LBL-31377

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

EARTH SCIENCES DIVISION

Imaging Earth's Interior: Tomographic Inversions for Mantle P-wave Velocity Structure

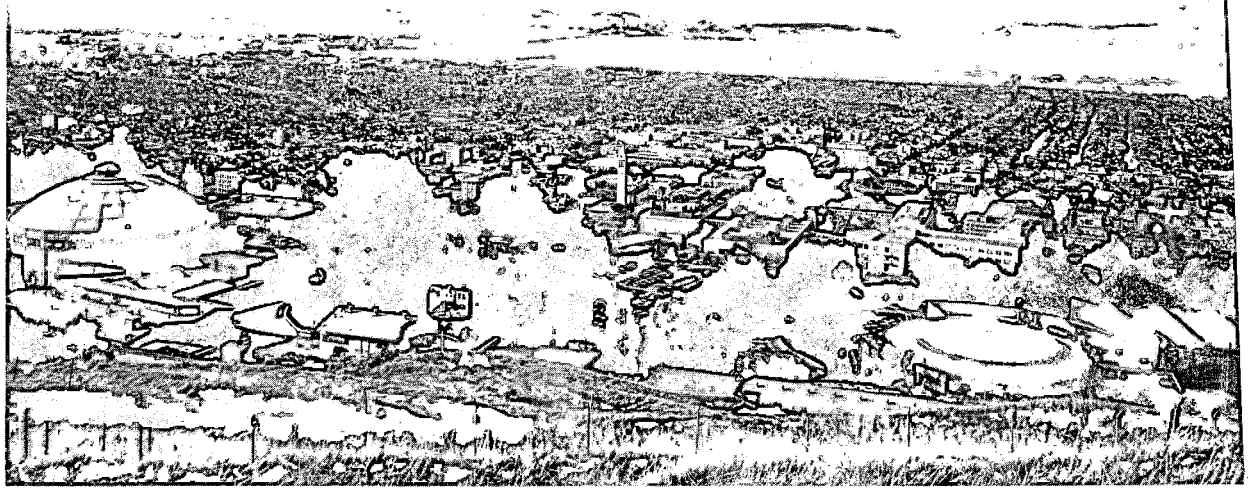
R.J. Pulliam
(Ph.D. Thesis)

July 1991

U. C. Lawrence Berkeley Laboratory
Library, Berkeley

FOR REFERENCE

Not to be taken from this room



LBL-31377
Copy 1
Bldg. 50 Library.

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. Neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California and shall not be used for advertising or product endorsement purposes.

Lawrence Berkeley Laboratory is an equal opportunity employer.

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

LBL-31377

Imaging Earth's Interior: Tomographic Inversions for Mantle P-wave Velocity Structure

Robert Jay Pulliam

(Ph.D. Thesis)

Department of Geology and Geophysics
University of California

and

Earth Sciences Division
Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

July 1991

This work was supported by the Director, Office of Energy Research,
Office of Basic Energy Sciences, Engineering and Geosciences Division,
of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Copyright © 1991 by Robert Jay Pulliam

The United States Department of Energy has the right to use this thesis
for any purpose whatsoever, including the right to reproduce all or any part thereof.

Imaging Earth's Interior: Tomographic Inversions for Mantle P-wave Velocity Structure

by

Robert Jay Pulliam

Abstract

A formalism is developed for the tomographic inversion of seismic travel time residuals. The travel time equations are solved both simultaneously, for velocity model terms and corrections to the source locations, and progressively, for each set of terms in succession. The methods differ primarily in their treatment of source mislocation terms. Additionally, the system of equations is solved directly, neglecting source terms. The efficacy of the algorithms is explored with synthetic data as we perform simulations of the general procedure used to produce tomographic images of Earth's mantle from global earthquake data.

The patterns of seismic heterogeneity in the mantle that would be returned reliably by a tomographic inversion are investigated. We construct synthetic data sets based on real ray sampling of the mantle by introducing spherical harmonic patterns of velocity heterogeneity and perform inversions of the synthetic data.

Inversions of real data, supplied by the ISC, are also performed. We use P arrival data from January 1964 through January 1987 and our inversion algorithms to solve for three-dimensional P velocity models of the mantle and source mislocations. The three-dimensional velocity model is presented in conjunction with the resolution estimates produced by an inversion for a checkerboard test pattern. Covariance is

estimated by averaging results from inversions of realistic errors and by a jackknife procedure.

The three-dimensional velocity model shows a fast anomaly in the lower mantle beneath the Tonga-New Hebrides subduction zone to a depth of 1670 km and another fast anomaly beneath the Japanese Island arc and eastern Asia reaching nearly to the core-mantle boundary. Continuity between these anomalies and shallower fast anomalies is not clear. A fast anomaly extending from 670 km to 2070 km depth appears beneath the eastern United States, Caribbean Sea, and Central South America. In addition, a number of slow anomalies associated with hotspots extend through the upper mantle but are extinguished in the lower mantle by our resolution weighting. Mid-ocean ridges are associated with moderately slow anomalies in the top 400 km of our model. The transition zone between depths of 400 and 670 km shows large l=1, 2, and 3 spherical harmonic components.

Acknowledgments

Many people helped to make my studies at Cal and my life in Berkeley my most fun and rewarding years. I find myself the grateful victim of a conspiracy of generosity and kindness. My peers in the seismology group and in the Department of Geology and Geophysics were helpful and supportive from beginning to end. At the outset, Jonathan Scheiner, Dave Tralli, Don Vasco, John Peterson, Phil Cummins, Bob Darragh, Dan O'Connell, Mary Templeton and Norm Abrahamson all made me feel welcome and helped with advice, ideas, and computer programs. On top of that they all sat me down at various times and explained a lot of things to me. Later, Joel Ita, Ann Kirkpatrick, Ali Tura, and Bill Foxall joined the fray at LBL. Sadly, none proved to be as apt or interested a table tennis player as their predecessors, so our late nights tended to be filled with seismology and other geophysical games. At the Earth Science Division of the Lawrence Berkeley Laboratory, Tom Daley, Eleni Karageorgi, and Rich Clymer all brightened my life at LBL and made sure I was exposed to work far removed from my own. All along the way, Mike Leonard and Alberto Michelini coaxed and cajoled me to completion. No words can say how much I appreciate their being here. Most recently, Paul Earle provided invaluable technical support, help, and scientific suggestions. I thank him for showing me more respect at work than he does on a basketball court.

Wes Bethel and Nancy Johnston of the Information and Computer Sciences Division of LBL gave profoundly of their technical support with computer graphics. Their efforts led to the most fun I had in the course of this work: looking at models upside down, inside out, and sideways.

Working closely with Don Vasco the last two years has been the most valuable and productive experience of my graduate career. I thank him for his help and encouragement and friendship.

Special thanks go to Tom McEvilly, Ernie Majer, and Lane Johnson for their efforts and success at setting up and maintaining the Center for Computational Seismology at LBL. These are world-class facilities without which the projects documented in this thesis would have been impossible to carry out. In my time here I witnessed the arrival and departure of several generations of computers and the replacement or addition of all types of support facilities. This upgrading of machines and expansion of computing capacity, along with the helpful Division staff and financial support offered to students indicates the seriousness with which Tom, Ernie, Lane, and many others approach the task of trying to create a good environment for research and learning. They have succeeded in this task and I appreciate their efforts. More than for the physical facilities, I am grateful for the faith Lane Johnson has shown in me and my abilities through the entire ordeal.

I thank the Associated Western Universities and the Department of Energy, ARCS, and Chevron for the financial support which made my stay in graduate school as comfortable as can be expected.

Lane Johnson, Tom McEvilly, Barbara Romanowicz, and Philip Stark, members of my thesis committee, all read my manuscript carefully and their suggestions greatly improved the content and readability of the product. However, any shortcomings in the work that follows are entirely my responsibility.

I owe debts to all the people I met in in graduate school at Berkeley and elsewhere for the ideas they shared with me and the skills they taught me. I expect these ideas and skills to make my professional life successful. However, I owe my most fundamental skills, values, and ideas to my parents and sisters, whose support and love sustain me through greater challenges than this thesis. Though my family cannot understand it as can many of my peers in geophysics, this thesis is for them anyway. They can appreciate it just as much, perhaps more, because they understand the well from which it springs.

Table of Contents

Chapter 1: Introduction 1
1.1 Review of Previous Work 2
1.2 Topics Investigated in This Thesis 8
1.3 References 11

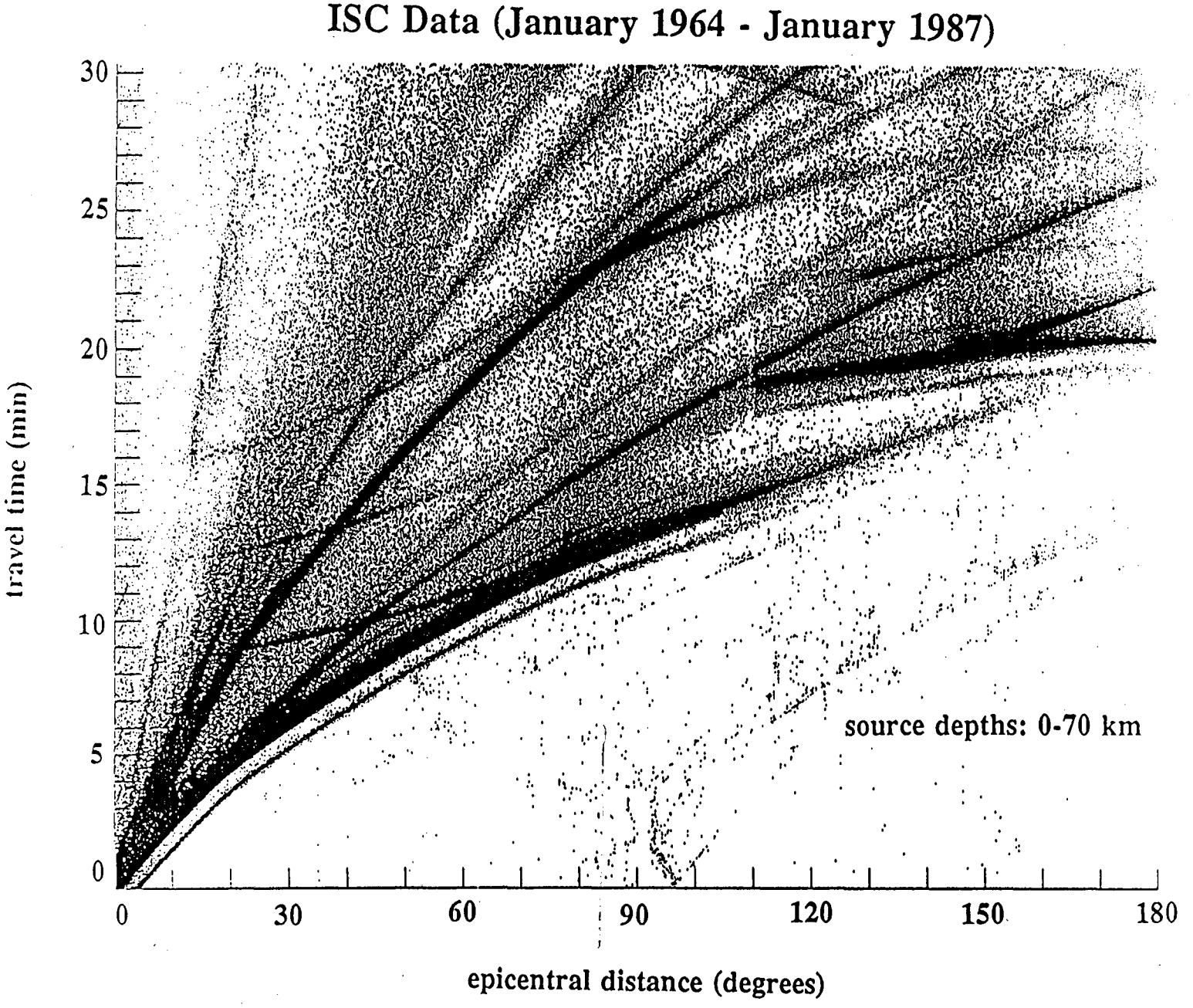
Chapter 2: A Study of the Effects of Source Mislocation in Mantle Travel Time Tomography 15
2.1 Introduction 15
2.2 Mathematical Development 16
2.2.1 Contributions to the Travel Time Discrepancy 17
2.2.2 Progressive Inversion 23
2.2.3 Simultaneous Inversion 29
2.2.4 Incorporating *a priori* Information 30
2.2.4.1 Conditioning the Data Space 30
2.2.4.2 Conditioning the Parameter Space 31
2.2.5 LSQR 35
2.3 Synthetic Tests of the Algorithms 36
2.3.1 Source Mislocations 38
2.3.2 Velocity Model 41
2.3.3 Summary of Results 42
2.4 Inversions of Real Data 43
2.4.1 Data Selection 43
2.4.2 Model 45
2.4.3 Inversion Results 46
2.4.4 Comparison of Small-scale Model Features 49
2.4.5 Comparison of Large-scale Model Features 51
2.4.6 Source Corrections 54
2.4.7 Discussion 56
2.5 Conclusions 57
2.6 References 58
2.7 List of Tables 61
2.8 List of Figures 63
2.9 Tables 67
2.10 Figures 72

Chapter 3: What Patterns of Heterogeneity in Earth's Mantle Can be Revealed by Seismic Tomography? 103
3.1 Introduction 103
3.2 Synthetic Experiments 106
3.3 Results and Discussion 108
3.3.1 Single Harmonics 110
3.3.2 A Higher-order l = 2 Pattern in the Upper Mantle 117
3.3.3 Coherent Patterns in the Upper Mantle 118
3.3.4 Lower Mantle 121
3.3.5 Lowermost Mantle 123
3.4 Conclusions 124
3.5 References 126
3.6 List of Tables 128
3.7 List of Figures 128
3.8 Table 134
3.9 Figures 135

Chapter 4: A Simultaneous, Tomographic Inversion of ISC Travel Time Residuals for Mantle P velocity, Source Mislocations, and Station Corrections 187
4.1 Introduction 187
4.2 Method 189
4.2.1 Model Parameterization 189
4.2.2 Simultaneous vs. Progressive Inversion 194
4.2.3 Simultaneous Inversion 195
4.3 Data 198
4.3.1 Data Selection 198
4.3.2 Summary Rays 200
4.3.3 Ray Coverage 201
4.3.4 Stochastic vs. Deterministic Analysis 202
4.4 Resolution and Uncertainty 203
4.4.1 Approximating Resolution 203
4.4.2 Approximating Covariance 208
4.5 Results and Discussion 211
4.5.1 Source Mislocations and Station Corrections 211
4.5.2 Velocity Model 212
4.5.3 Correlations with Surface Tectonics 213
4.5.3.1 Rift Zones 213
4.5.3.2 Subduction Zones 214
4.5.3.3 Continental Shields 217
4.5.3.4 Hotspots 217
4.5.4 Transition Zone 218
4.5.5 Mid mantle 219

4.5.6 Lower Mantle 220
4.5.7 Continuity of Features 221
4.5.8 Spherical Harmonic Expansion 222
4.6 Conclusions 219
4.7 References 225
4.8 List of Tables 230
4.9 List of Figures 230
4.10 Table 235
4.11 Figures 236

Chapter 5: Conclusions 257
5.1 Summary and Discussion 257
5.2 Recommendations for Further Studies 260
5.3 References 265


ISC Data (January 1964 - January 1987)
travel time (min)
30
25
20
15
10
5
0
0
30
60
90
120
150
180
epicentral distance (degrees)
source depths: 0-70 km

Chapter 1

Introduction

In Earth's mantle, lateral variations in a given material property generally amount to only a few percent of the property's value over thousands of kilometers while radial variations typically reach one hundred percent over similar distances. Accordingly, much effort and progress was made toward the elucidation of spherically-symmetric Earth structure in early seismological and geophysical studies and studies of these radially-varying properties remain important today. However, systematic lateral variations do exist in the mantle and crust, at least, and with the advent of plate tectonics as a framework to help describe large-scale lateral variations, seismologists began studies of three-dimensional Earth structure on a global scale. The collection of data worldwide from sets of standardized instruments and the development of practical and fast computers facilitated these studies.

This thesis documents attempts to image the three-dimensional seismic P velocity heterogeneity in Earth's mantle. The approaches considered here involve the tomographic inversion of body wave travel time residuals and differ primarily in their treatment of the source location problem. The seismic inverse problem to determine structural parameters of the medium (Earth) and parameters that describe the source is a complicated animal. The two sets of parameters are inextricably linked. Attempts to retrieve one or both sets typically concentrate on minimizing the influence of one set on the determination of the other set. This is the general approach we pursue here. We are most interested in the accurate estimation of Earth structure and will first try to simulate the general procedure by which tomographic inversions find models of Earth with body wave travel time data. We then find the effects of inaccurate source locations on the velocity model estimate and show how the velocity model estimate may

be improved by consideration of the source location.

1.1 Review of Previous Work

Previous attempts to retrieve models of P velocity in Earth's mantle with body waves differ in their approaches to parametrizing the model mantle and in their formulation and solution of the constraining equations. Popular model parametrizations include regionalization of the crust and mantle based on geographical association with surface tectonic processes [e.g., *Toksoz and Anderson*, 1966; *Okal*, 1977; *Tralli and Johnson*, 1986], spherical harmonic series expansion of the anomalous velocity field [e.g., *Dziewonski*, 1984; *Morelli and Dziewonski*, 1985, 1986], cubic splines [e.g., *Hovland et al.*, 1981] and division into a number of non-overlapping blocks [e.g., *Clayton and Comer*, 1983; *Inoue et al.*, 1990]. No one parametrization has been demonstrated to be clearly superior to the others. While the bias inherent in a tectonic regionalization renders such a scheme inappropriate for studies of the lower mantle, incomplete ray coverage makes it useful for studies of the upper mantle. Spherical harmonic expansions and cubic splines require fewer terms to describe a model to the same level of detail as a block parametrization, but do not offer the blocks' geometrical simplicity. With independent block parameters, one may examine the ray sampling, resolution, and covariance of a geographical location more easily.

In addition to differences in model parametrization, studies differ in the their construction of the system of equations to be solved and the numerical methods employed to solve them. Early efforts to map the three-dimensional velocity structure of Earth [*Aki et al.*, 1977; *Dziewonski et al.*, 1977; *Sengupta and Toksoz*, 1976], all of which use a block parametrization, were limited in their detail of structure by numerical methods that calculate the explicit inverse of the coefficient matrix. These methods allow the formal calculation of covariance and resolution matrices in order to evaluate the reliability of the model, but severely restrict the number of parameters available to

describe the model due to limitations of existing computers. *Aki et al.* [1977] use 1496 teleseismic P arrival times to find the 3-D slowness structure beneath the Norsar array, parametrized into 405 blocks, using both the generalized inverse formed from the products of singular value decomposition (SVD) of the coefficient matrix and the stochastic inverse of damped least squares. Their results are elegantly presented with the diagonal elements of the resolution matrix for direct evaluation of the models. On a global scale, *Dziewonski et al.* [1977] use nearly 700,000 P arrivals to find velocity perturbations for 120 blocks by means of the generalized inverse. These values, associated with individual blocks, are then expanded into spherical harmonics up to angular degree 3. While the resolution is easily evaluated, precision of the block model is limited to wavelengths on the order of 4000 km.

Dziewonski [1984] follows *Dziewonski et al.* [1977] in using about 500,000 P arrival times from 5,000 shallow ($h \leq 50$ km) events from the International Seismological Centre (ISC) catalog (1964-1979). All arrivals were teleseismic, in the epicentral distance range $25° \leq \Delta \leq 100°$. Similar observations, while not averaged, are weighted so that all distinct raypaths receive the same weight in the normal equations, regardless of the actual number of observations. Sources are initially relocated and travel times corrected for ellipticity and elevation. The solution is based on the generalized inverse for 245 coefficients of spherical harmonics expanded with a depth function. After each solution, sources are relocated using corrections to the travel times through the one-dimensional model, PREM, caused by the new 3-D model, and the process is repeated. Station corrections are calculated at each iteration. Standard errors for each model term (spherical harmonic coefficient) are calculated, but resolution and covariance are not formally assessed. The model's maximum anomalies, reaching 1.5%, are found at 670 km depth, the top of the model, and at the core-mantle boundary.

Attempts to obtain more detailed models of mantle velocities have turned to iterative, backprojection methods that do not explicitly find the inverse of the coefficient

matrix. *Sengupta et al.* [1981] extend the work of *Sengupta and Toksoz* [1976] to find velocities for 3008 out of a possible 3888 blocks spanning the entire mantle, with 3842 P arrivals. The block size, $10° \times 10° \times 500$ km in depth, brings the model precision down to roughly 1000 km. They use a method of successive approximation in which the velocity of each block is found one at a time by a weighted least squares procedure, accounting for the contributions to the travel time residual from the other blocks. This algorithm is iterated, considering the blocks in a random order, until convergence is reached. This method does not allow the calculation of resolution and covariance matrices. Their results for the first layer correlate qualitatively with surface tectonic features. Maximum velocity anomalies reach somewhat more than 2%.

Clayton and Comer [1983, and presented in *Hager and Clayton*, 1988] use a Simultaneous Reconstruction Technique (SIRT), which is similar to a Jacobi iteration technique for solutions to problems involving sparse matrices, in conjunction with 1.7 million ISC arrivals for the decade 1970-1980 to find 48,604 block parameters for the mantle. Observations are restricted to the distance range $25° \leq \Delta \leq 95°$, leaving the upper mantle relatively poorly sampled. A summarizing procedure for similar rays reduces the dimension of the coefficient matrix to 110,000 rows. The dimensions of the blocks are $5° \times 5° \times 100$ km in depth, about 550 km on a side, and represent a significant increase in precision over previous models. The cost of this increased precision is that resolution and covariance could not be calculated formally and presented alongside the mantle model. Several means to approximate resolution and covariance were developed in conjunction with, although they are not limited to, SIRT backprojections. For example, *Humphreys and Clayton* [1988] calculate a "point spread function," a column of the non-symmetric resolution matrix, and invert model error distributions to investigate the propagation of errors in the data through the algorithm to the solution. *Grand* [1987] introduces an inversion of a set of anomalies, each of which extends beyond the bounds of a single voxel, distributed throughout the model as a

means to evaluate resolution. *Clayton and Comer's* [1983; *Hager and Clayton*, 1988] results show maximum slowness perturbations of ±1% after layer averages are removed. Long wavelength features generally agree with *Dziewonski's* [1984] model, with some notable exceptions.

Tralli and Johnson [1986] use a tectonic regionalization derived from *Jordan* [1981] to find mantle P velocity anomalies associated with surface processes. Their study incorporates over 1.25 million ISC (1964-1981) P arrival times from shallow sources ($h \leq 70$ km) for $10° \leq \Delta \leq 100°$ and divides the upper mantle into $5° \times 5°$ cells that are assigned to one of seven distinct types of regions, depending on a cell's dominant or characteristic tectonic activity. Tau functions, from which velocity or slowness may be obtained immediately, are estimated as continuous profiles for each region. Maximum velocity perturbations reach about -2% for the top of the region representing young oceans.

Spakman and Nolet [1988] use the conjugate-gradient variant LSQR algorithm, due to *Paige and Saunders* [1982], to find a 3-D P velocity model for the upper mantle beneath the Mediterranean. Their procedure solves simultaneously for 9360 model blocks ($1° \times 1°$ with variable thickness), 105 station corrections, and 10,604 source relocation parameters, for a total of 20,069 free parameters. The total data amounts to over 480,000 rays from the ISC catalog (1964-1982) for the distance range $0° \leq \Delta \leq 90°$, which is reduced to slightly more than 300,000 composite rays by an averaging procedure. Theoretical and practical concerns about the nonlinearity of the travel time vs. delta function in the epicentral distance range $18° \leq \Delta \leq 25°$ (due to the triplication caused by refractions at the 400 km and 670 km discontinuities), the validity of a ray description of the wavefield in this range, and the data quality due to the nonuniqueness of arrivals in this range prove surmountable with judicious weighting of residuals based on their reliability. Maximum velocity perturbations reach over 3%.

Inoue et al. [1990] use a block parametrization, with cells $5° \times 5°$ and variable thickness depending on depth, and over 2 million arrival times from the ISC catalog (1964-1985) in the range $0° \leq \Delta \leq 95°$ to find P velocity perturbations for 32,768 blocks. Starting from relocations of events in a 1-D model based on *Herrin et al.* [1968], an iterative algorithm is performed on subsets of the data in which sources are relocated in an updated 1-D model calculated in the previous iteration, a new 3-D model is generated by means of the LSQR algorithm, and the 1-D model is refined. Given the epicentral distance range of data used, the authors are able to produce a model for the entire mantle. Resolution and covariance are evaluated approximately with synthetic test inversions of checkerboard patterns and by inverting distributions of Gaussian errors. Slowness perturbations exceed 4% in the upper mantle.

All the studies mentioned in this chapter employ finite-dimensional parametrizations. Consequently, assumptions are made regarding the smoothness of the medium to be imaged or that seismic rays sample only the smooth parts of the medium's structure. These assumptions follow from the fact that seismic observations are made at finite frequencies, which implies that the rays are not infinitely thin but actually sample a finite volume of Earth and contain information about the average properties of that volume. For example, for a wavelength of 10 km, the maximum ray width varies from 36 km for a ray of 1000 km length to 112 km for a ray of 10000 km length [*Nolet*, 1987]. When inferences are made about parts of the medium which are unsampled (or are assumed to be unsampled), an implicit assumption about the continuity of measured properties is made. Typically smoothness is imposed on the model by the inclusion of a roughness penalty in the inversion, in addition to the finite dimension of the parameterization. There are two complications with this approach. First, the degree of smoothness is both arbitrary at the outset and indeterminate at the conclusion of the inversion. Second, the estimation of uncertainties and resolution is made more difficult and the estimates produced bear only indirectly on the problem of

imaging the real earth. The uncertainty in the image depends directly on the smoothness assumed, but this dependence is difficult to evaluate with the techniques developed so far.

One alternative to a basis which spans Earth entirely is the "natural pixels" basis advocated by *Michelena and Harris* [1991]. Rather than parametrizing all parts of the medium under investigation, they estimate the actual sampling of the "fat" rays implied by finite frequency observations and find model values only for regions sampled. Explicit smoothing operators may then be used to interpolate between model values if a complete model is required. The "natural pixels" basis is not orthogonal, but is more flexible and rigorously correct in its representation of the reconstructed image. Perhaps more importantly, it allows a model to be parametrized with far fewer terms than are required with any of the orthogonal parametrizations mentioned above. This reduces the computational demands of the inversion and thus allows greater flexibility in the modeling process.

While the studies cited here generally vary in their methods for obtaining model solutions and in their approaches to parametrizing the model, the data involved in each study of P velocity are nearly identical. The International Seismological Center collects seismic arrival times from around the world. They employ these times in a sophisticated procedure in which times are associated into "events" and the events are located with P arrival times and the one-dimensional, Jeffreys-Bullen (J-B) travel time tables [*Jeffreys and Bullen*, 1940] with a standard least-squares technique [*Adams et al.*, 1982]. The system is far from perfect. Arrival times at stations around the world are read locally by different individuals from seismic records produced by different instruments. This variability surely propagates into the data. For the purposes of seismic imaging and accurate location of events, the geographical distribution of sources (earthquakes and large explosions) and seismographic stations, that so far are located almost exclusively on continents, is unfortunate. With the oceans nearly empty

of stations, large portions of Earth, particularly in the southern hemisphere, are undersampled by recorded seismic energy and event locations are inadequately constrained geometrically. Also, using the one-dimensional J-B model, that has known deficiencies in its representation of the spherically averaged structure of Earth in addition to its inability to account for lateral velocity variations, produces inaccurate event locations. Providing the means to correct this inaccuracy is one goal of our study. Another goal is to provide modeling constraints for studies of Earth's composition and dynamic processes, in order to investigate possible means by which Earth coalesced and evolved to its current state.

Wielandt [1987] offers a note of caution relevant to all the studies mentioned in this chapter, including ours. The ray-theoretic approximation assumed to be valid in tomography does not hold in the presence of diffracted or laterally refracted waves. With a set of synthetic experiments, *Wielandt* [1987] shows that such waves should be commonly included in seismic observations and will often hinder the identification of direct phases. The effects of employing the linearity assumption required by tomographic inversions would be to overestimate the size of positive (fast) velocity anomalies and to underestimate the amplitude of negative anomalies. He finds that the ray approximation is inadequate for negative anomalies in excess of 4% at 200 km distance, 2% at 500 km, and 1% at 1000 km. In each case a diameter of 100 km is assumed for the anomaly [*Wielandt*, 1987].

1.2 Topics Investigated in This Thesis

In chapter 2 we develop and investigate the performance of three different schemes for inverting seismic travel time residuals. The first scheme solves simultaneously for corrections to the source locations and for three-dimensional perturbations to the one-dimensional starting model. The second scheme solves the inverse problem progressively, for each set of terms in succession. We also solve the system of

equations directly, neglecting source terms, as a third approach that is not truly distinct from the simultaneous inversion. The efficacy of the inversion schemes is explored with both synthetic and real data. First, we perform simulations of the general procedure used to produce tomographic images of Earth's mantle from global earthquake data. Next, we invert real data supplied by the ISC. Results are summarized from both the simulations and real inversions in terms of what they tell us about the problem of ambiguous source locations and implications for contamination in our velocity models. These results are emphasized above correlations with tectonic features and geological and geophysical interpretation of the models. Here our intent is to isolate the effects of source mislocation and a complete treatment of the tomographic inverse problem is not attempted.

In chapter 3 we investigate which patterns of seismic velocity heterogeneity in the mantle would be returned reliably by a tomographic inversion in which the model mantle is parametrized by a set of discrete, non-overlapping voxels. We construct synthetic data sets based on real ray sampling of the mantle by introducing spherical harmonic patterns of velocity heterogeneity and perform inversions of the synthetic data. We expand the resulting voxel model in spherical harmonics and compare the power at each degree and in each model layer with the input spherical harmonics in order to determine which patterns produced by inversions of real data may be deemed reliable and to identify patterns that must be viewed with skepticism.

In chapter 4 we present a more detailed model of P velocity in the mantle than the ones presented in chapter 2. The model mantle is parametrized by approximately equal-area blocks which are $5° \times 5°$, rather than the $10° \times 10°$ blocks which are used previously. Also in contrast to the study described in chapter 2, individual rays which sample similar parts of Earth are averaged together to form summary rays. More than 3 million rays from 46,000 shallow events satisfying selection criteria are averaged according to $2° \times 2° \times 10$ km deep bins to construct nearly 726,000 summary rays for

the inversion. The construction of summary rays tends to average out contributions to the travel time residuals that arise due to heterogeneity of a scale too small to be resolved by our model blocks, and also averages out contributions due to mislocated sources. For the sake of completeness and in order to constrain as many degrees of freedom as possible, we solve simultaneously for summary source mislocation terms and demonstrate that these corrections are small. In this study we also solve for summary station corrections.

Resolution and co-variance are evaluated by approximate methods. Resolution is estimated by the inversion of a synthetic, checkerboard test pattern and the calculation of point spread functions for selected voxels. Covariance is estimated by averaging results from inversions of realistic errors and by a jackknife procedure. We present our three-dimensional velocity model in conjunction with the resolution estimates produced by our checkerboard test. Normalized checkerboard output values, ranging from 0 to 1, are used to modify each voxel's red-blue velocity value from full color saturation, indicating good resolution, to white, which indicates no resolution. This presentation aids us in a detailed interpretation of the correlation between our model and tectonic features at the Earth's surface and an evaluation of the significance of features deep in the mantle.

Finally, we summarize the findings of these three separate studies in chapter 5 and make suggestions for further work on these topics.

The progressive inversion technique in chapter 2 was developed after a suggestion by Lane Johnson, my committee chairman. Advice and guidance provided by him and by Dan O'Connell contributed greatly to my understanding of the technique. Lane Johnson's programs to trace rays in one and three dimensions were important aids to the work in this thesis. The work described in chapter 4 was undertaken in conjunction with Don Vasco and Lane Johnson. Their contributions are integral to the study and cannot be identified individually. I thank them for the time they spent with me

and for sharing their ideas unselfishly. For the sake of consistency throughout the thesis, I use the terms "we" and "our" rather than "I" and "my" to refer to the work carried out in these projects.

1.3 References

Adams, R.D., A.A. Hughes, and D.M. McGregor, Analysis procedures at the International Seismological Centre, *Phys. Earth. Planet. Int., 30,* 85–93, 1982.

Aki, K., A. Christoffersson, and E.S. Husebye, Determinations of the three-dimensional seismic structure of the lithosphere, *J. Geophys. Res., 82,* 277–296, 1977.

Clayton, R.W., and R.P. Comer, A tomographic analysis of mantle heterogeneities from body wave travel times, *EOS Transactions AGU, 64,* 776, 1983.

Dziewonski, A.M., Mapping the lower mantle, Determination of lateral heterogeneity in P velocity up to degree and order 6, *J. Geophys. Res., 89,* 5929–5952, 1984.

Dziewonski, A.M., B.H. Hager, and R.J. O'Connell, Large–scale heterogeneities in the lower mantle, *J. Geophys. Res., 82,* 239–255, 1977.

Grand, S., Tomographic inversion for shear velocity beneath the North American Plate, *J. Geophys. Res., 92,* 14065–14090, 1987.

Hager, B.H., and R.W. Clayton, Constraints on the structure of mantle convection using seismic observations, flow models, and the geoid, in *Mantle Convection,* W. R. Peltier (Ed.), Gordon and Breach, New York, 657–763, 1989.

Herrin, E., W. Tucker, J.N. Taggert, D.W. Gordon, and J.L. Lobdell, Estimation of surface focus P–travel times, *Bull. Seism. Soc. Am., 58,* 1273–1291, 1968.

Hovland, J., D. Gubbins, and E.S. Husebye, Upper mantle heterogeneities beneath Central Europe, *Geophys. J. R. Astron. Soc., 66,* 261–284, 1981.

Humphreys, E., and R.W. Clayton, Adaptation of back projection tomography to seismic travel time problems, *J. Geophys. Res., 93,* 1073–1085, 1988.

Inoue, H., Y. Fukao, K. Tanabe, and Y. Ogata, Whole mantle P–wave travel time tomography. *Phys. Earth and Planet. Int., 59,* 294–328, 1990.

Jeffreys, H., and K.E. Bullen, *Seismological Tables*. British Association for the Advancement of Science, London, 1940.

Jordan, T.H., Global tectonic regionalization for seismological data analysis, *Bull. Seis. Soc. Am., 71,* 1131–1141, 1981.

Morelli, A., and A.M. Dziewonski, Stability of aspherical models of the lower mantle, *EOS Transactions AGU, 66,* 975, 1985.

Morelli, A., and A.M. Dziewonski, 3D structure of the Earth's core inferred from travel–time residuals, *EOS Transactions AGU, 67,* 311, 1986.

Nolet, G., Seismic wave propagation and seismic tomography, in *Seismic Tomography,* G. Nolet (ed.), Reidel, Dordrecht, pp. 1–23, 1987.

Okal, E.A., The effect of intrinsic oceanic upper–mantle heterogeneity on regionalization of long–period Rayleigh–wave phase velocities, *Geophys. J. R. Astron. Soc., 49,* 357–370, 1977.

