shallower and more conductive than it actually is. The conductance plots of both cases (b3) and (c3) in Figure 4.22 show that the conductances of the resistive and conductive side can be recovered quite well. The only part that is underestimated is beneath the fault zones where more than 60 % of the conductance is lost for the third and fourth models, and about 30 % for the second model.

Next, we add a conductive block to the east of the 2-km and 10-km fault zone (d1) and (e1) of Figure 4.22) about 3 km east of the fault. The joint inversions reveal two conductors. The shape of the conductor to the east is distorted in both cases. The smoothing makes it thicker (d2) or wider (e2) than the actual shape. However, the blocks are pretty much located correctly. Surprisingly, the fault zone conductors appear to be estimated better than for the cases without the extra conductor to the east. The conductance at the fault trace of the 10-km-deep fault is still underestimated. Conductances for the eastern conductor are reasonable on average, though peak values are somewhat high.

We conclude here that even in this difficult environment (with a large resistivity contrast across the fault), the inversion is able to recover both the conductive structure located to the east, and the fault zone conductor. However, the conductance of both structures might not be as well-resolved as desired. The inversion tends to make conductivity higher than the actual value if the depth of the conductor is not well-resolved in order to match the conductance. By simulating the

models similar to the field data with the forward modeling and inversion, we have learned a great deal, particularly about the behavior of the smooth inversion. We hope that we will be able to distinguish the artifacts of the inversions from the real resistivity structures. However, our case studies are based on the 2-D structures. For the field data, 3-D effects (e.g. the coastline, the shapes of the Great Valley, or any local 3-D structures) might play more important roles.

4.6. Hypothesis and Sensitivity Tests

Because the inverse problem has a non-unique solution, here, we generate a range of models consistent with the data (For instance, inverting with different prior models, or freezing hypothesized structures) to better constrain which features are required by the data. We use the overall misfit of the preferred models (Figure 4.15 for MM1, and Figure 4.20 for CV), as the target misfit for the different tests. A closer look at fits to specific sites near the features of interest will be used when ambiguity occurs, as advocated by Park et al (1996). Note that most tests use all data (TM+TE and tipper) to compute the fit.

4.6.1. Hypothesis Tests on MM1 profile

The preferred model for MM1 shows that the resistivity beneath the fault zone increases with depth (around 30-100 $\Omega - m$ from 3-13 km depth), and varies laterally across the fault zone. There is no indication of a conductive fault zone at depth. Our case studies show that it is not easy for the smooth inversion to resolve a narrow deep conductive fault zone. Therefore, numerous inversion runs are performed to find a range of possible solutions, and to allow a direct comparison with the results from the case studies. Our first test is to stick a sharp vertical

boundary right beneath the surface fault trace, extending to different depths to disrupt the horizontal smoothness across the fault.

Figure 4.24 shows four models from inversion with different depths of a vertically sharp boundary: a) from the surface to 3 km depth, b) from the surface to 8 km depth, c) from 2 km to 8 km depth and d) from 2 km to 16 km depth. All of these models fit the data at the same overall level. With the vertical boundary located from the surface to 3 km depth, conductivity is high around the boundary on both sides down to 2 km depth, indicating no resistivity contrast to this depth. With the boundary extended to 8 km, there is again no resistivity contrast in the near surface (to 1.5 km depth). However, below this depth there is a sharp change in resistivity. This feature also shows up in the other two tests, with vertical boundaries from 2-8 km depth and 2-16 km depth. The change in resistivity across the boundary is around 2 orders of magnitude, from around $300 \Omega - m$ to around $3 \Omega - m$. This is not surprising since MT is sensitive to a volumetric average of the conductivity. Note that the circle shape of conductivity features is probably a result from the smoothing used in REBOCC (for more details see Siripunvaraporn and Egbert, 1999; Chapter 3). So, what are these features really showing ?

We therefore go back to our case study models, and invert the synthetic data with a vertical sharp boundary from the surface to 16 km depth, located exactly at the middle of the 500-m-wide fault zone. Figure 4.25 gives the results of inversions for three different models; (a) no conductive fault zone, with only the resistivity contrast and the eastern conductive block, (b) as in (a) but with a 2-km-deep fault zone and (c) as in (a) but with a 10-km-deep fault zone. The inversion results without allowing for the sharp boundary are in Figure 4.22(d) and Figure 4.22(e) for cases (b) and (c) respectively. In case (a), the sharp boundary helps the inversion to recover the other parts of the model accurately. For the 2-km-deep fault zone

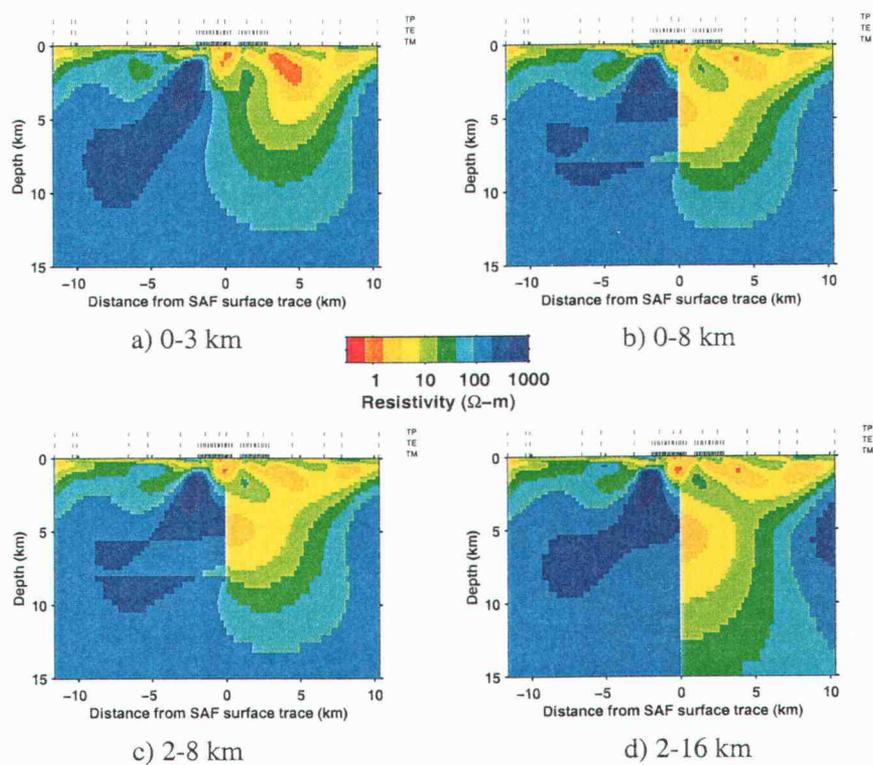


FIGURE 4.24. Results from inversion based on different vertically sharp boundaries; (a) from surface to 3 km depth, (b) from surface to 8 km depth, (c) from 2 km depth to 8 km depth, (d) from 2 km depth to 8 km depth.

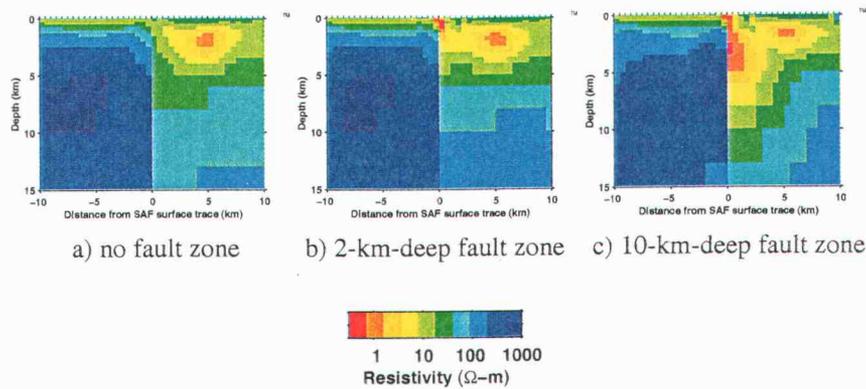


FIGURE 4.25. Results from inversion based on different vertically sharp boundaries on the case study models showing in Figure 4.22; (a) no fault zone conductor but only the eastern conductive block, (b) with 2-km-deep fault zone and eastern conductive block, (c) same as (b) but 10-km-deep fault zone.

model, because of the boundary from the surface to 16 km depth the inversion makes the near surface (< 1 km) very conductive on both sides. Below this the low resistivity continues on the east side of the boundary to about 3-4 km depth, and connects to the eastern conductive unit. The most interesting model is the inversion from the 10-km-deep fault zone. Similar to the 2-km fault zone, the inversion makes the near surface conductive, and below this the high conductivity continues on the east side of the boundary down to around 10-km depth. Its shape is very broad in a (half) circular shape. Its radius is around 3 km and the conductivity peaks at about 3-km depth.

The model from the case study (especially Figure 4.25c) and the models from the field data (Figures 4.24b, c and d) show significant similarity, particularly the (half) circular shape occurring on the east of the boundary. One interpretation of this similar pattern is that there is a narrow-deep conductive fault zone similar to that of the case study at Middle Mountain. However, its geometry (width and depth) and resistivity value are still unclear. We therefore use constrained inversion

to further test the possibility that a deep conductive fault zone could fit the MM1 data.

The preferred model (Figure 4.15), also the models from Figure 4.24 show us that near the surface trace of the fault (to 2-3 km depth) the data require a broad zone of high conductivity. Seismic observations of trapped modes in the SAF fault zone from earthquakes near Parkfield reveal a narrow fault zone core of 100-150 meters width with low shear wave velocities of 1.1 to 1.8 km/s. This zone extends from surface to a depth of at least 10 km (Li et al., 1990). From similar observations, Leary and Ben-Zion (1992) concluded that trapped waves near Parkfield require a 200-m-wide low velocity zone as an upward channel for energy to reach the surface. Based on these observation, we started our experiments with a vertical 236-m-wide (due to model discretization) fault zone extending from surface to a depth of 5 km. Resistivities from 0.1, 0.3, 1.0, 3.0 and 10.0 $\Omega - m$ were considered. We also experimented with wider faults, 382 and 555 meters. Table 4.1 lists all the parameters used in the tests. The rightmost column is the overall misfit difference (in %) relative to the overall misfit of the preferred model (Figure 4.15). If the inversion can find models with misfit at the level of the preferred model, the inversion is terminated.

Table 4.1 shows that for the narrow fault zone (236 m width and 5 km depth) tests, only three resistivity values (1, 3 and 10 $\Omega - m$; Figure 4.26) are allowed. With the low-resistivity (0.1 and 0.3 $\Omega - m$) faults, the inversions fails to reach an acceptable misfit. For the high resistivity case (100 $\Omega - m$), the inversion almost reaches the target level (only 6 % above). As the fault gets wider to 382 m and 555 m, we found that only the 3 $\Omega - m$ fault zone allows a reasonable fit to the data. Focusing on the overall misfit can be misleading. With the same overall misfit, one model might fit the data better at specific stations or period ranges than

Case	Depth (km)	Width (m)	Resistivity ($\Omega - m$)	Overall RMS Diff.*
1	5	236	0.1	80 %
2	5	236	0.3	58 %
3	5	236	1.0	0 %
4	5	236	3.0	0 %
5	5	236	10.0	0 %
6	5	236	100.0	6 %
7	5	382	1.0	10 %
8	5	382	3.0	0 %
9	5	555	0.3	72 %
10	5	555	1.0	28 %
11	5	555	3.0	3 %
12	5	555	10.0	13 %
*relative to the overall misfit of the preferred model (Figure 4.15)				

TABLE 4.1. Lists of fault parameters (depths, widths and resistivities) used as a fixed structure in the inversion tests.

another model, while fit might be worse at other locations. In our test case, stations are located very densely near the fault zone. A 5 km deep conductive fault should definitely affect the data for all of the stations near the fault. A closer look at the responses near the fault zones might give us a better idea of differences between the three different resistivities that allow equal fits to the data. Figure 4.27 displays the responses at station number 24 (near the west boundary of the fault) and station number 26 (right above the fault). This figure shows that the calculated responses