Paige, C. C., and M.A. Saunders, LSQR, An algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Software, 8,* 43–71, 1982.

Sengupta, M.K., and M.N. Toksoz, Three–dimensional model of seismic velocity variation in the Earth's mantle. *Geophys. Res. Letters, 3,* 84–86, 1976.

Sengupta, M.K., R.E. Hassell, and R.W. Ward, Three–dimensional seismic velocity structure of the Earth's mantle using body wave travel times from intra–plate and deep–focus earthquakes, *J. Geophys. Res., 86,* 3913–3934, 1981.

Spakman, W., and G. Nolet, Imaging algorithms, accuracy and resolution in delay time tomography, in *Mathematical Geophysics,* N.J. Vlaar, G. Nolet, M.J.R. Wortel, and S.A.P.L. Cloetingh (eds.), Reidel, Dordrecht, pp. 155–187, 1987.

Toksoz, M.N., and D.L. Anderson, Phase velocities of long–period surface waves and structure of the upper mantle, 1. Great–circle Love and Rayleigh wave data, *J. Geophys. Res., 71,* 1649–1658, 1966.

Tralli, D.M., and L.R. Johnson, Lateral variations in mantle P velocity from

tectonically regionalized tau estimates, *Geophys. J. R. Astr. Soc., 86,* 475–489, 1986.

Wielandt, E., On the validity of the ray approximation for interpreting delay times, in *Seismic Tomography,* G. Nolet (ed.), Reidel, Dordrecht, pp. 85–98, 1987.

Chapter 2

A Study of the Effects of Source Mislocation in Mantle Travel Time Tomography

2.1 Introduction

In this chapter a formalism is developed for the tomographic inversion of seismic travel time residuals. Once the tomographic system of travel time equations is constructed, two methods are presented for its solution: simultaneously for both velocity model terms and corrections to the source locations and progressively, for each set of terms in succession. Both algorithms perform least-squares inversions that minimize the l^2 norm of the residuals. The methods differ primarily in their treatment of source mislocation terms. Additionally, we solve the system of equations directly, neglecting source terms. The efficacy of the algorithms is explored in conjunction with synthetic data as we perform simulations of the general procedure used to produce tomographic images of Earth's mantle from global earthquake data. A data set is constructed in a way that mimics the practice of the International Seismological Centre (ISC) as it collects observations world-wide, associates observations with seismic events, locates the events, and distributes the codified data to interested researchers. These data consist of arrival times at reporting stations and estimates of earthquake locations calculated in a one-dimensional Earth model. Because of the three-dimensional nature of Earth, the ISC locations are only approximations to the true earthquake locations, so we investigate the effects of mislocations on the velocity model obtained in an inversion, and the ability of our simultaneous and progressive inversion techniques to correct mislocated earthquakes and produce an accurate velocity model. To simplify the problem and highlight the effects of source mislocation in our controlled simulations, we keep the numbers of data and model parameters small.

Finally, we invert real data supplied by the ISC. We use P arrival data from January 1964 through January 1987 and our inversion algorithms to solve for three-dimensional P velocity models of the mantle and source mislocations. The model mantle is parametrized by approximately equal-area blocks: 10° × 10° and generally 200 km in depth. Nearly 345,000 rays from more than 3,000 shallow events satisfying selection criteria are included in the inversions. The data are weighted by the inverse variance of travel time residuals as a function of epicentral distance; model parameters are weighted by a measure of the quality of sampling in each model block, or voxel. A roughness penalty is included in the inversions.

We summarize results from both the simulations and real inversions in terms of what they tell us about the problem of ambiguous source locations and implications for contamination of our velocity models. These results are emphasized above correlations with tectonic features and geological and geophysical interpretation of the models. In this chapter we seek to isolate the effects of source mislocation and do not attempt a complete treatment of the tomographic inversion problem. We may still obtain valuable insight into the structure of the problem, the inherent interdependence of the parameters and limitations as we try to disentangle them.

2.2 Mathematical Development

The ith arrival time from event j, that is recorded at a station k, may be represented as

$$(t_i)_j = \tau_j + T_i(\mathbf{r}_j, \mathbf{r}_k', c(\mathbf{r})) + (\varepsilon_i)_j, \tag{2.1}$$

where

τ_j = origin time for event j,

$T_i(\mathbf{r}_j, \mathbf{r}_k', c(\mathbf{r}))$ = travel time through the medium, $c(\mathbf{r})$, from event location, $\mathbf{r}_j$, to station location, $\mathbf{r}_k'$,

$(\varepsilon_i)_j$ = reading error associated with ray i from event j,

and

$$i = 1,2,\ldots,m_j, \qquad j = 1,2,\ldots,m_e, \qquad k = 1,2,\ldots,n_s;$$

where

m_j = number of arrival times reported for the jth event,

m_e = number of events in the data set,

n_s = number of stations reporting arrivals in the data set.

In the general case, we have collected the observations $(t_i)_j$ but do not know any of the terms on the right hand side of equation (2.1). If we assume we know the velocity structure of the medium to within a few percent of the actual velocity, $c(\mathbf{r})$, we may take a first-order Taylor expansion about our model, call it $\hat{c}(\mathbf{r})$, and try to estimate the error in our model by reconciling the perturbation terms of the expansion with the deviations of observed arrival times from arrival times calculated through the reference velocity model. *Wielandt* [1987] carries out a set of synthetic experiments to investigate the validity of the linearity assumption inherent in this ray-theoretic formulation. Performing the Taylor expansion and discarding higher terms we get

$$(t_i)_j = \hat{\tau}_j + \delta\tau_j + \hat{T}_i(\mathbf{r}_j,\mathbf{r}_k',c(\mathbf{r})) + \delta T_i(\mathbf{r}_j,\mathbf{r}_k',c(\mathbf{r})) + (\varepsilon_i)_j, \tag{2.2}$$

where $\hat{\tau}_j$ is an estimate of event origin time calculated using the starting velocity model.

2.2.1 Contributions to the Travel Time Discrepancy

$$\text{Let } \hat{T}_i(\mathbf{r}_j,\mathbf{r}_k',c(\mathbf{r})) = T_i(\mathbf{r}_j,\mathbf{r}_k{}',c(\mathbf{r}))|_{\hat{r}_j,\hat{r}_k',\hat{c}(\mathbf{r})} \tag{2.3}$$

be the travel time for ray i connecting $\hat{\mathbf{r}}_j$ and $\hat{\mathbf{r}}_k'$ through velocity model $\hat{c}(\mathbf{r})$, then the first-order term of the Taylor expansion may be expressed as the sum of three

terms:

$$\delta T_i(\mathbf{r}_j,\mathbf{r}_k{}',c(\mathbf{r})) = \delta T_i(\mathbf{r}_j,\mathbf{r}_k{}',c(\mathbf{r}))|_{\hat{r}_k',\hat{c}(\mathbf{r})} + \delta T_i(\mathbf{r}_j,\mathbf{r}_k{}',c(\mathbf{r}))|_{\hat{r}_j,\hat{r}_k'} + \delta T_i(\mathbf{r}_j,\mathbf{r}_k{}',c(\mathbf{r}))|_{\hat{r}_j,\hat{c}(\mathbf{r})}. \tag{2.4}$$

The first term on the right hand side of equation (2.4) represents the perturbation in the travel time due to a perturbation in the location of the earthquake's hypocenter. The second term represents the travel time perturbation due to perturbations in the velocity model. The third term represents contributions to the travel time anomaly that are unique to a particular station. Strictly, this term represents travel time discrepancies due to poorly known station locations, but in practice the term serves to isolate the effects of velocity anomalies occurring in the vicinity of a station on a scale too small to be resolved by our model parametrization. Errors in observed travel time residuals resulting either from incorrect observations, such as instrument errors and systematic phase mispicks or misidentifications at a particular station, are also described by this "station" term.

We define the travel time residual to be the observed arrival time minus a predicted arrival time,

$$\delta(t_i)_j = (t_i)_j - (\hat{\tau}_j + T_i(\mathbf{r}_j,\mathbf{r}_k{}',c(\mathbf{r}))|_{\hat{r}_j,\hat{r}_k',\hat{c}(\mathbf{r})}). \tag{2.5}$$

Substituting equation (2.5) into equation (2.2) gives

$$\begin{aligned}\delta(t_i)_j &= \delta\tau_j + \delta T_i(\mathbf{r}_j,\mathbf{r}_k',c(\mathbf{r})) + (\varepsilon_i)_j,\\ &= \delta\tau_j + \delta T_i(\mathbf{r}_j,\mathbf{r}_k{}',c(\mathbf{r}))|_{\hat{r}_k',\hat{c}(\mathbf{r})}\\ &\quad + \delta T_i(\mathbf{r}_j,\mathbf{r}_k{}',c(\mathbf{r}))|_{\hat{r}_j,\hat{r}_k'} + \delta T_i(\mathbf{r}_j,\mathbf{r}_k{}',c(\mathbf{r}))|_{\hat{r}_j,\hat{c}(\mathbf{r})} + (\varepsilon_i)_j,\end{aligned} \tag{2.6}$$

The perturbation to the origin time, $\delta\tau_j$, may be viewed as a fourth hypocenter term. Then

$$\delta t_j^{hypocenter} = \delta\tau_j + \delta T_i(\mathbf{r}_j,\mathbf{r}_k{}',c(\mathbf{r}))|_{\hat{r}_k,\hat{c}(\mathbf{r})}$$

$$= \frac{\partial T_i}{\partial h_1}\Big|_{\hat{\mathbf{r}}_k',\hat{c}(\mathbf{r})}\delta h_{1_j} + \frac{\partial T_i}{\partial h_2}\Big|_{\hat{\mathbf{r}}_k',\hat{c}(\mathbf{r})}\delta h_{2_j} + \frac{\partial T_i}{\partial h_3}\Big|_{\hat{\mathbf{r}}_k',\hat{c}(\mathbf{r})}\delta h_{3_j} + \frac{\partial t_i}{\partial h_4}\Big|_{\hat{\mathbf{r}}_k',\hat{c}(\mathbf{r})}\delta h_{4_j}, \tag{2.7}$$

where

$$h_1 = \tau,\ \ h_2 = \theta,\ \ h_3 = \phi,\ \ h_4 = z.$$

In matrix notation,

$$\delta \mathbf{t}_j = \mathbf{H}_j \delta \mathbf{h}_j, \tag{2.8}$$

where

$\delta \mathbf{t}_j$ = vector of travel time residuals for event j,

$(H_{il})_j = \frac{\partial T_i}{\partial h_l}\Big|_{\hat{\mathbf{r}}_k',\hat{c}(\mathbf{r})}$ = matrix of source mislocation partial derivatives (l=1,2,3,4),

$\delta \mathbf{h}_j = (\delta\tau, \delta\theta, \delta\phi, \delta z)_j$ = vector of hypocenter perturbations for event j.

The second term on the right hand side (RHS) of equation (2.6) represents the deviation of our starting velocity model, $\hat{c}(\mathbf{r})$ from the actual velocity structure, $c(\mathbf{r})$,

$$\delta t_i^{model} = \delta T_i(\mathbf{r}_j, \mathbf{r}_k', c(\mathbf{r}))|_{\hat{r}_j,\hat{r}_k'}. \tag{2.9}$$

The travel time along a ray, S_i, is given by

$$t_i = \int_{S_i} \frac{ds}{c(\mathbf{r})}, \tag{2.10}$$

where $c(\mathbf{r})$ is the velocity of the medium. Our task is to determine $c(\mathbf{r})$ from a set of travel time observations t_i, $i = 1,2,...M$. This task is made more difficult by the implicit dependence of the ray path, S_i, on the velocity model, $c(\mathbf{r})$. Once again, we assume that our starting velocity model is within a few percent of the true structure and seek to reconcile the discrepancy by solving for the perturbation term. Let

$$\delta t_i^{model} = \delta T_i(\mathbf{r}_j, \mathbf{r}_k', c(\mathbf{r}))|_{true} - \delta T_i(\mathbf{r}_j, \mathbf{r}_k', c(\mathbf{r}))|_{\hat{r}_k',\hat{c}(\mathbf{r})}, \tag{2.11}$$

$$= \int_{S_i} \frac{ds}{c(\mathbf{r})} - \int_{\hat{S}_i} \frac{ds}{\hat{c}(\mathbf{r})}.$$

Fermat's principle justifies the assumption that the raypath persists relatively unchanged in the presence of small three-dimensional velocities anomalies. This allows us to perform the line integrals as one integral along the initial raypath, i.e.,

$$\delta t_i^{model} = \approx \int_{\hat{S}_i} \left(\frac{1}{c(\mathbf{r})} - \frac{1}{\hat{c}(\mathbf{r})}\right) ds,$$

$$\approx -\int_{\hat{S}_i} \frac{\delta c(\mathbf{r})}{\hat{c}^2(\mathbf{r})} ds, \tag{2.12}$$

where $\hat{S}_i$ is the path of the ith ray through the starting velocity model, $\hat{c}(\mathbf{r})$.

In order to represent the function of velocity perturbations over the medium of interest, we must choose a set of basis functions. Two approaches are popular. The first divides the medium under investigation into non-overlapping volume elements, or voxels. Following *Nolet* [1987], let

$$f_k(\mathbf{r}) = \begin{cases} \nu_k^{-1/2} & \text{if } \mathbf{r} \text{ is in cell } k \\ 0 & \text{elsewhere} \end{cases}, \tag{2.13}$$

where ν_k is the volume of cell k. The functions f_k form a basis that spans a subspace of the Hilbert space of all possible velocity models, $c(\mathbf{r})$. Since the cells do not overlap,

$$\int_{volume} f_k(\mathbf{r}) f_l(\mathbf{r}) d^3\mathbf{r} = \delta_{kl}. \tag{2.14}$$

A second popular set of basis functions consist of solid spherical harmonics [e.g., *Dziewonski*, 1984; *Morelli and Dziewonski*, 1986],

$$f_i = f_k(r) Y_l^m(\theta,\phi),$$

where $f_k(r)$ is a set of orthogonal functions in radius.

Our choice of a local basis is arbitrary in many respects. A block parametrization allows a more accurate assessment of ray sampling of Earth and the resulting coefficient matrix is quite sparse. This sparseness may be exploited to solve the matrix problem efficiently. Fewer terms are required to describe the model to the same level of resolution with the global spherical harmonic basis (fewer by up to an order of magnitude), but the coefficient matrix in the spherical harmonic case is dense. It is important to note that choosing a model parametrization represents an opportunity to introduce bias into the inversion. Depending on the geometry of the inverse problem, a particular model parameterization may or may not allow the accurate reconstruction of interesting features of the real earth, or it may require an inaccurate (i.e., smeared or aliased) estimation of the model simply because of limitations in its representation of features. *Michelena and Harris* [1991] suggest a way to make the model parametrization more flexible and complete in its representation of model anomalies sampled by a set of data. Their representation acknowledges the finite width of the zone sampled by a given seismic ray and seeks to construct a solution in terms of the portions of Earth sampled by these "fat" rays. They call this representation a parametrization based on "natural pixels."

Choosing the set of functions described in equation (2.13), we may represent the function of velocity perturbations as a linear combination of basis functions,

$$\delta c(\mathbf{r}) = \sum_{k=1}^{n} \gamma_k f_k(r). \tag{2.15}$$

Substituting equation (2.15) into equation (2.12) results in an expression for the travel time perturbations in terms of velocity perturbation basis functions,

$$\delta t_i^{model} = \sum_{k=1}^{n} -\int_{\hat{S}_i} \frac{\gamma_k f_k(\mathbf{r})}{\hat{c}(\mathbf{r})^2} ds = \sum_{k=1}^{n} A_{ik}\gamma_k, \tag{2.16}$$

where

$$A_{ik} = -\int_{S_i} \frac{f_k(\mathbf{r})}{\hat{c}(\mathbf{r})^2} ds.$$

In matrix form,

$$\delta \mathbf{t}^{model} = \mathbf{A}\boldsymbol{\gamma}. \tag{2.17}$$

We express the "station" term of equation (2.4) as

$$\delta \mathbf{t}^{station} = \mathbf{S}\boldsymbol{\mu}, \tag{2.18}$$

where

$$S_{ik} = \begin{cases} 1 & \text{if } k = \text{station number} \\ 0 & \text{if } k \neq \text{station number'} \end{cases}$$

$\mu_k =$ the station correction for the kth station.

Substituting equations (2.8), (2.17), and (2.18) into equation (2.6) for all rays ($i = 1,2,...,m_j$) of all events ($j = 1,2,...,m_e$) we find the problem we wish to solve is now

$$(\delta t_i)_j = (\delta t_i^{model})_j + (\delta t_i^{hypocenter})_j + (\delta t_i^{station})_j, \tag{2.19}$$

or

$$\delta \mathbf{t} = \mathbf{A}\,\boldsymbol{\gamma} + \mathbf{H}\,\delta \mathbf{h} + \mathbf{S}\,\delta \boldsymbol{\mu}, \tag{2.20}$$

where

$\delta \mathbf{t} \in R^{M\times 1}$ = vector of travel time residuals,

$\mathbf{A} \in R^{M\times n_p}$ = matrix of ray segments in voxels,

$\boldsymbol{\gamma} \in R^{n_p\times 1}$ = vector of coefficients in the expansion of perturbations to the starting model,

$\mathbf{H} \in R^{M\times 4n_e}$ = matrix of partial derivatives for all events,

$\delta \mathbf{h} \in R^{4n_e\times 1}$ = vector of perturbations to the hypocenters,

$\mathbf{S} \in R^{M\times n_s}$ = matrix of partial derivatives for stations,

$\delta\mu \in R^{n_s \times 1}$ = vector of station corrections,

M = number of data (reported arrivals),

n_e = number of events,

n_p = number of model blocks,

n_s = number of reporting stations.

2.2.2 Progressive Inversion

At this point we could combine matrices and solve for all parameters simultaneously, but there are two reasons why we may choose not to do so. First, since relocating the hypocenter of each event consists of estimating four terms: origin time plus three spatial coordinates, the number of hypocentral parameters totals $4m_e$, where m_e is the number of events in the data set. The combined matrix would therefore have dimensions $M \times (n_p + 4m_e)$, resulting in considerable demands for core memory and mixing different classes of parameters. More importantly, it turns out that we may exploit the natural separation of the parameters to solve for each set of parameters in a step-wise fashion. This approach follows *Pavlis and Booker* [1980], *Spencer and Gubbins* [1980], *Jordan and Sverdrup* [1981], and *O'Connell and Johnson* [1991], among others, and allows a more detailed analysis of resolution and uncertainty in the determination of mislocation terms than would be practical otherwise. The idea is to find an orthogonal transformation that will rotate the first coefficient matrix, in our case **H**, so that only the first four elements of the travel time residual vector have non-zero projections into the parameter space. Actually, the number of independent data providing information to the specification of parameters is equal to the rank of the original, unrotated matrix **H**, where $0 \leq rank(\mathbf{H}) \leq 4$. *Pavlis and Booker* [1980] call this orthogonal transformation an "annulling transformation" because its effect is to separate the problem involving two (or more) different classes of parameters into two

problems, the second of which is independent of the first class of parameters. The independent problem involves data that have been "annulled" with respect to the first parameter class.

The orthogonal transformation we choose to employ comes from the singular value decomposition (SVD) of the matrix of hypocenter mislocation partial derivatives, **H**. Any matrix may be factored into the form, $\mathbf{H} = \mathbf{USV}^T$ [*Lawson and Hanson*, 1974]. If **H** is an $m \times n$ matrix of rank k, then **U** is an $m \times m$ orthogonal matrix, **V** is an $n \times n$ orthogonal matrix, and **S** is an $m \times n$ diagonal matrix of singular values in which k values are strictly non-zero. The orthogonal matrix $\mathbf{U}^T$ may serve as an annulling transformation matrix, when used to pre-multiply through equation (2.20). A heuristic proof of this annulling property follows.

Note that only k entries of the diagonal matrix S are non-zero, and that these non-zero elements are all positive. Since

$$\mathbf{H} = \mathbf{USV}^T,$$

then

$$\begin{aligned}
\mathbf{U}^T\mathbf{H} &= \mathbf{U}^T\mathbf{USV}^T, \\
&= \mathbf{SV}^T, \\
&= \begin{bmatrix} s_1 & & & & 0 \\ & \ddots & & & \\ & & s_k & & \\ & & & 0 & \\ & & & & \ddots \\ 0 & & & & 0 \end{bmatrix} \mathbf{V}^T.
\end{aligned}$$

Only the first k rows of $\mathbf{U}^T\mathbf{H}$ contain non-zero elements, therefore only the first k rows will project onto a non-zero (range) space.

By partitioning the data set into individual events and disregarding the station term, equation (2.20) may be rewritten as

$$\begin{bmatrix} (\delta t)_1 \\ (\delta t)_2 \\ \cdot \\ \cdot \\ \cdot \\ (\delta t)_{m_e} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_1 & 0 & \dots & 0 \\ 0 & \mathbf{H}_2 & \dots & 0 \\ \cdot & 0 & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & \mathbf{H}_{m_e} \end{bmatrix} \begin{bmatrix} (\delta h)_1 \\ (\delta h)_2 \\ \cdot \\ \cdot \\ \cdot \\ (\delta h)_{m_e} \end{bmatrix} + \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{A}_{m_e} \end{bmatrix} \gamma \tag{2.21}$$

where

$$(\delta t)_j \in R^{m_j \times 1},$$

$$\mathbf{H}_j \in R^{m_j \times 4},$$

$$\delta \mathbf{h}_j \in R^{4 \times 1},$$

$$\mathbf{A}_j \in R^{m_j \times n_p},$$

$$m_j = \text{number of data for the } j^{th} \text{ event},$$

$$j = 1,2,\dots,m_e.$$

The effects of the station term could be dealt with in a fashion similar to the treatment of the source term, but the large numbers of data and parameters involved in our whole-mantle inversions render the full problem unwieldy. We expect the deleterious effects of the station errors on our retrieval of velocity parameters to be small compared to the effects of source mislocation. Later we will test this assumption when we solve for subsets of the three parameters classes with the simultaneous inversion algorithm.

The orthogonal transformation matrix now has the form

$$\begin{bmatrix} \begin{bmatrix} \mathbf{U}_R^T \\ \mathbf{U}_N^T \end{bmatrix}_1 & 0 & \cdots & 0 \\ 0 & \begin{bmatrix} \mathbf{U}_R^T \\ \mathbf{U}_N^T \end{bmatrix}_2 & \cdots & 0 \\ \cdot & 0 & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \begin{bmatrix} \mathbf{U}_R^T \\ \mathbf{U}_N^T \end{bmatrix}_{m_e} \end{bmatrix}, \tag{2.22}$$

where

$(\mathbf{U}_R^T)_j \in R^{4\times4}$ = range space of hypocenter partial derivatives for the jth event,

$(\mathbf{U}_N^T)_j \in R^{(m_j-4)\times4}$ = null space of hypocenter partial derivatives for the jth event,

$j = 1,2,\ldots,m_e$.

Applying the transformation matrix (2.21) to equation (2.22) we get

$$\begin{bmatrix} \begin{bmatrix} \delta\mathbf{t}_R \\ \delta\mathbf{t}_N \end{bmatrix}_1 \\ \begin{bmatrix} \delta\mathbf{t}_R \\ \delta\mathbf{t}_N \end{bmatrix}_2 \\ \vdots \\ \begin{bmatrix} \delta\mathbf{t}_R \\ \delta\mathbf{t}_N \end{bmatrix}_{m_e} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \mathbf{H}_R \\ 0 \end{bmatrix}_1 & 0 & \cdots & 0 \\ 0 & \begin{bmatrix} \mathbf{H}_R \\ 0 \end{bmatrix}_2 & \cdots & 0 \\ \cdot & 0 & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \begin{bmatrix} \mathbf{H}_R \\ 0 \end{bmatrix}_{m_e} \end{bmatrix} \begin{bmatrix} \delta h_1 \\ \delta h_2 \\ \vdots \\ \delta h_{m_e} \end{bmatrix} + \begin{bmatrix} \begin{bmatrix} \mathbf{A}_R \\ \mathbf{A}_N \end{bmatrix}_1 \\ \begin{bmatrix} \mathbf{A}_R \\ \mathbf{A}_N \end{bmatrix}_2 \\ \vdots \\ \begin{bmatrix} \mathbf{A}_R \\ \mathbf{A}_N \end{bmatrix}_{m_e} \end{bmatrix} \boldsymbol{\gamma}, \tag{2.23}$$

where

$(\delta\mathbf{t}_R)_j = (\mathbf{U}_R^T)_j \delta t \in R^{4\times1}$ = travel time residual in range space of hypocenter partial derivatives,

$(\mathbf{H}_R)_j = (\mathbf{U}_R^T)_j \mathbf{H}_j \in R^{4\times4}$ = rotated matrix of hypocenter mislocation partial

derivatives,

$\delta \mathbf{h}_j \in R^{4\times 4}$ = matrix of hypocenter mislocation partial derivatives for event i,

$(\mathbf{A}_R)_j = (\mathbf{U}_R^T)_j \mathbf{A}_j \in R^{4\times n_p}$ = portion of rotated $\mathbf{A}_N$ that contributes to $\delta \mathbf{t}_R$,

$(\delta \mathbf{t}_N)_j = (\mathbf{U}_N^T)_j \delta t \in R^{(m_j-4)\times 4}$ = travel time residual in null space of hypocenter partial derivatives,

$(\mathbf{A}_N)_j = (\mathbf{U}_N^T)_j \mathbf{A}_j \in R^{4\times n_p}$ = portion of rotated $\mathbf{A}_N$ that contributes only to $\delta \mathbf{t}_N$,

$j = 1,2,\ldots,m_e$.

Rearranging (2.22) gives

$$\begin{bmatrix} (\delta \mathbf{t}_R)_1 \\ (\delta \mathbf{t}_R)_2 \\ \cdot \\ \cdot \\ \cdot \\ (\delta \mathbf{t}_R)_{m_e} \\ - \\ (\delta \mathbf{t}_N)_1 \\ (\delta \mathbf{t}_N)_2 \\ \cdot \\ \cdot \\ \cdot \\ (\delta \mathbf{t}_N)_{m_e} \end{bmatrix} = \begin{bmatrix} (\mathbf{H}_R)_1 & 0 & \ldots & 0 \\ 0 & (\mathbf{H}_R)_2 & \ldots & 0 \\ \cdot & 0 & \ldots & \cdot \\ \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \ldots & \cdot \\ 0 & 0 & \ldots & (\mathbf{H}_R)_{m_e} \\ - & - & - & - \\ 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \ldots & \cdot \\ 0 & 0 & \ldots & 0 \end{bmatrix} \begin{bmatrix} \delta h_1 \\ \delta h_2 \\ \cdot \\ \cdot \\ \cdot \\ \delta h_{m_e} \end{bmatrix} + \begin{bmatrix} (\mathbf{A}_R)_1 \\ (\mathbf{A}_R)_2 \\ \cdot \\ \cdot \\ \cdot \\ (\mathbf{A}_R)_{m_e} \\ - \\ (\mathbf{A}_N)_1 \\ (\mathbf{A}_N)_2 \\ \cdot \\ \cdot \\ \cdot \\ (\mathbf{A}_N)_{m_e} \end{bmatrix} \gamma. \tag{2.24}$$

We may now separate the two problems

$$\delta \mathbf{t}_R = \mathbf{H}_R \ \delta \mathbf{h} \ + \mathbf{A}_R \ \gamma, \tag{2.25}$$

and

$$\delta \mathbf{t}_N = \mathbf{A}_N \ \gamma, \tag{2.26}$$

where

$$\delta \mathbf{t}_R \in R^{4m_e \times 1},$$

$$\mathbf{H}_R \in R^{4m_e \times 4m_e},$$

$$\delta \mathbf{h} \in R^{4m_e \times 4},$$

$$\mathbf{A}_R \in R^{4m_e \times n_p},$$

$$\delta \mathbf{t}_N \in R^{m \times 1},$$

$$\mathbf{A}_N \in R^{m \times n_p},$$

$$m = M - 4m_e = \text{(total number of data) - 4} \times \text{(the number of events).}$$

For an individual event, equation (2.25) becomes

$$(\delta \mathbf{t}_R)_j = (\mathbf{H}_R)_j \, (\delta \mathbf{h})_j + (\mathbf{A}_R)_j \, \boldsymbol{\gamma}, \tag{2.27}$$

where

$$(\delta \mathbf{t}_R)_j \in R^{4 \times 1},$$

$$(\mathbf{H}_R)_j \in R^{4 \times 4},$$

$$(\delta \mathbf{h})_j \in R^{4 \times 1},$$

$$(\mathbf{A}_R)_j \in R^{4 \times n_p},$$

$$j = 1,2,\ldots,m_e.$$

Equation (2.26) is independent of hypocenter mislocation, $\delta \mathbf{h}$. We will solve it first, then use the solution obtained for $\boldsymbol{\gamma}$ to substitute into equation (2.27) for all events and solve for $(\delta \mathbf{h})_j$. For the sake of standardization, note that equation (2.26) is of the general form

$$\mathbf{G}\mathbf{x} = \mathbf{b}, \tag{2.28}$$

where

$$\mathbf{G} = \mathbf{A}_N,$$

$$\mathbf{x} = \boldsymbol{\gamma},$$

$$\mathbf{b} = \delta\mathbf{t}_N.$$

2.2.3 Simultaneous inversion

We may also choose to combine the three coefficient matrices and solve for all parameters simultaneously, i.e.

$$\left[\mathbf{A}|\mathbf{H}|\mathbf{S}\right]\begin{bmatrix}\boldsymbol{\gamma}\\ \hline \delta\mathbf{h}\\ \hline \delta\boldsymbol{\mu}\end{bmatrix} = \delta\mathbf{t}, \tag{2.29}$$

or

$$\mathbf{G}\mathbf{x} = \mathbf{b}, \tag{2.30}$$

where

$$\mathbf{G} = \left[\mathbf{A}|\mathbf{H}|\mathbf{S}\right],$$

$$\mathbf{x} = \begin{bmatrix}\boldsymbol{\gamma}\\ \hline \delta\mathbf{h}\\ \hline \delta\boldsymbol{\mu}\end{bmatrix},$$

and

$$\mathbf{b} = \delta\mathbf{t}.$$

H and **S** are first scaled so that each row has the same euclidean norm as the same row of **A**.

Equations (2.28) and (2.30) present us with a classical linear inverse problem. Typically, the $M \times N$ coefficient matrix, **G**, will have many more data than parameters ($M \gg N$) and, given that errors are contained in the data, the equations will be inconsistent. We need to adopt a criterion for minimizing the misfit of parameters to data.

We choose to minimize the euclidean (l^2) norm, resulting in the least squares problem:

$$Min\,\|\mathbf{Gx} - \mathbf{b}\|^2 = Min\,(\mathbf{Gx} - \mathbf{b})^T(\mathbf{Gx} - \mathbf{b}). \tag{2.31}$$

Differentiating equation (2.30) and setting the result equal to zero yields the normal equations

$$\mathbf{G}^T\mathbf{G} = \mathbf{G}^T\mathbf{b}. \tag{2.32}$$

These normal equations will commonly be numerically close to singular. A direct solution will produce either no parameter values at all, or will produce a set of large-magnitude, grossly disparate parameters that delicately offset each other nominally to satisfy the least squares criterion. We may direct the solution of the normal equations (2.32) toward a particular solution by appending additional equality constraints to equation (2.28). This is often called "ridge regression" or "damped least squares" and may be performed by appending the additional equations $\mathbf{Ix} = \xi$ to $\mathbf{Gx} = \mathbf{b}$. This expresses a preference for a solution vector, $\mathbf{x}$, that is close to the vector ξ, but leaves the degree of this preference to be determined implicitly by the relative magnitude of the elements of $\mathbf{I}$ and $\mathbf{G}$. To express the degree of preference explicitly, we introduce a scaling factor, λ. Equation (2.28) becomes

$$\begin{bmatrix}\mathbf{b}\\ \lambda\xi\end{bmatrix} = \begin{bmatrix}\mathbf{G}\\ \lambda\mathbf{I}\end{bmatrix}\mathbf{x}, \tag{2.33}$$

and the minimization we must perform is of the norm

$$\|\mathbf{Gx} - \mathbf{b}\|^2 + \lambda^2\|\mathbf{x} - \xi\|^2. \tag{2.34}$$

2.2.4 Incorporating a priori Information

2.2.4.1 *Conditioning the Data Space*

Solving equation (2.28) directly involves the implicit assumption that all the data have equal significance. In the absence of explicit weighting, all the rows of $\mathbf{G}$ are treated equally. Should we have greater confidence in some of the data, and wish to

avoid allowing these better observations to be overwhelmed by those of poorer quality, we may add a weighting matrix to the scheme. If, for example, we are able to estimate *a priori* the covariance matrix of the data, $\mathbf{C}_d$, we may multiply both sides of equation (2.28) by the weighting matrix $\mathbf{W}_d = \mathbf{C}_d^{-1/2}$. Here $\mathbf{W}_d$ represents the inverse matrix of standard errors of the data. In practice this left-multiplication serves as a row-scaling operation. Equation (2.33) becomes

$$\begin{bmatrix} \mathbf{W}_d\mathbf{b} \\ \lambda\xi \end{bmatrix} = \begin{bmatrix} \mathbf{W}_d\mathbf{G} \\ \lambda\mathbf{I} \end{bmatrix}\mathbf{x}, \tag{2.35}$$

and the least squares solution for $\mathbf{x}$ requires the minimization of

$$\|\mathbf{W}_d(\mathbf{G}\mathbf{x} - \mathbf{b})\|^2 + \lambda^2\|\mathbf{x} - \xi\|^2. \tag{2.36}$$

2.2.4.2 *Conditioning the Parameter Space*

If we recognize that an unwanted bias exists in the elements of $\mathbf{B}$ or have reasonable estimates of the uncertainties in the elements of ξ as an *a priori* estimate of $\mathbf{x}$ we may attempt to correct this pre-existing bias with a right-multiplication of $\mathbf{G}$ by a weighting matrix $\mathbf{W}_x$. For example, we might have an estimate of the *a priori* covariance matrix of the model, $\mathbf{C}_m$. In this case $\mathbf{W}_x$ would be the matrix of inverse model standard errors: $\mathbf{W}_x = \mathbf{C}_m^{-1/2}$. Our intention is to transform the solution vector, $\mathbf{x}$, to a vector in which all elements have approximately equal uncertainty and zero bias. We replace equation (2.35) with

$$\begin{bmatrix} \tilde{\mathbf{b}} \\ 0 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{G}} \\ \lambda\mathbf{I} \end{bmatrix}\tilde{\mathbf{x}}, \tag{2.37}$$

where

$$\tilde{\mathbf{G}} = \mathbf{W}_d\mathbf{G}\mathbf{W}_x,$$

$$\tilde{\mathbf{b}} = \mathbf{W}_d(\mathbf{b} - \mathbf{G}\xi),$$

$$\tilde{\mathbf{x}} = \mathbf{W}_x^{-1}(\mathbf{x} - \xi),$$

and, assuming both $\mathbf{W}_d$ and $\mathbf{W}_x$ are diagonal, minimize

$$\|\tilde{\mathbf{G}}\tilde{\mathbf{x}} - \tilde{\mathbf{b}}\|^2 + \lambda^2\|\tilde{\mathbf{x}}\|^2 = \|\mathbf{W}_d(\mathbf{G}\mathbf{x} - \delta\mathbf{t})\|^2 + \lambda^2\|\mathbf{W}_x^{-1}(\mathbf{x} - \xi)\|^2. \tag{2.38}$$

In our case, our best estimate of the solution vector, $\mathbf{x}$, is the zero vector, $\xi = 0$, which assumes that our starting model is correct. We do not attempt to estimate the uncertainties contained in $\xi = 0$ as an *a priori* estimate of $\mathbf{x}$, as does *Spakman* [1988]. Instead, we assume that all these uncertainties are unity, so (at this point) $\mathbf{W}_x = \mathbf{C}_m = \mathbf{I}$.