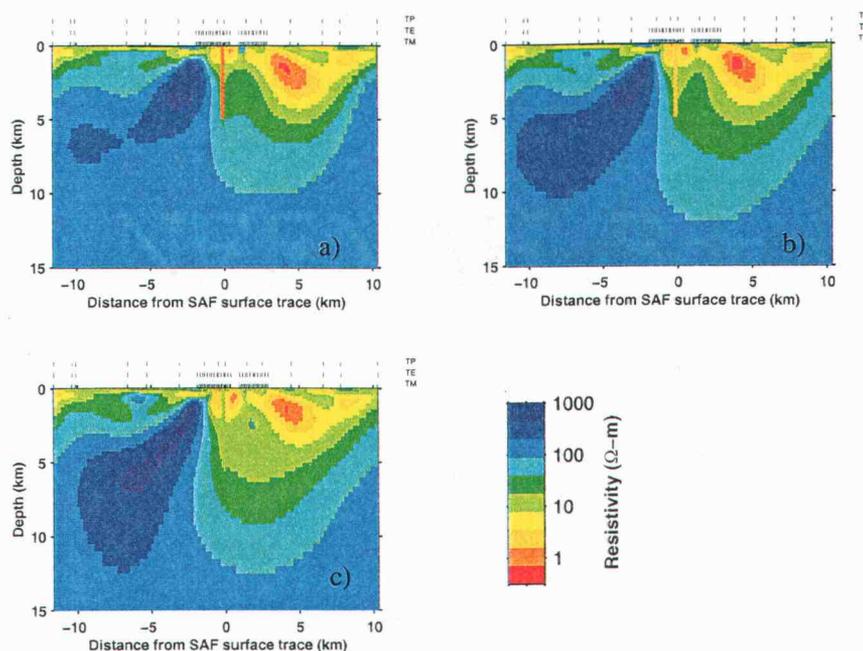


FIGURE 4.26. Results after inversion based on fixed fault zone with different resistivities; (a) $1 \Omega - m$, (b) $3 \Omega - m$ and (c) $10 \Omega - m$.

of the $3 \Omega - m$ fixed fault zone fit the data at these stations at about as well as the preferred model, while other fixed fault zone resistivities (1 and $10 \Omega - m$) result in larger misfits, especially at short periods (e.g. TM and TE ρ_a at station number 26). This suggests that near the surface a resistivity of $3 \Omega - m$ is most preferred.

Figure 4.26 shows that the sizes of the conductive prisms on either side of the fault vary with the resistivity assumed. These are smaller for a more conductive fault. The major difference between these three models (Figure 4.26) and the preferred model (Figure 4.15) is in the area east of the fault. With the fixed fault zone, the conductive zones in the east become shallower and more conductive which might be the result of overfitting the data.

From these tests, we find that it is possible to have a narrow (approximate 250-m-wide and 5-km-deep) conductive fault zone and still fit the data. The possible

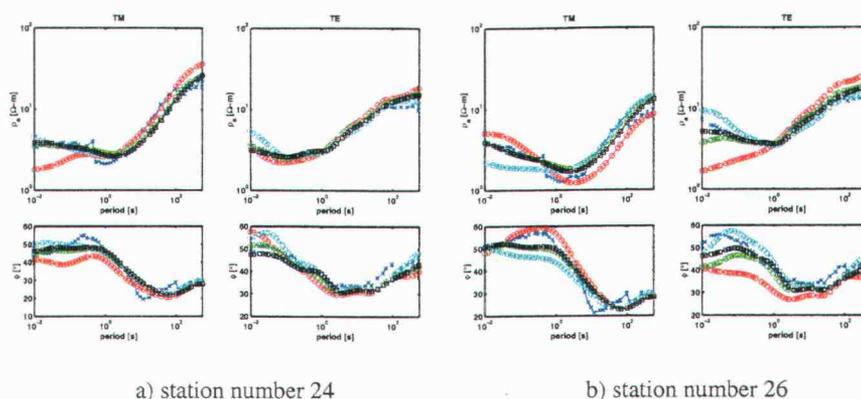


FIGURE 4.27. Plot of observed responses (blue) and the calculated responses from various models (red for $1 \Omega - m$, green for $3 \Omega - m$, cyan for $10 \Omega - m$ and black for the preferred model) at (a) station number 24 (west boundary of the fault), and (b) station number 26 (right above the fault trace).

resistivity values range from 1 to $10 \Omega - m$, but a $3 \Omega - m$ fault zone best fits the data at stations near the fault. Increasing the width of the fault zone tends to increase the misfit to the data, but wider fault zone are permissible. From seismicity, the fault zone is observed to extend to great depth. Other seismic investigations (e.g. Li et al., 1990) also suggests that there are low velocities in the core of the fault zone to a depth of at least 10 km. Do the MT data allow a conductive fault zone this deep? To investigate this possibility, we extend the fault zone 5 km to a total depth of 10 km, with the $3 \Omega - m$ fault zone in the upper 5 km as a base model. Different widths (236, 555 and 994 m) and resistivities (0.1, 0.3, 1, 3, and 10) for the next 5 km block are used as listed in Table 4.2. These fault structures are frozen to let the inversion search for the model that fits the data subject to this constraint.

With the continuation of the narrow (236-m) fault zone to great depth, the resistivity of the deeper fault zone (5-10 km depth) can be as low as $0.3 \Omega - m$. If the width of this segment is increased, the minimum resistivity that the inversion allows becomes higher (e.g., to $1 \Omega - m$ if 994-m-wide). Most of the stations near the

Case	Depth (km)	Width (m)	Resistivity ($\Omega - m$)	Overall RMS Diff.*
1	10	236/236	3.0/0.1	10 %
2	10	236/236	3.0/0.3	0 %
3	10	236/236	3.0/3.0	0 %
4	10	236/236	3.0/1.0	0 %
5	10	236/236	3.0/10.0	0 %
6	10	236/555	3.0/0.1	28 %
7	10	236/555	3.0/0.3	0 %
8	10	236/555	3.0/1.0	0 %
10	10	236/555	3.0/3.0	0 %
11	10	236/994	3.0/0.3	12 %
12	10	236/994	3.0/1.0	0 %
13	10	236/994	3.0/3.0	0 %
14	15	236/236/236	3.0/3.0/3.0	0 %
15	22	236/236/236/236	3.0/3.0/3.0/3.0	0 %
*relative to the overall misfit of the preferred model (Figure 4.15)				

TABLE 4.2. Lists of fault parameters (depths, widths and resistivities) used as a fixed structure in the inversion tests.

fault are little affected by a high conductivity fault zone at great depth. Overall, these models fit the data to about the same level as the 5-km-deep, 236-m-wide fault zone, in Figure 4.26b. When the conductivity becomes higher or the fault zone wider, a slight effect can be seen at periods longer than 100 seconds for most stations. However, if the fault is too conductive (e.g. $0.1 \Omega - m$ from 5-10 km

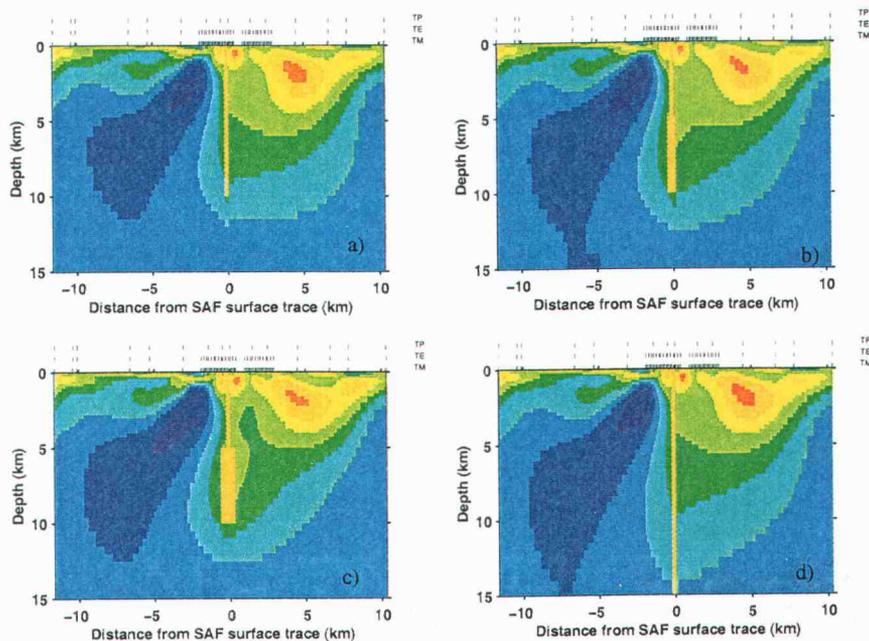


FIGURE 4.28. Results after inversion based on fixed fault zone with different shapes of fault zones ($3 \Omega - m$); (a) 10-km depth and 236-m width, (b) 10-km depth, but wider to 555-m for the deep segment, (c) same as (b) but wider to 994-m, and (d) same as (a) but deeper to 15 km. These models fit the data almost the same level.

depth), the inversion cannot find a model that fits the data above 100 second for the stations near fault. Fits are even worse in the eastern remote sites for periods above 10 second. Because of the high conductivity in the deep fault zone, the inversion cannot put conductive material anywhere else, e.g. in the upper 5 km of the eastern crust, which degrades fits east of the SAF. This shows that using overall fit to assess the model is reasonable, in this case. Figure 4.28 shows some of the models resulting from inversion with the constraints of Table 4.2. These models fit the data equally well in a global sense. Note that by making the fault zone deeper, the conductivity east of the fault zone is lower.

4.6.2. Hypothesis Tests on CV profile

As in the synthetic case studies, excessively high conductivity near the fault zone could mean that the fault zone might actually be deeper, but less conductive. A test similar to that used for the MM1 profile is therefore performed on this profile. First, we inserted a sharp vertical boundary from 2 to 10 km depth at two different locations, beneath the SAF Cholame segment and beneath the SAF Parkfield segment (Figure 4.29). We found that at both locations there is no continuation of the high conductivity to greater depths, as occurred for MM1 (Figure 4.24). Note also that the high conductivity beneath the fault zones remains in these models. Comparing these results to the synthetic case studies (Figure 4.25b), suggests that the fault zone here is probably not conductive to great depth. Similar to the MM1 profile, we inserted fixed fault zone of various depths, widths and resistivities. We found that a fixed fault zone with a width of 240-m, resistivity of $3 \Omega - m$, and depth of 5 or 10 km fit the data as well as the preferred model (Figure 4.30). A conductive fault zone is thus allowed by the data.

4.7. Discussion

The MT profiles at different locations across the SAF reveal distinct resistivity structures from the north (Middle Mountains) to the south (Cholame Valley), about 20 km separation. These models are consistent with the seismic velocity models derived from the 3-D seismic tomography by Eberhart-Phillips and Michael (1993). The models from MM1 and CV with seismic velocity plotted on top are given in Figures 4.31, and 4.32 respectively.

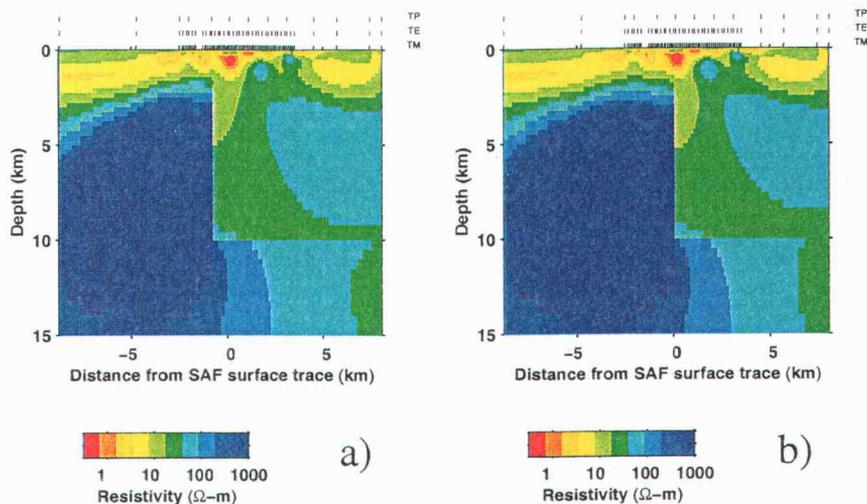


FIGURE 4.29. Results from inversion based on vertically sharp boundary from 2 to 10 km depths at two different locations as prior models; (a) beneath the Cholame segment, and (b) beneath the Parkfield segment.

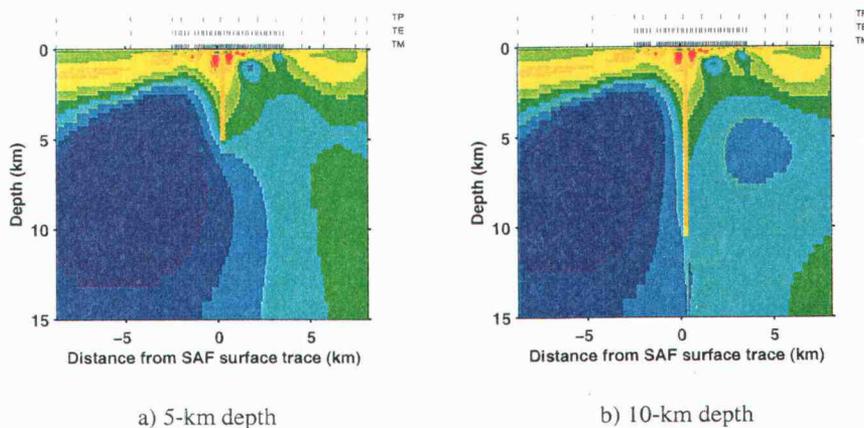


FIGURE 4.30. Results from inversion based on fixed fault structure with $1 \Omega - m$ and 244-m-wide (a) from 0.2 to 5 km depth, (b) from 0.2 to 10 km depth. Both models fit the data the same level.

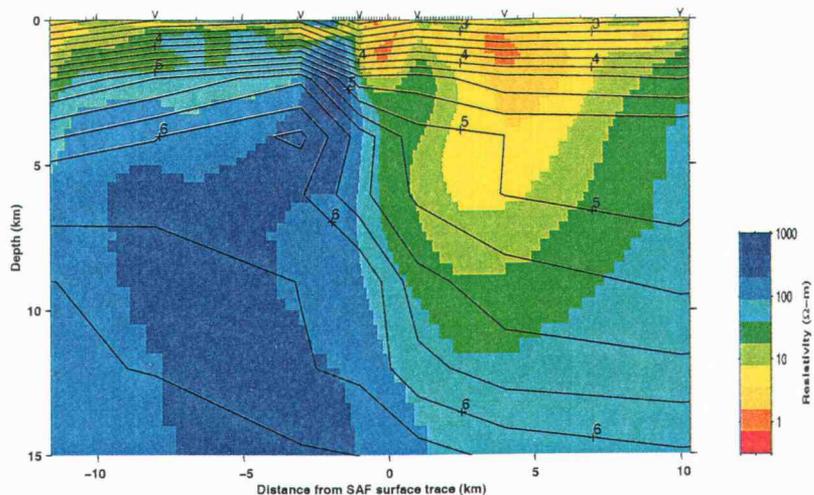


FIGURE 4.31. The MM1 preferred model. On top of it is the seismic velocity model of Eberhart-Phillip and Michael (1993) at the same locations. The seismic stations are given at the surface with 'V' shape, while MT stations are a small tick mark.

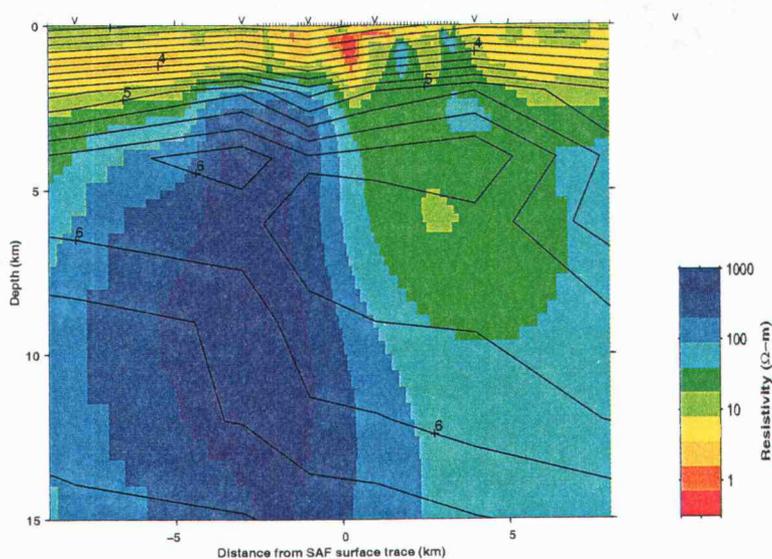


FIGURE 4.32. Same as Figure 4.31, but for CV profile.

At MM1, west of the SAF both seismic velocity and electrical resistivity increase with depth. A velocity of 5.75 km/s is estimated to be the top of the granite rocks (Holbrook et al., 1992). This is in close agreement to the transition to more resistive ($> 100\Omega - m$) rocks at about 4 km depth. Close to the fault, the top of the resistive unit becomes shallower, rising to 1 km depth. This could be an artifact of the inversion. However, forward modeling tests suggest that the data (especially the TM data) require a more resistive unit at shallow depth close to the fault. A comparable feature is not seen in the seismic velocity model, but station spacing for the seismic data was too coarse to resolve details near the fault. The sharp drop of resistivity in the upper 5 km about 2 km from the west end of the profile is most likely an artifact of the inversion.

Beneath the surface fault trace, the seismic velocity model does not contain a sharp change in velocity comparable to the prism of low resistivity observed in the MT data. Again, this is probably due to large station gaps near the fault in the seismic data. However, at greater depths (> 2 km) the tomography reveals a sharp velocity contrast across the fault which coincides with the resistivity contrast imaged by the MT. To the east, the MT model shows a low resistivity core ($\approx 1-3 \Omega - m$) at a depth of 1-5 km. This feature is not present in the seismic image. However, at greater depth, the 5.25 km/s contour of seismic velocity corresponds with the $10 \Omega - m$ lower boundary of the low resistivity zone. The low velocity zone continues about 13 km east from the SAF, while the low resistivity zone appears to be limited to lie within the profile. However, we believe that the high resistivity feature found beneath the easternmost station (around 10 km) may be a local feature. This feature is required only by the TM data, and if we invert without the last station, the low resistive zone continues further to the east. This local resistive feature could

possibly be interpreted as a localized fragment of the volcanic rock seen throughout the Franciscan in this area.

The 3-D seismic velocity model of Eberhart-Phillips and Michael (1993) shows that the low velocity feature in the Franciscan is roughly 40-km long, 12-km wide and at least 10-km deep. It approximately lies parallel to the SAF but, is most prominent adjacent to Middle Mountain. It continues further to the south for about 10 km before disappearing. This is also consistent with our resistivity models. The anomalously low resistive zone in the Franciscan is presented from MM1 to MM3, but is absent in the CV profile. However, the location and size of this low resistive feature is probably not well-constrained along the strike in the MM profiles, because the data errors on the east side had to be increased to allow for 3-D effects. Nevertheless, we find that this low resistivity zone must be located outside of the dense profile to the east (i.e., between 3 and 7 km from surface trace), and within 5 km depth, with a resistivity from $1-4 \Omega - m$.

Beneath the Cholame Valley (Figure 4.32), a localized-narrow high velocity zone around 6.2 km/s (Figure 4.32) at about 5 km depth is observed. This corresponds to a slight reduction in resistivity to about $30 \Omega - m$. Similar to the MM1 profile, the 5.75 km/s velocity contour coincide with the top of the resistive unit ($> 100 \Omega - m$) on the west. Beneath the Cholame segment fault trace (at about -0.8 km) there is no significant conductive feature associated with the fault zone as observed in the MM profile. However, a distinctive high-conductive fault zone is observed to the east of the Parkfield segment fault trace. This high conductivity is coincident with highly coherent reflectors observed in the high-resolution seismic reflection profile, which were interpreted as lateral-deformed clastic sedimentary strata overlying the Franciscan complex by Shedlock et al (1990). These reflection data also shows that at about 1.50 km east of the Cholame fault trace there are many

shallow faults, which coincide with the western end of the high conductivity fault zone. Also, the resistivity model shows a very thin conductive layer just east of the fault. This is very similar to the average thickness of 350 m of sediments overlying the Franciscan complex in wells drilled in southern Cholame Valley (Shedlock et al., 1990).

4.7.1. Sources of Low Resistivity

There are several possible sources for the low resistivity zones. Serpentine has relatively low resistivity and has been found in this area. Because of the large magnetic signature associated with this mineral, aeromagnetic data of Jachens (published in Eberhart-Phillips and Michael, 1993) has been interpreted to form an airfoil-like body centered northeast of and trending parallel to the SAF. The thickness of the inferred body is around 0.5-2 km and its top is at 3 km depth. Though its depth and thickness can be varied, its shape and location are well constrained (Eberhart-Phillips and Michael, 1993). Serpentine resistivities have been measured *in situ* in Quatipuru, Brazil to be $90 \Omega - m$ (Palacky, 1987). Its general resistivity value is around 10 to $100 \Omega - m$ (Palacky, 1987). It is thus unlikely that the low resistivities (as low as $1 \Omega - m$) observed in and east of the fault zone result only from serpentine.

Clay minerals could be another explanation of the low resistivity zones. Wang (1984) proposed that the SAF zone in central California is composed of saturated, clay-rich fault gouge throughout seismogenic depths. The effects of clay on rock resistivity is dramatic because it can form an interconnected coat on grain surfaces that binds fluids. Because of its high cation exchange capability, the bound water can pass along charges to surrounding fluids leading to an increase in conduction

paths and a significant reduction in rock resistivity (Eberhart-Phillips et al, 1995). The typical resistivity of the clay-rich fault gouge is around $3-10 \Omega - m$ (Palacky, 1987). However, clay is not generally observed at that great depth, and if it was, the fault would be expected to be aseismic because the frictional sliding in the presence of clay is always stable (Scholz, 1990).

Another possible source for the low resistivity is saline fluid, which has been documented in this area, e.g. at the Varian well (J. Thordsen, personal communications, 1999). The primary factors controlling the electrical resistivity of rocks containing a saline aqueous pore fluid are the resistivity of the fluid, the porosity, and the degree of pore interconnection. Because the difference in resistivity between the solid and fluid phases is about six order of magnitude (Hyndman and Shearer, 1989), the degree of pore interconnection plays a major role in controlling the bulk resistivity. A few percent difference of the porosity can also have a significant effect on other physical properties. For example seismic velocity can be reduced by 10 % with a 1 % increase in porosity (Hyndman et al., 1993). The empirical Archie's law given as

$$\sigma_B = a\sigma_w W^m, \quad (4.4)$$

is normally used to relate electrical resistivity to porosity, where σ_B is the bulk conductivity of the rock, σ_w is the fluid conductivity, W is the porosity, and a and m are empirically determined parameters which account for connectivity of the fluid. Keller (1989) suggests that reasonable values for these two parameters for sedimentary rocks where the pore space is the volume left over between grains, are $a = 0.6$ to 1.0 and $m = 1.6$ to 2.0 . Higher values of m correspond to more poorly interconnected pore spaces. Appropriate values for crystalline rock are $a = 1.4$ to 2.0 and $m = 1.3$ to 1.6 .

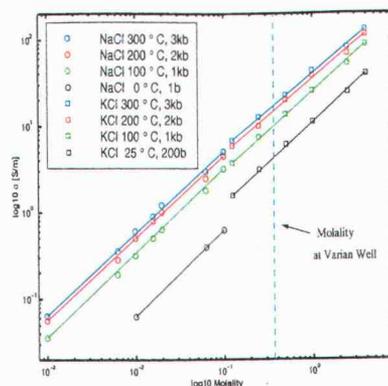


FIGURE 4.33. Plots of experimentally determined conductivity of NaCl (Quist and Marshall, 1968) and KCl (Hwang et al., 1970) solutions at different molality, temperature and pressure (redrawn from Merzer and Klemperer, 1997). The dash line indicates the molality estimated to be about 0.375 M at Varian Well (1.4 km from SAF).