A closer examination of the procedure used to construct the coefficient matrix, $\mathbf{A}$, leads us to conclude that its elements contain bias from at least three sources. First, although the original voxels were constructed so as to have approximately equal surface areas, the voxels have widely varying volumes. Larger voxels will generally have longer ray segments, given a random or uniform sampling of voxels, than small voxels. This ultimately produces an $\mathbf{A}$ matrix that weights large voxels more heavily than small voxels. *Nolet* [1987] offers an elegant demonstration of the deleterious effect of differing volumes on the velocities obtained from inversion. Second, since the ray sampling is not uniform, the $\mathbf{A}$ matrix will tend to over-weight more heavily sampled voxels. A third source of bias is the non-uniform distribution of directions of rays propagating through a voxel. Geometrically, a set of parallel rays will produce weaker constraints on a voxel's velocity than a set of rays well-distributed over the three orthogonal directions. We attempt to reduce the sampling bias inherent in the formulation of $\mathbf{A}$ by right-multiplying equation (2.28) by a matrix representing the quality of sampling of Earth afforded by the ISC data set. Our $\mathbf{W}_x$ is diagonal, so right-multiplication serves as a column-scaling operation to balance the euclidean norms of the columns of $\mathbf{A}$. Now $\mathbf{W}_x$ consists of the elements

$$W_{x_{kk}} = \begin{cases} \|s_k\|^{-1} & \text{if } \|s_k\| \neq 0 \\ 0 & \text{if } \|s_k\| = 0 \end{cases},$$

where the vector s, of length n_p, is a measure of sampling quality. Examples of such weighting schemes include (a) the vector of voxel volumes (in which case the norm of the kth column of $\mathbf{W}_x$ is simply the inverse volume of the kth voxel), (b) the vector of voxel "hits", (c) the vector of norms of **A**'s column vectors, (d) a vector of factors indicating the distribution of directions of ray segments propagating through the voxels. The weighting scheme we employ consists of ratios of average sampling of a particular voxel to the size of the voxel, i.e.

$$W_{x_{kk}} = \frac{\frac{1}{n_k}\sum_i^{n_k} l_i}{\mathrm{v}_k^{1/3}}, \tag{2.39}$$

where

l_i = the length of the ith ray segment in voxel k,

v_k = the volume of the kth voxel,

n_k = the number of ray segments that sample the kth voxel.

In equation (2.37), **I** may be replaced by by a more general "conditioning" matrix, **B**, to influence the character of the solution vector $\tilde{\mathbf{x}}$. Appending equations $\mathbf{B}\tilde{\mathbf{x}} = 0$ to equations (2.37) is equivalent to right-multiplying $\mathbf{W}_d\mathbf{G}$ by **B**. Because the matrix **B** is full-rank (i.e., non-singular), the set of vectors $\mathbf{x} = \mathbf{B}\tilde{\mathbf{x}} + \xi$ where $\tilde{\mathbf{x}}$ minimizes $\|\tilde{\mathbf{b}} - \tilde{\mathbf{G}}\tilde{\mathbf{x}}\|$ is the same as the set of vectors **x** which minimizes $\|\mathbf{b} - \mathbf{G}\mathbf{x}\|$. However, because **B** is not normalized, the condition number of $\tilde{\mathbf{G}}$ will generally differ from that of **G**. The pseudorank of **G** (rank of $\tilde{\mathbf{G}}$) may be less than the rank of **G** and the minimization of $\|\tilde{\mathbf{x}}\|$ alters the norm by which we determine the "minimum length" vector. By minimizing $\|\tilde{\mathbf{x}}\|$ $(= \|\mathbf{B}^{-1}(\mathbf{x} - \xi)\|)$ instead of **x** we will generally choose a different vector from the set that satisfy equation (2.28) than we would choose from the set that satisfy equation (2.17) [*Lawson and Hanson*, 1974]. For clarity, and to keep the number of floating point operations to a minimum in our row-active

implementation, we use the format of appended rows for the smoothing operation and right-multiplication to equalize the *a priori* bias contained in columns of $\tilde{\mathbf{G}}$.

Following *Lees and Crosson* [1989], we wish to minimize the variation of velocity between adjacent voxels and seek to minimize the inverse of a discrete representation of a three-dimensional Laplacian operator applied locally in the neighborhood of each voxel. To do this, we append the rows $\mathbf{B}\tilde{\mathbf{x}} = 0$, where $\mathbf{B}$ is an $n_p \times n_p$ matrix in which, for the kth row of $\mathbf{B}$, $B_{kk} = 1$ and the columns corresponding to all adjacent voxels in the same layer contain elements equal to (# of adjacent voxels)$^{-1}$. Note that our model parametrization varies the size, in degrees, of the voxels as we move from the equator to the poles, so the number of immediate neighbors a given voxel has will depend on its location in the model. The kth row of the equation $\mathbf{Bx} = 0$ will be

$$x_k - \sum_n \frac{x_{neighbor}}{n} = 0,$$

where

n = the number of voxels adjacent to the kth voxel.

The equation we solve is

$$\begin{bmatrix} \tilde{\mathbf{b}} \\ 0 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{G}} \\ \lambda \mathbf{B} \end{bmatrix} \tilde{\mathbf{x}}, \tag{2.40}$$

where

$$\tilde{\mathbf{G}} = \mathbf{W}_d \mathbf{G} \mathbf{W}_x,$$

$$\tilde{\mathbf{b}} = \mathbf{W}_d \mathbf{b},$$

$$\mathbf{x} = \mathbf{W}_x \tilde{\mathbf{x}}$$

subject to the minimization of $||\tilde{\mathbf{x}}||^2$. It is important to recognize that once again we are minimizing an altered functional, so that the particular solution we choose will, in general, be different from the one we would choose if $\mathbf{B} \equiv \mathbf{I}$. One might minimize the

norm (2.20) to find $\mathbf{x}$ and then apply the matrix $\mathbf{B}$ to find a smoothed solution $\mathbf{x}_{smooth} = \mathbf{B}\mathbf{x}$, but this procedure is *not* equivalent to minimizing the conditioned norm $\|\mathbf{B}^{-1}\mathbf{x}\|^2$.

2.2.5 LSQR

Equation (2.29) is difficult and time-consuming to solve. Because the coefficient matrix, $\hat{\mathbf{G}}$, is $O(10^5 \times 10^4)$, most common methods are not practical. Most computers cannot accommodate the entire matrix in core memory, so I/O operations to retrieve each row of the matrix from a disk file slow the inversion process even more. We can take advantage of the sparseness of $\hat{\mathbf{G}}$, however, and store only the non-zero elements in a collapsed vector format. Still, memory requirements are daunting and row-active methods, such as LSQR (a variant of the conjugate gradient method), Algebraic Reconstruction Techniques (ART), and Simultaneous Iterative Reconstruction Techniques (SIRT), must be used.

Theoretically LSQR will converge to the true least squares solution in n iterations, where n is the size of the model space. Roundoff errors will interfere with this convergence property in practice, but we normally deal with numerically singular matrices of very large size, so we stop the algorithm after relatively few iterations. LSQR resembles Singular Value Decomposition in that it constructs its solution in a subspace of the model space that it generates by finding, at each iteration, one search direction vector that is orthogonal to all the vectors found previously. After p iterations, the solution is the vector $\mathbf{x}$ in the p-dimensional subspace that minimizes $\|\mathbf{G}\mathbf{x} - \mathbf{t}\|^2$, while also minimizing $\|\mathbf{x}\|^2$. Complete descriptions of the LSQR algorithm and its properties can be found in *Paige and Saunders* [1982] and *Nolet* [1985]. *Spakman and Nolet* [1988] and *van der Sluis and van der Vorst* [1987] conduct detailed comparisons of SIRT and LSQR algorithms.

2.3 Synthetic Tests of the Algorithms

In order to test the effectiveness of our algorithm in retrieving both velocity structure and source mislocations we conduct a controlled simulation of the general procedure used to produce tomographic images of the mantle from synthetic global earthquake data, similar to the data supplied by the ISC. These data consist of arrival times at reporting stations and estimates of earthquake locations calculated in a one-dimensional Earth model. Because of the three-dimensional nature of Earth, the ISC location estimates are only approximations to the true locations, so we investigate the effects of mislocating the earthquakes on the velocity model obtained in an inversion, and the ability of our simultaneous and progressive inversion techniques to correct mislocated earthquakes and produce an accurate velocity model.

The steps we follow to perform these simulations are outlined in figure 2.1. To produce synthetic data such as those provided by the ISC we distribute sources and receivers around a model Earth (step 1). We want to address the problem of source mislocation in a three-dimensional medium, and not the problem of poor ray coverage of Earth, so we intend to distribute sources and receivers adequately to allow accurate retrieval of velocity anomalies given "true" source locations (i.e., the starting source locations). To this end, we distribute nine sources around Earth, located at depths ranging from 40 to 180 km, and a total of 207 stations, for an average of 45 reporting stations per event (see figure 2.2). To check the adequacy of the geometrical constraint placed on the source location by the ray coverage we immediately re-locate the introduced sources using a damped least-squares procedure and the one-dimensional Jeffreys-Bullen (J-B) P velocity model (step 1a). At this point, no velocity anomalies are present in the model. The standard errors on these direct re-locations tell us the best we can expect to do later, when we correct the deliberately mislocated sources.

Next we introduce four velocity anomalies (step 2) and calculate travel times through the new 3-D model (step 3). The raytracing performed here is for a fully 3-D

medium. Travel times are calculated through the 3-D model by a shooting method involving the direct numerical integration of the eikonal equations that uses a Newton-type search for the solution to the two-point boundary value problem. We parametrize the model Earth with approximately equal-area voxels, 30° x 30°, at the equator and six layers, which makes each layer about 500 km thick and gives a total of 276 voxels. Figures 2.3 and 2.4 show the ray coverage for layers two and three. Highlighted in light gray and black are the voxels in which positive and negative velocity anomalies, respectively, are introduced. All four anomalous voxels are located in these two layers. Magnitudes of the anomalies range from 1.5% to 2.0% of the local velocity. These anomalies are located in reasonably well-sampled voxels, but not the most heavily sampled.

Our choice of a block model parametrization causes problems for the 3-D raytracing required to produce synthetic data. The eikonal equations can only be solved practically for a reasonably smooth model. Instead of a smooth model, our blocks confront the raytracer with an overwhelming set of discontinuities in both lateral and radial directions that cause unwarranted and physically implausible complexity. We smooth an introduced, "spike" anomaly by placing the anomalous velocity at the center point of its assigned voxel and requiring the value to decrease linearly toward the voxel boundaries. Thus the velocity experienced by each ray that visits an anomalous voxel will be well below the peak value located at the voxel center. This is a technical point that affects only the velocity perturbations returned by our inversion scheme and not the relative values as they are altered by the effects of source mislocation. To find the absolute values, we perform an inversion based on the travel times through the 3-D model from the true source locations (step 3a). In practice, these true locations are never known. The purpose of this exercise is to construct a controlled simulation in which we isolate the effects of just one type of error. Here we retain control of the velocity problem and seek to isolate the effects of source mislocation.

At this point, we have synthesized data analogous to those provided to the ISC by observers located around the world. Next we use the travel times through the 3-D model as arrival times to re-locate the sources in a 1-D model (step 4). Again we use the J-B model. Now we have a set of data analogous to the catalog provided by the ISC to researchers world-wide. Next we calculate residuals by subtracting the synthetic arrival times from the travel times through the J-B model from the new, re-located source (step 5) and invert these residuals in three ways:

1.) directly, neglecting source mislocation terms (step 6a),

2.) simultaneously for velocity and source terms (step 6b), and

3.) by means of the progressive inversion scheme, in which the velocity problem is separated from the source mislocation problem and solved separately (step 6c).

In each of these cases we use the conjugate-gradient projection method LSQR. Gaussian noise with mean and variance matching those of the residual distribution is added to the synthetic data. Finally, we compare the corrected source locations to the true locations and the estimates of the four velocity anomalies produced in each inversion, along with smearing and artifact anomalies (step 7).

2.3.1 Source Mislocations

Tables 2.1-2.9 show the source corrections resulting from both the simultaneous and progressive inversions. The tables show, for each source parameter, the "a priori" standard error in the first column. This is the standard error from the first computed location of the sources introduced to the J-B model and "located" with the FORTRAN program BERQLY (by Lane Johnson) in the J-B model with no anomalies present. These standard errors represent the best our algorithm can hope to achieve with the given ray coverage. The second column contains the initial parameter offset. For each parameter, these are the amounts the source re-located in the 3-D model differs from the true source location. Depending on the proximity of the introduced anomalies to the earthquake hypocenter, a given hypocenter will be moved a great deal (e.g., events

4 and 5) or only slightly (e.g., events 1, 2, and 7). Column 3 contains the results after source corrections obtained from simultaneous inversion have been applied to the initial parameter offset and it shows how far away the corrected location is from the true source location. The fourth column shows how much the simultaneous inversion improved the source location. Columns 5 and 6 present the same information as columns 3 and 4, but for corrections that emerge from the progressive inversion. To interpret these results, compare the second column to the first column of each table to see if the improvements indicated are significant. Is the initial offset greater than the "a priori" standard error? If so, does the correction applied reduce the parameter offset or increase it?

For example, the origin time, latitude, and longitude parameters for Events 1 and 7 are not significantly offset. Therefore the resulting corrections may be misleading. Events 2, 3, and 9 have unusually small initial offsets, locations this accurate would not require corrections anyway. For both the simultaneous and progressive inversions, event 4 shows significant improvement of an initially poorly located source. Perhaps the latitude and longitude terms are less significant. Event 5 has the most dramatic results. All parameters are initially offset a significant amount and for the progressive inversion, all but the longitude term were corrected to well within the "a priori" standard error. Particularly with respect to the origin time and source depth parameters, the progressive inversion performed better than the simultaneous inversion for these two most significant events, 4 and 5, as it generally did throughout these tests. Of particular concern is the poor estimation by the simultaneous inversion of the corrections to origin time and depth. Events 6 and 8 show good, though mixed, results. In all cases in which a parameter is offset an amount greater than the standard error, the correction produced by the progressive inversion reduces the offset to within the standard error. The simultaneous inversion produced just one exception to this rule (the longitude correction for event 8). In cases in which the initial offset is still within the

standard error the correction usually reduces the offset further, but it may also result in a greater offset from the true source. In all but one of the cases in which the offset rendered by the progressive inversion is an increase over the initial offset, the final offset is still within the parameter's "a priori" standard error. This is true for all but four of the offsets rendered by the simultaneous scheme. Apparently once an offset is within the standard error, attempts to decrease the mislocation further result in a waffling about within a range of the true value roughly bounded by the standard error. This "loose" bound, rather than than a "hard", inviolable bound is expected for the standard error. We tested this further by performing another iteration of the imaging scheme in which our source location estimates are updated by applying the corrections resulting from the first iteration and the entire algorithm is repeated, based on the new locations. Indeed, in every case in which the first iteration's offset still lay outside the "a priori" standard error, the second iteration improved the offset to within this standard error. In addition, for the progressive inversion five more parameters corrected after the first iteration to within the standard error jumped outside of the standard error after the second iteration. Perhaps if the initial offsets were larger, and generally more significant according to the "a priori" standard errors, the second iteration would be warranted and helpful. In our test case, the second iteration produced negligible improvement and, in fact, resulted in a degradation of source location estimates as often as improvement. The important result, however, is that the "a priori" standard error estimates allow the reliable determination of the significance of a particular correction. These standard errors are supplied by the ISC along with their location estimates.

Ideally, sources would be re-located in a three-dimensional model rather than with corrections produced as a by-product of an inversion for velocity. Both the location and velocity estimation problems are nonlinear and should be approached with an optimization scheme. But an iterative scheme for a fully three-dimensional Earth that

incorporates enough data and parameters to constrain interesting features of Earth is beyond our computational capacity at present. At this stage our greatest interest is in finding an inversion scheme that decouples the source location and velocity problems as much as is possible. In the next section we explore the consequences of removing the contributions of source mislocations from the travel time residual for our retrieval of a 3-D velocity model.

2.3.2 Velocity model

Table 2.10 shows estimates for the four input anomalies and a fifth entry for the next largest value emerging from the inversion. This fifth entry is the largest artifact anomaly and does not represent the same voxel across the bottom row of the table. Column 2 shows the number of rays sampling each model block. The most-sampled voxel had 96 hits; several voxels had more than 72 hits. The 3-D anomalies we introduced to the J-B model are indicated as "peak" anomalies in column 3. However, in order to trace rays through the three-dimensional model, given the model parametrization into discrete voxels, we first smooth the input velocity model. To smooth the input model, we place a "peak" anomaly at the center point of a voxel and constrained the anomaly to decrease linearly toward the voxel's boundaries. The average velocity encountered by each ray is therefore well below the "peak" anomaly. To find the actual image we are trying to recover, neglecting the effects of imperfect ray coverage (i.e. to assess the effects of smoothing the four input "spikes"), we invert residuals calculated by subtracting the synthetic travel times from travel times through the J-B model from the *true* source locations. This result, listed in column 4 of Table 2.10, contains the effects of imperfect ray coverage, which introduces a skewed average velocity depending on what parts of each voxel are sampled by rays and the type of function employed to smooth the input velocity "spikes". Columns 5 and 6 show the results of directly inverting the data from mislocated sources. Neglecting the effects of source offset results in underestimating the velocity anomalies by over 50% in some

cases,

and produces artifact anomalies with absolute values greater than estimates for true values. Columns 7 and 8 show the same information for the simultaneous estimation of source and velocity terms. Note the overshoot in two cases, gross underestimate in one case, and the large artifact anomaly. The last columns of Table 2.10 show the results of our progressive inversion scheme. Here we retrieve a much greater portion of the velocity, with slight overshoot in one case, and with artifacts attaining a maximum magnitude of about one quarter the smallest true anomaly.

2.3.3 Summary of Results

With the simulations described in this section we discover that ignoring the effects of source mislocation in a tomographic inversion results in underestimating velocity anomalies by up to 50%, creating smeared anomalies in adjacent voxels with values up to 50% of the retrieved velocity of its neighbor, and creating anomalies elsewhere in the mantle with values greater than those estimated for true anomalies. Simultaneous inversion for corrections to the source location and for a velocity model usually improves source locations when initial offsets are "significant" in the sense that ray coverage is distributed in azimuth and distance well enough to constrain the source location to a range smaller than the offset. Velocity estimates are generally accurate, though the magnitudes of the anomalies are less reliable. Also, entirely inaccurate anomalies, produced as artifacts of the inversion, reach disturbingly high values. Progressive inversion improves source locations 60-80% and successfully retrieves velocity anomalies after one iteration for velocity anomalies of 1-2%. The largest ghost image is small compared to the smallest true anomaly.

The success of these tests in correcting the source mislocation and in retrieving the overwhelming portion of the anomalous velocity is probably due to the small

source offsets produced by our four velocity anomalies. These small offsets are due, in part, to the relatively good constraints provided by our source-receiver geometry and, in part, to the small number of low-amplitude anomalies introduced to the velocity model.

These results were obtained with tomographic imaging based on raytracing through a one-dimensional velocity model. For larger anomalies, more iterations and three-dimensional raytracing may be necessary. However, computational requirements may not be feasible for such a scheme and better results are not guaranteed. When we perform a second iteration of our algorithms in which the source corrections are applied and rays are calculated from the new locations through the same 1-D model we started with initially, results for both source corrections and velocity terms are mixed. This is probably due to the success of the first iteration. The remaining offsets are small with regard to the standard errors of the first computed source locations.

In our row-active implementation, the progressive inversion scheme used 40% more CPU time than the direct LSQR in vectorized mode. Requirements for disk space (or core memory if the application is small enough to allow the coefficient matrix to be stored in core) is about 5 times the requirements of the simultaneous inversion. As the projections are performed in the progressive scheme, columns of the previously sparse coefficient matrix are filled in, resulting in a more dense matrix.

2.4 Inversions of Real Data

2.4.1 Data selection

The data inverted in this study were obtained from the catalog of the International Seismological Centre (ISC) for the period January 1964 through January 1987 (frontispiece). To avoid contamination of our mantle phases by Earth's core we limit the range of our coverage to epicentral distances between 0° and 96°. The scatter caused by refractions from the 400 km and 670 km discontinuities, at about 15° to 25°, is

dealt with in the inversion process by weighting each summary ray by the inverse of the standard error of travel time residuals as a function of delta. To ensure that sources are well-located, each event must have a minimum of forty reporting stations, and source depths, as reported by the ISC, must be greater than 0 km and less than 70 km. In addition, maximum standard errors for the ISC locations must be 1 sec for origin time, 0.1° for both latitude and longitude, and 10 km for depth. We discard all events located by the ISC at Earth's surface, but retain events located at the other default depths. To ensure adequate and reasonably uniform ray coverage of Earth we keep a maximum of twenty-five events in each voxel. The set of events retained for each voxel always includes the events with the most observations. In this study we do not form summary rays. All observations are corrected for ellipticity by numerical integration along the raypath of the travel time perturbation arising from deviations of a hydrostatic figure from a sphere. Lengths of ray segments in voxels are found by integrating distance along the curved raypath and finding the intersections of rays with voxel boundaries. Rays associated with residuals greater than seven seconds are discarded. Approximately 345,000 rays satisfy these criteria. Figures 2.6 and 2.7 show the locations of the selected events and seismographic stations, respectively.

Figure 2.9 shows a histogram of travel time residuals binned in 1° intervals associated with sources located by the ISC at depths between 0 and 70 km, inclusive. Poorly constrained events are assigned by the ISC to default depths of 0, 5, 10, 15, and 33 km. We examined histograms of travel time residuals associated with events assigned by the ISC to these five different default depths, and compared the residual distributions for these events to the residual distribution of remaining events. The travel time residuals associated with sources located at 0 km depth (figure 2.10) show a much different distribution than that of the remaining residuals. Due to a problem with our FORTRAN subroutine, ISC records in which the source depth was left blank defaulted to zero source depth. Although the residual distribution shows a clear

bimodal pattern, we are unable to distinguish reliably between true, zero-depth locations and defaulted locations after the data are been extracted from the ISC master set. For this study we discard all events with source depth equal to zero. Histograms for the remaining default depths are nearly identical to the histogram for all remaining events, so we cannot justify culling events with source depths of 5, 10, 15, or 33 km. A histogram of the winnowed data set, along with the first four moments of the travel time residual distribution is shown in figure 2.11.

2.4.2 Model

The starting model used in this study is a one-dimensional, spherically-symmetric P velocity model modified from *Jeffreys* [1960]. Modifications to the Jeffreys velocity model are necessary to obtain a model consistent with the Jeffreys-Bullen [*Jeffreys and Bullen*, 1940] travel time tables. These modifications are small but important because they remove a systematically slow trend for the mid-mantle from the model published by Jeffreys and make the model more consistent with the tables, that were used by the the ISC to find source locations originally.

The model mantle is divided into 14 layers, approximately 200 km thick, with radial boundaries located at Earth's major discontinuities. Each layer contains 406 approximately equal area voxels, 10° x 10° at the equator, for a total of 5,684 model parameters. The exceptions to the 200 km thick layers occur in the upper mantle, in order to place a radial boundary at the 670 km discontinuity (resulting in a 270 km thick layer) and above the core-mantle boundary, where the lowermost layer is 228 km thick. One layer of our model parametrization is shown in figure 2.8.

Figures 2.12a-h show the ray coverage of the mantle provided by the approximately 345,000 observations included in our data set. Sampling is described in terms of the number of rays that traverse each voxel. The most-sampled voxel has over 50,000 samples. Only 166 of 5684 voxels are unsampled. Figure 2.12a shows the

clear demarcation of plate boundary source regions that, along with Asia, North America, Europe, and Australia, are well-sampled. In contrast, other regions tend to be quite poorly-sampled. There are also large oceanic areas in the first few layers that are completely unsampled by our data set. These voxels do not enter into the inversion. The next depth layer, figure 2.12b, shows a broadening of the well-sampled regions and a slight reduction of the unsampled oceanic areas. At 400-670 km and 670-870, figures 2.12c and d, these trends continue, and by the mid-mantle, figures 2.12e and f, virtually all voxels are sampled. In general, sampling becomes more homogeneous with depth and at the bottom of the mantle, figures 2.12g and h, the sampling is much more uniform than in the first layer. Note in all eight figures the strong bias toward the northern hemisphere, in general, and toward continents in particular. However, in absolute numbers the sums of ray segments in voxels decrease with depth, even as more voxels are sampled in each layer. Table 2.11 details the average number of hits for sampled voxels in each layer along with the the average sum of ray segments in a voxel at a given depth and the number of voxels sampled in each depth interval. These averages include only voxels that have non-zero sampling. The trends in Table 2.11 show that while homogeneity of sampling increases with depth, voxels tend to be less frequently and less heavily traversed by recorded seismic rays.

2.4.3 Inversion Results

Inversions of the ISC data were performed directly, neglecting source terms, simultaneously for source mislocation and velocity terms, and progressively for each set of terms. The resulting models are named ISC10_direct, ISC10_sim, and ISC10_pro, respectively. The weights applied in each inversion are identical and the LSQR algorithm is performed for 20 iterations in each case. Convergence was determined by the relative change of the residual norm after each iteration. At 20 iterations, each model produces a slightly different variance reduction of the travel time residual distribution. For model ISC10_direct the variance reduction is 12%, for

model ISC10_sim it is 14%, and for model ISC10_pro it is 16%. Minimum and maximum velocity perturbations are [−1.8%, 2.0%] for ISC10_direct, [−1.5%, 1.9%] for ISC10_sim, and [−2.1%, 2.2%] for ISC10_pro.

Figures 2.13a-f, 2.14a-f, and 2.15a-f show six of the fourteen layers for each inversion. Despite the large voxel size, the top layers, 0-200 km depth (figures a) and 200-400 km (figures b), show quite strong correlations with surface tectonics. All models show fast anomalies in the Asian, Australian, and North American shield regions. The Indian subcontinent and southern Africa are consistently fast in all the models' top layers. Also in the top layers, a ring of slow anomalies surrounds the Pacific basin, though the ring is not as continuous a feature of the progressive model. Nevertheless, the Central American subduction zone, Nazca Plate, Galapagos hotspot, northwestern South America and all of the North American Great Basin and Range Province, including the Yellowstone and Raton, New Mexico hotspots, are covered by a broad, unusually slow anomaly. All models share this feature in the 0-200 km layer. In the 200-400 km depth range ISC10_direct and ISC10_sim show an intruding fast anomaly that extends across northern Mexico and Baja California while ISC10_pro remains slow, consistent with the layer above. The first two models show this same fast anomaly in the 400-670 km depth range while ISC10_pro remains slow.

Elsewhere around the Pacific, slow backarc basins appear to compete with fast subducting lithosphere to claim the dominant anomaly for a particular region. From southern Alaska westward along the Aleutian island arc all the models begin with a fast anomaly and switch to a slow anomaly as the backarc basin comprises a larger portion of the next voxel to the west. Still further west, the next voxel also includes parts of the Kurile arc as well as the Aleutian arc and model ISC10_pro returns a positive anomaly while ISC10_direct and ISC10_sim are marked by slow anomalies. Similar differences between the first two models and the progressive model appear in the northern Japanese, Mariana, Philippine, Micronesian, Tongan, and Chilean

subduction zones.

Further similarities between all three models include slow anomalies in East Africa, which are associated with a broad slow anomaly that persists through the first three layers of each model, the Mediterranean Sea, and the Hawaiian/Emperor hotspot. Another common feature is the abrupt change at the 400-670 km layer in each model of the sign of the anomaly associated with continental shields. Only the Australian shield remains largely a fast region. In the northern hemisphere, slow anomalies have displaced the fast anomalies in the continents' southern portions.

Further differences between the models arise in a comparison of mid-ocean ridges. Models ISC10_direct and ISC10_sim are fairly consistently marked by slow anomalies in the top layers, with exceptions arising almost exclusively in the southern hemisphere, where ray coverage is relatively poor. ISC10_pro shares the same negative sign for most anomalies, but ridges in the southern hemisphere are more consistently slow and the mid-Atlantic ridge is not marked by the same broad slow anomaly as in the first two models.

Figures 2.13d-f, 2.14d-f, and 2.15d-f present the three models for the 1270-1470 km, 1470-1670 km, and 2470-2670 km depth ranges, respectively. In the mid-mantle (figures d and e), models ISC10_direct and ISC10_sim show larger-scale anomalies, less broken by small-scale intrusions, than does model ISC10_pro. Surprisingly, continental regions in the northern hemisphere are generally associated with fast anomalies and oceanic regions are generally associated with slow anomalies. Continents in the southern hemisphere are not marked by fast anomalies. Most striking are fast anomalies beneath eastern North America, the Caribbean, and northwestern South America, and the fast features beneath Tonga and Japan/eastern Asia. The fast anomaly beneath eastern North America and the Caribbean appears in the same location as a large S-velocity anomaly reported by *Grand* [1987]. Similar features for P velocity appear in the inversions performed by *van der Hilst* [1990] for the Caribbean and

Central American region. The fast feature beneath Tonga broadens and continues to dip to the west to a depth of 1670 km. Beneath Japan and eastern Asia the fast anomaly is diffuse but extends all the way to the core-mantle boundary.

The 2470-2670 km depth layer (figures f) shows a more broken, fast pattern beneath the Pacific basin at the mantle's bottom. A ring of slow anomalies around the Pacific is emerging, but is not nearly as strong as the ring observed by *Dziewonski* [1984], *Morelli and Dziewonski* [1985, 1986], and *Clayton and Comer* [1983; *Hager and Clayton*, 1988] in the lowermost mantle. Seeking to avoid contamination of our data by diffractions at the core-mantle boundary, we impose an epicentral distance limit of 96° on our observations. The resulting ray coverage does not allow us to be confident of our results for the lowermost layer (D"). Regardless of differences between our models and models produced previously by others, it is clear that differences between the upper mantle layers of our three models are greater than are differences between layers of the lower mantle.

2.4.4 Comparison of Small-scale Model Features

The upper mantles of our three models apparently differ from each other more than do the lower mantles. To test this observation more rigorously at the scale of individual model blocks we employ a statistical correlation technique. Because we do not know the probability distribution function from which our sample model values are drawn we prefer a non-parametric procedure, and since we already know that our models generally differ in the amplitudes of individual model values we are most interested in a technique that compares the heterogeneity patterns of two models rather than the individual values of heterogeneity. For these reasons we choose to evaluate model layer correlations with a non-parametric rank-order correlation procedure. When comparing a given layer of two models we replace each velocity value from the first model with its rank among the $N-1$ other values in the same layer, and do

likewise with values from the same layer of the second model. Now the series to be examined for correlation consist of integers, 1 to N, that are drawn from a perfectly known distribution. If some of the velocity values are identical, they are assigned rank equal to the mean of the rank they would have had were they distinct. This assigned rank will not, and need not, necessarily be an integer. Regardless, the sum of all assigned ranks will equal the sum of the numbers 1 to N. What results is two sets of rankings, generally the integers 1 to N, for which statistics have been invented and well-used. As the most straight-forward of the common rank-order correlation statistics we choose to employ the Spearman statistic, which is defined as

$$r_s = \frac{\sum_{i=1}^{N}(R_i-\bar{R})(S_i-\bar{S})}{\left[\sum_{i=1}^{N}(R_i-\bar{R})^2\right]^{1/2}\left[\sum_{i=1}^{N}(S_i-\bar{S})^2\right]^{1/2}} \tag{2.41}$$

where

R = series indicating the rankings of the first model's velocity values for a given layer,

S = series indicating the rankings of the second model's velocity values for the same layer.

Figure 2.16 shows the Spearman rank-order correlations between layers of each set of two models. Clearly the models ISC10_direct and ISC10_sim are quite similar even at the scale of individual model blocks. Still, the small differences that do exist tend to be located in the upper portions of the models. These difference disappear with depth. Correlations between models ISC10_sim and ISC10_pro are fairly consistent throughout the lower mantle, though a small peak appears again at the 1270-1470 km depth layer. In the upper mantle, however, the top two layers display a marked decrease in correlation. This concentration of differences in the top layer, in which all

our sources are located, and the next lower layer indicates that our inversion procedures' different treatment of the source location terms portends important differences only for these layers at short wavelengths.

2.4.5 Comparison of Large-scale Model Features

To serve as a low-pass filtering procedure, we calculate surface spherical harmonic series expansions to degree 10 by integration around the globe for each coefficient, rather than by fitting coefficients to model values by least squares. The associated Legendre polynomials are fully normalized, i.e.,

$$p_l^m(\theta) = \left[(2-\delta_{m,0})(2l+1)\frac{(l-m)!}{(l+m)!}\right]^{1/2} P_l^m(\cos\theta). \tag{2.42}$$

Figures 2.17, 2.18, and 2.19 show the total power in the series expansions for each model plotted as a function of depth. The distribution of power with depth is quite similar for ISC10_direct and ISC10_sim, though the total power contained in the direct model is greater than that in ISC10_sim. ISC10_pro shows a slightly different pattern. Unlike the first two models, the most heterogeneous layer is the topmost, 0-200 km. The anomalously low power in the 200-400 km layers of all our models is probably due to the fact that rays bottoming in this layer, which emerge at the epicentral distance range $15° \leq \Delta \leq 20°$, have the largest variance of all the travel time residuals. These rays are the most sensitive to velocity perturbations in the 200-400 km layer, but in our inversion their influence on the final model is downweighted by the inverse of the residuals' standard errors. The transition zone, 400-670 km, contains the highest power in models ISC10_direct and ISC10_sim, indicating the greatest heterogeneity in these models occurs at these depths. Note that the absolute magnitudes of these transition zone power totals are comparable to, and do not exceed, the power in the ISC10_pro transition zone. Because our starting model does not contain discontinuities, our theoretical ray coverage of the transition zone is more uniform than

is the case for models of the mantle that include discontinuities at 400 and 670 km depth. As a result, we are probably mapping more power into the transition zone for all three models than is justified. Deeper in the mantle a relative peak appears at about 1300 km depth and is followed by diminished heterogeneity at greater depths. This increase in power at the 1270-1470 km depth layer may be attributed to anomalously large $l = 1$, 2, and 3 components in all three models. At the bottom of the mantle, ISC10_direct and ISC10_sim show a dramatic increase in heterogeneity while the ISC10_pro shows only a modest increase. The drop in power from the 2470-2670 km layer to the lowermost layer, 2670-CMB, is probably due to the poor ray coverage in this bottom layer that results from our epicentral distance limit of 96°.

Figures 2.20, 2.21, and 2.22 show, for each model, the power in series expansions of each layer as a function of angular degree. As figures 2.17-19 would lead us to expect, power at all degrees in the top two layers of models ISC10_direct and ISC10_sim is smaller than the power contained in the top two layers of ISC10_pro. Particularly striking are the large values of the $l = 5$ and $l = 6$ components and the consistent importance of the $l = 6$ component throughout the upper third of the mantle. The finding of a large $l = 2$ component in the transition zone confirms previous reports, but a prominent $l = 3$ also appears in all models. The progressive model shows a large $l = 6$ harmonic as well. In the mid-mantle, 1070-1670 km, the $l = 2$ and 3 components rise above the higher-degree harmonics with nearly the same pattern for all models. More differences arise in the lowermost mantle, where the dominant heterogeneity of model ISC10_pro is concentrated in the $l = 3$ term, while the first two models show anomalous $l = 1$ components.

Since sign information is not included in power calculations, figures 2.20-2.22 do not offer any clues as to how the distribution patterns for all layers combine constructively or destructively to form a pattern for the whole mantle. Figures 2.23-2.25 show the power in the spherical harmonic expansions for the respective models averaged

through the whole mantle and through the upper and lower mantle separately. The averaging is performed on the the individual harmonic coefficients, weighted at each layer by the square of the layer mid-point's radius, which normalizes the power in each layer to the layer's surface area. Here, differences between the models appear most dramatically. For the upper mantle the ISC10_pro (figure 2.25) power spectrum shows a dominant $l = 6$ component, along with prominent $l = 2$, 5, and 8 terms. In contrast, the ISC10_direct (figure 2.23) and ISC10_sim (figure 2.24) power spectra show no constructive patterns other than a quite prominent $l = 2$ pattern. When averaged over the lower mantle alone the three models show quite similar patterns, apart from a slightly more prominent $l = 2$ harmonic in the progressive model. Apparently there is some type of compensation at work, either numerical tradeoff between layers of our computed models or physical compensation of velocity heterogeneity in the real Earth. When individual layer series are averaged over the entire mantle, the prominent patterns of the ISC10_pro upper mantle nearly disappear and the components of the ISC10_direct and ISC10_sim upper mantles decrease in power. Only the large $l = 2$ term survives the whole mantle average.

In an effort to find the location in the mantle of the primary long-wavelength differences between our three models, we calculate correlation coefficients between spherical harmonic expansions of a given layer for two models at a time. We employ a correlation coefficient, r, defined as

$$r = \frac{\sum_{m=0}^{l} (R^{lm1}S^{lm1}+R^{lm2}S^{lm2})}{[\sum_{m=0}^{l} (R^{lm1^2}+R^{lm2^2})]^{1/2}[\sum_{m=0}^{l} (S^{lm1^2}+S^{lm2^2})]^{1/2}}, \tag{2.43}$$

where

R^{lmi} = first model harmonics (superscript i = 1 refers to cosine term, i = 2 refers to sine term),

S^{lmi} = second model harmonics.

Figures 2.26, 2.27, and 2.28 show the correlations between spherical harmonic degrees as functions of depth between models ISC10_direct and ISC10_pro. The differences found at small-scales (figure 2.16) are not duplicated exactly here, but some trends are similar. Most degrees show relatively low correlation in the first two layers, 0-200 km and 200-400 km, which improves in the upper part of the lower mantle. Correlations in the lowermost mantle are erratic, though the higher degrees (l = 8, 9, and 10), show a gradual decrease in correlation with depth. Exceptions to these general trends include l = 4 (figure 2.26) and l = 6 (figure 2.27). Results in the lowermost mantle do not reflect the results found for the small-scale correlations, but simulations performed elsewhere (see chapter 3) with long-wavelength patterns indicate that higher_degree harmonics are not resolved reliably in the lower mantle. Most important to our purpose here is the observation that correlations between models are generally weak in the upper mantle.

2.4.6 Source Corrections

Figures 2.29 and 2.30 show source correction vectors produced by the simultaneous and progressive inversions, respectively, for the same 400 events. These relocations are representative of the corrections required by each inversion procedure for the 3077 events used in this study. In each case the starting location (found by the ISC) is indicated with either an asterisk or hexagon and a scaled vector points in the direction of the correction required by the latitude and longitude adjustments. For each event the direction of the depth correction is indicated by the type of symbol marking the relocation vector's endpoint. Events that receive a shallowing correction are marked with asterisks; events that are relocated deeper are marked by a hexagon. Origin time corrections are not shown. Both figures 2.29 and 2.30 are scaled to the same maximum vector length.

The most striking feature of these figures is that the corrections produced by the progressive inversion (figure 2.30) are clearly larger than the corrections produced by the simultaneous inversion (figure 2.29). Overall the progressive corrections are generally two to four times the simultaneous corrections, although in some cases the progressive corrections are far greater. The corrections required in remote regions, where we might expect the constraints provided by the station distribution to be relatively weak, are not generally the largest in either case. The simultaneous inversion, particularly, produces small corrections in such remote regions as the Carlsberg and Indian Ocean ridges. The progressive inversion produces much larger corrections, but these corrections are not generally larger than the corrections produced in well-instrumented regions such as western North America, the eastern Mediterranean, and the Japanese subduction zone. The progressive corrections do not claim to be unerringly correct, they simply find that more of the travel time residual can be explained by moving the source than does the simultaneous, which finds a best-fitting location. Recall that poorly-located events, as determined by the standard errors of the ISC locations, were culled from our data set originally. Each source used here is one of the best-constrained of the ISC events located in its vicinity.

Though the corrections from the two inversions differ in size, some recognizable patterns and similarities emerge from a comparison of figures 2.29 and 2.30. In the northern parts of Japan, both sets of corrections are overwhelmingly oriented northward and slightly west of north. These corrections are some of the largest in each set. Whether this direction is correct, or at least expected from what we know of the location of lithosphere subducting underneath Japan, depends on whether the events actually occur predominantly on one side of the descending slab or are well-distributed throughout the slab. The depth corrections are moderate ($5 < \delta h \leq 10$) for both inversions. Along the west coast of North America an interesting pattern shows up in both sets of corrections. Proximate events, in regions well-covered by seismic

instrumentation, are corrected large amounts in nearly opposite directions. This result, along with non-systematic and unexpected corrections elsewhere, such as at Hawaii, the Mediterranean Sea and Middle East, lead us to the expected conclusion that the progressive inversion did not find a realistic source correction.

2.4.7 Discussion

Our visual comparison of models ISC10_direct, ISC10_sim, and ISC10_pro reveals that upper mantle differences are located overwhelmingly in source regions, implying that the model differences result from the differences in our treatment of the source terms. That the differences between models ISC10_direct and ISC10_sim are smaller than the differences between ISC10_sim and ISC10_pro is surprising, but is probably due to the fact that the progressive scheme actually extracts all of the travel time residual that may be attributed to source mislocation, not just the portion that is independent of the velocity terms. Despite the concentration of model differences in source regions, the effects of different treatments of the source appear in the velocity models at both long and short wavelengths in the upper mantle. Differences between the models diminish with depth.

Our comparison of the source corrections produced by simultaneous and progressive inversion support this interpretation. Source corrections emerging from the progressive inversion are generally two to four times greater than the simultaneous inversion's corrections. Some events get extremely large corrections with the progressive inversion, but only moderate corrections with the simultaneous inversion.

Progressive inversion is intended to remove as much of the travel time residual as can possibly be attributed to source mislocation and then use the remainder to find a velocity model. While this procedure might be expected to remove the contamination due to source mislocation from the velocity inversion, it should also tend to overestimate source corrections. Included in the source terms are the effects of vagaries in ray

coverage and some portion of the noise in the data. These effects show up most noticeably in the estimates of origin times and source depths. The direct P phase is notorious for its poor control of these two parameters. Due to the downward takeoff angles of P phases from the source, there is generally a strong tradeoff between them. The addition of pP and/or S phases to the inversion better constrains these two parameters [*O'Connell and Johnson*, 1991].

2.5 Conclusions

For simulations of a global tomographic inversion scheme, ignoring the effects of source mislocation results in underestimating velocity anomalies by up to 50%, creating smeared anomalies in adjacent voxels with values up to 50% of the retrieved velocity of its neighbor, and creating anomalies elsewhere in the mantle with values greater than those estimated for true anomalies. We would expect these results to be even more exaggerated in the real case, in which source mislocations and velocity contrasts are greater than in our simulations. Clearly, careful treatment of the source location problem is critical to the accurate retrieval of three-dimensional velocity variations. The progressive inversion developed here generally produces more accurate source corrections and velocity anomaly estimates than does an inversion scheme in which both source corrections and velocity terms are found simultaneously. These results are superior particularly with respect to the suppression of artifact anomalies in the velocity estimation. The success of the progressive scheme, and the satisfactory performance of the simultaneous scheme, may be attributed to the strong geometrical constraints provided by our set of stations on the source locations and the relatively small amplitudes of the introduced velocity anomalies. There is no reason to expect that corrections to source locations provided by a progressive inversion will be more accurate that those provided by a simultaneous inversion for cases in which reporting stations are not well-distributed around a source or additional systematic errors are present in the data. In such cases, the progressive inversion will generally

overestimate the source correction in its effort to find a velocity model free of contamination. Looking at the source mislocation terms themselves reveals that the sizes of corrections emerging from the progressive inversion are generally two to four times greater than corrections produced by the simultaneous inversion. In some cases the progressive corrections are far greater. In the simultaneous case, the relative scaling of velocity and source terms is critical. With a judicious choice of weights, one could emphasize fitting the source terms at the expense of the fit to the model, but this is not the same as first extracting the entire portion of the travel time residual that may be explained by source mislocation. In the first case, the simultaneous inversion with hypocentral partial derivatives more heavily weighted than velocity coefficients, the algorithm is still trying to find a best-fitting solution to the source mislocation problem. In the second case, the progressive inversion, the algorithm is not finding a solution for the source corrections initially. It simply considers the geometry of the problem, including the strength of the constraints on hypocenter location, and finds the maximum travel time discrepancy that might be accounted for by moving the source location.

As our inversions of real data demonstrate, the two inversion schemes produce clearly different velocity models. Moreover, these differences are concentrated near the models' surfaces, in general, and in source regions, in particular. The bulk of the models' differences are therefore due to the algorithms' differences in their treatment of source terms, and, as the simulations performed at the beginning of this chapter attest, the progressive inversion is more successful at retrieving accurate estimates of velocity anomalies.

2.6 References

Clayton, R.W., and R.P. Comer, A tomographic analysis of mantle heterogeneities from body wave travel times, *EOS Transactions AGU, 64,* 776, 1983.

Dziewonski, A.M., Mapping the lower mantle, Determination of lateral heterogeneity

in P velocity up to degree and order 6, *J. Geophys. Res., 89,* 5929–5952, 1984.

Grand, S., Tomographic inversion for shear velocity beneath the North American Plate, *J. Geophys. Res., 92,* 14065–14090, 1987.

Hager, B.H., and R.W. Clayton, Constraints on the structure of mantle convection using seismic observations, flow models, and the geoid, in *Mantle Convection,* W. R. Peltier (Ed.), Gordon and Breach, New York, 657–763, 1989.

Jeffreys, H., and K.E. Bullen, *Seismological Tables.* British Association for the Advancement of Science, London, 1940.

Jeffreys, H., *The Earth,* Cambridge University Press, London, 1960.

Jordan, T.H., and K.A. Sverdrup, Teleseismic location techniques and their application to earthquake clusters in the South–Central Pacific, *Bull. Seis. Soc. Am., 71,* 1105–1130, 1981.

Kennett, B.L.N., and P.R. Williamson, Subspace methods for large–scale inversion, in *Mathematical Geophysics,* N.J. Vlaar, G. Nolet, M.J.R. Wortel, and S.A.P.L. Cloetingh (eds.), Reidel, Dordrecht, pp. 139–154, 1988.

Lawson, C.L., and R.J. Hanson, *Solving Least Squares Problems.* Prentice–Hall, Englewood Cliffs, N.J., 340 pp., 1974.

Lees, J.M., and R.S. Crosson, Tomographic inversion for three–dimensional velocity structure at Mount St. Helens using earthquake data, *J. Geophys. Res., 94,* 5716–5728, 1989.

Michelena, R.J., and J.M. Harris, Tomographic traveltime inversion using natural pixels, *Geophysics, 56,* 635–644, 1991.

Morelli, A., and A.M. Dziewonski, Stability of aspherical models of the lower mantle, *EOS Transactions AGU, 66,* 975, 1985.

Morelli, A., and A.M. Dziewonski, 3D structure of the Earth's core inferred from travel–time residuals, *EOS Transactions AGU, 67,* 311, 1986.

Nolet, G., Seismic wave propagation and seismic tomography, in *Seismic Tomography,*

G. Nolet (ed.), Reidel, Dordrecht, pp. 1–23, 1987.

Nolet, G., and R. Snieder, Solving large linear inverse problems by projection, *Geophys. Jour. Int., 103,* 565–568, 1990.

O'Connell, D.R.H., Seismic velocity structure and microearthquake source properties at the Geysers, California, geothermal area, *Ph.D. Dissertation,* University of California, Berkeley, 1986.

O'Connell, D.R.H., and L.R. Johnson, Progressive Inversion for Hypocenters and P wave and S wave Velocity Structure, Application to the Geysers, California, Geothermal Field, *J. Geophys. Res., 96,* 6223–6236, 1991.

Paige, C. C., and M.A. Saunders, LSQR, An algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Software, 8,* 43–71, 1982.

Pavlis, G.L., and J.R. Booker, The mixed discrete–continuous inverse problem, application to the simultaneous determination of earthquake hypocenters and velocity structure, *J. Geophys. Res., 88,* 4801–4810, 1980.

Spakman, W., Upper mantle delay time tomography, *Ph.D. Dissertation,* University of Utrecht, The Netherlands, 1988.

Spakman, W., and G. Nolet, Imaging algorithms, accuracy and resolution in delay time tomography, in *Mathematical Geophysics,* N.J. Vlaar, G. Nolet, M.J.R. Wortel, and S.A.P.L. Cloetingh (eds.), Reidel, Dordrecht, pp. 155–187, 1987.

Spencer, C., and D. Gubbins, Travel–time inversion for simultaneous earthquake location and velocity structure determination in laterally varying media, *Geophys. J. R. Astron. Soc., 63,* 95–116, 1980.

van der Sluis, A., and H.A. van der Vorst, Numerical solution of large, sparse linear systems arising from tomographic problems, in *Seismic Tomography,* G. Nolet (ed.), Reidel, Dordrecht, pp. 53–87, 1987.

Wielandt, E., On the validity of the ray approximation for interpreting delay times, in *Seismic Tomography,* G. Nolet (ed.), Reidel, Dordrecht, pp. 85–98, 1987.

2.7 List of Tables

Table 2.1 The source corrections resulting from the simultaneous and progressive inversions for Event 1, located under the Kamchatka Peninsula. For each source parameter the first column indicates the "a priori" standard error. This is the standard error from the initial location of the sources introduced into the J-B model and located in the J-B model with no anomalies present. These standard errors represent the best any algorithm can hope to achieve with the given ray coverage. The second column contains the initial parameter offset. For each parameter, these are the amounts the sources re-located in the 3-D model differ from the true source locations. Column three contains the results after source corrections obtained from the simultaneous inversion have been applied to the initial parameter offset and it shows how far away the corrected source are from the true sources. The next column shows how much the simultaneous inversion improved the source locations. Columns five and six contain information similar to columns three and four, but for corrections that emerge from the progressive inversion.

Table 2.2 The information contained in Table 2.2 is similar to the information contained in Table 2.1, but for Event 2, located in the Chilean Subduction Zone.

Table 2.3 The information contained in Table 2.3 is similar to the information contained in Table 2.1, but for Event 3, located in the Mariana Subduction Zone.

Table 2.4 The information contained in Table 2.4 is similar to the information contained in Table 2.1, but for Event 4, located in Mongolia, China.

Table 2.5 The information contained in Table 2.5 is similar to the information contained in Table 2.1, but for Event 5, located in the Aleutian Islands.

Table 2.6 The information contained in Table 2.6 is similar to the information contained in Table 2.1, but for Event 6, located in the Central American

Subduction Zone.

Table 2.7 The information contained in Table 2.7 is similar to the information contained in Table 2.1, but for Event 7, located in the South China Sea.

Table 2.8 The information contained in Table 2.8 is similar to the information contained in Table 2.1, but for Event 8, located in the Himalayas.

Table 2.9 The information contained in Table 2.9 is similar to the information contained in Table 2.1, but for Event 9, located in the South Pacific Ocean.

Table 2.10 Estimates for the four input anomalies and a fifth entry for the next largest value emerging from the inversion. This fifth entry is the largest artifact anomaly and does not represent the same voxel across the bottom row of the table. Column 2 shows the number of rays visiting each anomalous block. We place a "peak" anomaly at the center point of a voxel and constrain the anomaly to decrease linearly toward the voxel's boundaries. To find the actual image we are trying to recover, neglecting the effects of imperfect ray coverage, we invert residuals calculated by subtracting the synthetic travel times from travel times through the J-B model from the *true* source locations. This result is listed in column 4. Columns 5 and 6 show the results of directly inverting the data from mislocated sources. Columns 7 and 8 show the same information for the simultaneous estimation of source and velocity terms. The last two columns show the results of our progressive inversion scheme.

Table 2.11 Details of the model parametrization and the sampling provided by our data set. Include are the average number of hits for sampled voxels in each layer along with the the average sum of ray segments in a voxel at a given depth and the number of voxels sampled in each depth interval. These averages include only voxels that have non-zero sampling.

2.8 List of Figures

Figure 2.1 Flow chart outlining the steps performed in the tomography simulations.

Figure 2.2 Locations of sources (large gray spheres) and stations (small black dots) used to construct the synthetic data set for the simulations. The data set consists of 9 events and a total of 207 stations. An average of 45 arrivals are calculated for each event.

Figure 2.3 Model parametrization and input anomalies for layer 2: 483-966 km. The model mantle is parametrized as voxels, 30° × 30° at the equator and 500 km thick, for a total of 276 model parameters. Voxels in a given layer have approximately equal surface area. The introduced velocity anomalies are highlighted in gray (0.15 km/s) and black (–0.20 km/s).

Figure 2.4 Input anomalies for layer 3, 966-1449 km depth, are 0.30 km/s and 0.10 km/s.

Figure 2.5 (a-f) Ray coverage of the mantle is indicated in terms of the number of rays that sample each voxel. The most-sampled voxel has 96 hits; several voxels has more than 72 hits. All six depth layers of our model are shown: (a) 0-483 km, (b) 483-966 km, (c) 966-1449 km, (d) 1449-1932 km, (e) 1932-2415 km, (f) 2415-2898 km.

Figure 2.6 Locations of sources used in the inversions of real data. The data set consists of about 3,000 shallow events located by the ISC for the time period January 1964 - January 1987. A minimum of forty observations was required to include an event.

Figure 2.7 Locations of seismographic stations that reported the observations used in this study.

Figure 2.8 The model mantle is parametrized as voxels, 10° × 10° at the equator and generally 200 km thick, for a total of 5684 model parameters. Voxels in a

given layer have approximately equal surface area.

Figure 2.9 Histogram of approximately 409,000 travel time residuals for sources located by the International Seismological Centre (ISC) at depths between 0 and 70 km, inclusive. Also shown are the first four moments of the distribution.

Figure 2.10 Histogram of about 65,000 travel time residuals for sources located by the ISC at 0 km, which is an ISC default depth. The bimodality of the distribution arises from a programming problem in which a blank depth field in the ISC catalog is read as 0 km.

Figure 2.11 Histogram of the nearly 345,000 travel time residuals associated with the events included in this study, Source depths are greater than 0 km and less than 70 km.

Figure 2.12 (a-h) Ray coverage of the mantle provided by the approximately 345,000 observations included in our data set is shown in terms of the number of rays that sample each voxel, The most-sampled voxel has over 50,000 samples. Only 166 of 5684 voxels are unsampled. The layers shown are: (a) 0-200 km, (b) 200-400 km, (c) 400-670 km, (d) 670-870 km, (e) 1270-1470 km, (f) 1470-1670 km, (g) 2470-2670 km, and (h) 2670-2898 km.

Figure 2.13 (a-f) Six depth layers of model ISC10_direct, the direct inversion that neglects source terms: (a) 0-200 km, (b) 200-400 km, (c) 400-670 km, (d) 1270-1470 km, (e) 1470-1670 km, and (f) 2470-2670 km. Each layer's mean has been removed. Velocity perturbations grade from red (slow) to blue (fast). Maximum and minimum velocity variations are

Figure 2.14 (a-f) Six depth layers of model ISC10_sim, the simultaneous inversion for both source and velocity terms. The layers shown are the same as in figure 2.13.

Figure 2.15 (a-f) Six depth layers of model ISC10_pro, the progressive inversion for source and velocity terms in succession. The layers shown are the same as in figure 2.13.

Figure 2.16 The Spearman rank-order correlations between layers of each set of two models as functions of depth. When comparing a given layer of two models we replace each velocity value from the first model with its rank among the $N-1$ other values in the same layer, and do likewise with values from the same layer of the second model. This statistic allows us to examine correlations between small-scale features of the models.

Figure 2.17 Power contained in surface spherical harmonic series expansions of model ISC10_direct as a function of depth.

Figure 2.18 Power contained in surface spherical harmonic series expansions of model ISC10_sim as a function of depth.

Figure 2.19 Power contained in surface spherical harmonic series expansions of model ISC10_pro as a function of depth.

Figure 2.20 Power in the spherical harmonic expansions for each depth interval of model ISC10_direct as a function of angular degree. All values are normalized to the maximum value appearing in the figure. Numbers on the right refer to the maximum power for each layer.

Figure 2.21 Power in the spherical harmonic expansions for each depth interval of model ISC10_sim as a function of angular degree. Figure conventions are the same as in figure 2.20.

Figure 2.22 Power in the spherical harmonic expansions for each depth interval of model ISC10_pro as a function of angular degree. Figure conventions are the same as in figure 2.20.

Figure 2.23 Power in spherical harmonic series generated by averaging ISC10_direct layer expansions through the whole mantle and through the upper and lower mantle separately. Figure conventions are the same as in figure 2.20.

Figure 2.24 Power in spherical harmonic series generated by averaging ISC10_sim layer expansions through the whole mantle and through the upper and lower mantle separately. Figure conventions are the same as in figure 2.20.

Figure 2.25 Power in spherical harmonic series generated by averaging ISC10_pro layer expansions through the whole mantle and through the upper and lower mantle separately. Figure conventions are the same as in figure 2.20.

Figure 2.26 Correlations between spherical harmonic degrees as functions of depth between models ISC10_direct and ISC10_pro. Shown are $l = 2$, 3, and 4. These correlations allow us to compare large-scale features of the models.

Figure 2.27 Correlations between spherical harmonic degrees as functions of depth between models ISC10_direct and ISC10_pro. Shown are $l = 5$, 6, and 7.

Figure 2.28 Correlations between spherical harmonic degrees as functions of depth between models ISC10_direct and ISC10_pro. Shown are $l = 8$, 9, and 10.

Figure 2.29 Source correction vectors produced by the simultaneous inversion for 400 of the 3077 events used in this study. In each case the starting location is indicated with either an asterisk or hexagon and a scaled vector points in the direction of the correction required by the latitude and longitude adjustments. For each event the direction of the depth correction is indicated by the type of symbol marking the relocation vector's endpoint. Size of the depth correction is indicated by the size of the symbol. Events that require a shallowing correction are marked with asterisks; events that are relocated deeper are marked by a hexagon. Origin time corrections are not shown. Both figures 2.29 and 2.30

are scaled to the same maximum vector length.

Figure 2.30 Source correction vectors produced by the progressive inversion for the same 400 events shown in figure 2.29. Figure conventions are the same as for figure 2.29.

Event 1: Kamchatka Peninsula (53.0°N, 160.0°E, 73.9 km)

Table 2.1

Source parameter	"a priori" standard error	Initial parameter offset	After correction: Simultaneous parameter offset	Simultaneous % improved	Progressive parameter offset	Progressive % improved
origin time (sec)	0.00	0.00	-0.08		0.04	
latitude (deg)	0.163	-0.008	0.003	63%	0.002	77%
longitude (deg)	0.420	-0.080	-0.008	90%	-0.034	58%
depth (km)	0.329	2.70	-0.75	72%	0.07	97%

Event 2: South American Subduction Zone (31.1°S, 67.9°W, 72.2 km)

Table 2.2

Source parameter	"a priori" standard error	Initial parameter offset	After correction: Simultaneous parameter offset	Simultaneous % improved	Progressive parameter offset	Progressive % improved
origin time (sec)	0.14	0.00	-0.09		0.00	
latitude (deg)	0.022	-0.006	0.004	33%	0.002	75%
longitude (deg)	0.029	0.006	0.004	33%	0.001	90%
depth (km)	0.18	0.10	1.00	-900%	0.16	-67%

Event 3: Mariana Subduction Zone (18.9°N, 144.8°E, 41.0 km)

Table 2.3

Source parameter	"a priori" standard error	Initial parameter offset	After correction: Simultaneous parameter offset	Simultaneous % improved	Progressive parameter offset	Progressive % improved
origin time (sec)	0.26	0.18	-0.03	83%	0.01	92%
latitude (deg)	0.019	0.031	-0.024	23%	0.003	92%
longitude (deg)	0.015	-0.012	-0.001	92%	0.001	90%
depth (km)	2.82	0.30	1.75	-483%	0.55	-85%

Event 4: Mongolia
(50.0°N, 110.0°E, 180.0 km)

Table 2.4

Source parameter	"a priori" standard error	Initial parameter offset	After correction: Simultaneous parameter offset	Simultaneous % improved	Progressive parameter offset	Progressive % improved
origin time (sec)	0.00	-0.05	-0.02	60%	0.00	100%
latitude (deg)	0.031	0.005	0.003	40%	0.001	80%
longitude (deg)	0.047	0.041	0.015	63%	0.016	61%
depth (km)	0.31	8.70	4.58	47%	0.01	100%

Event 5: Aleutian Islands
(51.0°N, 178.0°W, 50.0 km)

Table 2.5

Source parameter	"a priori" standard error	Initial parameter offset	After correction: Simultaneous parameter offset	Simultaneous % improved	Progressive parameter offset	Progressive % improved
origin time (sec)	0.69	2.09	2.04	2%	0.08	96%
latitude (deg)	0.024	0.024	-0.005	79%	-0.006	74%
longitude (deg)	0.036	-0.159	-0.022	86%	-0.066	58%
depth (km)	6.57	24.30	23.81	2%	0.55	98%

Event 6: Central American Subduction Zone
(9.5°N, 84.1°W, 66.6 km)

Table 2.6

Source parameter	"a priori" standard error	Initial parameter offset	After correction: Simultaneous parameter offset	Simultaneous % improved	Progressive parameter offset	Progressive % improved
origin time (sec)	0.23	0.00	0.15		0.09	
latitude (deg)	0.013	-0.023	-0.017	26%	-0.009	63%
longitude (deg)	0.014	-0.012	0.024	-100%	0.003	76%
depth (km)	2.52	-3.10	-0.71	77%	-1.46	53%

Event 7: South China Sea (18.8°N, 111.9°, 53.0 km)

Table 2.7

Source parameter	"a priori" standard error	Initial parameter offset	After correction: Simultaneous parameter offset	Simultaneous % improved	Progressive parameter offset	Progressive % improved
origin time (sec)	0.12	-0.01	-0.13	-1200%	-0.10	-900%
latitude (deg)	0.208	0.029	0.01	66%	0.003	90%
longitude (deg)	0.228	0.010	0.003	70%	0.004	63%
depth (km)	15.84	2.50	0.46	82%	1.30	48%

Event 8: Himalayas (30.5°N, 79.4°E, 88.0 km)

Table 2.8

Source parameter	"a priori" standard error	Initial parameter offset	After correction: Simultaneous parameter offset	Simultaneous % improved	Progressive parameter offset	Progressive % improved
origin time (sec)	0.28	-0.12	-0.05	42%	-0.06	50%
latitude (deg)	0.022	0.022	0.004	82%	0.000	100%
longitude (deg)	0.027	0.038	0.084	-120%	0.011	71%
depth (km)	2.84	0.20	1.45	-625%	-0.13	36%

Event 9: South Pacific (20.8 126.9°W, 87.8 km)

Table 2.9

Source parameter	"a priori" standard error	Initial parameter offset	After correction: Simultaneous parameter offset	Simultaneous % improved	Progressive parameter offset	Progressive % improved
origin time (sec)	0.45	0.00	-0.08		-0.01	
latitude (deg)	0.022	0.001	0.061	-6000%	-0.002	-50%
longitude (deg)	0.020	-0.007	0.005	29%	0.001	80%
depth (km)	4.33	0.60	0.12	80%	0.23	61%

Velocity Anomalies

Table 2.10

			Inversions						
			No source offset	With source offset					
				Direct		Simultaneous		Progressive	
anomaly	# hits	peak	km/sec	km/sec	%	km/sec	%	km/sec	
1	18	0.100	0.063	0.038	60%	0.066	105%	0.064	101%
2	56	0.300	0.187	0.168	90%	0.170	91%	0.184	98%
3	72	0.150	0.079	0.036	46%	0.087	110%	0.079	100%
4	38	-0.200	-0.088	-0.058	66%	-0.055	63%	-0.080	91%
next		0.000	0.010	-0.042		0.051		0.017	

Table 2.11: Model layer depths and average voxel sampling

Layer number	Depth range (km)	Average velocity (km/s)	Average voxel volume (10^8 km^3)	Number of voxels sampled	% of voxels sampled	Average number hits	Average column sum of A (10^4 km)
1	0-200	7.94	2.43	333	82	2375	51.25
2	200-400	8.58	2.28	352	87	2122	47.66
3	400-670	9.88	2.85	383	94	2115	66.96
4	670-870	10.95	1.94	394	97	1900	53.72
5	870-1070	11.40	1.81	402	99	1712	51.23
6	1070-1270	11.68	1.67	406	100	1460	42.23
7	1270-1470	11.96	1.55	406	100	1293	37.54
8	1470-1670	12.24	1.43	406	100	1134	32.74
9	1670-1870	12.51	1.31	406	100	1008	29.31
10	1870-2070	12.76	1.20	406	100	908	27.19
11	2070-2270	13.01	1.09	406	100	768	24.22
12	2270-2470	13.24	0.99	406	100	587	18.75
13	2470-2670	13.47	0.89	406	100	430	14.64
14	2670-2898	13.64	0.91	406	100	164	5.78

Figure 2.1

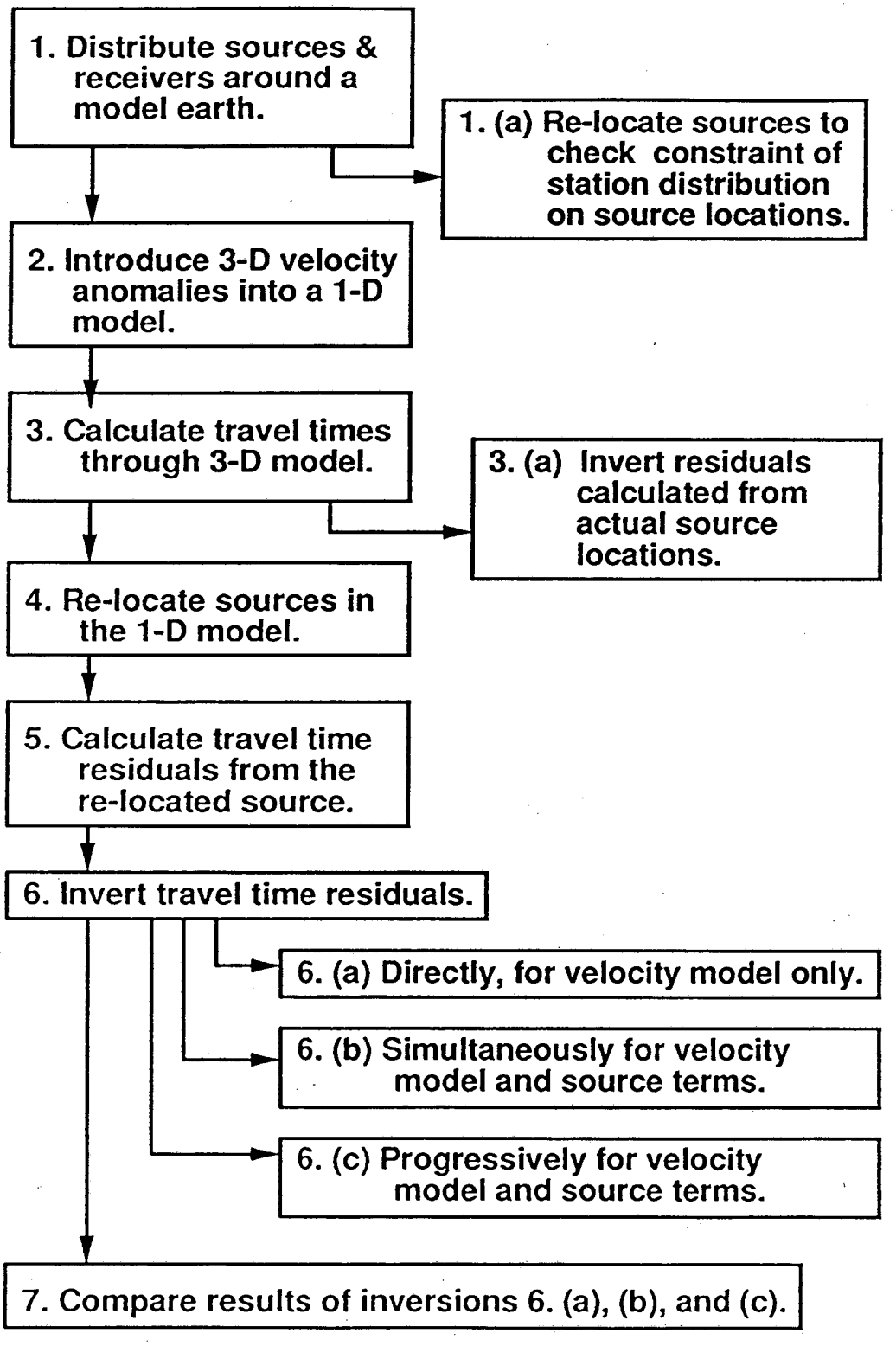

1. Distribute sources & receivers around a model earth.
1. (a) Re-locate sources to check constraint of station distribution on source locations.
2. Introduce 3-D velocity anomalies into a 1-D model.
3. Calculate travel times through 3-D model.
3. (a) Invert residuals calculated from actual source locations.
4. Re-locate sources in the 1-D model.
5. Calculate travel time residuals from the re-located source.
6. Invert travel time residuals.
6. (a) Directly, for velocity model only.
6. (b) Simultaneously for velocity model and source terms.
6. (c) Progressively for velocity model and source terms.
7. Compare results of inversions 6. (a), (b), and (c).