The main factors that control the resistivity of the fluid are the salinity and the temperature and pressure conditions. Sea-water is about 0.5 M NaCl, where M is the molality (concentration in moles per kg of solvent). Salinity has been observed to generally increase with depth. In the Canadian Shield, Frapet and Fritz (1987) found that the salinity increased from 0.1 M above 500 m to more than 2.5 M at 1500 m depth. In the deep borehole studies of the German Continental Deep Drilling Program (KTB), salinity of the fluid increased with depth from fresh water at the surface to about 1 M NaCl at 4 km depth, and to 3 M CaCl_2 at a depth of about 9 km (Huenges et al., 1997). Figure 4.33 shows the relationship between the salinity and resistivity at different temperatures and pressures. As the salinity increases, the resistivity decreases (conductivity increases). At constant salinity, increasing the temperature (pressure) increases the conductivity.

A recent study by Nesbitt (1993) showed typical crustal fluids vary in resistivity from 10 to 0.01 $\Omega - m$. The decrease of the resistivity of the crustal fluid results

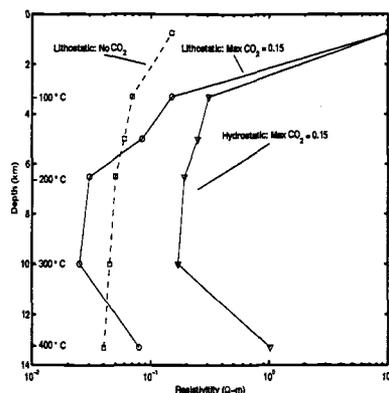


FIGURE 4.34. Resistivity of 3.8 % saline (KCl) with depth in the crust at different conditions; lithostatic pressure and no CO_2 , lithostatic pressure with CO_2 and hydrostatic pressure with CO_2 . Assumed a temperature gradient of $30^\circ \text{C}/\text{km}$ (redrawn from Nesbitt, 1993). The resistivity estimated with NaCl is expected to be very similar.

from increased concentrations of chloride and bicarbonate salts, and increases in temperature and density of the fluids. As shown in Figure 4.34, at 3.8 weight % ($\approx 0.5 \text{ M}$) KCL, the lowest resistivity appears around $0.03 \Omega - m$ at the depth of 6-10 km ($200^\circ\text{C} - 300^\circ\text{C}$), assuming lithostatic pressure and a 30°C temperature gradient. The sharpest decrease occurs at shallow depths ($< 8 \text{ km}$) and low temperatures ($< 200^\circ\text{C}$).

In the Parkfield area, at the 1.5 km deep Varian well, south of the MM2 profile and about 1.4 km off the fault trace to the east, saline fluids of 0.37 M NaCl were found with an average conductivity of $3.5 \text{ S}/\text{m}$ (J. Thordsen, USGS, personal communication, 1999). The fluid pressure presented in the well was high, about 12 MPa above hydrostatic (Johnson and McEvilly, 1995). The heat flow study of Sass et al (1997) reveals a mean heat flow of around 74 mW m^{-2} . Based on this Sass et al (1997) estimated the temperature to be about 400°C at 15 km (Figure 4.35a), which is the approximate base of the seismogenic layer.

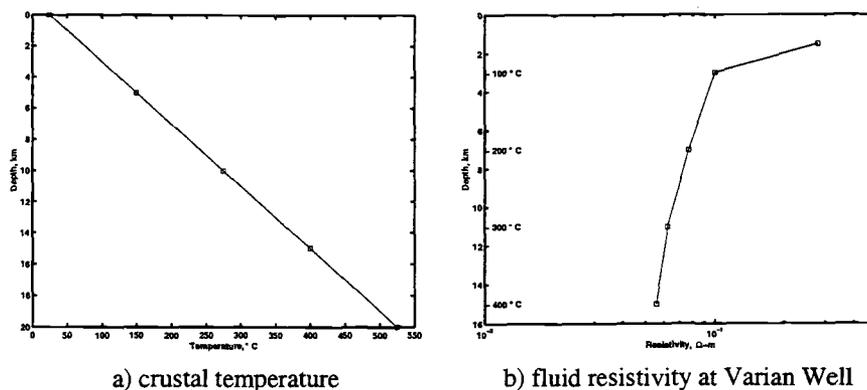


FIGURE 4.35. a) Plot of steady state crustal temperature profile at Parkfield-Cholame region (redrawn from Sass et al, 1997). b) Fluid resistivity at Varian Well, estimated at lithostatic pressure (No CO_2) and temperature gradient of a). Water is predominantly NaCl at 0.37 M.

From the crustal temperature profile (Figure 4.35a), and the experimental determined conductivity of NaCl and KCl as a function of temperature, at a given salinity (e.g., at 0.375 M) an estimate of fluid resistivity at Parkfield can be plotted as a function of depth, Figure 4.35b. Here we assume lithostatic pressure and no CO_2 . There is a sharp drop of resistivity within the upper 3 km from around $0.28 \Omega - m$ (3.5 S/m) at shallow depth to $0.1 \Omega - m$ (10 S/m) at 100°C or 3 km depth. There is a slightly change beneath this, similar to the case shown in Figure 4.34, for lithostatic pressure with no CO_2 . There have been reports of the mantle-derived water and CO_2 moving through some major fault zones (Rice, 1992). Based on geochemical constraints Wang (1984) suggested that it is possible for the SAF zone in the central California to have aqueous fluids which are rich in CO_2 . If this is the case, resistivity below 6 km could be lower by a factor of 2-3 (see Figure 4.34).

Figure 4.36 shows plots of porosity W against the fluid conductivity σ_w for various bulk conductivities σ_B , and for two different values of m (1.5 and 2.), assuming $a = 1$. The dashed lines parallel to the y-axis indicate the estimated average

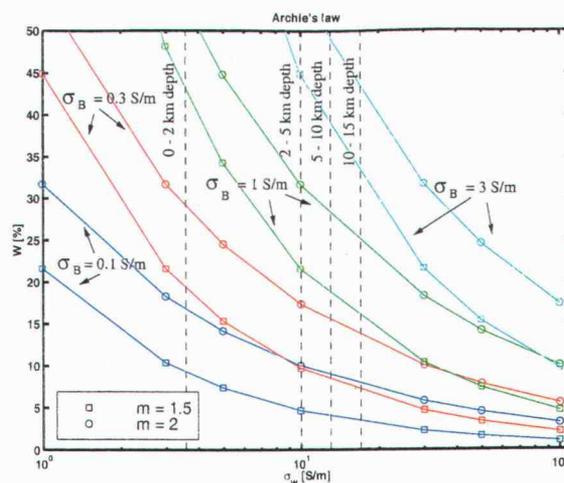


FIGURE 4.36. Plots of porosity required for different values of saline water conductivity to achieve bulk conductivities of 0.1, 0.3, 1 and 3 S/m. Dash lines indicate the fluid conductivities estimated for Parkfield at various depths from Figure 4.35.

fluid conductivities at different depths from Figure 4.35b. Note that if the effects of CO_2 and the possible increases of salinity with depth are taken into account in estimating the fluid resistivity, the estimated porosities plotted in Figure 4.36 would be lower.

4.7.2. Conductive Fault Zones at Middle Mountain

Fluid may play an important role in controlling the mechanism of earthquakes along the Parkfield segment (Johnson and McEvilly, 1995). If this is the case, MT data should be able to detect the low resistivity of the saturated zone. W. D. Stanley's model (published in Eberhart-Phillips, 1993) from the inversion of MT data from the Middle Mountain area does not contain any conductors beneath the fault trace. However, Stanley's station spacing is quite large, on the order of about 4 kilometers. In an earlier interpretation of the dense MM2 profile, Unsworth et

al (1997) found that there is a vertical low resistivity zone ($< 5\Omega - m$) about 500-m wide extending to a depth of around 4 km beneath Middle Mountain. The conductive fault zone had the appearance of positive flower structure, developed by fault-normal compression, with a strong central conductor, similar to the 3-km wide and 5-6 km deep disrupted zone imaged with seismic reflection data (McBride and Brown, 1986),

However, there are concerns that the conductive fault zone imaged on MM2 could be the result of near surface, local inhomogeneities. Our results show that all three MT profiles at the Middle Mountain (MM1, MM2 and MM3) (Figure 4.17) reveal a similar zone of high conductivity beneath the surface fault trace. The positions of the conductors all appear to be slightly west of the surface trace of the fault, with some small variations along strike. Although static distortions still play a major role in the interpretation, the consistency between profiles is a good indication that beneath the surface fault trace the fault zone is indeed very conductive, relative to the surrounding rocks.

For the MM1 profile, the average resistivity of the fault zone (Figure 4.31) is around $3\Omega - m$ in the upper 2 km, with a width of 1.3 km centered around the fault trace. Using an Archie's law and the fluid resistivity of the 1.5 km deep Varian Well (Figure 4.36), this resistivity would correspond to a porosity of around 20-30 % for $m = 1.5 - 2$, and $a = 1$. The average resistivity from 2 to 15 km depths is around $50\Omega - m$. This would correspond to a porosity of around 1 to 3 %, with average fluid conductivity of 15 S/m.

A detailed gravity survey across the SAF near Bear Valley (about 50 km north of Parkfield) revealed a density contrast around 200 kg/m^3 between the fault gouge and the surrounding rocks (Wang et al., 1986). If fluid is assumed to cause the density contrast, an average porosity of the fault gouge (about 2-3 km wide and

15 km deep) of around 12 % is inferred. With the assumption of a narrower fault zone the gravity observation would imply greater porosity. At the porosity of about 12 % and with fluid conductivity increasing with depth, the resistivity can be as low as $1 \Omega - m$ at greater depth, using high fluid conductivity, Figure 4.36.

Such a high porosity would definitely cause a drop in seismic velocity, and should be observed. An estimation of the P-wave velocity reduction from many studies (Feng and McEvelly, 1983; Healy and Peake, 1975; Trehu and Wheeler, 1987) across the SAF is around 0.3 to 1.0 km/sec over 2-4 km width. From the guided waves study, Li et al (1990) revealed a narrow low velocity fault zone with a width of 100-150 m, extending to great depth. The S-wave velocity reduction across the fault is around 0.6 - 1.1 km/sec. Using an empirical relationships provided by Eberhart-Phillips et al. (1989), the relationships between the P and S wave velocity reductions and the porosity can be written as $\Delta V_p/\Delta W = -6.94$ and $\Delta V_s/\Delta W = -4.94$, respectively. Using this empirical relationships, we can estimate the ranges of porosity that cause the velocity reductions to be around 4 to 14 % for the P-wave, and 12 to 22 % for the S-wave velocity.

These evidences show that the porosity estimated from the model in Figure 4.31 is not consistent with such a high porosity suggested by other geophysics techniques. However, the inversion tests show that the data can be fit to the same acceptable level with a conductive fault zone extending to greater depth. Out of many tests, our "preferred" resistivity and width of the fault zone are around $3 \Omega - m$ and 250-m. If this low resistive zone exists at depth, it would require porosities in the vicinity of 8-18 % for $m = 1.5$ and 2, respectively (see Figure 4.36). Note that this range of porosity could be reduced by a few percent due to increase of salinity with depth, and also the presence of CO_2 . This porosity range is in good agreement with other geophysical observations.

4.7.3. Low Resistivity in the Franciscan, East of the SAF

The large low resistivity zone in the Franciscan east of SAF is supported by previous MT interpretations. W. D. Stanley's model (published in Eberhart-Phillips, 1993) shows a thick low resistive zone of $3-10 \Omega - m$ extending from the surface to a depth of around 8 km. Stanley interpreted this low resistivity unit as Great Valley and Franciscan formation with Tertiary marine units. Furthermore, his model includes an unusually low resistivity feature ($< 1 \Omega - m$) at a depth of around 4 to 6 km, embedded in the $3-10 \Omega - m$ structure. This low resistivity extends from the fault zone to about 12 km east, where the last of his ten stations is located. He interpreted this unusually low resistivity as high pore pressure brine.

Also, in the earlier interpretation of MM2 profile when no remote sites were deployed, Unsworth et al (1999) found that to adequately fit both TM and TE data, a conductive layer of $1 \Omega - m$, 1-km thick fixed in the model at the depth of 2-km and located 4-km to the east of the SAF was required. The low resistivity crust is coincident with low P velocity and but not low gravity (Eberhart-Phillips and Michael, 1993). Eberhart-Phillips and Michael (1993) concluded that one explanation for the low resistivity and the low velocity with little effect on the gravity would be high pore pressure fluids. The serpentinite body inferred from the aeromagnetic data could provide a low permeability barrier that keeps fluids trapped.