Figure 2.2

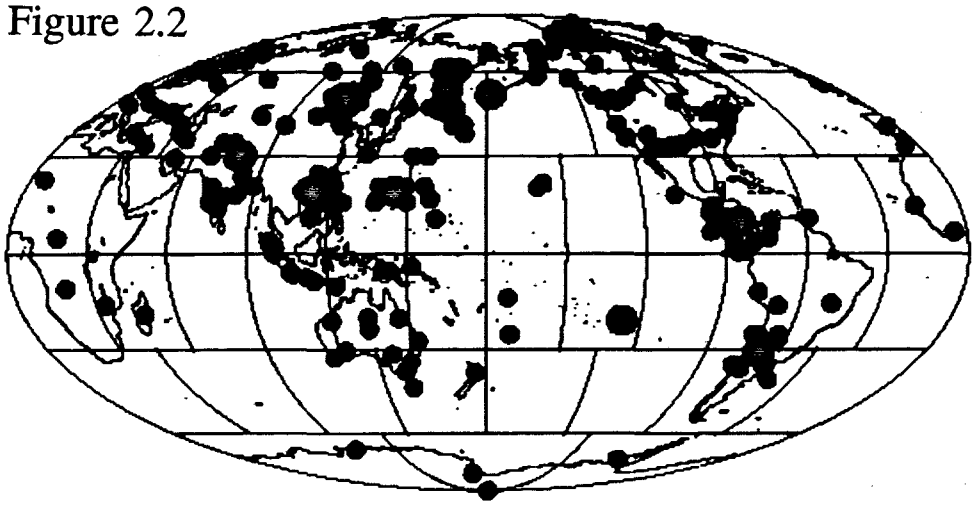


Figure 2.3

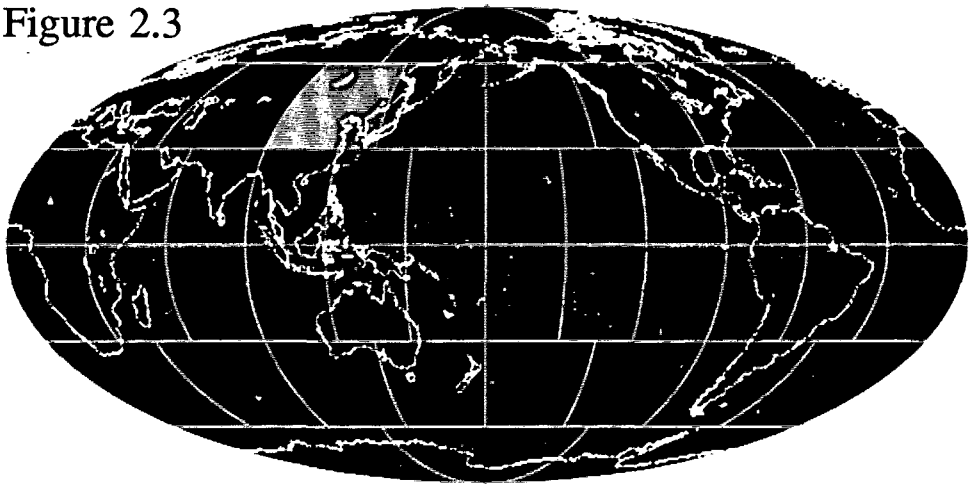
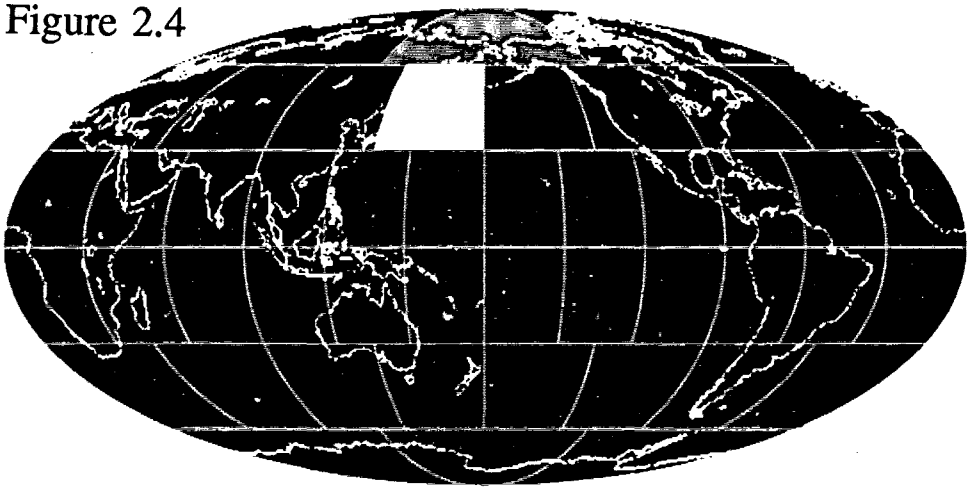


Figure 2.4



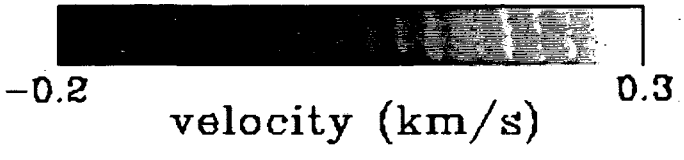


Figure 2.5

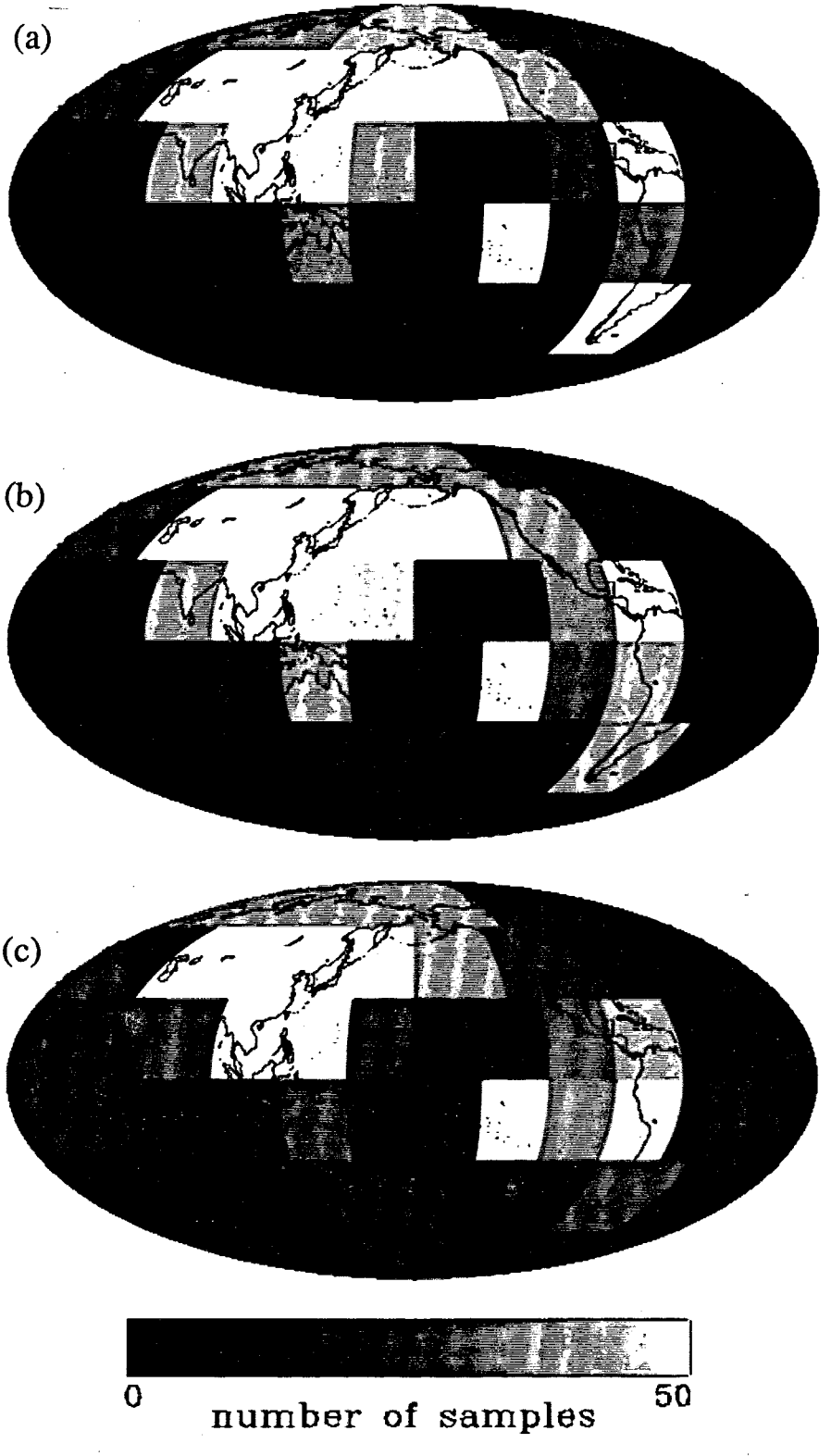
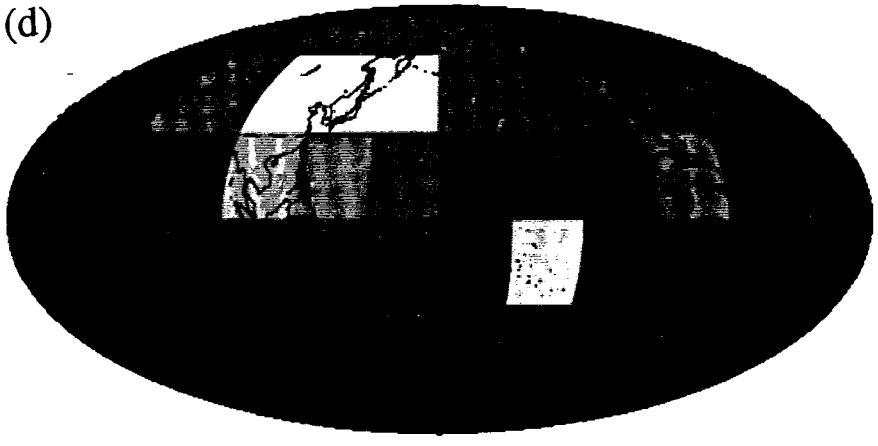
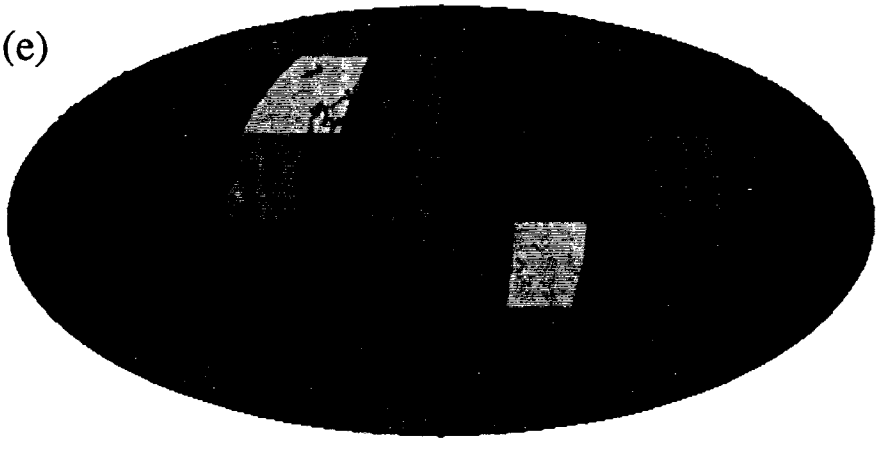


Figure 2.5

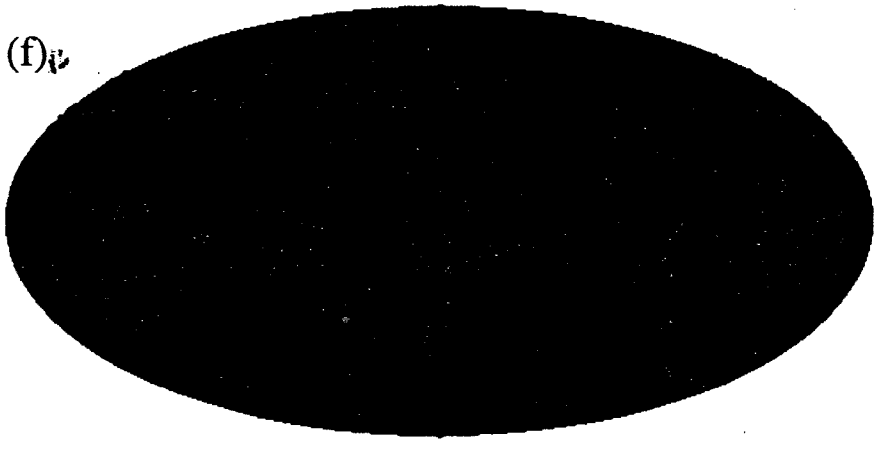
(d)



(e)



(f)



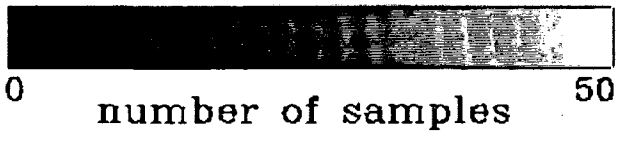


Figure 2.6

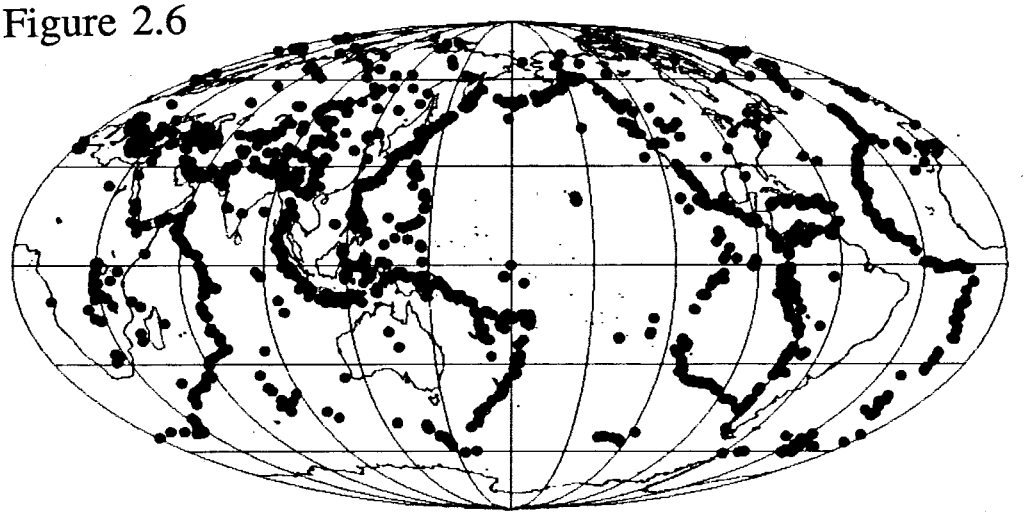


Figure 2.7

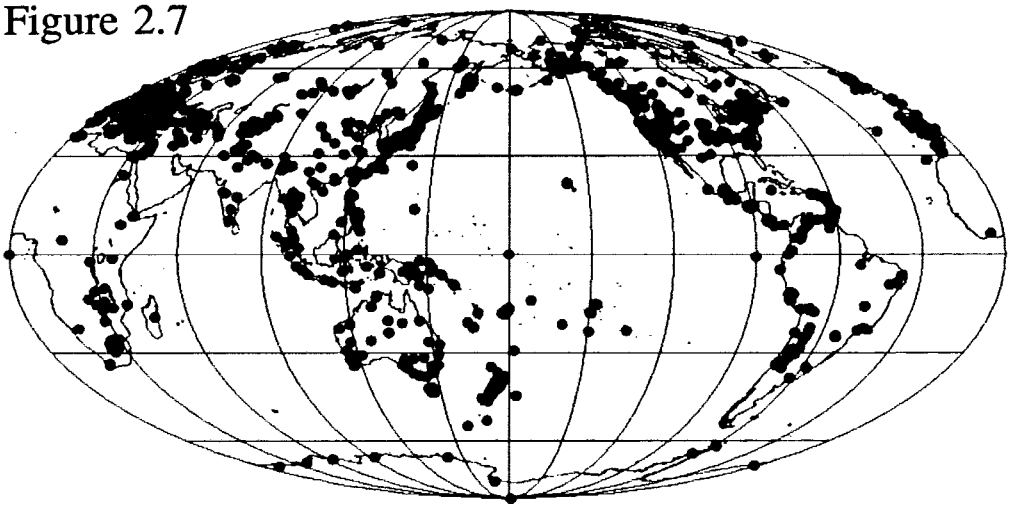


Figure 2.8

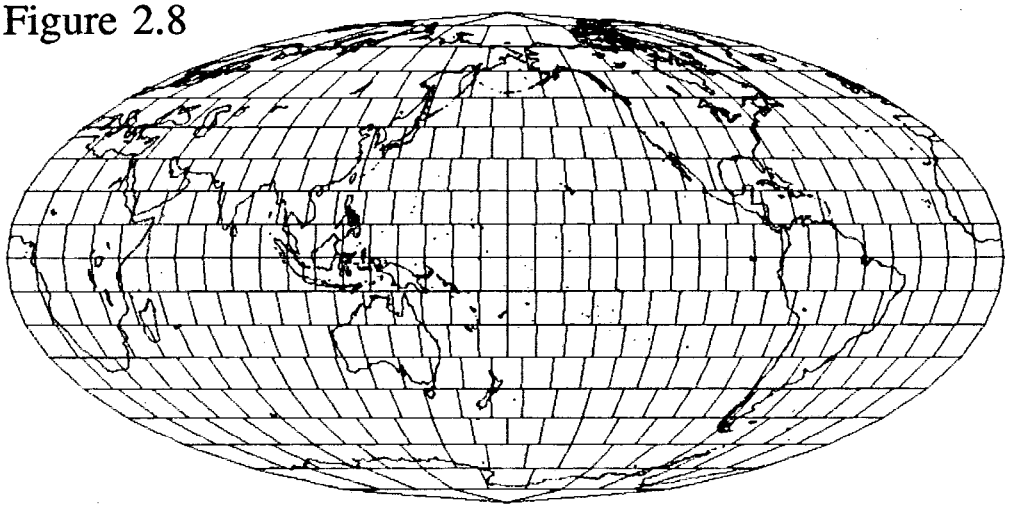


Figure 2.9

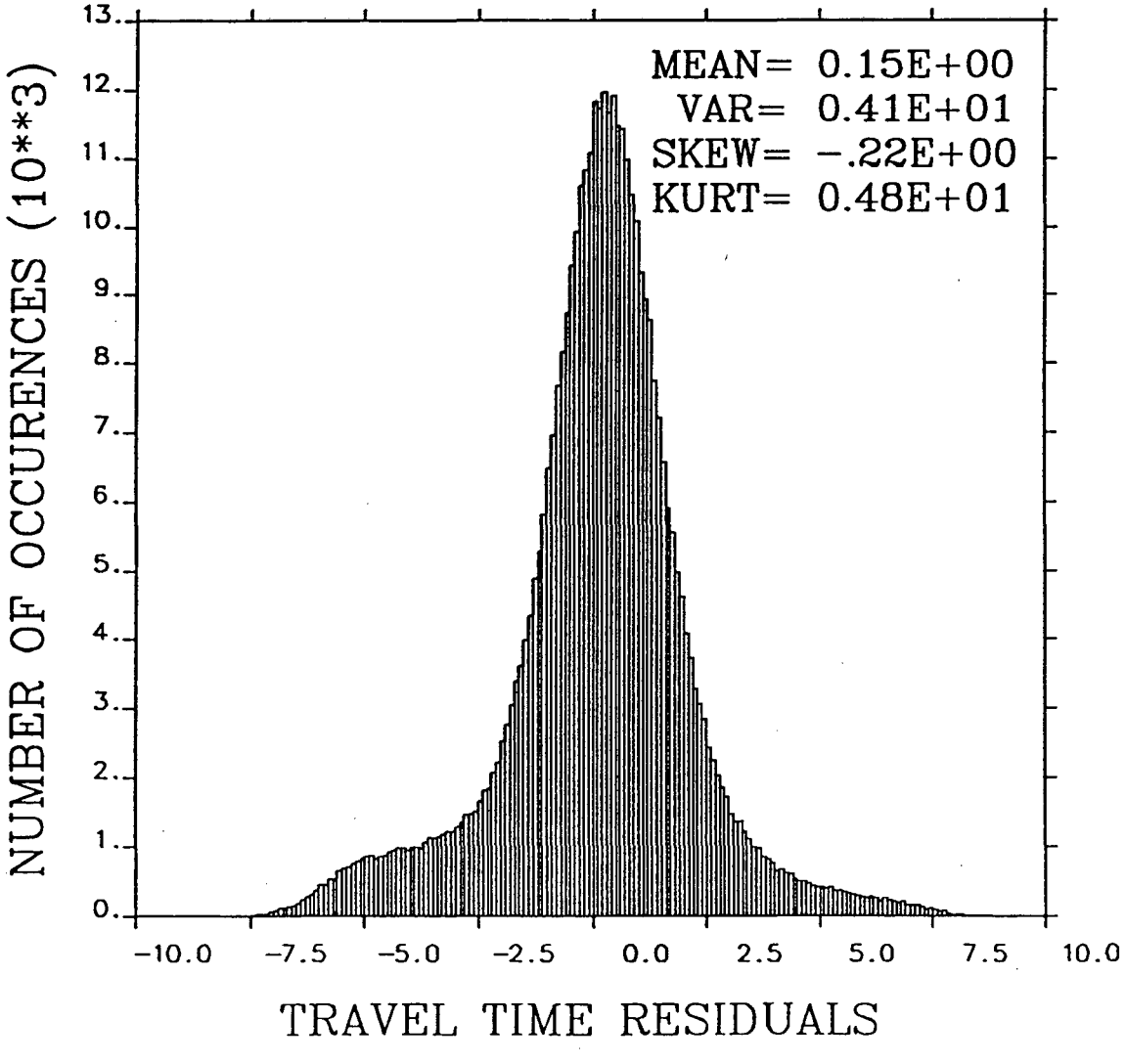

MEAN= 0.15E+00
VAR= 0.41E+01
SKEW= -.22E+00
KURT= 0.48E+01
NUMBER OF OCCURENCES (10**3)
13.
12.
11.
10.
9.
8.
7.
6.
5.
4.
3.
2.
1.
0.
-10.0
-7.5
-5.0
-2.5
0.0
2.5
5.0
7.5
10.0
TRAVEL TIME RESIDUALS

Figure 2.10

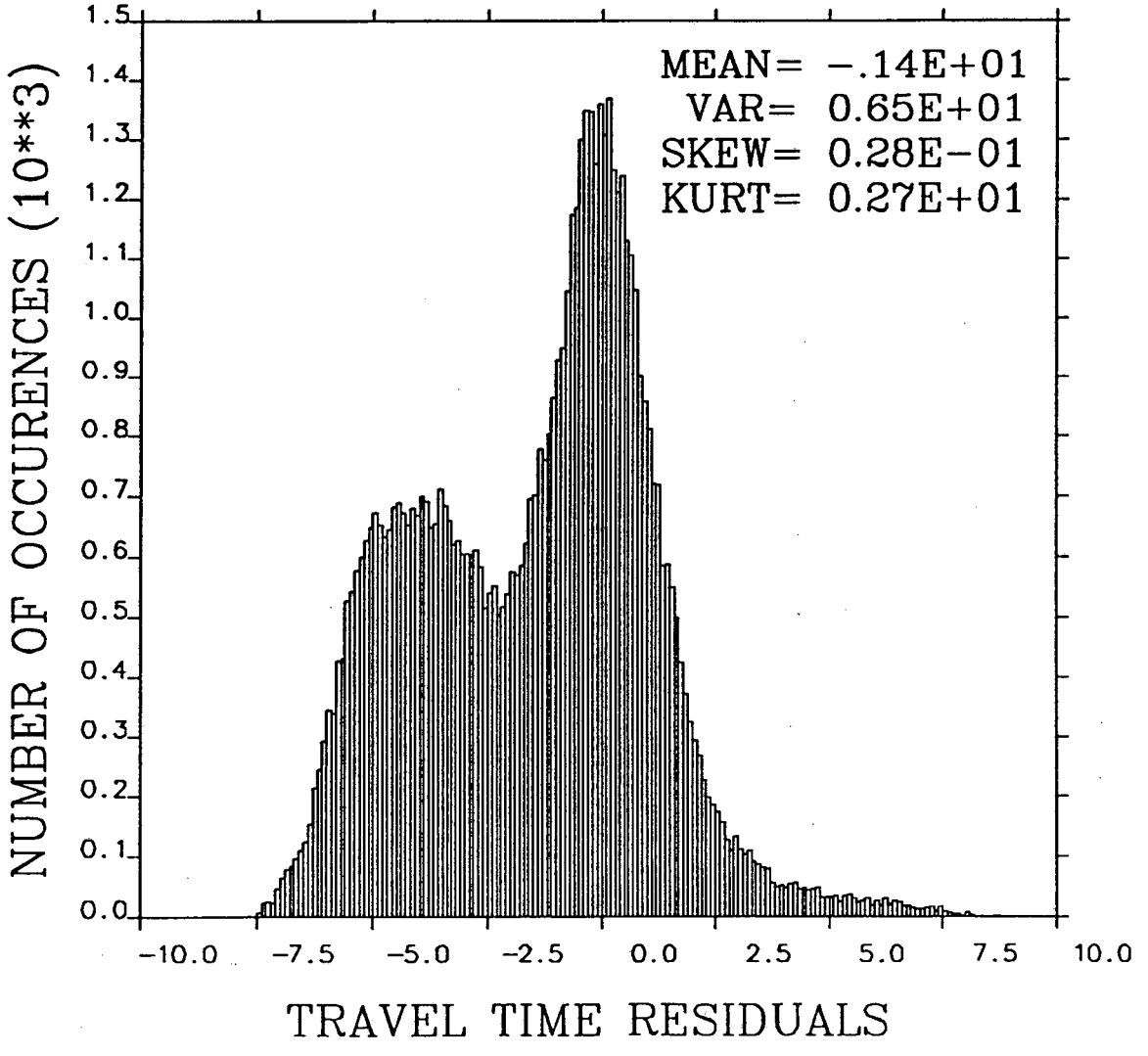

MEAN= -.14E+01
VAR= 0.65E+01
SKEW= 0.28E-01
KURT= 0.27E+01
NUMBER OF OCCURENCES (10**3)
1.5
1.4
1.3
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-10.0
-7.5
-5.0
-2.5
0.0
2.5
5.0
7.5
10.0
TRAVEL TIME RESIDUALS

Figure 2.11

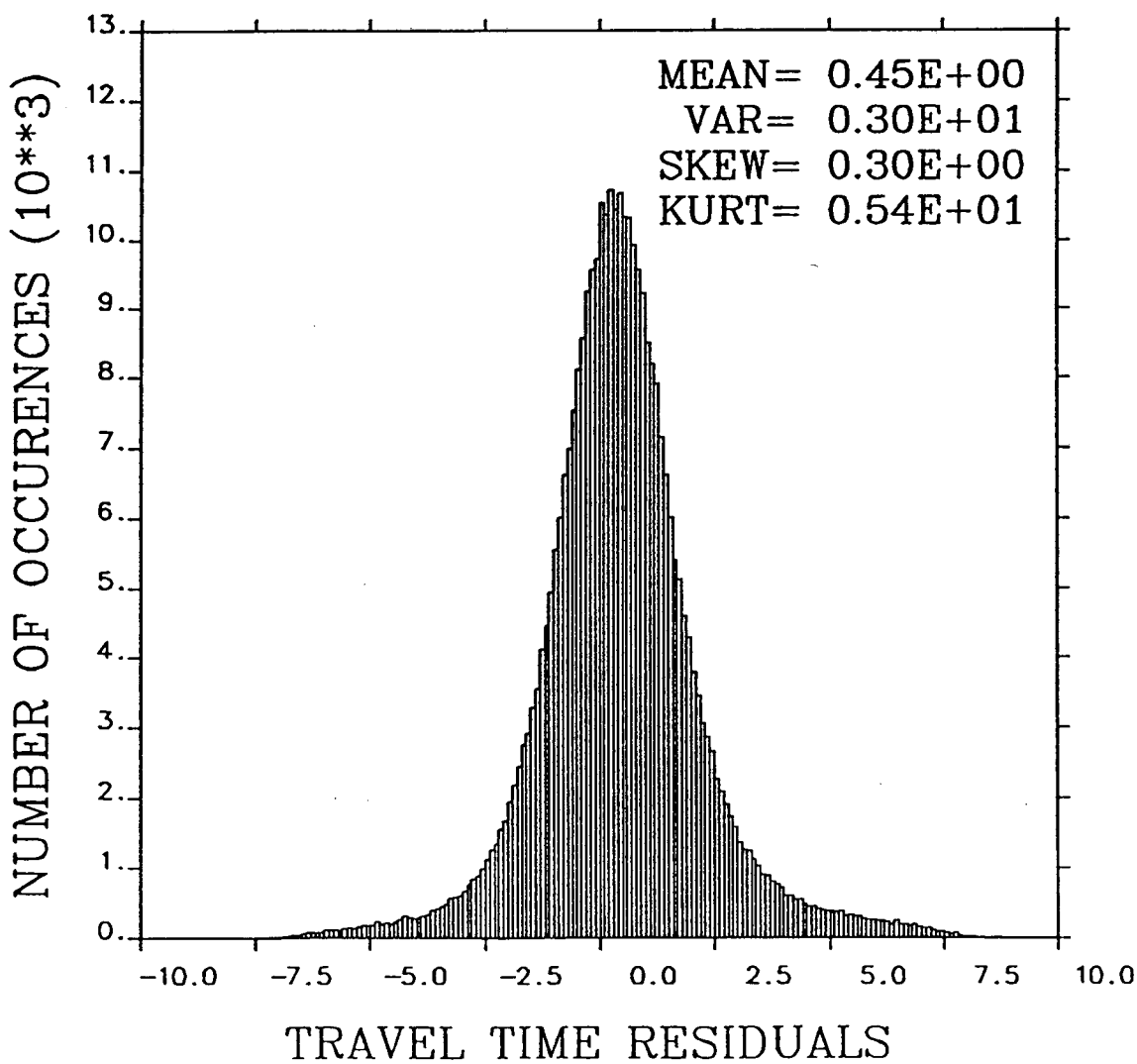

MEAN= 0.45E+00
VAR= 0.30E+01
SKEW= 0.30E+00
KURT= 0.54E+01
NUMBER OF OCCURENCES (10**3)
13.
12.
11.
10.
9.
8.
7.
6.
5.
4.
3.
2.
1.
0.
-10.0
-7.5
-5.0
-2.5
0.0
2.5
5.0
7.5
10.0
TRAVEL TIME RESIDUALS

Figure 2.12

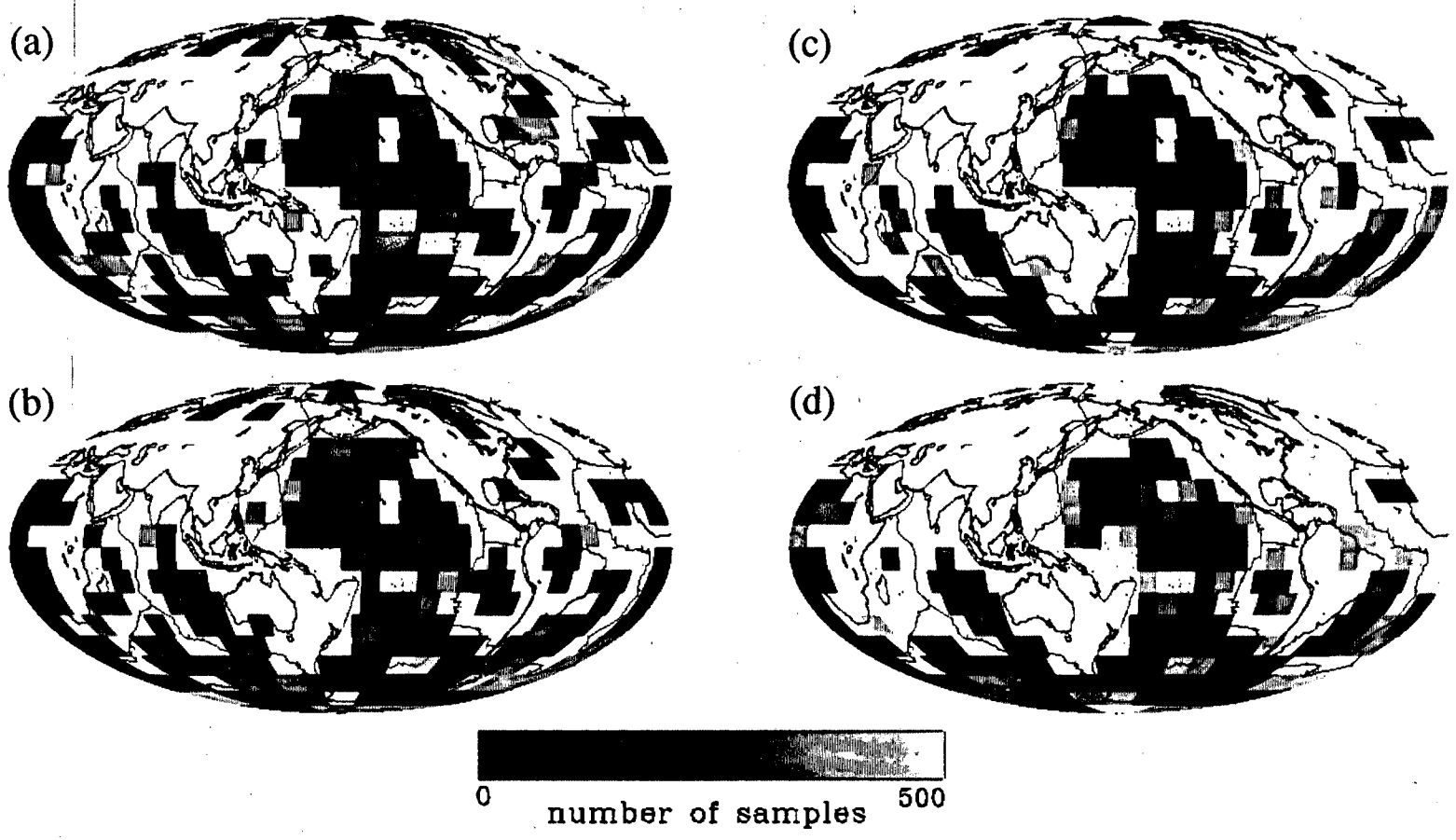


Figure 2.12

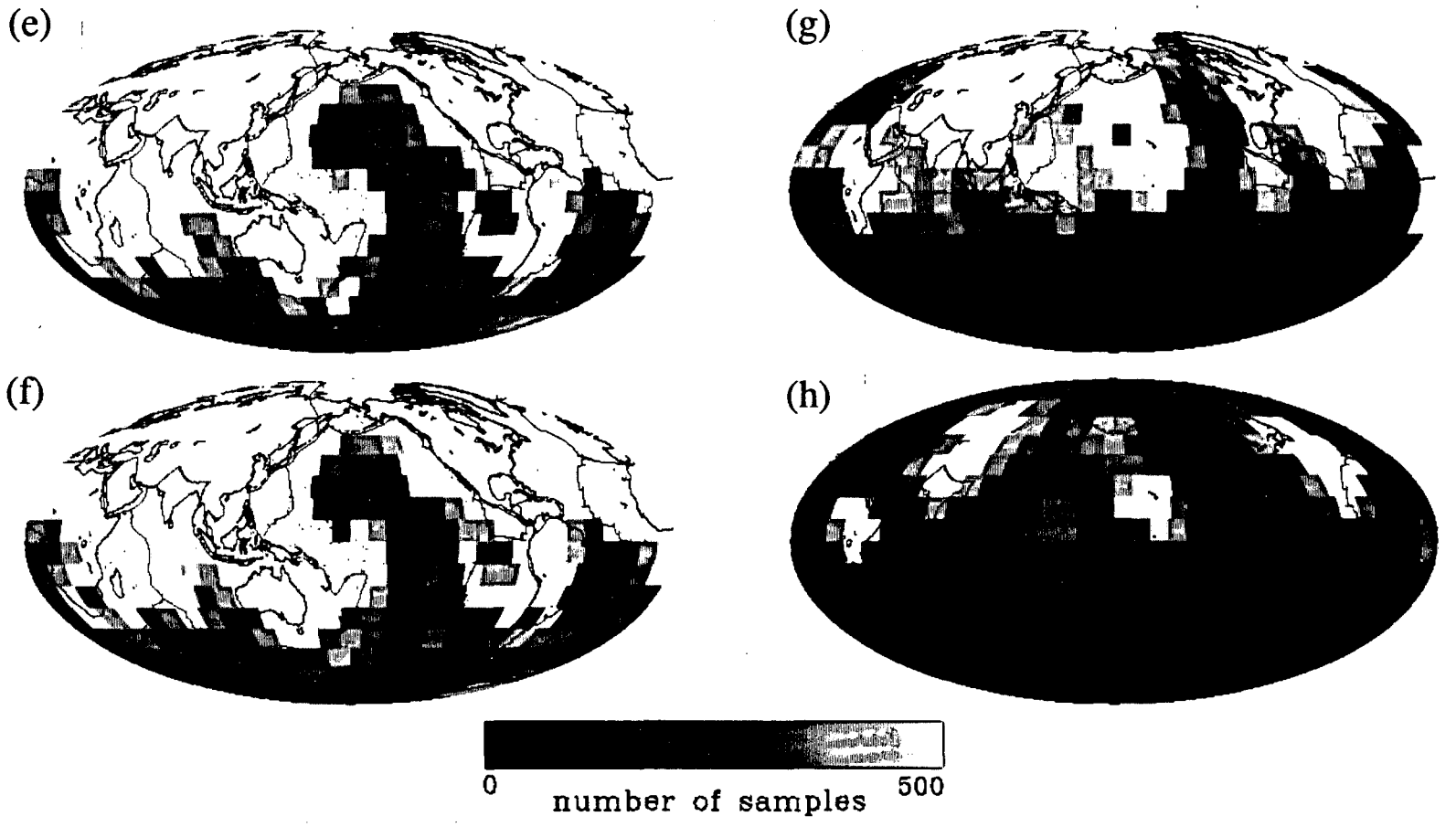


Figure 2.13

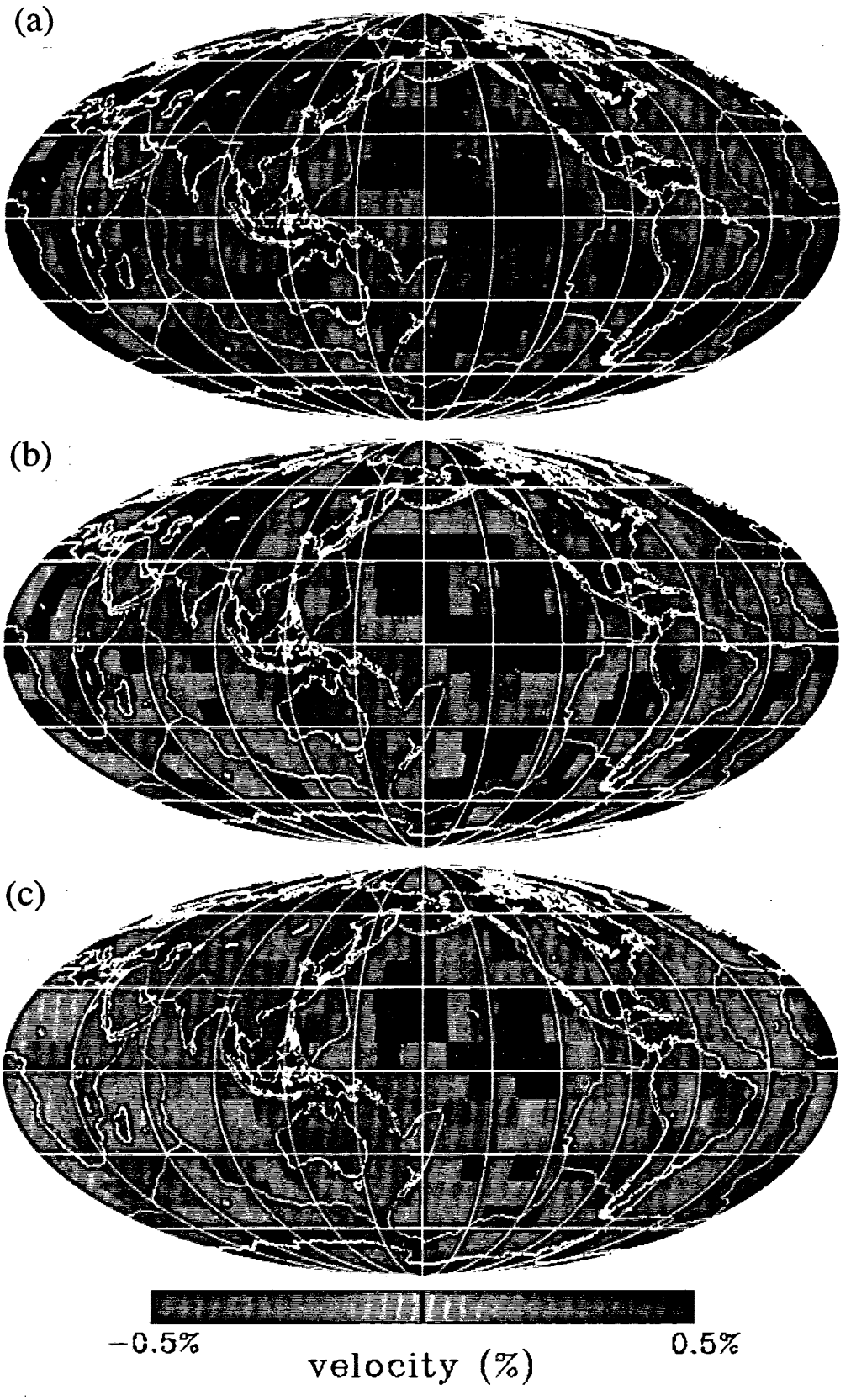
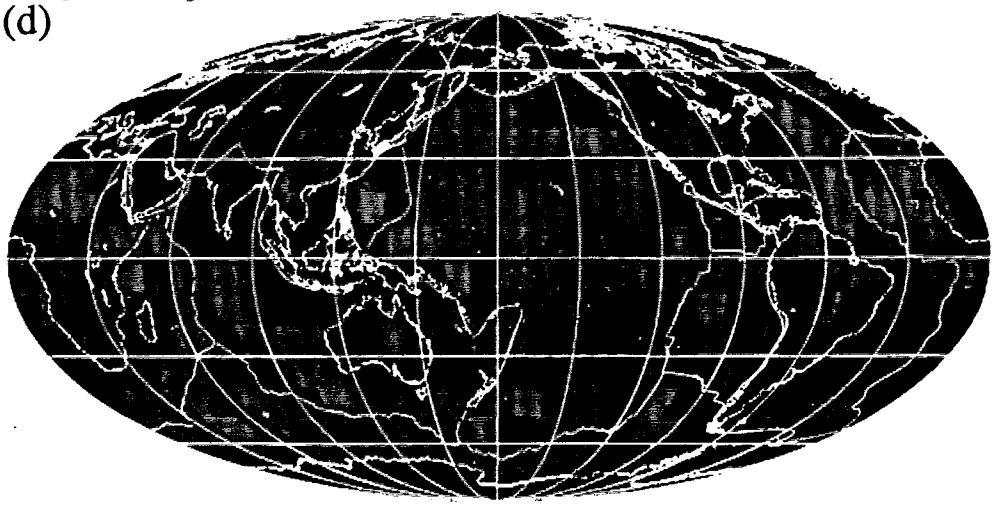
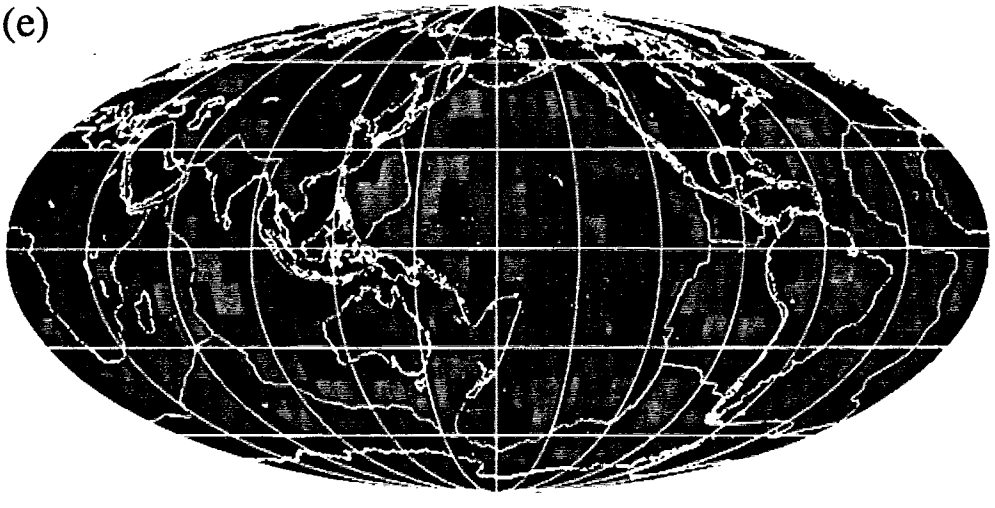


Figure 2.13

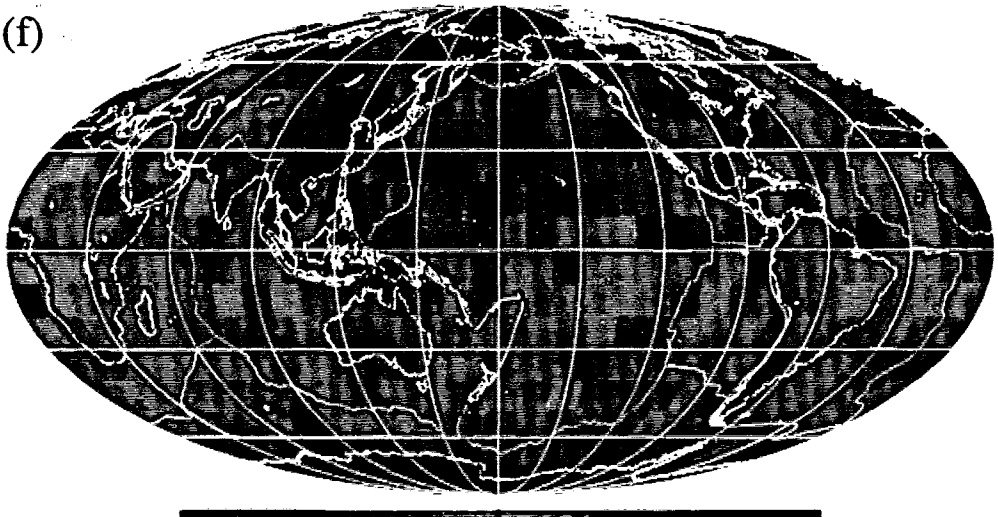
(d)



(e)



(f)



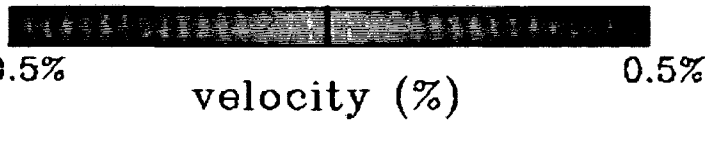


Figure 2.14

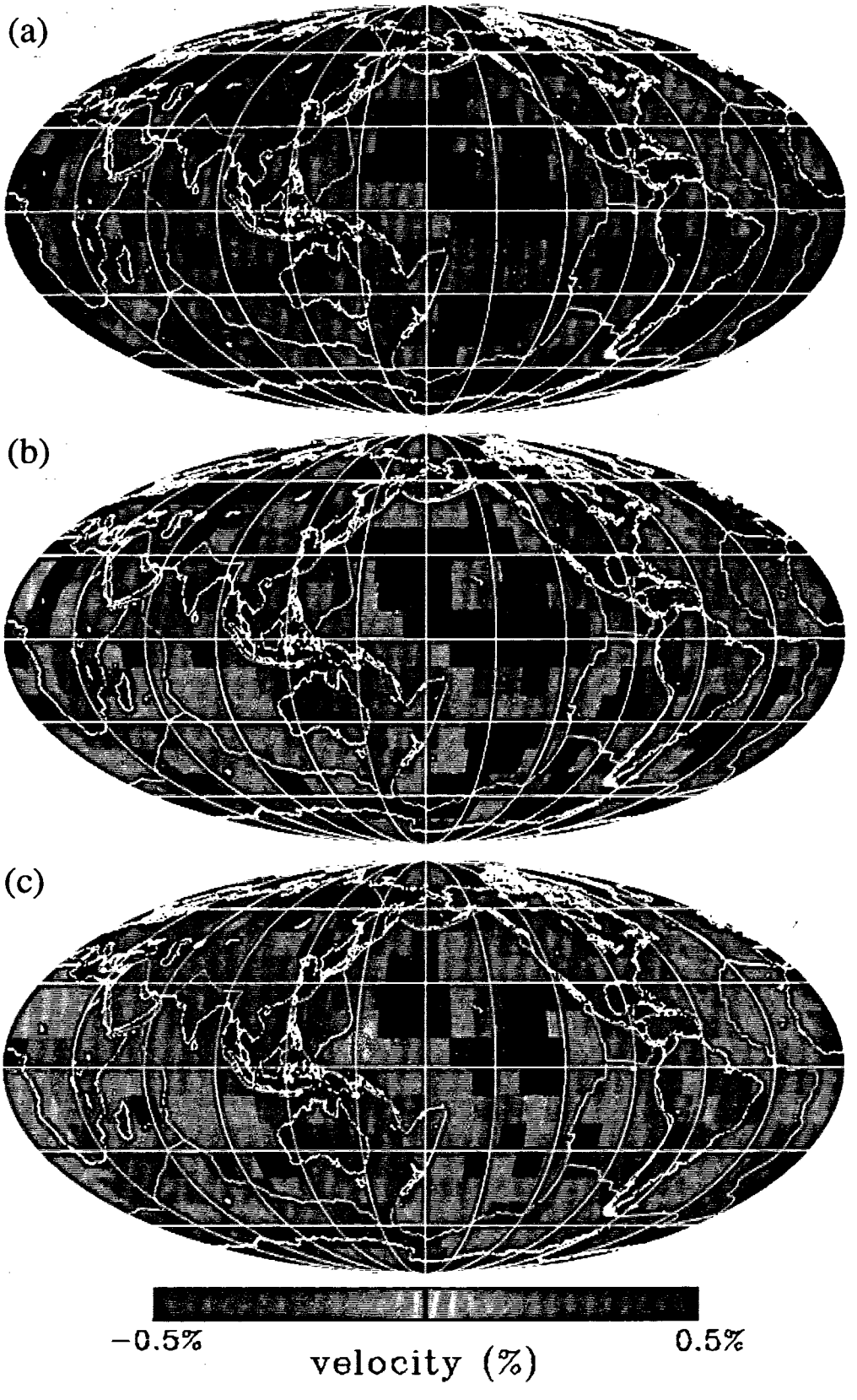


Figure 2.14

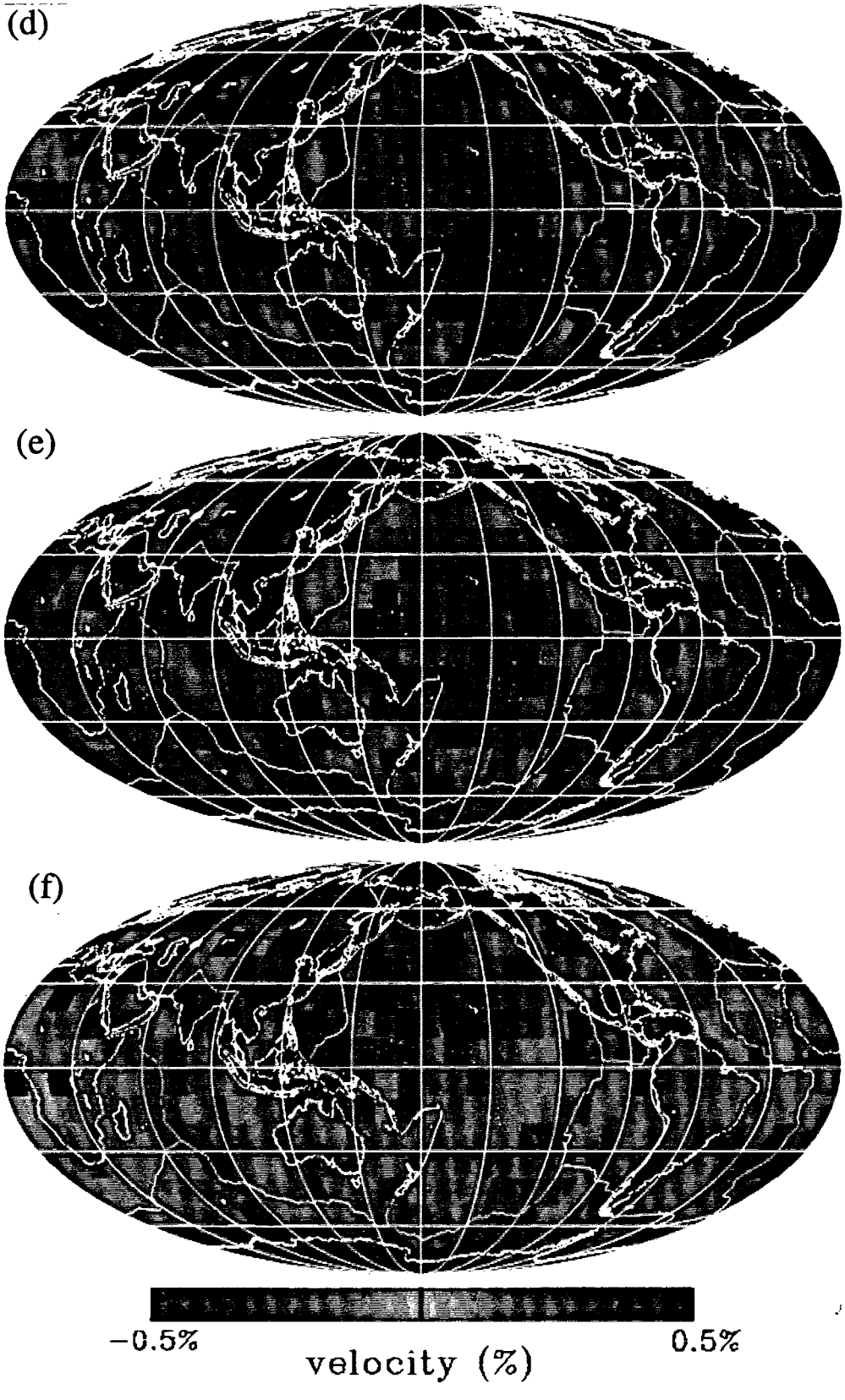


Figure 2.15

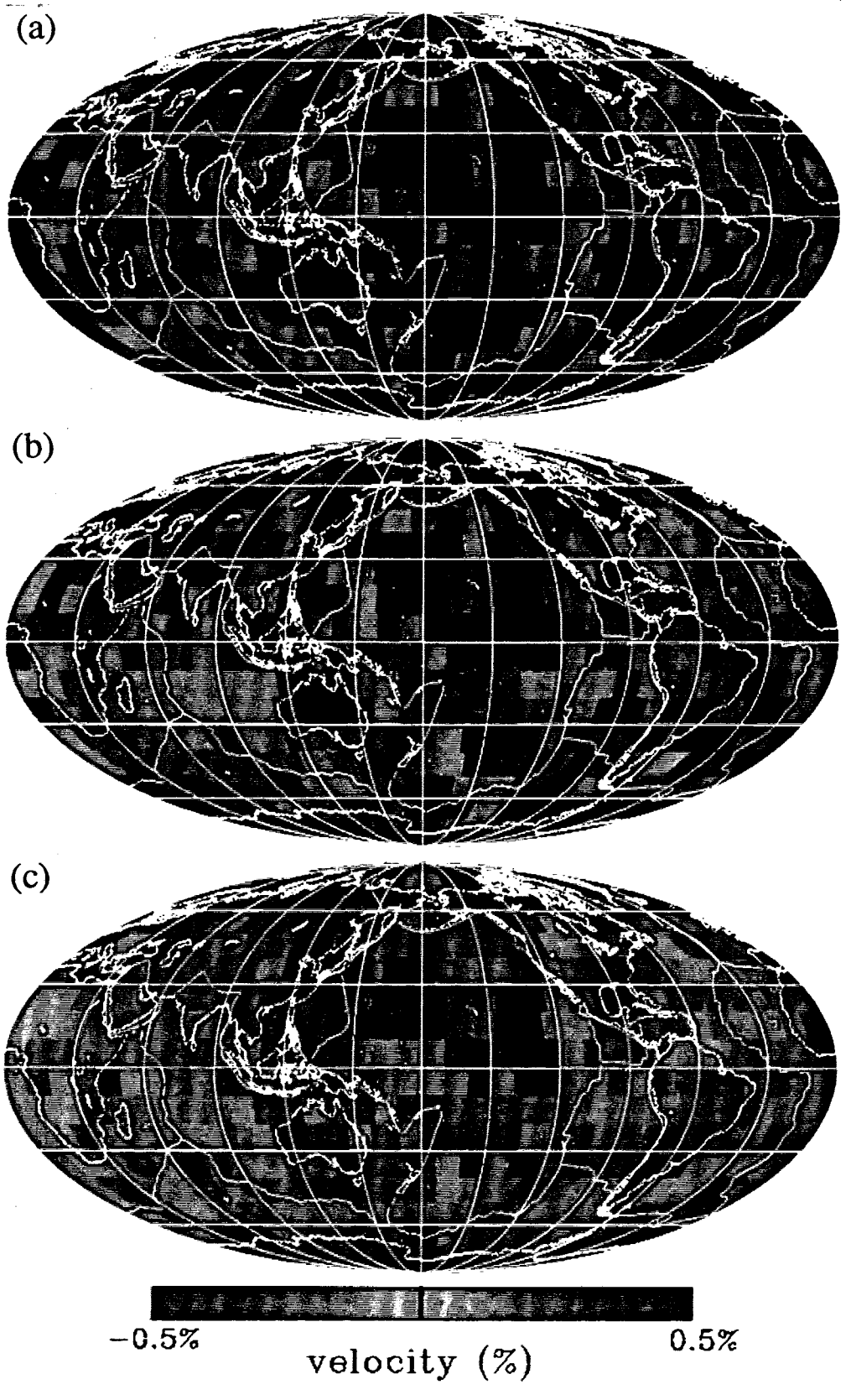


Figure 2.15

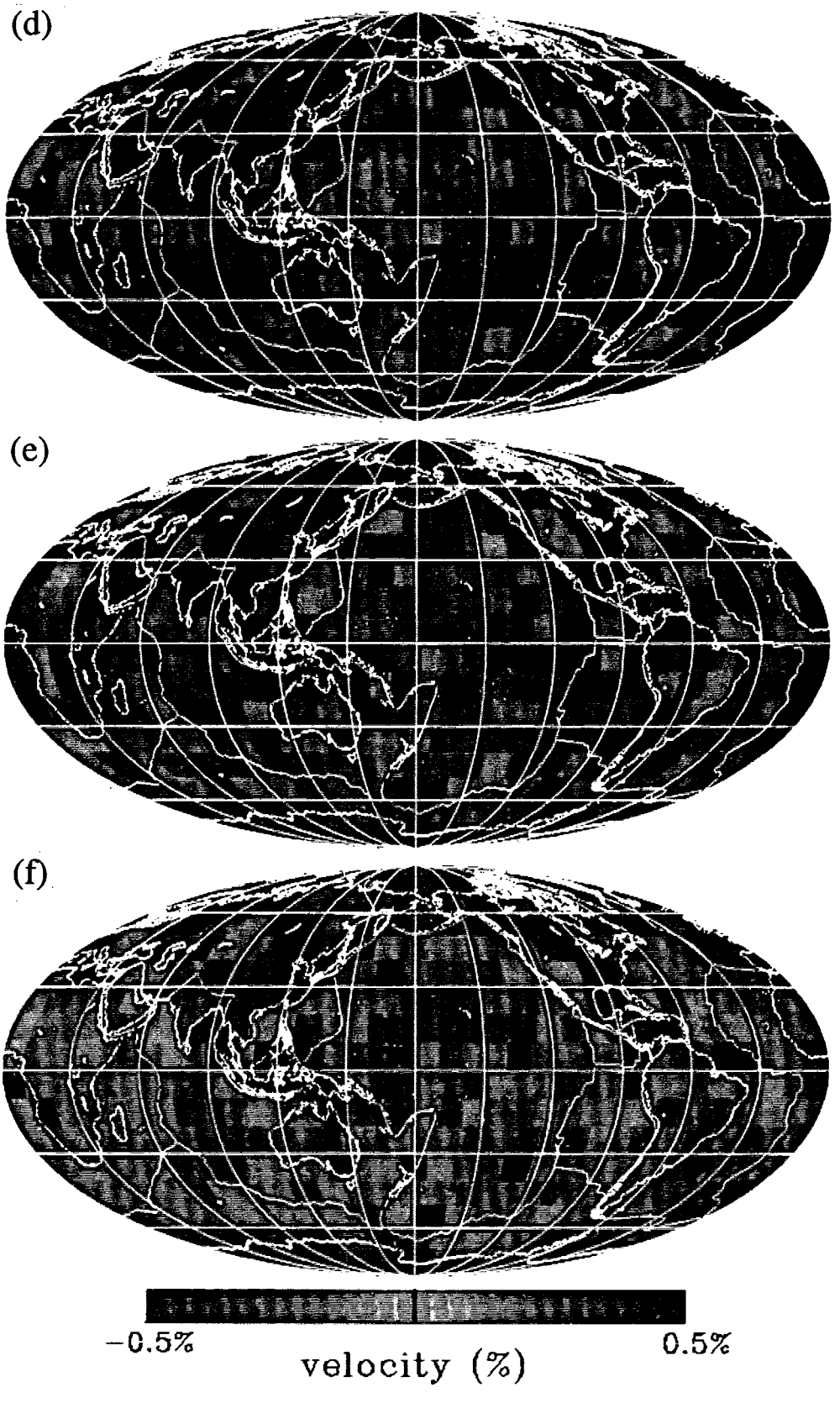


Figure 2.16

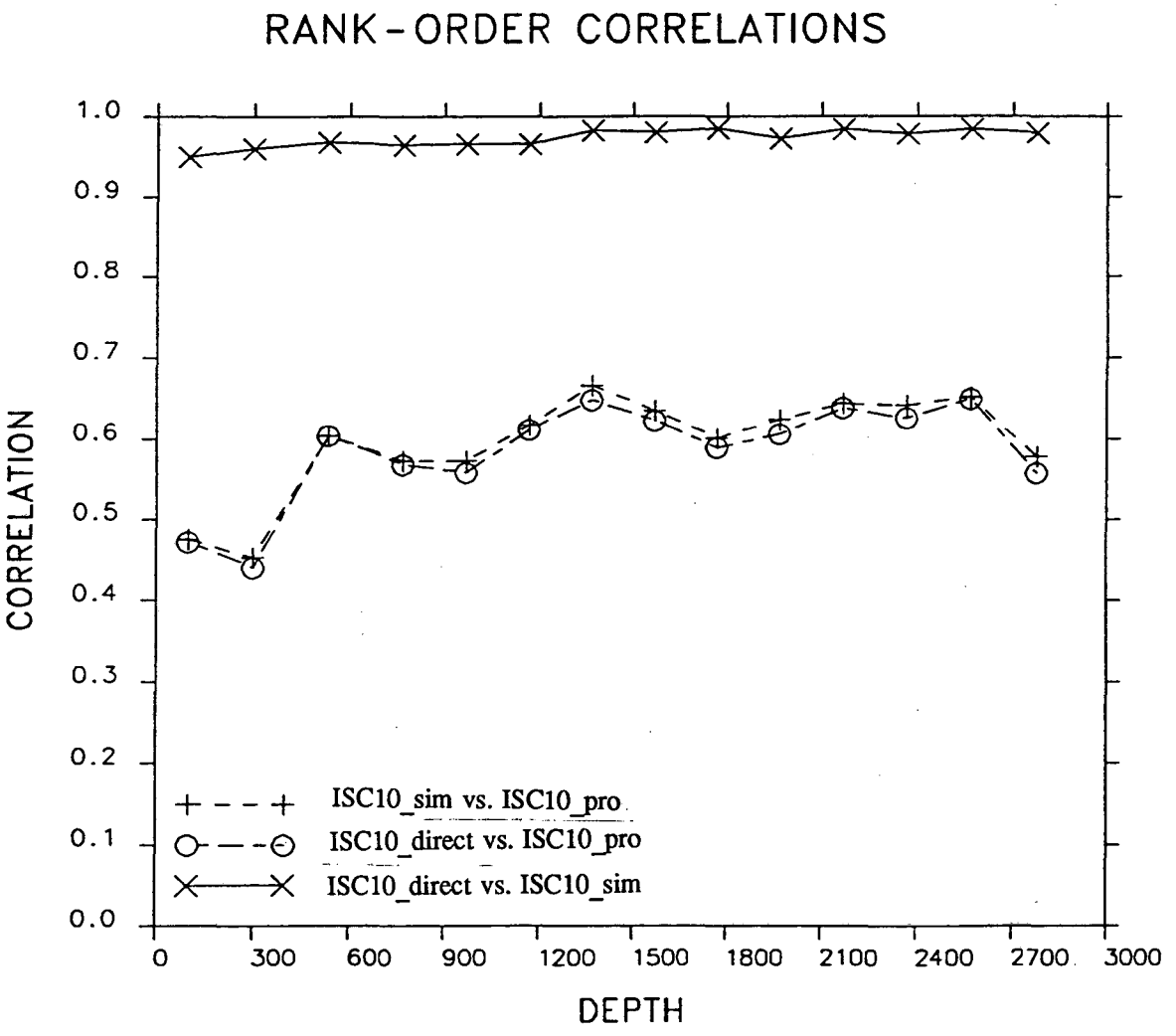

RANK-ORDER CORRELATIONS
CORRELATION
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
300
600
900
1200
1500
1800
2100
2400
2700
3000
DEPTH
ISC10_sim vs. ISC10_pro
ISC10_direct vs. ISC10_pro
ISC10_direct vs. ISC10_sim

Figure 2.17

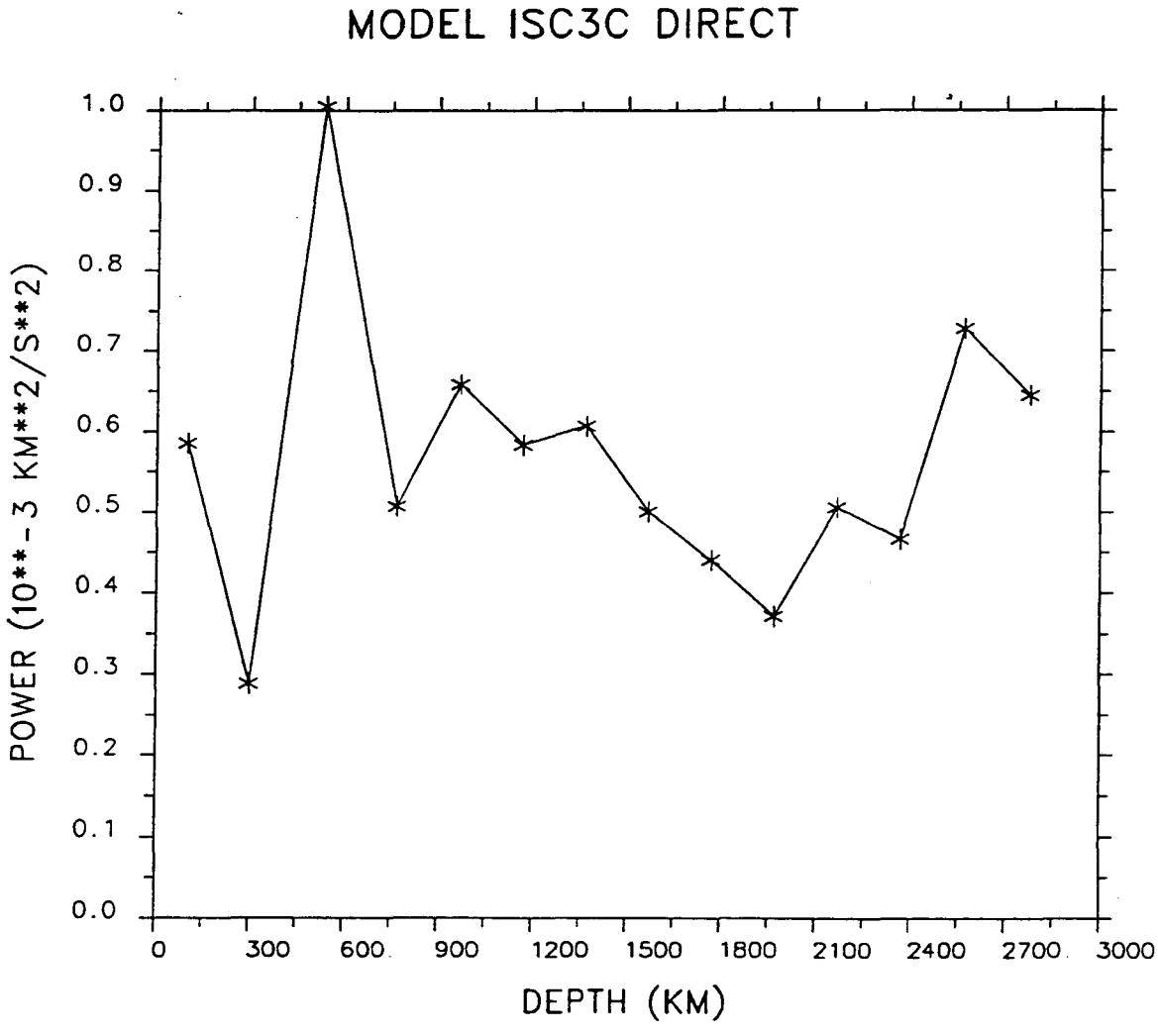

MODEL ISC3C DIRECT
POWER (10**-3 KM**2/S**2)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
300
600
900
1200
1500
1800
2100
2400
2700
3000
DEPTH (KM)