The resistivity models from the Middle Mountain profiles (MM1, MM2 and MM3) (Figure 4.17) all have large volumes of conductive material ($1-10 \Omega - m$) in the Franciscan east of the SAF. Several inversions including the single mode inversions and many constrained inversion tests, indicate that the data indeed require this low resistivity feature. In the case of constrained inversions with fixed-narrow and deep fault zones beneath the fault trace (e.g., Figure 4.28), this low resistivity appears

to be located at somewhat shallower depth, between 1 and 5-km, entirely outside of the dense profile between the 2-km to 7-km marks. On the other hand, with no constraints on the fault zone, this low resistivity structure extends as deep as 10 km.

For the preferred model (Figure 4.15), the low resistivity zone has an average resistivity of $3.1 \Omega - m$ at shallow depths (0-3 km). If the conductivity is to be explained by fluids, porosity of around 20-30 %, similar to the estimated porosity in the near surface fault zone would be required. This high porosity should be observable in the gravity data or in significant reductions of seismic velocity. Evidences of both sorts presented by Eberhart-Phillips and Michael (1993) do not support such high porosities at these depths.

However, the aeromagnetic profile conducted across the SAF about 20 km north of MM1 shows a very strong magnetic object with magnetization value of 3.3 A/m, extended from the SAF to the Great Valley, with its top at around 1-2 km depth (Griscom and Jachens, 1990). This observation is similar to the aeromagnetic data analyzed by R.C. Jachens (published in Eberhart-Phillips and Michael, 1993). They interpreted this magnetic object as serpentinite. Michelini and McEvelly (1991) also observed a high V_p/V_s ratio (around 2) from east of the SAF to eastern end of profile (4.5 km), and at shallow depths.

Serpentinite can have low resistivity, but the general range of resistivity is around 10 to 100 $\Omega - m$ (Palacky, 1987) considerably higher than the resistivities observed by the MT data. Because of the 3-D complications in this area, along with the poor quality of data at station number 52 (first remote station on the east, directly above the shallow low resistivity zone), we believe that this shallow resistivity zone (resistivity in between 1-4 $\Omega - m$) may result from overfitting the data, at shallow depths. On the other hand, the geological map of Sims (1990)

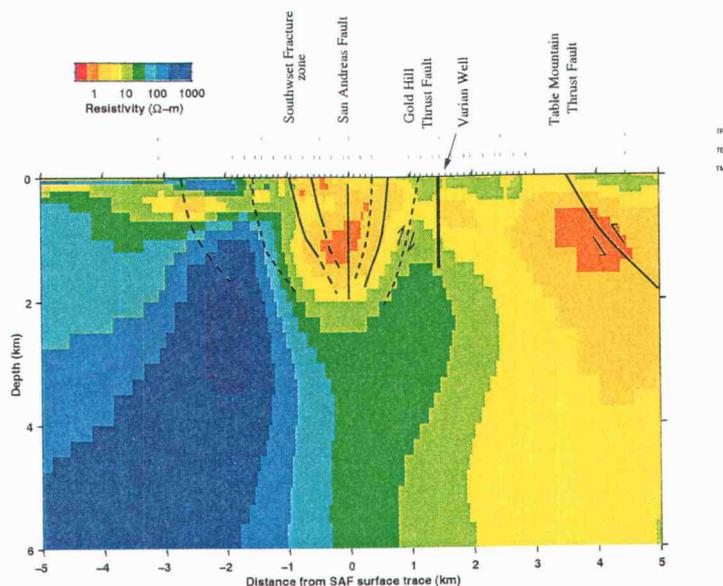


FIGURE 4.37. Resistivity model of MM1 profile with subsurface faults (Sims, 1990) plotted on top.

shows that about 3-4 km east of the SAF the Table Mountain thrust fault exposes the Franciscan assemblage to the surface (Figure 4.37). A saline fluid trapped in the thrust fault zone might be a factor for the shallow low resistivity in this region, similar to the SAF fault zone in the near surface.

The average resistivity over a wider area (3-7 km marks) and greater depth (0-10 km), is about $7 \Omega - m$. This resistivity would require porosity of about 5 to 10 % for $m = 1.5$ and 2, (Figure 4.36). This porosity might be more compatible with the low velocity of Eberhart-Phillips and Michael (1993). As discussed above, the data on the east have a stronger 3-D effect, and therefore is relatively less well constrained. The data can be fit with the low resistivity zone concentrated at a shallower depth (see Figure 4.26 and Figure 4.28), or spread out over great depths (Figure 4.31). Nevertheless, the data require a conductivity to the east of the SAF. The 3-D effects on the data must be further investigated.

4.8. Conclusions

The high resolution MT data collected across the Middle Mountain and the Cholame Valley are exhibits some 3-D effects. Overfitting the data in the area of the 3-D complication can result in geologically unreasonable interpretation. By increasing the error bars as suggested by the dimensionality analysis, we found that three main features are required by the data at Middle Mountain; the resistive Salinian granite overlain by a thin conductive sediment to the west, the large volume of low resistivity to the east consistent with the low velocity model, and the prism fault zone conductor with an average resistivity around $3 \Omega - m$, and a lateral gradient of resistivity which consistent with the lateral gradient of velocity.

Because of the difficulty of the inversion to resolve the deep and narrow conductive fault zone, constrained inversions with fixed conductive fault zone are tests. We found that at great depth the fault zone resistivity can be in the order of $3 \Omega - m$, and about 250-m width. The low resistivity fault zone, if interpreted to result from fluid, would requires a porosity in the range of of 8-18 %, consistent with other geophysical observations.

Three major observable differences between the Middle Mountain profiles and the Cholame Valley profile are the higher conductive fault zone in the Cholame Valley, the absence of the deep-low resistivity zone ($< 10\Omega - m$) in the east beneath the Cholame Valley, and the thicker near surface conductive sediment layer west of the SAF on top of the resistive Salinian unit. The conductivity of the fault zone at the Cholame Valley is unreasonable high, which might be the result from strong 3-D effect.

4.9. References

- [1] Atwater T., 1970, Implication of plate tectonics for the Cenozoic tectonic evolution of western North America, *Geol. Soc. Am. Bull.*, **81**, 4381-4399.
- [2] Atwater T., and P. Molnar, 1973, Relative motion of the Pacific and North American plates deduced from sea-floor spreading in the Atlantic, Indian, and South Pacific Oceans, *in*
- [3] Bakun W. H., and T. V. McEvelly, 1984, Recurrence models and Parkfield, California, earthquakes, *J. Geophys. Res.*, **89**, 3051-3058.
- [4] Brune J. N., T. L. Henyey and R. F. Roy, 1969, Heat flow, stress, and rate of slip along the San Andreas Fault, California, *J. Geophys. Res.*, **74**, 3821-3827.
- [5] Byerlee J., 1990, Friction, overpressure and fault normal compression, *Geophys. Res. Lett.*, **17**, 2109-2122.
- [6] Byerlee J., 1993, Model for episodic flow of high-pressure water in fault zones before earthquakes, *Geology*, **21**, 303-306.
- [7] Champion D. E., D. G. Howell, C. S. Gromme, 1984, Paleomagnetic and geologic data indicating 2500 km of northward displacement for the Salinian and related terranes, California, *J. Geophys. Res.*, **89**, 7736-7752.
- [8] Constable, C. S., Parker, R. L., and Constable, C.G., 1987, Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics*, **52**, 289-300.
- [9] Dickinson W. R., 1970, Relations of andesite, granites, and derived sandstones to arc-trench tectonics, *Revin Geophysics and Space Physics*, **8**, 813-860.
- [10] Eberhart-Phillips D., and A. J. Michael, 1993, Three-Dimensional Velocity Structure, Seismicity, and Fault Structure in the Parkfield Region, Central California, *J. Geophys. Res.*, **98**, 15737-15758.
- [11] Eberhart-Phillips D., W. D. Stanley, B. D. Rodriguez, and W. J. Lutter, 1995, Surface seismic and electrical methods to detect fluids related to faulting, *J. Geophys. Res.*, **100**, 12919-12936.
- [12] Egbert, G. D., and J. R. Booker, 1986, Robust estimation of geomagnetic transfer functions, *Geophys. J. R. astr. Soc.*, **87**, 173-194.
- [13] Egbert, G. D., 1997, Robust multiple station magnetotelluric data processing, *Geophys. J. Int.*, **130**, 475-496.
- [14] Egbert, G. D., M. Eisel, O. S. Boyd, and H. F. Morrison, 1999, Pc3's and DC Trains: A Causal Connection, *Geophys. Res. Lett.*, submitted.

- [15] Eisel M. and Egbert, 1999, On the Stability of Magnetotelluric Transfer Functions and the Reliability of their Variances, *Geophysics*, submitted.
- [16] Eisel M., G. D. Egbert, W. Siripunvaraporn and M. Unsworth, 1999, A Dimensionality Study of the Parkfield High Resolution EM Profiling Data: Implications for two-dimensional Inversions, *Geophys. Res. Lett.*, submitted.
- [17] Frapre S. K. and P. Fritz, 1987, Geochemical trends for groundwaters from the Canadian Shield. in *Saline Water and Gases in Crystalline Rocks* (eds. Fritz, P., and Frapre S.K.), *Geol. Ass. Can. Spec. Paper*, **33**, 19-38.
- [18] Fuis G. S., and W. D. Mooney, 1990, Lithospheric Structure and Tectonics from Seismic-Refraction and Other data in *The San Andreas Fault System, California* by R. W. Wallace, *U.S. Geol. Surv. Prof. Pap.*, **1515**, 206-236.
- [19] Fyfe W. S., N. J. Price, A. B. Thompson, 1978, *Fluids in the Earth's Crust*, Elsevier, New York.
- [20] Gamble, T. D., W. M. Goubau and J. Clarke, 1979, Magnetotellurics with a remote magnetic reference, *Geophysics*, **44**, 53-68.
- [21] Griscom A., and R. C. Jachens, 1990, Crustal and Lithospheric Structure from Gravity and Magnetic Studies in *The San Andreas Fault System, California* by R. W. Wallace, *U.S. Geol. Surv. Prof. Pap.*, **1515**, 238-259.
- [22] Groom R. W., and R. C. Bailey, 1989, Decomposition of Magnetotelluric Impedance Tensors in the Presence of Local Three-Dimensional Galvanic Distortion, *J. Geophys. Res.*, **94**, 1913-1925.
- [23] Hill D. P., J. P. Eaton and L. M. Jones, 1990, Seismicity, 1980-86, in *The San Andreas Fault System, California* by R. W. Wallace, *U.S. Geol. Surv. Prof. Pap.*, **1515**, 115-151.
- [24] Holbrook W. S., W. D. Mooney, and N. I. Christensen, 1992, The seismic velocity structure of the deep continental crust, *in* Fountain, D. M., Arculus, R. J., and Kay, R. W., Eds., *Continental Lower Crust*: Elsevier Science Publ. Co., Inc., 1-43.
- [25] Huenges E., J. Erzinger, J. Kuck, B. Engeser, and W. Kessels, 1997, The permeable crust: Geohydraulic properties down to 9101 m depth, *J. Geophys. Res.*, **102**, 18255-18265.
- [26] Hyndman R. D. and R. M. Shearer, 1989, Water in the lower continental crust: modelling magnetotelluric and seismic reflection results, *Geophys. J. Int.*, **98**, 343-365.