Figure 2.18

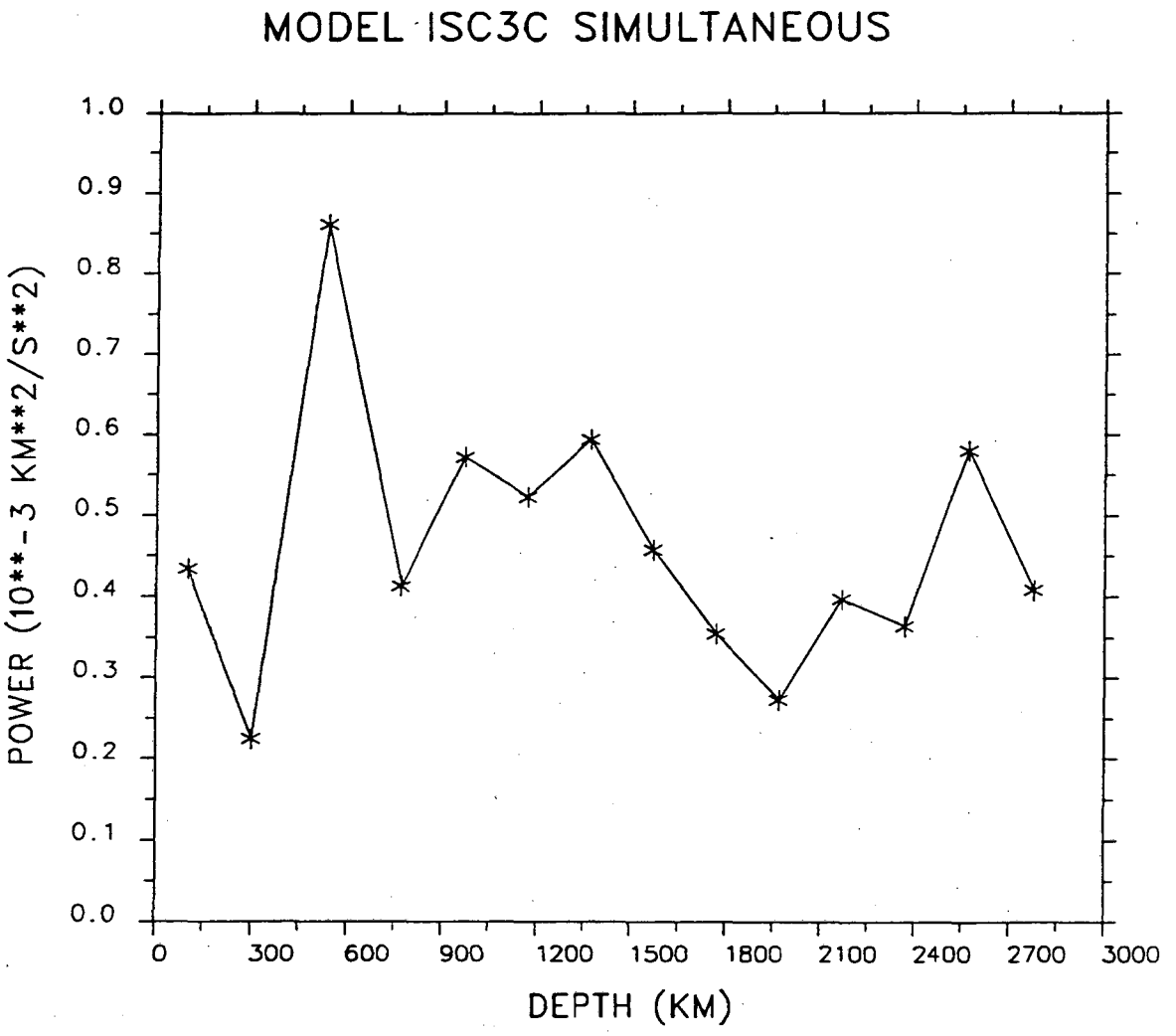
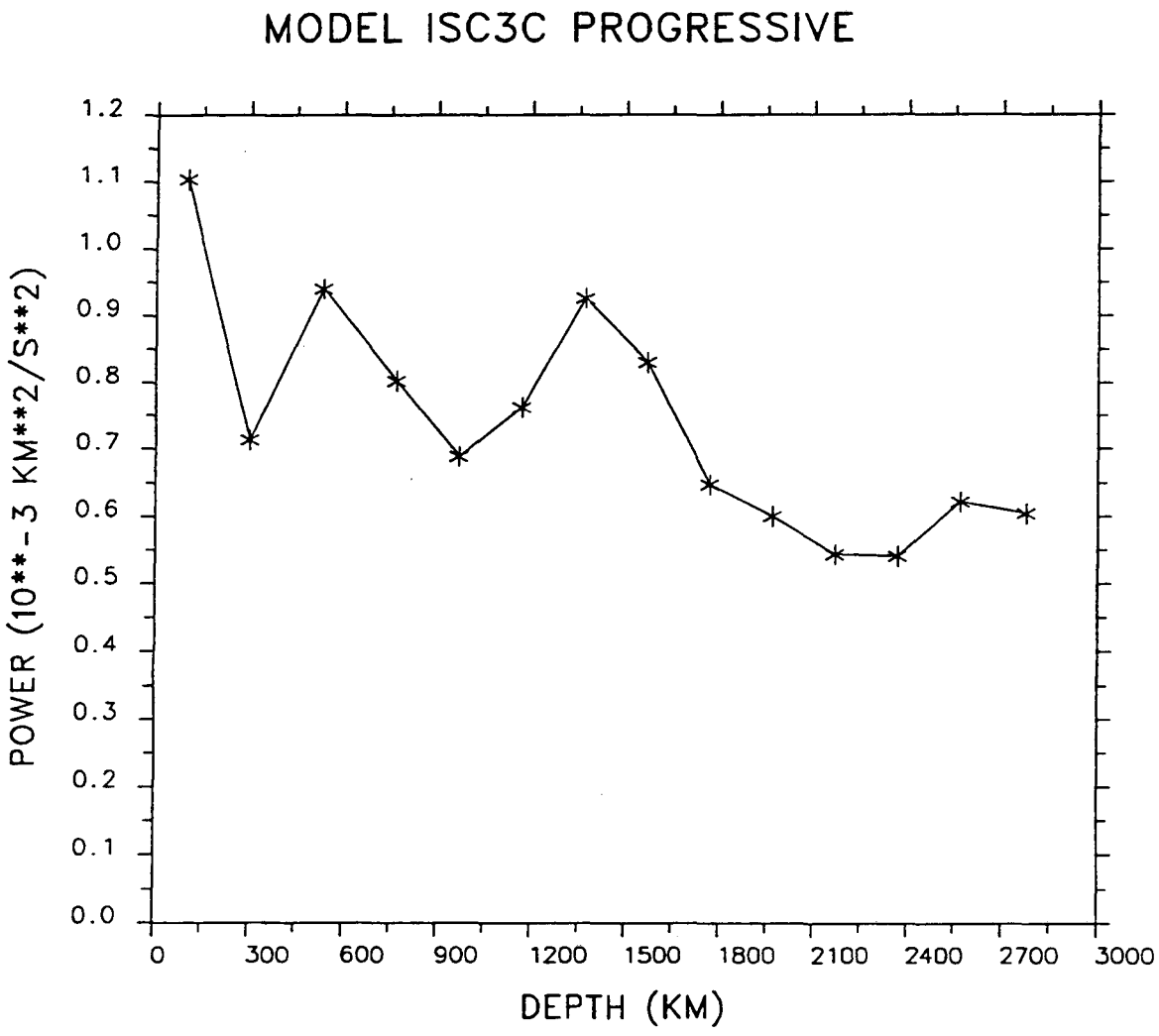

MODEL ISC3C SIMULTANEOUS
POWER (10**-3 KM**2/S**2)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
300
600
900
1200
1500
1800
2100
2400
2700
3000
DEPTH (KM)

Figure 2.19


MODEL ISC3C PROGRESSIVE
POWER (10**-3 KM**2/S**2)
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
300
600
900
1200
1500
1800
2100
2400
2700
3000
DEPTH (KM)

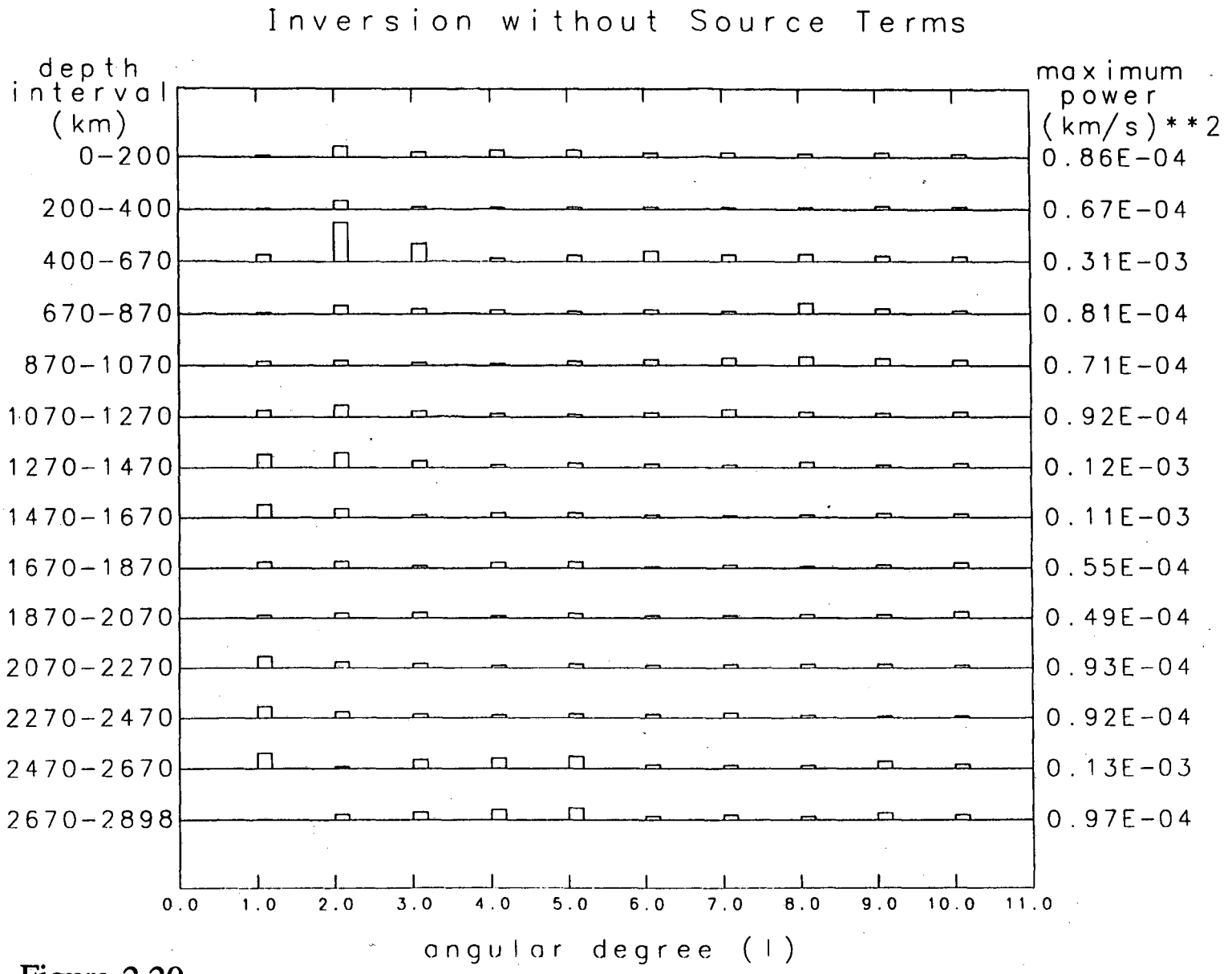

Inversion without Source Terms
depth interval (km)
maximum power (km/s)**2
0-200
0.86E-04
200-400
0.67E-04
400-670
0.31E-03
670-870
0.81E-04
870-1070
0.71E-04
1070-1270
0.92E-04
1270-1470
0.12E-03
1470-1670
0.11E-03
1670-1870
0.55E-04
1870-2070
0.49E-04
2070-2270
0.93E-04
2270-2470
0.92E-04
2470-2670
0.13E-03
2670-2898
0.97E-04
0.0 1.0 2.0 3.0 4.0 5.0 6.0 7.0 8.0 9.0 10.0 11.0
angular degree (l)

Figure 2.20

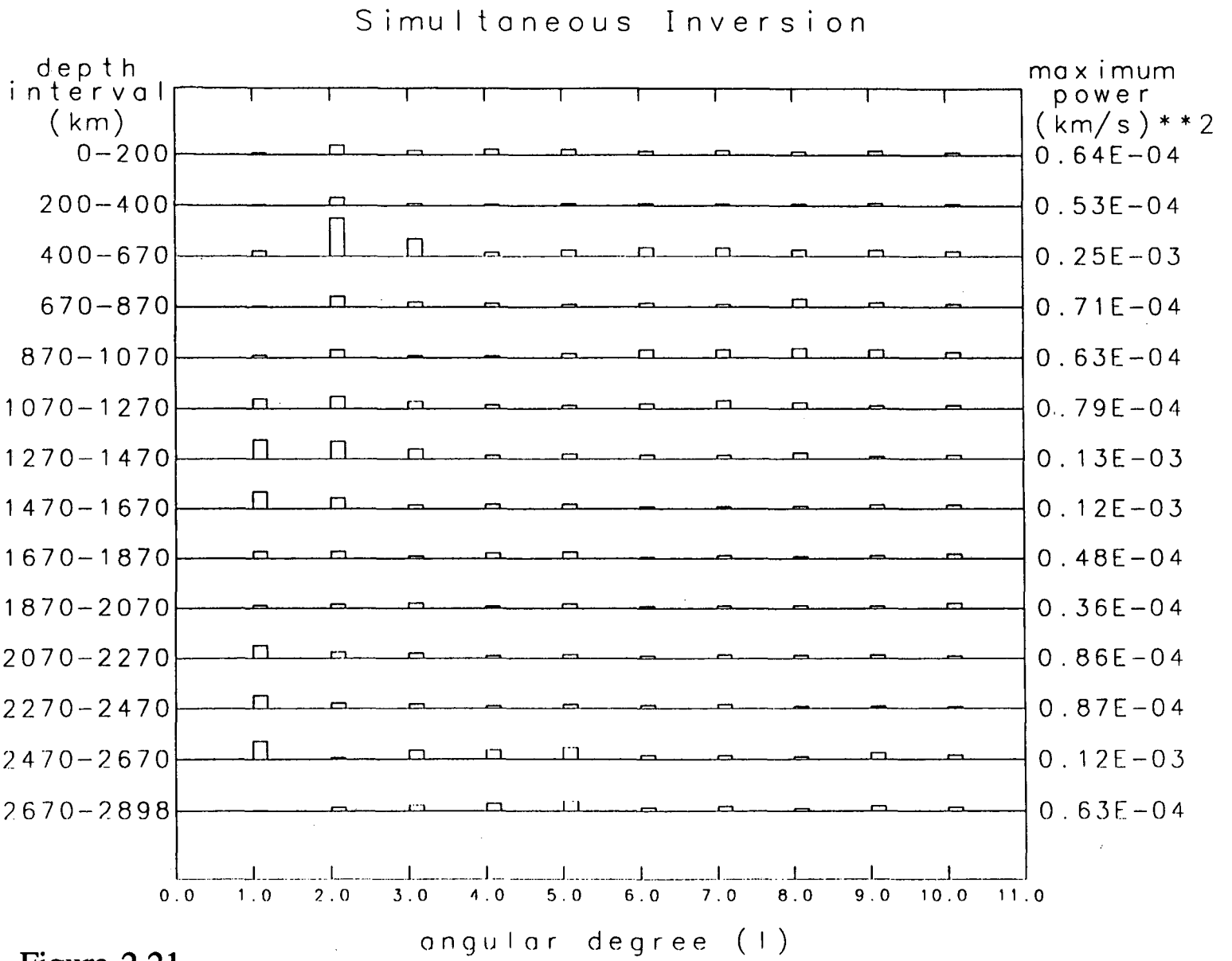

Simultaneous Inversion
depth interval (km)
maximum power (km/s)**2
0-200
0.64E-04
200-400
0.53E-04
400-670
0.25E-03
670-870
0.71E-04
870-1070
0.63E-04
1070-1270
0.79E-04
1270-1470
0.13E-03
1470-1670
0.12E-03
1670-1870
0.48E-04
1870-2070
0.36E-04
2070-2270
0.86E-04
2270-2470
0.87E-04
2470-2670
0.12E-03
2670-2898
0.63E-04
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
11.0
angular degree (l)

Figure 2.21

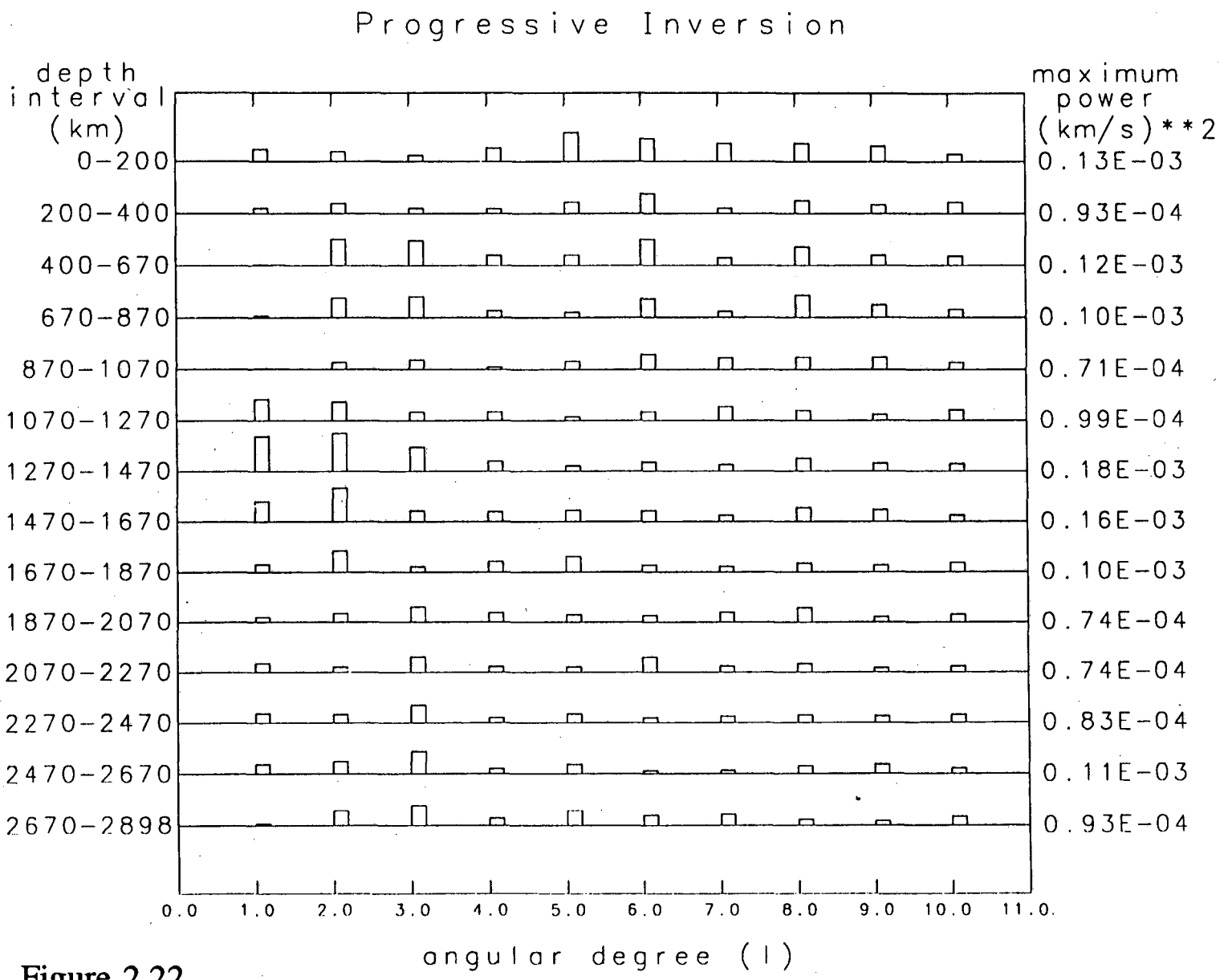

Progressive Inversion
depth interval (km)
maximum power (km/s)**2
0-200
0.13E-03
200-400
0.93E-04
400-670
0.12E-03
670-870
0.10E-03
870-1070
0.71E-04
1070-1270
0.99E-04
1270-1470
0.18E-03
1470-1670
0.16E-03
1670-1870
0.10E-03
1870-2070
0.74E-04
2070-2270
0.74E-04
2270-2470
0.83E-04
2470-2670
0.11E-03
2670-2898
0.93E-04
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
11.0
angular degree (l)

Figure 2.22

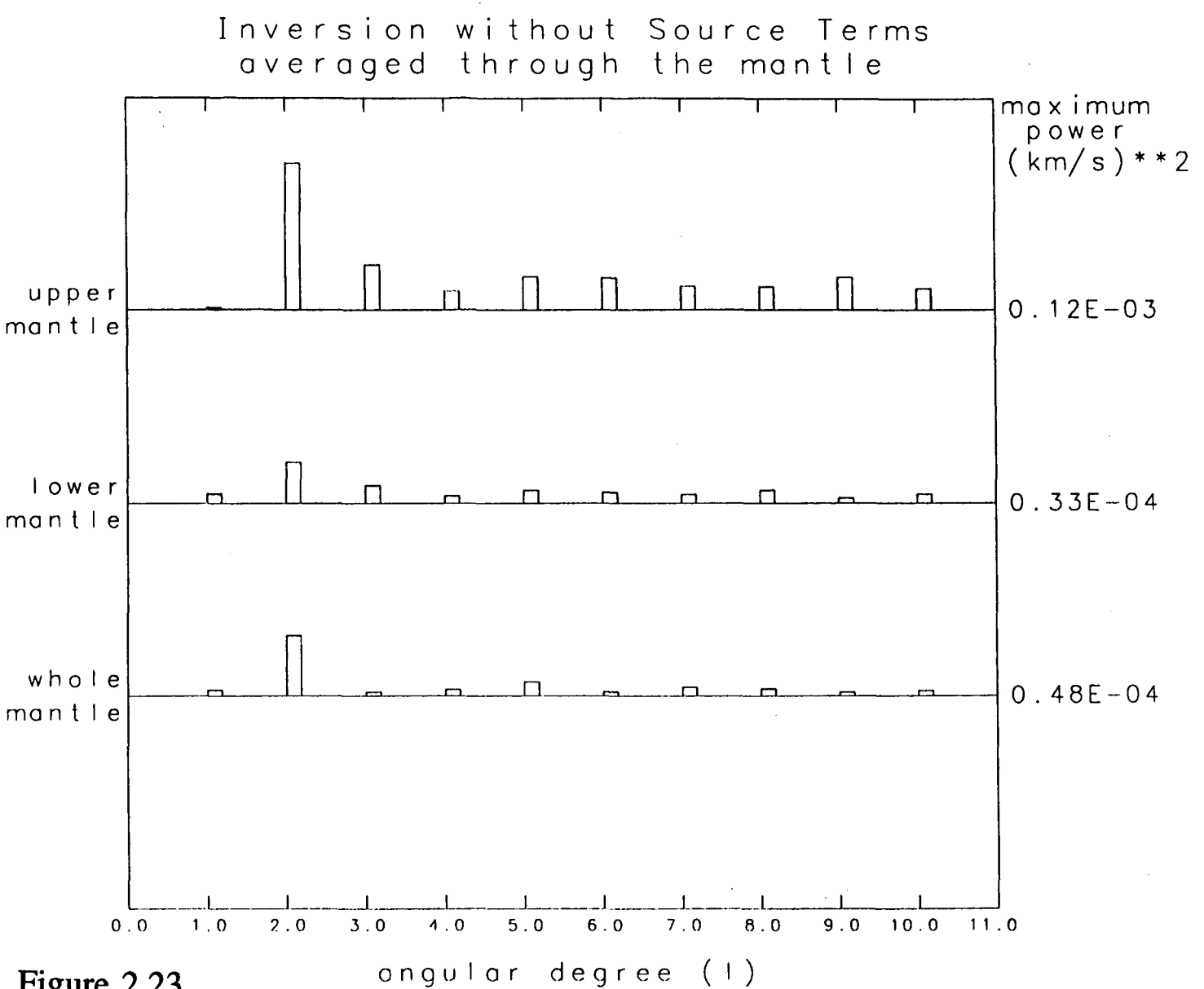

Inversion without Source Terms
averaged through the mantle
maximum
power
(km/s)**2
upper
mantle
0.12E-03
lower
mantle
0.33E-04
whole
mantle
0.48E-04
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
11.0
ongular degree (l)

Figure 2.23

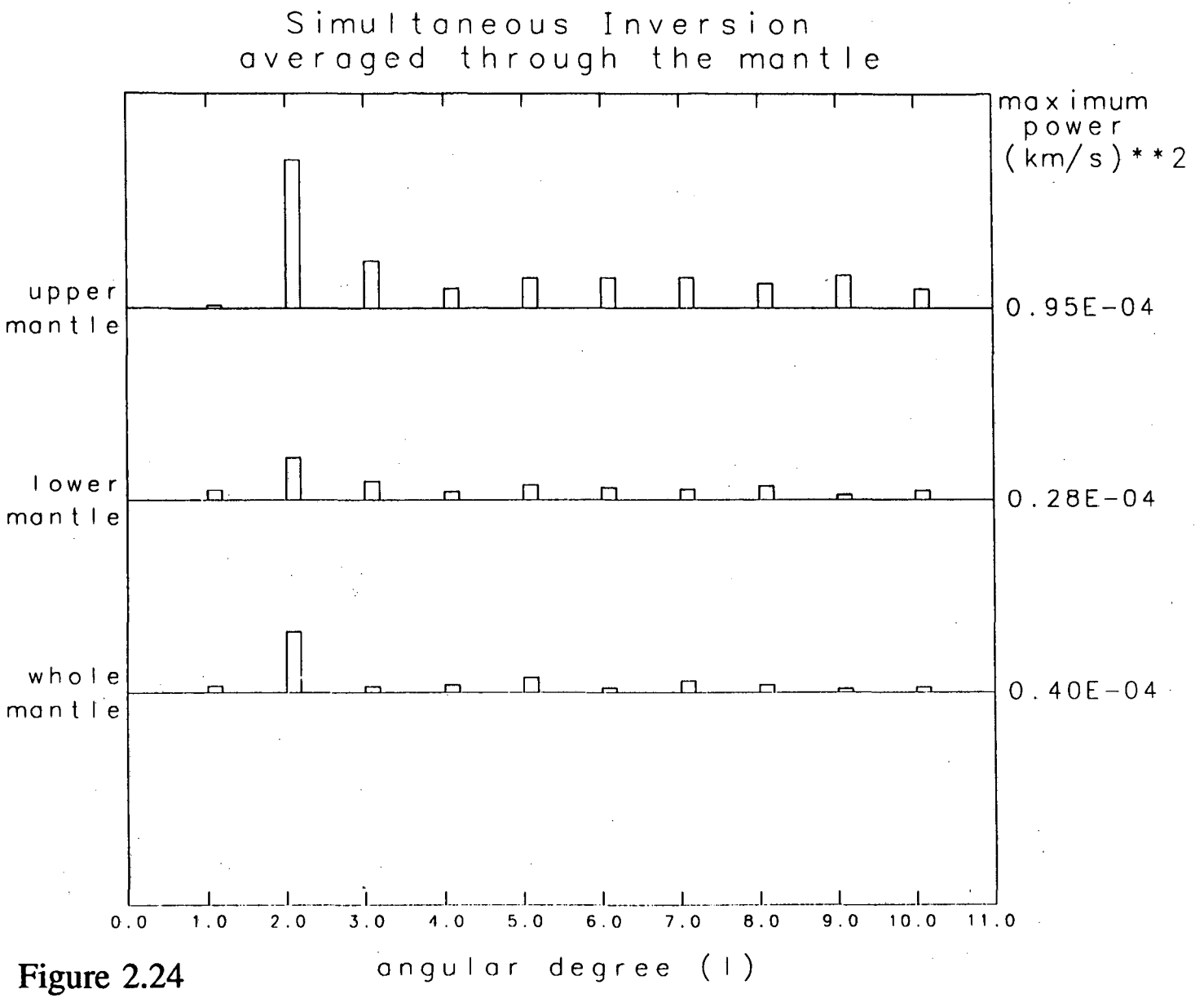

Simultaneous Inversion
averaged through the mantle
maximum
power
(km/s)**2
upper
mantle
0.95E-04
lower
mantle
0.28E-04
whole
mantle
0.40E-04
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
11.0
angular degree (l)

Figure 2.24

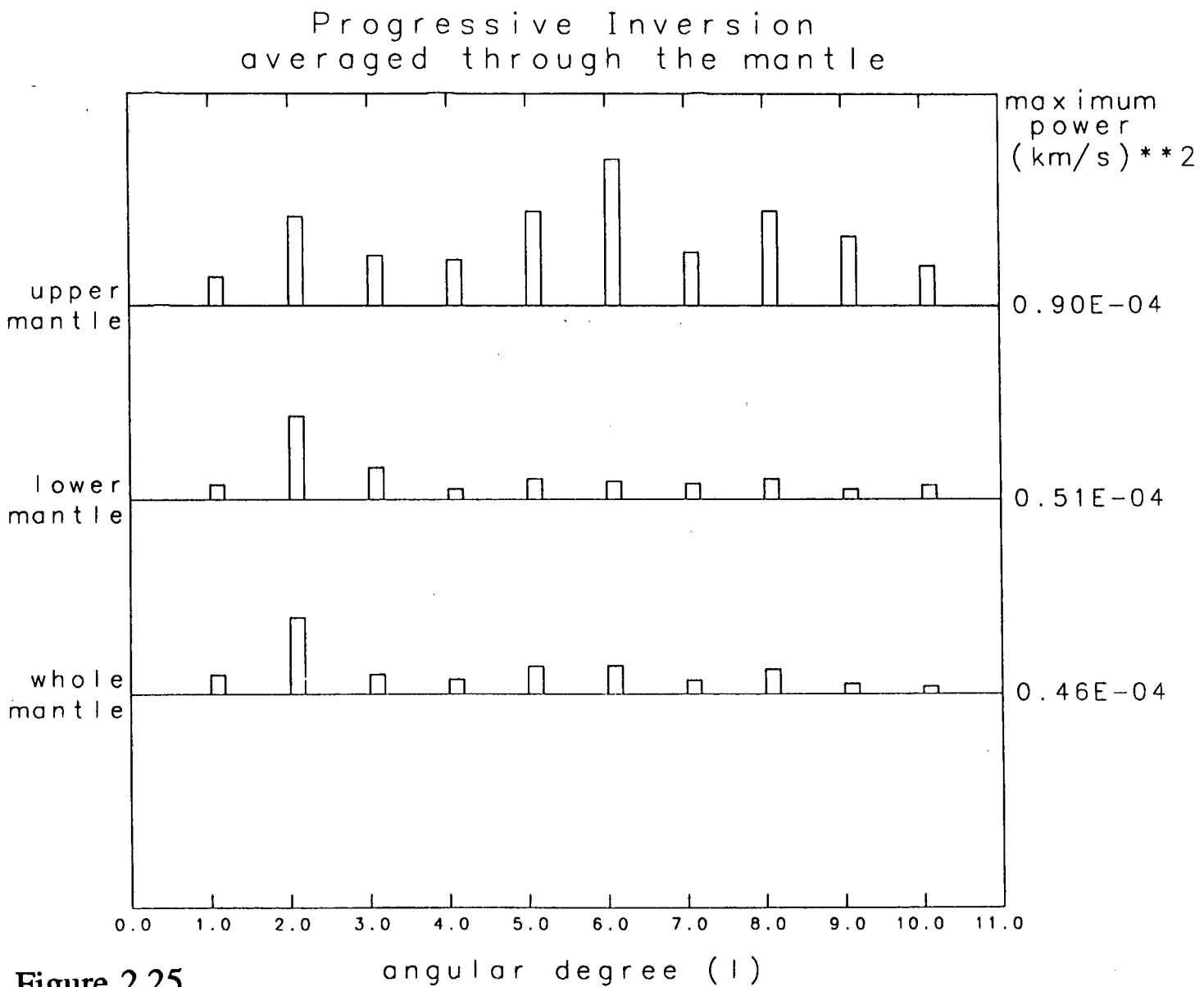

Progressive Inversion
averaged through the mantle
maximum
power
(km/s)**2
upper
mantle
0.90E-04
lower
mantle
0.51E-04
whole
mantle
0.46E-04
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
11.0
angular degree (l)

Figure 2.25

Figure 2.26

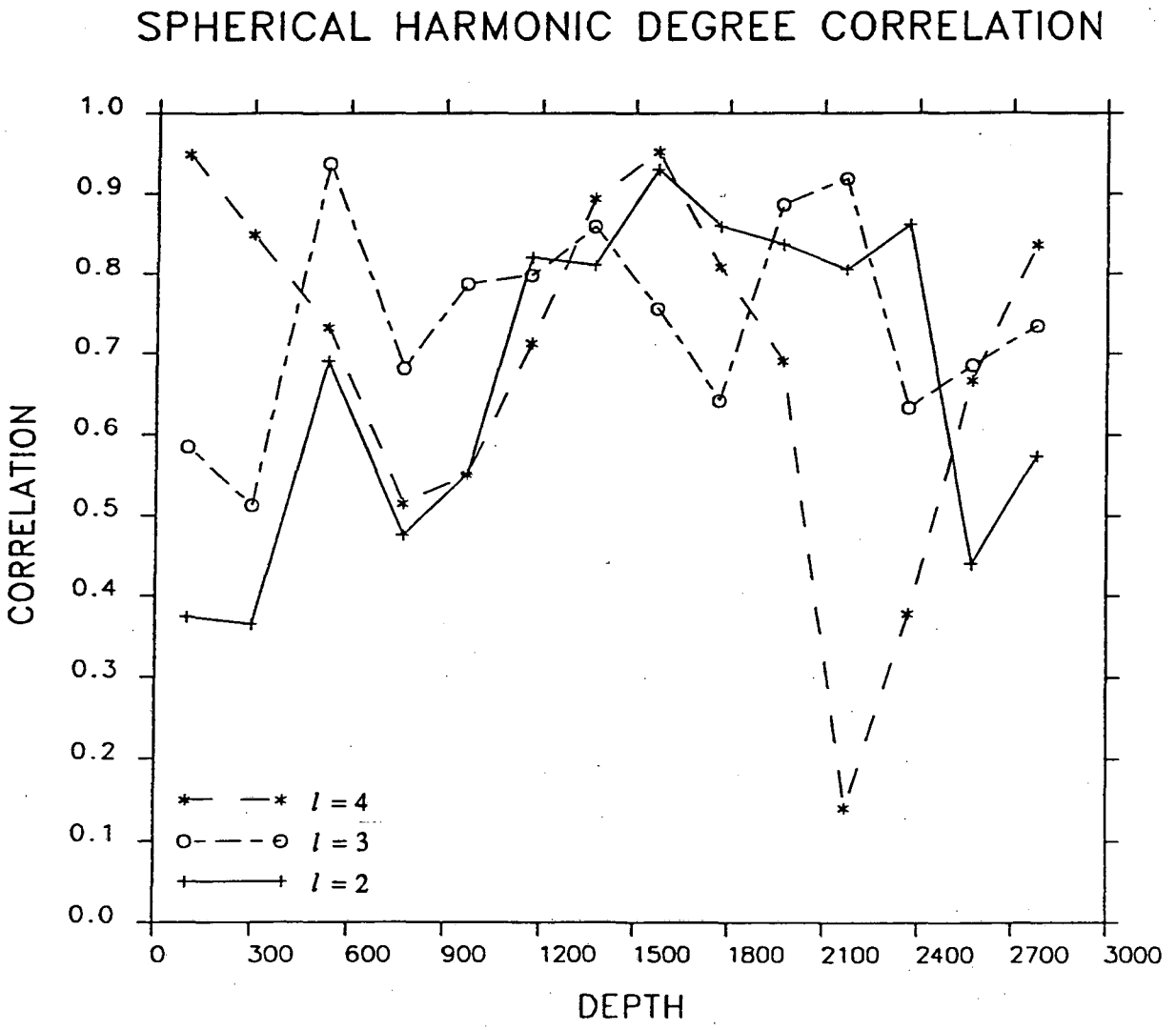

SPHERICAL HARMONIC DEGREE CORRELATION
CORRELATION
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
300
600
900
1200
1500
1800
2100
2400
2700
3000
DEPTH
l = 4
l = 3
l = 2