- [27] Hyndman R. D., L. L. Vanyan, G. Marquis, and L. K. Law, 1993, The origin of electrically conductive lower continental crust: saline water or graphite?, *Phys. Earth Planet. Inter.*, **81**, 325-344.
- [28] Irwin W. P., 1990, Geology and Plate-Tectonic Development, in *The San Andreas Fault System, California* by R. W. Wallace, U.S. Geol. Surv. Prof. Pap., **1515**, 61-80.
- [29] Johnson P. A., and T. V. McEvilly, 1995, Parkfield seismicity: Fluid-driven?, *J. Geophys. Res.*, **100**, 12937-12950.
- [30] Keller G. V., 1989, Electrical structure of the crust and upper mantle beneath the United States; Part 1, Methods for determining the conductivity profile, in *Geophysical framework of the continental United States* by Pakiser L. C., and Mooney, W.D., *Geolo. Soc. Am. Memoir* 172, 71-90.
- [31] Kurtz R. D., J. M. DeLaurier, and J. C. Gupta, 1986, A magnetotelluric sounding across Vancouver Island sees the subducting Juan de Fuca plate, *Nature*, **321**, 596-599.
- [32] Lachenbruch A. H., 1980, Frictional heating, fluid pressure and the resistance to fault motion, *J. Geophys. Res.*, **85**, 6097-6112.
- [33] Larsen J. C., 1977, Removal of local surface conductivity effects from low frequency mantle response curves, *Acta Geodæt. Geophys. Mont.*, **12**, 183-186.
- [34] Larson R. L., H. W. Menard, and S. M. Smith, 1968, Gulf of California: A result of ocean-floor spreading and transform fault, *Science*, **161**, 781-784.
- [35] Leary P. and Y. Ben-Zion, 1992, A 200 m Wide Fault Zone Low Velocity Layer on the San Andreas Fault at Parkfield: Results from Analytic Waveform Fits to Trapped Wave Groups, *Seis. Res. Lett.*, **63**, 62.
- [36] Li Y. G., P. C. Leary, K. Aki and P. E. Malin, 1990, Seismic Trapped Modes in the Oroville and San Andreas Fault Zones, *Science*, **249**, 763-766.
- [37] Lindh A. G., and D. M. Moore, 1981, Control of rupture by fault geometry during the 1966 Parkfield earthquake, *Bull. Seis. Soc. Am.*, **71**, 95-116.
- [38] Mackie, R. L, B. R. Bennett and T. R. Madden, 1988, Long-period magnetotelluric measurements near the central California coast: a land-locked view of the conductivity structure under the Pacific Ocean, *Geophys. J.*, **95**, 181-194.
- [39] Merzer M. and S. L. Klemperer, 1997, Modeling Low-frequency Magnetic-field Precursors to the Loma Prieta Earthquake with a Precursory Increase in Fault-zone Conductivity, *Pure and Appl. Geophys.*, **150**, 217-248.

- [40] McBride J. H., and L. D. Brown, 1986, Reanalysis of the COCORP deep seismic reflection profile across the San Andreas fault, Parkfield, California, *Bull. Seis. Soc. Am.*, **76**, 1668-1686.
- [41] McKenzie D. P., and W. J. Morgan, 1969, Evolution of triple junctions, *Nature*, **224**, 125-133.
- [42] Michelini A., and T. V. McEvilly, 1991, Seismological studies at Parkfield, I, Simultaneous inversion for velocity structure and hypocenters using cubic b-splines parameterization, *Bull. Seis. Soc. Am.*, **81**, 524-552.
- [43] Mooney W. D. and C. S. Weaver, 1989, Regional crustal structure and tectonics of the Pacific Coastal States; California, Oregon, and Washington, in *Geophysical framework of the continental United States* by Pakiser L. C., and Mooney, W.D., *Geol. Soc. Am. Memoir* 172, 129-160.
- [44] Nesbitt B. E., 1993, Electrical Resistivities of Crustal Fluids, *J. Geophys. Res.*, **98**, 4301-4310.
- [45] Page B. M., 1981, The Southern Coast Ranges, in *The Geotectonic Development of California*, Rubey Vol. I, by W. G. Ernst, 329-417, Prentice-Hall, Englewood Cliffs, N.J.
- [46] Palacky G. J., 1987, Resistivity characteristics of geologic targets, in *Electromagnetic Methods in Applied Geophysics—Theory*, vol. 1, by M. N. Nabighian, Society of Exploration Geophysicists, Tulsa, Okla, 53-129.
- [47] Park S. K., B. Hirasuna, G. R. Jiracek and C. Kinn, 1996, Magnetotelluric evidence of lithospheric mantle thinning beneath the southern Sierra Nevada, *J. Geophys. Res.*, **101**, 16241-16255.
- [48] Ranganayaki, R. P. and T. R. Madden, 1980, Generalized thin sheet analysis in magnetotellurics, and extension of Price's analysis, *Geophys. J. R. astr. Soc.*, **60**, 445-457.
- [49] Rice J. R., 1992, Fault stress states, pore pressure distributions, and the weakness of the San Andreas Fault, in *Fault Mechanics and Transport Properties of Rocks* by B. Evans and T.F. Wong, Academic Press, New York, 475-503.
- [50] Rodi, W. L., and Mackie, R. L., 1999, Nonlinear Conjugate Gradients Algorithm for 2-D Magnetotelluric Inversion, *Geophysics*, submitted.
- [51] Shedlock K. M., T. M. Brocher, and S. T. Harding, 1990, Shallow Structure and Deformation along the San Andreas Fault in Cholame Valley, California, Based on High-Resolution Reflection Profiling, *J. Geophys. Res.*, **95**, 5003-5020.

- [52] Smith J. T, and Booker, J. R., 1991, Rapid Inversion of Two- and Three-Dimensional Magnetotelluric Data, *J. Geophys. Res.*, **96**, 3905-3922.
- [53] Smith J. T, 1995, Understanding telluric distortion matrices, *Geophys. J. Int.*, **122**, 219-226.
- [54] Unsworth M. J., P. E. Malin, G. D. Egbert, and J. R. Booker, 1997, Internal structure of the San Andreas fault at Parkfield, California, *Geology*, **25**, 359-362.
- [55] Unsworth M. J., P. Bedrosian, J. R. Booker, G. D. Egbert, M. Eisel, W. Siripunvaraporn, 1998, Imaging Fault Zone Structure with Electromagnetic Exploration techniques, *EOS, Trans. Am. Geophys. Union*, **79**, 45, Fall Meet. Suppl., F227.
- [56] Unsworth M. J., G. D. Egbert, and J. R. Booker, 1999, High Resolution electromagnetic imaging of the San Andreas fault in Central California, *J. Geophys. Res.*, **104**, 1131-1150.
- [57] Sass J. H., C. F. Williams, A. H. Lachenbruch, S. P. Galanis Jr, and F. V. Grubb, 1997, Thermal regime of the San Andreas fault near Parkfield, California, *J. Geophys. Res.*, **102**, 27575-27585.
- [58] Scholz C. H., 1990, *The Mechanics of Earthquakes and Faulting*, Cambridge University Press.
- [59] Sims J. D., 1988, Geologic map of the San Andreas fault in the Cholame Valley and Cholame Hills quadrangles, Monterey and San Luis Obispo counties, California, scale 1:24,000, U.S. Geol. Surv. Misc. Field Stud. Map, MF-1995.
- [60] Sims J. D., 1990, Geologic map of the San Andreas fault in the Parkfield 7.5-minute quadrangle, Monterey and Fresno counties, California, scale 1:24,000, U.S. Geol. Surv. Misc. Field Stud. Map, MF-2115.
- [61] Siripunvaraporn, W., G. Egbert, H. F. Morrison and B. Wong, 1993, Effects of DC Electric Railway Noise on MT Measurements in the San Francisco Bay Area, XII Workshop on Electromagnetic Induction in the Earth, 104.
- [62] Siripunvaraporn, W., Egbert, G. D., Eisel, M., and Unsworth, M., 1998, A High Resolution EM Survey of the San Andreas Fault (SAF): Local Conductivity Structure in a Regional Context, *EOS, Trans. Am. Geophys. Union*, **79**, 45, Fall Meet. Suppl., F227.
- [63] Siripunvaraporn, W. and G. Egbert, 1999, REBOCC: An Efficient Data-Subspace Inversion for Two-Dimensional Magnetotelluric Data, *Geophysics*, submitted.

- [64] Sleep N. H. and M. L. Blanpied, 1992, Creep, compaction and the weak rheology of major faults, *Nature*, **359**, 687-692.
- [65] Thatcher, W., 1977, Secular deformation, episodic movements and relative plate motion in southern California, *EOS, Trans. Am. Geophys. Union*, **58**, 496.
- [66] Trehu A. M., and I. W. H. Wheeler, 1987, Possible evidence for subducted sedimentary materials beneath central California, *Geology*, **15**, 254-258.
- [67] Vozoff, K., 1972, The Magnetotelluric Method in the Exploration of Sedimentary Basins, *Geophysics*, **37**, 98-141.
- [68] Wang C. Y., 1984, On the Constitution of the San Andreas Fault Zone in Central California, *J. Geophys. Res.*, **89**, 5858-5866.
- [69] Wang C. Y., F. Rui, Y. Zhengsheng, and S. Xingjue, 1986, Gravity Anomaly and Density Structure of the San Andreas Fault Zone, *Pure and Appl. Geophys.*, **124**, 127-140.
- [70] Wannamaker P. E., G. W. Hohmann and S. H. Ward, 1984, Magnetotelluric responses of three-dimensional bodies in layered earths, *Geophysics*, **49**, 1517-1533.
- [71] Zoback M. D., M. L. Zoback, V. S. Mount, J. Suppe, J. P. Eaton, J. H. Healy, D. Oppenheimer, P. Reasenber, L. Jones, C. B. Raleigh, I. G. Wong, O. Scotti, and C. Wentworth, 1987, New Evidence on the State of Stress of the San Andreas Fault System, *Science*, **238**, 1105-1110.

5. CONCLUSION

In this dissertation, we present two main studies of the MT method. The first major study is the developments of numerical algorithms: the development of the 3-D forward modeling on a parallel computer and the development of a new 2-D inversion technique. These numerical tools were then used for interpreting the high resolution MT data collected across the San Andreas Fault at Parkfield, California.

An iterative 3-D MT modeling program with divergence correction has been implemented and developed on a massively parallel computer (CM-500e) which is available at the COAS. This results that large size problems can then be solved, which accurate results can be achieved within a few hundred of iterations. The algorithm is based on solving the second order of Maxwell's equation with the preconditioned conjugate gradient method, along with the divergence correction which dramatically improve an accuracy of the solutions. Different parallel preconditioners are applied to speed up the convergent rate of the relaxation method.

To improve the computational speed and reduce the memory required for 2-D MT inversions, we develop a new technique called REBOCC (reduced basis OCCAM's inversion), which has practically shown to be effective in term of convergence and computational costs. To a given data set, its speed is comparable to RRI but significantly faster than other methods, NLCCG and OCCAM. However, approximate approaches (e.g., RRI) which is very fast, but in practice often fail to converge without significant expert user intervention.

REBOCC scheme is based on transforming the linearized inverse problem from the M -dimensional model space to the N -dimensional data space, by expressing the solution as a linear combination of rows of the sensitivity matrix smoothed by the model covariance (the representers). Since generally $N \ll M$, this transformation

by itself can result in significant computational saving. More importantly the data space formulation suggests a simple approximate method for constructing the inverse solution. Since MT data are smooth and "redundant", a subset of the representers is typically sufficient to form the model without significant loss of detail. Computations required for constructing sensitivities and the size of matrices to be inverted can be significantly reduced by this approximation. The basic idea behind REBOCC should be more broadly applicable, in particular to 3-D MT inversion.

In the last chapter, it is the analysis and interpretation of the high resolution MT data collected across the Middle Mountain and the Cholame Valley at the Parkfield area. We found that three main features are required by the data at Middle Mountain; the resistive Salinian granite overlain by a thin conductive sediment to the west, the large volume of low resistivity to the east consistent with the low velocity model, and the prism fault zone conductor, and a lateral gradient of resistivity which consistent with the lateral gradient of velocity. Three major observable differences between the Middle Mountain profiles and the Cholame Valley profile are the higher conductive fault zone in the Cholame Valley, the absence of the deep-low resistivity zone ($< 10\Omega - m$) in the east beneath the Cholame Valley, and the thicker near surface conductive sediment layer west of the SAF on top of the resistive Salinian unit. The conductivity of the fault zone at the Cholame Valley is unreasonable high, which might be the result from strong 3-D effect. Because of the difficulty of the inversion to resolve the deep and narrow conductive fault zone, constrained inversions with fixed conductive fault zone are tests. We found that at great depth the fault zone resistivity can be in the order of $3\Omega - m$, and about 250-m width at the Middle Mountain. The low resistivity fault zone, if interpreted to result from fluid, would requires a porosity in the range of of 8-18 %, consistent with other geophysical observations.

BIBLIOGRAPHY

- [1] Atwater T., 1970, Implication of plate tectonics for the Cenozoic tectonic evolution of western North America, *Geolo. Soc. Am Bull.*, **81**, 4381-4399.
- [2] Atwater T., and P. Molnar, 1973, Relative motion of the Pacific and North American plates deduced from sea-floor spreading in the Atlantic, Indian, and South Pacific Oceans, *in*
- [3] Barret, R., M. Berry, T.F. Chan, J. Demmel, J.M. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine and H. van der Vorst, 1994. Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, *SIAM*, Philadelphia.
- [4] Bakun W. H., and T. V. McEvilly, 1984, Recurrence models and Parkfield, California, earthquakes, *J. Geophys. Res.*, **89**, 3051-3058.
- [5] Bennett, A. F., Chua, B. S., and Leslie, L. M., 1996, Generalized inversion of a global numerical weather prediction model, *Meteorology and Atmospheric Physics*, **60**, 165-178.
- [6] Brune J. N., T. L. Henyey and R. F. Roy, 1969, Heat flow, stress, and rate of slip along the San Andreas Fault, California, *J. Geophys. Res.*, **74**, 3821-3827.
- [7] Byerlee J., 1990, Friction, overpressure and fault normal compression, *Geophys. Res. Lett.*, **17**, 2109-2122.
- [8] Byerlee J., 1993, Model for episodic flow of high-pressure water in fault zones before earthquakes, *Geology*, **21**, 303-306.
- [9] Cagniard, L., 1953, Basic Theory of the Magnetotelluric Method of Geophysical Prospecting, *Geophysics*, **18**, 605-635.
- [10] Champion D. E., D. G. Howell, C. S. Gromme, 1984, Paleomagnetic and geologic data indicating 2500 km of northward displacement for the Salinian and related terranes, California, *J. Geophys. Res.*, **89**, 7736-7752.
- [11] Chen, L., Booker, J. R., Jones, A. G., Wu, N., Unsworth, M., Wei, W., and Tan, H., 1996, Electrically conductive crust in southern Tibet from INDEPTH magnetotelluric surveying, *Science*, **274**, 1694-1696.
- [12] Constable, C. S., Parker, R. L., and Constable, C.G., 1987, Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics*, **52**, 289-300.

- [13] deGroot-Hedlin, C., and Constable, S., 1990, Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data, *Geophysics*, **55**, 1613-1624.
- [14] deGroot-Hedlin, C., and Constable, S., 1993, Occam's Inversion and the North American Central Plains Electrical Anomaly, *J. Geomag. Geoelectr.*, **45**, 985-999.
- [15] Dickinson W. R., 1970, Relations of andesite, granites, and derived sandstones to arc-trench tectonics, *Revin Geophysics and Space Physics*, **8**, 813-860.
- [16] Dobrin, M.B. and C.H. Savit, 1988, *Introduction to Geophysical Prospecting* 4th, McGraw-Hill Inc.
- [17] Eberhart-Phillips D., and A. J. Michael, 1993, Three-Dimensional Velocity Structure, Seismicity, and Fault Structure in the Parkfield Region, Central California, *J. Geophys. Res.*, **98**, 15737-15758.
- [18] Eberhart-Phillips D., W. D. Stanley, B. D. Rodriguez, and W. J. Lutter, 1995, Surface seismic and electrical methods to detect fluids related to faulting, *J. Geophys. Res.*, **100**, 12919-12936.
- [19] Egbert, G. D., and J. R. Booker, 1986, Robust estimation of geomagnetic transfer functions, *Geophys. J. R. astr. Soc.*, **87**, 173-194.
- [20] Egbert, G. D., and J. R. Booker, 1989, Multivariate analysis of geomagnetic array data, 1, The response space, *J. Geophys. Res.*, **94**, 14227-14247.
- [21] Egbert, G.D. and J.R. Booker, 1992, Very Long Period Magnetotellurics at Tucson Observatory: Implications for Mantle Conductivity, *J. Geophys. Res.*, **97**, 15,099-15,112.
- [22] Egbert, G. D., A. F. Bennett and M. G. Foreman, 1994, TOPEX/POSEIDON tides estimated using a global inverse model, *J. Geophys. Res.*, **99**, 24821-24852.
- [23] Egbert, G. D., 1997, Robust multiple station magnetotelluric data processing, *Geophys. J. Int.*, **130**, 475-496.
- [24] Egbert, G. D., 1997, Tidal data inversion: interpolation and inference, *Prog. Oceanog.*, **40**, 53-80.
- [25] Egbert, G. D., M. Eisel, O. S. Boyd, and H. F. Morrison, 1999, Pc3's and DC Trains: A Causal Connection, *Geophys. Res. Lett.*, submitted.
- [26] Eisel M. and Egbert, 1999, On the Stability of Magnetotelluric Transfer Functions and the Reliability of their Variances, *Geophysics*, submitted.

- [27] Eisel M., G. D. Egbert, W. Siripunvaraporn and M. Unsworth, 1999, A Dimensionality Study of the Parkfield High Resolution EM Profiling Data: Implications for two-dimensional Inversions, *Geophys. Res. Lett.*, submitted.
- [28] EMSLAB, 1989, Special Issue of Journal of Geophysical Reserach devoted to the EMSLAB experiment (14 papers)., *J. Geophys. Res.*, **94**, 14093-14283.
- [29] Farquharson, C. G., and Oldenburg, D. W., 1996, Approximate sensitivities for the electromagnetic inverse problem, *Geophys. J. Int.*, **126**, 235-252.
- [30] Fletcher, R., 1975, Conjugate gradient methods for indefinite systems in Numerical Analysis by A. Dold and B. Eckmann, Springer, New York.
- [31] Frape S. K. and P. Fritz, 1987, Geochemical trends for groundwaters from the Canadian Shield. in *Saline Water and Gases in Crystalline Rocks* (eds. Fritz, P., and Frape S.K.), *Geol. Ass. Can. Spec. Paper*, **33**, 19-38.
- [32] Fuis G. S., and W. D. Mooney, 1990, Lithospheric Structure and Tectonics from Seismic-Refraction and Other data in *The San Andreas Fault System, California* by R. W. Wallace, *U.S. Geol. Surv. Prof. Pap.*, **1515**, 206-236.
- [33] Fyfe W. S., N. J. Price, A. B. Thompson, 1978, *Fluids in the Earth's Crust*, Elsevier, New York.
- [34] Gamble, T. D., W. M. Goubau and J. Clarke, 1979, Magnetotellurics with a remote magnetic reference, *Geophysics*, **44**, 53-68.
- [35] Griscom A., and R. C. Jachens, 1990, Crustal and Lithospheric Structure from Gravity and Magnetic Studies in *The San Andreas Fault System, California* by R. W. Wallace, *U.S. Geol. Surv. Prof. Pap.*, **1515**, 238-259.
- [36] Groom R. W., and R. C. Bailey, 1989, Decomposition of Magnetotelluric Impedance Tensors in the Presence of Local Three-Dimensional Galvanic Distortion, *J. Geophys. Res.*, **94**, 1913-1925.
- [37] Gupta, A., V. Kumar, and A. Sameh, 1995, Performance and Scalability of Preconditioned Conjugate Gradient Methods on Parallel Computers, *IEEE Trans. Parallel and Distrib. System.*, **6**, 455-469.
- [38] Hill D. P., J. P. Eaton and L. M. Jones, 1990, Seismicity, 1980-86, in *The San Andreas Fault System, California* by R. W. Wallace, *U.S. Geol. Surv. Prof. Pap.*, **1515**, 115-151.
- [39] Holbrook W. S., W. D. Mooney, and N. I. Christensen, 1992, The seismic velocity structure of the deep continental crust, *in* Fountain, D. M., *Arculus*,

- R. J., and Kay, R. W., Eds., *Continental Lower Crust*: Elsevier Science Publ. Co., Inc., 1-43.
- [40] Huenges E., J. Erzinger, J. Kuck, B. Engeser, and W. Kessels, 1997, The permeable crust: Geohydraulic properties down to 9101 m depth, *J. Geophys. Res.*, **102**, 18255-18265.
- [41] Hyndman R. D. and R. M. Shearer, 1989, Water in the lower continental crust: modelling magnetotelluric and seismic reflection results, *Geophys. J. Int.*, **98**, 343-365.
- [42] Hyndman R. D., L. L. Vanyan, G. Marquis, and L. K. Law, 1993, The origin of electrically conductive lower continental crust: saline water or graphite?, *Phys. Earth Planet. Inter.*, **81**, 325-344.
- [43] Irwin W. P., 1990, Geology and Plate-Tectonic Development, in *The San Andreas Fault System, California* by R. W. Wallace, U.S. Geol. Surv. Prof. Pap., **1515**, 61-80.
- [44] Jacobs, D. A. H., 1986, A generalization of the conjugate-gradient method to solve complex systems, *IMA J. Numer. Anal.*, **6**, 447-452.
- [45] Johnson P. A., and T. V. McEvelly, 1995, Parkfield seismicity: Fluid-driven?, *J. Geophys. Res.*, **100**, 12937-12950.
- [46] Jones, A. G., 1992, Electrical conductivity of the continental lower crust, *in* Fountain, D. M., Arculus, R. J., and Kay, R. W., Eds., *Continental Lower Crust*: Elsevier Science Publ. Co., Inc., 81-143.
- [47] Jones, A. G., A. D. Chave, G. Egbert, D. Auld, and K. Bahr, 1989, A Comparison of Techniques for Magnetotelluric Response Function Estimation, *J. Geophys. Res.*, **94**, 14201-14213.
- [48] Jupp, D. L. B., and Vozoff, K., 1975, Stable iterative methods for the inversion of geophysical data, *Geophys. J. R. astr. Soc.*, **42**, 957-976.
- [49] Keller G. V., 1987, Rock and Mineral Properties, in *Electromagnetic Methods in Applied Geophysics—Theory*, vol. 1, by M. N. Nabighian, Society of Exploration Geophysicists, Tulsa, Okla, 13-52.
- [50] Keller G. V., 1989, Electrical structure of the crust and upper mantle beneath the United States; Part 1, Methods for determining the conductivity profile, in *Geophysical framework of the continental United States* by Pakiser L. C., and Mooney, W.D., *Geolo. Soc. Am. Memoir*172, 71-90.

- [51] Kershaw, D.S., 1978, The Incomplete Cholesky-Conjugate Gradient Method for the Iterative Solution of Systems of Linear Equations, *J. Comput. Phys.*, **26**, 43-65.
- [52] Kurtz R. D., J. M. DeLaurier, and J. C. Gupta, 1986, A magnetotelluric sounding across Vancouver Island sees the subducting Juan de Fuca plate, *Nature*, **321**, 596-599.
- [53] Lachenbruch A. H., 1980, Frictional heating, fluid pressure and the resistance to fault motion, *J. Geophys. Res.*, **85**, 6097-6112.
- [54] Larsen J. C., 1977, Removal of local surface conductivity effects from low frequency mantle response curves, *Acta Geodæt. Geophys. Mont.*, **12**, 183-186.
- [55] Larson R. L., H. W. Menard, and S. M. Smith, 1968, Gulf of California: A result of ocean-floor spreading and transform fault, *Science*, **161**, 781-784.
- [56] Leary P. and Y. Ben-Zion, 1992, A 200 m Wide Fault Zone Low Velocity Layer on the San Andreas Fault at Parkfield: Results from Analytic Waveform Fits to Trapped Wave Groups, *Seis. Res. Lett.*, **63**, 62.
- [57] Li Y. G., P. C. Leary, K. Aki and P. E. Malin, 1990, Seismic Trapped Modes in the Oroville and San Andreas Fault Zones, *Science*, **249**, 763-766.
- [58] Lindh A. G., and D. M. Moore, 1981, Control of rupture by fault geometry during the 1966 Parkfield earthquake, *Bull. Seis. Soc. Am.*, **71**, 95-116.
- [59] Mackie, R. L., B. R. Bennett and T. R. Madden, 1988, Long-period magnetotelluric measurements near the central California coast: a land-locked view of the conductivity structure under the Pacific Ocean, *Geophys. J.*, **95**, 181-194.
- [60] Mackie, R.L., T.R. Madden and P.E. Wannamake, 1993, Three-dimensional magnetotelluric modeling using difference equations: Theory and comparisons to integral equation solutions, *Geophysics*, **58**, 215-226.
- [61] Mackie, R.L and T.R. Madden, 1993, Conjugate direction relaxation solutions for 3-D magnetotelluric modeling, *Geophysics*, **58**, 1052-1057.
- [62] Mackie, R.L., J.T. Smith and T.R. Madden, 1994, Three-dimensional electromagnetic modeling using finite difference equations: The magnetotelluric example, *Radio Science*, **29**, 923-935.
- [63] Marquardt, D. W., 1963, An Algorithm for least-squares estimation of nonlinear parameters, *J. Soc. Indust. Appl. Math.*, **11**, 431-441.

- [64] Merzer M. and S. L. Klemperer, 1997, Modeling Low-frequency Magnetic-field Precursors to the Loma Prieta Earthquake with a Precursory Increase in Fault-zone Conductivity, *Pure and Appl. Geophys.*, **150**, 217-248.
- [65] McBride J. H., and L. D. Brown, 1986, Reanalysis of the COCORP deep seismic reflection profile across the San Andreas fault, Parkfield, California, *Bull. Seis. Soc. Am.*, **76**, 1668-1686.
- [66] McKenzie D. P., and W. J. Morgan, 1969, Evolution of triple junctions, *Nature*, **224**, 125-133.
- [67] Michelini A., and T. V. McEvelly, 1991, Seismological studies at Parkfield, I, Simultaneous inversion for velocity structure and hypocenters using cubic b-splines parameterization, *Bull. Seis. Soc. Am.*, **81**, 524-552.
- [68] Mooney W. D. and C. S. Weaver, 1989, Regional crustal structure and tectonics of the Pacific Coastal States; California, Oregon, and Washington, in *Geophysical framework of the continental United States* by Pakiser L. C., and Mooney, W.D., *Geolo. Soc. Am. Memoir* 172, 129-160.
- [69] Nesbitt B. E., 1993, Electrical Resistivities of Crustal Fluids, *J. Geophys. Res.*, **98**, 4301-4310.
- [70] Oldenburg, D. W., and Ellis, R. G., 1991, Inversion of geophysical data using an approximate inverse mapping, *Geophys. J. Int.*, **105**, 325-353.
- [71] Oldenburg, D. W., McGillivray, P. R., and Ellis, R. G., 1993, Generalized subspace methods for large scale inverse problems, *Geophys. J. Int.*, **114**, 12-20.
- [72] Oldenburg, D. W., and Ellis, R. G., 1993, Efficient inversion of magnetotelluric data in two dimensions, *Phys. Earth Planet. Inter.*, **81**, 177-200.
- [73] Ogawa, Y. and Uchida, T., 1996, A two-dimensional magnetotelluric inversion assuming Gaussian static shift, *Geophysics*, **126**, 69-76.
- [74] Orange, A. S., 1989, Magnetotelluric Exploration for Hydrocarbons, *Proc. IEEE*, **77**, 287-317.
- [75] Page B. M., 1981, The Southern Coast Ranges, in *The Geotectonic Development of California*, Rubey Vol. I, by W. G. Ernst, 329-417, Prentice-Hall, Englewood Cliffs, N.J.
- [76] Palacky G. J., 1987, Resistivity characteristics of geologic targets, in *Electromagnetic Methods in Applied Geophysics—Theory*, vol. 1, by M. N. Nabighian, Society of Exploration Geophysicists, Tulsa, Okla, 53-129.

- [77] Park S. K., B. Hirasuna, G. R. Jiracek and C. Kinn, 1996, Magnetotelluric evidence of lithospheric mantle thinning beneath the southern Sierra Nevada, *J. Geophys. Res.*, **101**, 16241-16255.
- [78] Parker, R. L., 1980, The inverse problem of electromagnetic induction: existence and construction of solutions based upon incomplete data, *J. Geophys. Res.*, **85**, 4421-4425.
- [79] Parker, R. L., and Shure, L., 1982, Efficient Modeling of the Earth's Magnetic Field with Harmonics Splines, *Geophys. Res. Lett.*, **9**, 812-815.
- [80] Parker, R. L., 1994, *Geophysical Inverse Theory*: Princeton Univ. Press.
- [81] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., 1992, *NUMERICAL RECIPES in FORTRAN: The Art of Scientific Computing* - 2nd edition: Cambridge Univ. Press.
- [82] Ranganayaki, R. P. and T. R. Madden, 1980, Generalized thin sheet analysis in magnetotellurics, and extension of Price's analysis, *Geophys. J. R. astr. Soc.*, **60**, 445-457.
- [83] Rice J. R., 1992, Fault stress states, pore pressure distributions, and the weakness of the San Andreas Fault, in *Fault Mechanics and Transport Properties of Rocks* by B. Evans and T.F. Wong, Academic Press, New York, 475-503.
- [84] Rodi, W. L., 1976, A technique for improving the accuracy of finite element solutions for magnetotelluric data, *Geophys. J. R. astr. Soc.*, **44**, 483-506.
- [85] Rodi, W. L., and Mackie, R. L., 1999, Nonlinear Conjugate Gradients Algorithm for 2-D Magnetotelluric Inversion, *Geophysics*, submitted.
- [86] Sass J. H., C. F. Williams, A. H. Lachenbruch, S. P. Galanis Jr, and F. V. Grubb, 1997, Thermal regime of the San Andreas fault near Parkfield, California, *J. Geophys. Res.*, **102**, 27575-27585.
- [87] Scholz C. H., 1990, *The Mechanics of Earthquakes and Faulting*, Cambridge University Press.
- [88] Shedlock K. M., T. M. Brocher, and S. T. Harding, 1990, Shallow Structure and Deformation along the San Andreas Fault in Cholame Valley, California, Based on High-Resolution Reflection Profiling, *J. Geophys. Res.*, **95**, 5003-5020.
- [89] Sims J. D., 1988, Geologic map of the San Andreas fault in the Cholame Valley and Cholame Hills quadrangles, Monterey and San Luis Obispo counties, California, scale 1:24,000, U.S. Geol. Surv. Misc. Field Stud. Map, MF-1995.

- [90] Sims J. D., 1990, Geologic map of the San Andreas fault in the Parkfield 7.5-minute quadrangle, Monterey and Fresno counties, California, scale 1:24,000, U.S. Geol. Surv. Misc. Field Stud. Map, MF-2115.
- [91] Siripunvaraporn, W., G. Egbert, H. F. Morrison and B. Wong, 1993, Effects of DC Electric Railway Noise on MT Measurements in the San Francisco Bay Area, XII Workshop on Electromagnetic Induction in the Earth, 104.
- [92] Siripunvaraporn, W., Egbert, G. D., Eisel, M., and Unsworth, M., 1998, A High Resolution EM Survey of the San Andreas Fault (SAF): Local Conductivity Structure in a Regional Context, EOS, Trans. Am. Geophys. Union, **79**, 45, Fall Meet. Suppl., F227.
- [93] Siripunvaraporn, W. and G. Egbert, 1999, REBOCC: An Efficient Data-Subspace Inversion for Two-Dimensional Magnetotelluric Data, Geophysics, submitted.
- [94] Sleep N. H. and M. L. Blanpied, 1992, Creep, compaction and the weak rheology of major faults, *Nature*, **359**, 687-692.
- [95] Smith J. T, and Booker, J. R., 1988, Magnetotelluric inversion for minimum structure, *Geophysics*, **53**, 1565-1576.
- [96] Smith J. T, and Booker, J. R., 1991, Rapid Inversion of Two- and Three-Dimensional Magnetotelluric Data, *J. Geophys. Res.*, **96**, 3905-3922.
- [97] Smith J. T, 1995, Understanding telluric distortion matrices, *Geophys. J. Int.*, **122**, 219-226.
- [98] Smith, J.T., 1996, Conservative modeling of 3-D electromagnetic fields, Part I: Properties and error analysis, *Geophysics*, **61**, 1308-1318.
- [99] Smith, J.T., 1996, Conservative modeling of 3-D electromagnetic fields, Part II: Biconjugate gradient solution and an accelerator, *Geophysics*, **61**, 1319-1324.
- [100] Spies B. R., 1989, Depth of investigation in electromagnetic sounding methods, *Geophysics*, **54**, 872-888.
- [101] Thatcher, W., 1977, Secular deformation, episodic movements and relative plate motion in southern California, EOS, Trans. Am. Geophys. Union, **58**, 496.
- [102] Tikhonov, A.V., 1950, Determination of Electrical Characteristics of the Deep Strata of the Earth's Crust, *Dokl. Akad. Nauk*, **73**, 295.
- [103] Trehu A. M., and I. W. H. Wheeler, 1987, Possible evidence for subducted sedimentary materials beneath central California, *Geology*, **15**, 254-258.

- [104] Uchida, T., 1993, Smooth 2-D Inversion for magnetotelluric data based on statistical criterion ABIC, *J. Geomag. Geoelectr.*, **45**, 841-858.
- [105] Unsworth M. J., P. E. Malin, G. D. Egbert, and J. R. Booker, 1997, Internal structure of the San Andreas fault at Parkfield, California, *Geology*, **25**, 359-362.
- [106] Unsworth M. J., P. Bedrosian, J. R. Booker, G. D. Egbert, M. Eisel, W. Siripunvaraporn, 1998, Imaging Fault Zone Structure with Electromagnetic Exploration techniques, EOS, Trans. Am. Geophys. Union, **79**, 45, Fall Meet. Suppl., F227.
- [107] Unsworth M. J., G. D. Egbert, and J. R. Booker, 1999, High Resolution electromagnetic imaging of the San Andreas fault in Central California, *J. Geophys. Res.*, **104**, 1131-1150.
- [108] Vozoff, K., 1972, The Magnetotelluric Method in the Exploration of Sedimentary Basins, *Geophysics*, **37**, 98-141.
- [109] Vozoff, K., ed., 1986, Magnetotelluric Methods. Soc. Explor. Geophys. Reprint Ser. No. 5, Publ. by Soc. Expl. Geophys., Tulsa, Okla., ISBN 0-931830-36-2.
- [110] Wahba, G., 1990, Spline Methods for Observational Data, Soc. Ind. Appl. Math.
- [111] Wang C. Y., 1984, On the Constitution of the San Andreas Fault Zone in Central California, *J. Geophys. Res.*, **89**, 5858-5866.
- [112] Wang C. Y., F. Rui, Y. Zhengsheng, and S. Xingjue, 1986, Gravity Anomaly and Density Structure of the San Andreas Fault Zone, *Pure and Appl. Geophys.*, **124**, 127-140.
- [113] Wannamaker P. E., G. W. Hohmann and S. H. Ward, 1984, Magnetotelluric responses of three-dimensional bodies in layered earths, *Geophysics*, **49**, 1517-1533.
- [114] Wannamaker, P. E., Stodt, J. A., and Rijo, L., 1986, Two-dimensional topographic responses in magnetotelluric modeled using finite elements, *Geophysics*, **51**, 2131-2144.
- [115] Wannamaker, P. E., Booker, J. R., Filloux, J. H., Jones, A. G., Jiracek, G. R., Chave, A. D., Tarits, P., Waff, H. S., Egbert, G. D., Young, C. T., Stodt, J. A., Martinez, M. G., Lwaw, L. K., Yukutake, T., Segawa, J. S., White, A., and Green, A. W., Jr, 1994, Magnetotelluric Observations Across the Juan de Fuca Subduction System in the EMSLAB Project, *J. Geophys. Res.*, **94**, 14111-14125.

- [116] Weidelt, P., 1972, The inverse problem of geomagnetic induction, *Z. Geophys.*, **38**, 257-289.
- [117] Wu, N., Booker, J. R., and Smith, J. T., 1993, Rapid Two-Dimensional Inversion of COPROD2 Data, *J. Geomag. Geoelectr.*, **45**, 1073-1087.
- [118] Zhdanov M. S. and G. V. Keller, 1994, *The Geoelectrical Methods in Geophysical Exploration*, Elsevier.
- [119] Zoback M. D., M. L. Zoback, V. S. Mount, J. Suppe, J. P. Eaton, J. H. Healy, D. Oppenheimer, P. Reasenber, L. Jones, C. B. Raleigh, I. G. Wong, O. Scotti, and C. Wentworth, 1987, New Evidence on the State of Stress of the San Andreas Fault System, *Science*, **238**, 1105-1110.