Figure 2.27

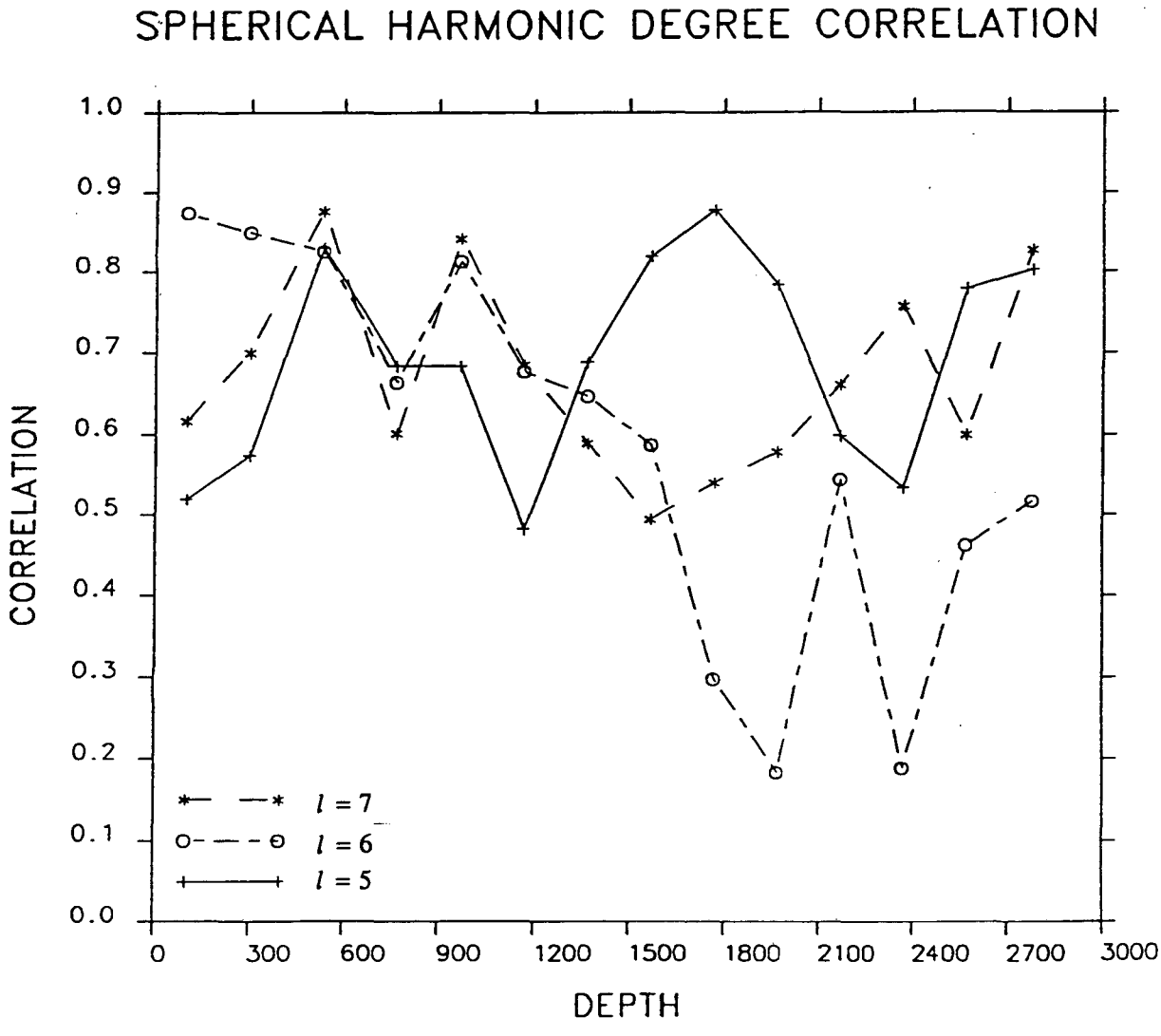

SPHERICAL HARMONIC DEGREE CORRELATION
CORRELATION
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
300
600
900
1200
1500
1800
2100
2400
2700
3000
DEPTH
l = 7
l = 6
l = 5

Figure 2.28

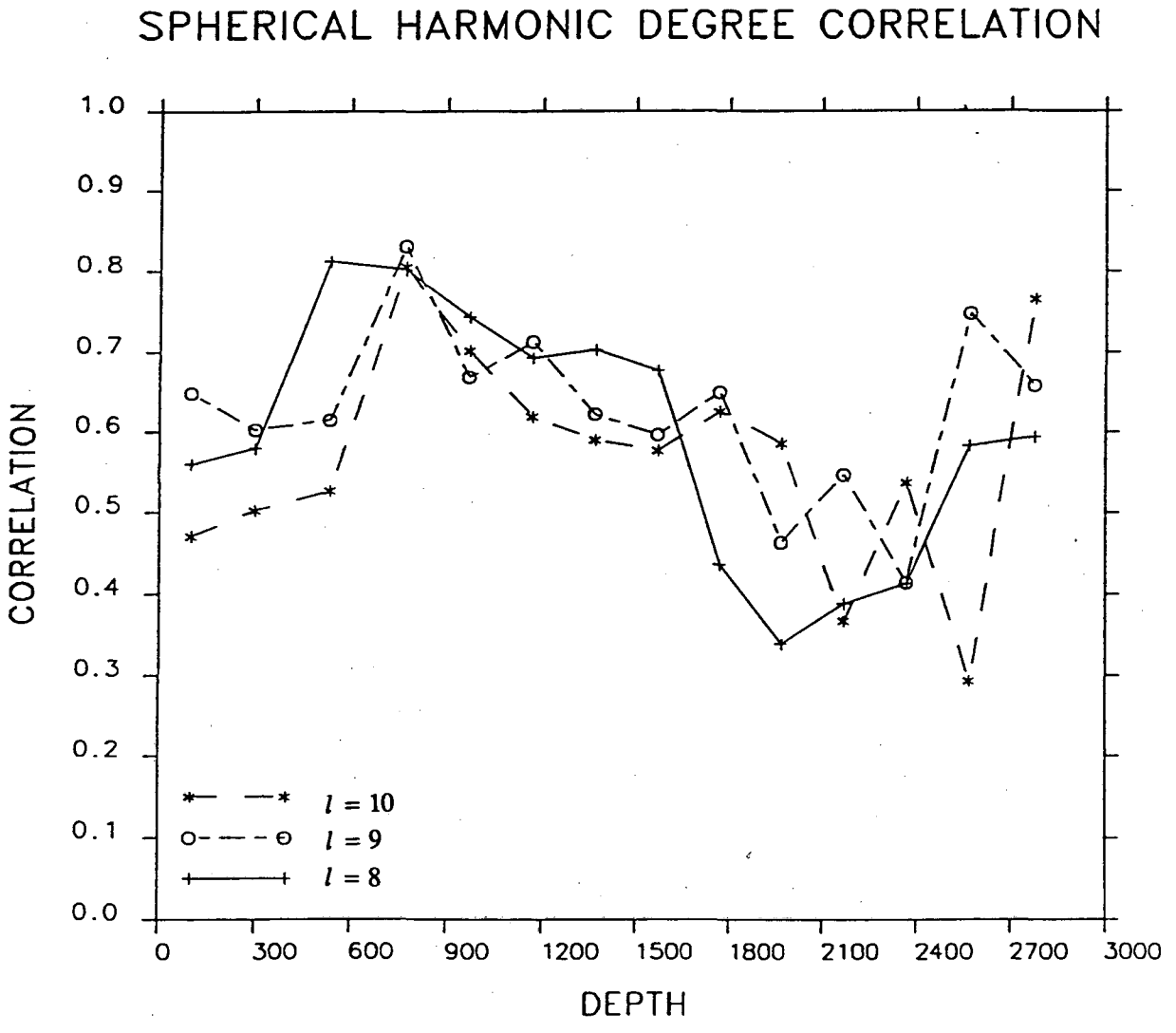
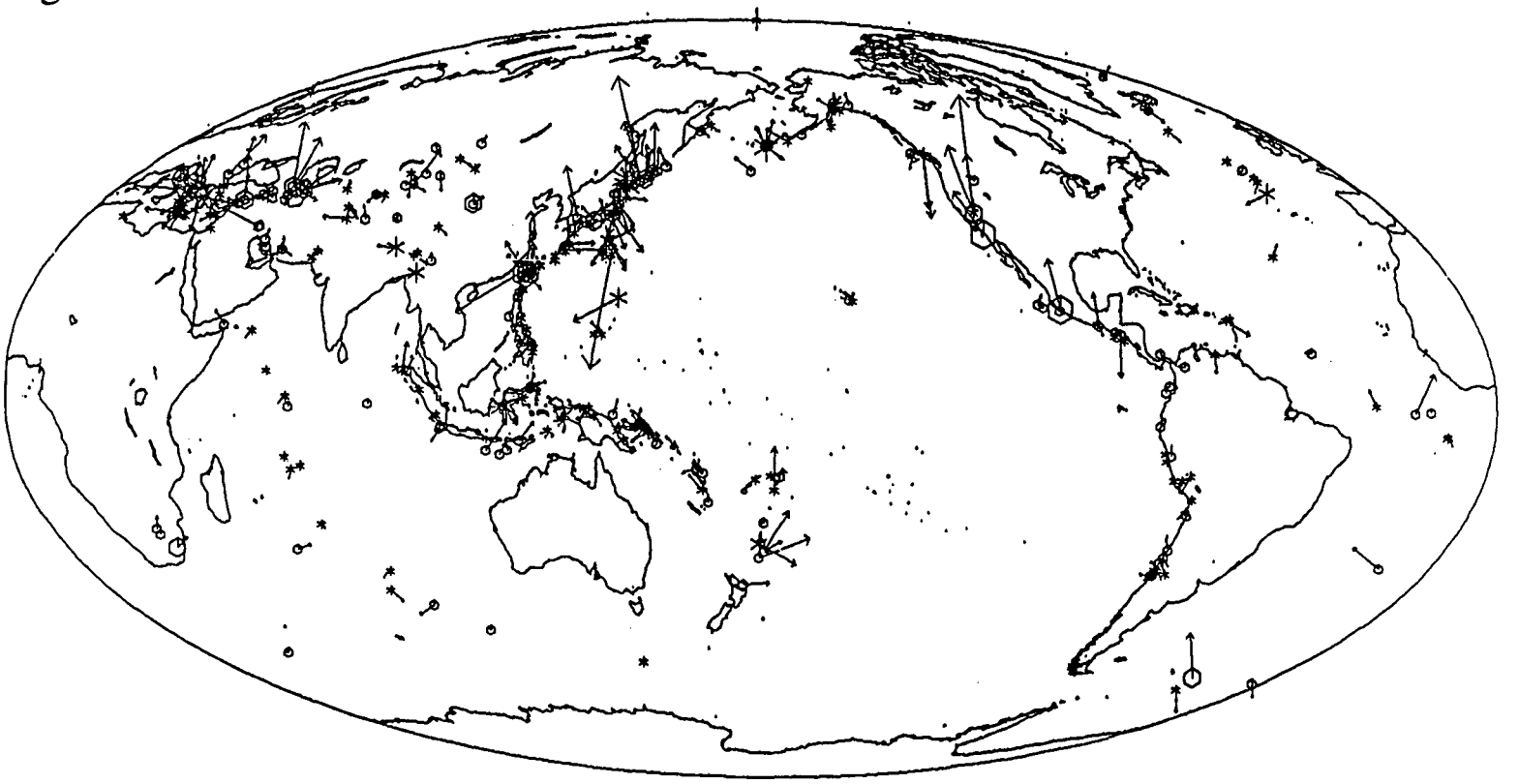
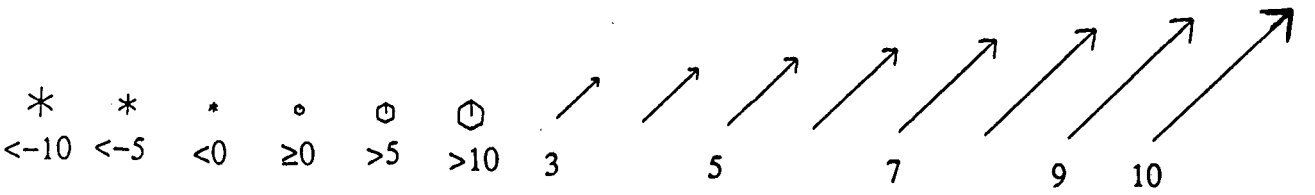

SPHERICAL HARMONIC DEGREE CORRELATION
CORRELATION
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
300
600
900
1200
1500
1800
2100
2400
2700
3000
DEPTH
l = 10
l = 9
l = 8

Figure 2.29

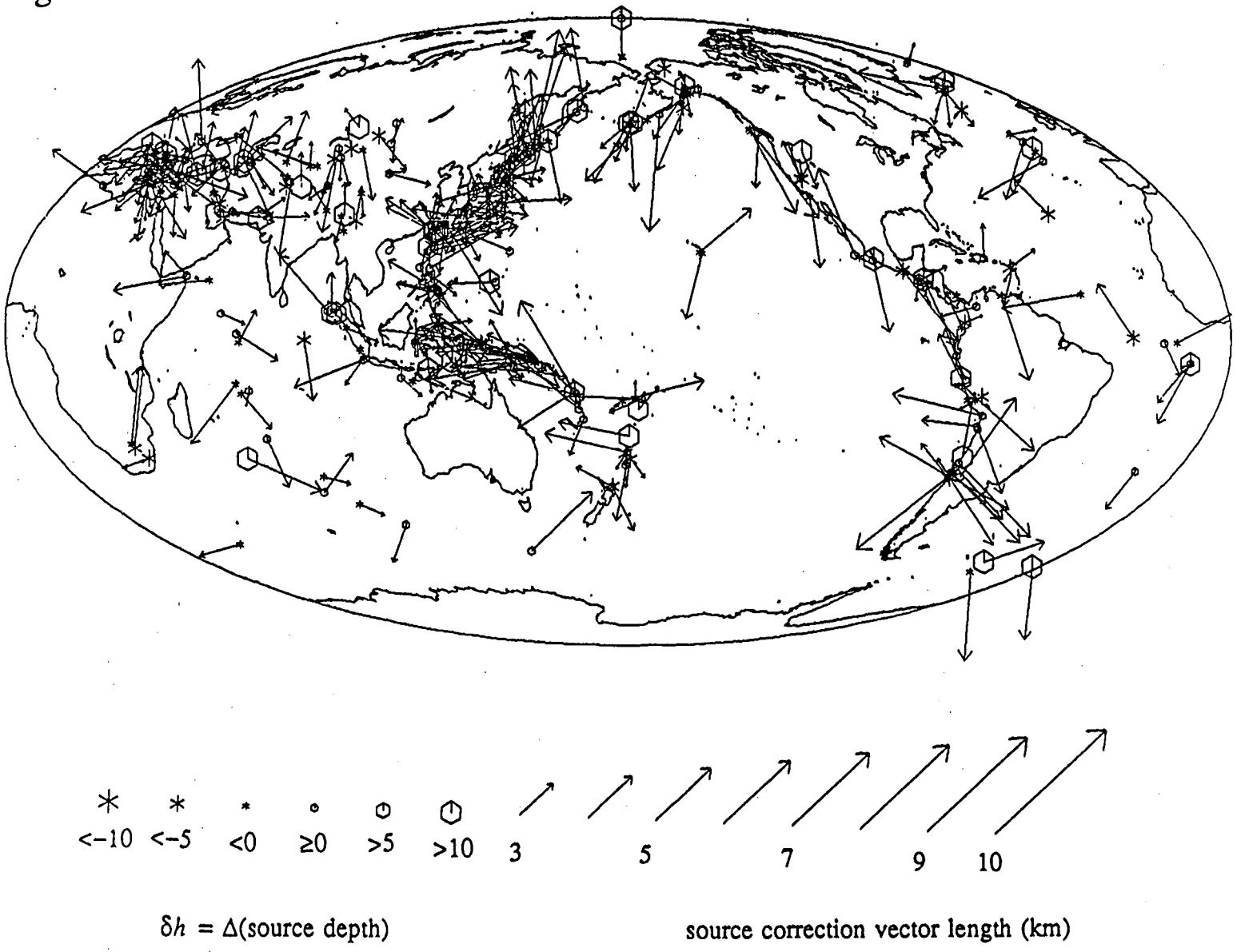




$\delta h = \Delta$(source depth)

source correction vector length (km)

Figure 2.30


<−10
<−5
<0
≥0
>5
>10
3
5
7
9
10
δh = Δ(source depth)
source correction vector length (km)

Chapter 3

What Patterns of Heterogeneity in Earth's Mantle Can be Revealed by Seismic Travel-Time Tomography?

3.1 Introduction

Global tomographic studies require massive volumes of data and huge numbers of model parameters to image Earth's interior to any helpful level of detail. That such studies are feasible at all is a tribute to innovative numerical techniques that solve the constraining equations iteratively and to the impressive speed of modern computers. However, our current computational capability has not quite reached the point where we can readily calculate resolution and covariance matrices to accompany our single, best-fitting solution (velocity model). To date, tomographic studies have bypassed the traditional formalism developed by geophysicists to evaluate the structure of the inverse problem and the reliability of its result. Approximate methods are employed instead.

Humphreys and Clayton [1988] explore the resolution of an inversion by means of a synthetic test in which a velocity perturbation is introduced to one or more voxels (volume elements) in a region of interest. Using Fermat's principle, which holds that travel times calculated through the three-dimensional Earth are insensitive to changes in raypath, one may calculate the travel time residuals that would be produced by the synthetic anomalies without tracing rays in the 3-D model. The inversion of the synthetic data may be performed and the cells adjacent to the perturbed voxel examined for smeared and "ghost" images that are artifacts of the inversion. The result may be thought of as the response of the algorithm to an impulse introduced to the system and forms one column of the resolution matrix, that is non-symmetric. *Humphreys and Clayton* [1988] call this vector the "point spread function", distinguishing it from the

"resolving kernel" that is the corresponding row of the resolution matrix. *Inoue et al.* [1990] show a way to approximate the resolving kernel for one model parameter, as well as the corresponding row of the covariance matrix. The drawback of calculating a single row or column of the resolution matrix is that each interesting feature must be examined individually and a separate inversion performed for each.

A similar, though more complete approach is to introduce a full model, so that a value is specified for every block, and invert the synthetic data generated through this model. With this method only a single resolution "value" is produced for each voxel. Each value may be regarded as a superposition of point spread functions. We are therefore unable to estimate a spread function, that would indicate the tradeoff in our resolution of one model voxel with others. *Inoue et al.* [1990] advocate a checkerboard pattern in which adjacent voxels alternate between two extreme values. The approximate model image may then be compared to the starting model to identify regions with poorly recovered values. *Spakman and Nolet* [1988] use a harmonic function instead of a checkerboard pattern. Both test patterns share the advantage that a more complete sense of resolution for the model may be presented with just a few figures. Unfortunately, they also share the disadvantage that separate inversions must be performed for various input models with different wavelengths and amplitudes. Both patterns are parametric in the sense that we assume before inversion that we have some idea of the spatial scale and amplitudes of interesting features of the real Earth. Impulse tests assume we know the location and amplitude of interesting features. Both methods require the same computation time as generating the best-fitting solution itself.

Nolet and Snieder [1990] suggest a less time-consuming means of producing a resolving kernel with a reduced basis, produced by the LSQR algorithm. It is common for coefficient matrices in tomographic problems to be numerically singular, so the matrix is rank-deficient and the space may be spanned by a basis that is considerably

smaller in dimension than the original matrix. For each iteration of the LSQR algorithm a single search direction vector is produced. One must decide when a sufficient number of vectors have been produced to represent the solution to the desired degree of accuracy. In the absence of the singular value spectrum, this decision presents a serious problem. *Scales* [1989] offers a way to obtain the singular values from the tri-diagonal matrix also produced by the LSQR algorithm, but he points out that numerical round-off errors can produce erroneous entries in the set of singular values.

Vasco [1991] presents an extremal bound approach to evaluating resolution and uncertainty in a tomographic inversion. Instead of finding the single model that is "best-fitting" in some sense, he finds the range of models that are consistent with the data. While this method is much different in its approach, it shares the computational drawbacks of "point spread function" methods mentioned previously. Each parameter must be considered individually and the computation time required for each voxel is comparable to the time required to find the entire best-fitting model. A subset of the model parameters could be examined but calculating bounds for every voxel is not yet feasible.

Even if all these techniques produce excellent approximations to resolution or covariance matrices, some basic questions remain unanswered. What are these tomographic models good for? Specifically, can they help us locate earthquakes more accurately? Can they help us constrain the composition of deep-Earth materials? Can they help us interpret the structure and scale of Earth's dynamic processes? In this paper we address only the last question. Rephrasing, we ask: If a particular pattern of flow exists in Earth's mantle and is reflected in the mantle's velocity structure, can that pattern be revealed by our tomographic imaging procedure?

3.2 Synthetic Experiments

There are many reasons why an existing pattern in velocity may not be observable seismically. The most obvious is that the amplitude of the anomaly is too small and is subsumed into the noise of the data. Second, approximations employed when parametrizing the model mantle and formulating the constraining equations might obscure a particular pattern. The finite size, shape, and distribution of cells in a voxel parametrization or pole orientation in a spherical harmonic series parametrization might not allow some patterns to be reproduced. Third, sampling of Earth by recorded seismic energy may be insufficient to show particular patterns. Plots of ray coverage for tomographic inversions show that large areas of Earth are either unsampled or under-sampled. In a sense, completely unsampled regions, provided they are small, are less of a problem than are regions in which sampling is inadequate to constrain model parameters sufficiently. Anomalies localized in one region or depth layer may be smeared into adjacent voxels or appear as "ghost" images elsewhere in the model as artifacts of the inversion. Some combination of these factors and others may conspire to obscure even the dominant patterns of heterogeneity in Earth. So what sorts of patterns can be revealed by seismic tomography and, equally important, what sorts of patterns are preferentially revealed by seismic tomography?

We investigate these questions with the set of experiments diagrammed in figure 3.1. We start by parametrizing our model mantle into approximately equal-area blocks of uniform velocity perturbation, $5° \times 5°$ at the equator and varying thickness (see table 1 and figure 3.2). To construct a synthetic data set based on real ray coverage of the mantle, we extract 45,000 events from the ISC data set, January 1964-January 1987, for each of which there are a minimum of 25 reported arrivals. We construct summary rays in order to minimize the effects of grossly redundant ray coverage and to filter out the effects of heterogeneity on a scale too small to be resolved by our voxels. After forming summary rays, the details of which are not relevant to this study,

our number of data are reduced from more than three million "actual" rays to about 726,000 summary rays. Further details of both data preparation and inversion procedure can be found in *Pulliam et al.* [1991]. Figure 3.3 shows the power contained in spherical harmonic expansions for each layer of the distribution of "hits", i.e., the number of times each cell is visited by rays, provided by the ray coverage of our data set. We calculate surface spherical harmonic series expansions by integrating around the globe for each coefficient, rather than by fitting coefficients to model values by least-squares. In this way, coefficients are independent of each other and coefficient values are independent of the point of truncation of the harmonic series. That is, coefficients do not change if the series expansion is calculated a second time with a different number of terms. The associated Legendre polynomials are fully normalized, i.e.,

$$p_l^m(\cos\theta) = \left[(2-\delta_{m,0})(2l+1)\frac{(l-m)!}{(l+m)!}\right]^{1/2} P_l^m(\cos\theta).$$

Figure 3.4 shows the power contained in spherical harmonic expansions of the sum of ray segments sampling each voxel. These series basically represent the coefficient matrix, which is the true weight applied to each model parameter in the inversion. The quantities expanded here are simply the column sums of the coefficient matrix.

We first introduce a simple spherical harmonic pattern into a single layer by allowing one term of the series to have a non-zero coefficient and recombining harmonics to find values for each voxel (step 1). At this point (step 1a) we check the effects of our model parametrization by immediately re-expanding our test model into spherical harmonics and examining the coefficients. Next (step 1b) we check the effects of our null coverage by zeroing out cells that are not sampled by rays in our data set and then expand the resulting model in spherical harmonics. Some smearing will necessarily occur and this simple result shows us the filter through which we view all of our results from global inversions with real data performed with our model

parametrization. The first re-expansion indicates degradation due to our block parametrization; another sort of velocity representation might suffer more or less than ours. The second re-expansion basically reflects the limitations of the ISC data set. Subsets extracted by other workers from the ISC master set will differ slightly, as individual decisions regarding data acceptability differ, but ultimately the ray coverage in each data subset will be similar.

In step 2 we calculate a synthetic data set of travel time residuals. Invoking Fermat's principle, which holds that travel times of seismic rays are insensitive to changes in raypath, to first order, we assume that the raypaths calculated through our starting, one-dimensional velocity model will persist unchanged in the presence of the three-dimensional anomaly. We may then calculate the perturbation in travel time that the introduced anomaly would contribute to each ray and replace the vector of real travel time residuals with the newly calculated vector and perform an inversion as detailed in *Pulliam et al.* [1991] (step 3). Finally, we expand the resulting voxel model in spherical harmonics and compare the power at each degree and in each model layer with the input spherical harmonics (step 4). We repeat these steps for more complex harmonic patterns, with and without added noise, using the results of previous seismic investigations of the mantle as guides to choosing patterns.

3.3 Results and Discussion

We choose test patterns based on the results of *Pulliam et al.* [1991]. These results are summarized, for the purposes of this paper, in figures 3.5 and 3.6. Figure 3.5 shows the power in series expansions of each layer as a function of angular degree. The top two layers appear relatively devoid of power at the lower degrees despite the strong concentration of ray coverage in the northern hemisphere and in continental regions, which are distributed primarily in l = 1, 2, and 4 patterns. However, in the 400-670 km layer l = 1, 2, 3, and 6 dominate. The finding of a large l = 2 component

confirms previous reports, but to our knowledge, no other study has shown the equally prominent $l = 1$ and 3 components. In the mid-mantle power is more or less evenly distributed across the harmonic terms. The exception is in the depth layer 1270-1470 km, where the $l = 1$ harmonic appears strongly and the $l = 2$ and 3 components rise above the higher-degree harmonics. These components clearly are responsible for the unexpected power total of this layer.

Since sign information is not included in power calculations, figure 3.5 does not offer any clues as to how the distribution patterns for all layers combine constructively or destructively to form a pattern for the whole mantle. Figure 3.6 shows the power in the spherical harmonic expansions averaged through the whole mantle and through the upper and lower mantle separately. The average is performed on the the individual harmonic coefficients, weighted at each layer by the square of the layer mid-point's radius, which normalizes the power in each layer to the layer's surface area. For the upper mantle the power spectrum shows a dominant $l = 6$ component, along with prominent $l = 2$, 5, 12, and 13 terms. The $l = 1$ power for the upper mantle is low, simply reflecting the results in figure 3.5 which show the $l = 1$ components of the first two layers to be small. However, in the lower mantle the same component is unexpectedly low, given the large values in several of the individual layers. This may mean the $l = 1$ component is poorly resolved in the lower mantle and trades off between layers in our model. In general, high power in harmonic degrees of the upper mantle coincide with high power in the same degrees of the lower mantle, though relative amplitudes once again point to the concentration of heterogeneity in the upper mantle. Degrees 2, 5, and 12 dominate the expansion averaged through the whole mantle. Surprisingly, degree 6 appears as a minimum even though its power dominates the top layers and contributes significantly to the lower mantle's total. From the relatively low power of the whole-mantle average compared to the separate lower- and upper-mantle averaged series, it is clear that either one part of the real mantle is

compensating for anomalies in the other part or our inversion scheme is trading off power between the lower and upper portions of our model mantle. The possibility of a numerical tradeoff in our inversion is investigated with these experiments.

3.3.1 Single Harmonics

When we include arrivals whose epicentral distances extend all the way back to 0° our intention is to allow better ray coverage of the upper mantle which, in turn, would permit us to image structure in that region. One problem with this strategy is that coverage is generally limited to regions which contain both sources and stations, typically these are tectonically-active continental regions. Although the ray coverage of the top layer is the least complete in the mantle, it is unique in the quality of its sampling. The many short rays which sample the top layer allow us to determine the contributions of individual voxels well, while the poorly sampled voxels contribute a moderate amount and the unsampled ones contribute not at all. Only in the top layer are some cells sampled individually, without being subject to trade off with other parameters, or sampled in small groups, trading off between just a few parameters and in many cases not trading off with any voxels below. Another problem arises from the extreme heterogeneity in the upper mantle and crust, which has been documented by other studies. Heterogeneity that has a coherent wavelength of less than about 550 km will not be seen in our results. This oversight should not affect our investigation of, and results for, longer wavelengths, but we should keep in mind that we may be missing the dominant forms of heterogeneity and concentrating on patterns that are less strong. The most serious problem arises from errors in travel time data for the upper mantle. The jumps in the mantle's spherically-symmetric velocity profiles at the 400 km and 670 km discontinuities cause triplications in the P travel time curve in the epicentral distance range 15° to 25°. The presence of more than one arrival on a seismogram at these distances frequently results in misidentifications of phases and contamination of the P database. This contamination is visible in increased variance of the

travel time residuals as a function of epicentral distance in the range 15° to 25°. Our purpose here is to explore the limitations in resolution due to the geometrical structure of the inverse problem, not to analyze the errors in the data. *Gudmundsson et al.* [1990] present an elegant analysis of the incoherency inherent in ISC P-arrival data.

In the top layer, 0-200 km, we examine the $l = 3$, $m = 0$ harmonic (figure 3.7a). The amplitude of the input pattern is quite large compared to amplitudes we might expect to find in the real Earth. For the moment we are most concerned with the redistribution of input anomalies and do not want these results to be obscured by random noise or systematic errors. In these first tests, of single harmonic input patterns, no noise has been added to the synthetic residuals. A set of experiments more realistic with respect to amplitudes follows the tests of single harmonics. The series terms combined on our model grid and re-expanded with contributions from all voxels (i.e., with "ideal" ray coverage) degrade the pattern very slightly (figure 3.7b). When voxels with no ray samples are zeroed out ("null coverage"), the re-expansion degrades quite a bit further but is still more than adequate to identify the pattern (figure 3.7c). Only 60% of the input power is contained in blocks that have nonzero sampling. This value represents the maximum we can hope for our inversion to return; anomalies placed in voxels that are not sampled cannot be recovered.

Figures 3.8a-c show the results of an inversion performed with a set of synthetic travel time residuals produced by estimating the delays caused by the introduced $l = 3$ velocity anomaly. Though the pattern in the Pacific and Atlantic oceans and the Southern hemisphere is poorly retrieved, this simple pattern is still recognizable in the top layer (figure 3.8a). Portions of this same pattern are just barely discernible in the second layer (figure 3.8b), to which no anomaly was introduced. By the third layer the pattern is gone, though traces of stray anomalies are observed.

Though visually compelling as representations of the input pattern's distortion, these figures do not facilitate a quantitative analysis of the redistribution of power by

our inversion scheme. Figures 3.9a and 3.9b are more helpful for this purpose. Figure 3.9a, analogous to figure 3.7a, shows the power input to the model's first layer and the re-expansions with ideal and null ray coverage. Figure 3.9b shows that the power returned by our test inversion is overwhelmingly concentrated in the $l = 3$ term, but significant aliasing to other harmonics occurs. The largest-amplitude aliased harmonic is the $l = 4$ term, immediately adjacent to the anomalous term on the short-wavelength side of the power spectrum. This is a tendency we observe throughout these tests. The largest-amplitude smearing is generally to the next larger degree, followed by smaller-amplitude smearing to the next larger degree and so on. The $l = 3$ harmonic is also smeared downward slightly to adjacent layers, though not significantly. Note that the $l = 1$ harmonic returns nearly devoid of power, despite the strong bias of our ray sampling to that term (see figure 3.4). The total power returned in our inversion's top layer comprises 0.136 $(km/sec)^2$ of the 0.151 $(km/sec)^2$ introduced to sampled voxels. Of this layer's total power, just over a third is contained in the $l = 3$ component. However, the lower layers contain .022 $(km/sec)^2$. Thus the inversion slightly overestimates the power in the input model, which is unusual for tomographic inversions and is not the case for any but the top layer. We will discuss the problem of power estimation later.

The $l = 10$ test for the first layer (figure 3.10a) illustrates the effects of our model parametrization and incomplete ray coverage on inversion results for shorter wavelength features. When re-expanded with ideal ray coverage, the input power of the $l = 10$ term is underestimated by nearly 20%, though the total power in the re-expanded series comprises 95% of the input. The expansion of just the sampled voxels contains only 60% of the input power. Figure 3.10b shows results similar to those of the $l = 3$ test, including aliasing to other harmonics in the same layer and slight smearing to the $l = 10$ harmonic in lower layers. The total power contained in the $l = 10$ test inversion comprises 75% of the sampled blocks' input.

An entirely different story is told by tests of our model's second layer, 200-400 km. In the expansion using only sampled voxels the $l = 2$ harmonic alone contains 64% of the input power, though the entire series contains 75% of the input (figure 3.11a). Inversion results show significantly less aliasing of power to other terms in the same layer, but significantly more power is smeared to the layers above and below the anomalous region. Testing the $l = 4$ harmonic in the second layer (figures 3.12a and b) gives results similar to the $l = 2$ test. The $l = 4$ input harmonic is of the same amplitude as the $l = 2$ anomaly (figure 3.12a), but the re-expansion of the combined series with "ideal" ray coverage shows that even at this relatively low degree, the uneven nature of our model parametrization has a deleterious effect. As one moves from the equator to Earth's poles the voxels subtend greater and greater longitudinal angles in order to have approximately the same surface area as other voxels in the same layer (see figure 3.1). Ultimately this construction of voxels will limit the wavelength that our parametrization can resolve. Aliasing to other harmonics in layer two is more extreme than with $l = 2$, but the smearing to layers above and below is not quite as strong. An explanation for this smearing might be that the conditioning we apply to rows of the coefficient matrix and vector of travel time residuals down-weights the rays that bottom in the second and third layers. This conditioning is the inverse variance of the travel time residuals as a function of epicentral distance and reduces the effective weight given to rays emerging at about 15° to 25° because of these residuals' large variance. However, were it not for results in layers one and three, which show far less smearing in depth, the results for layer two would be expected and even encouraging. Poor depth control in the upper mantle with body-wave tomography can be accepted and explained easily by geometrical arguments involving crossing versus parallel rays. One might argue that our inclusion of rays all the way back to 0° gains us little because we are forced to downweight these contributions based on the scatter of residuals, but our results for layer one and the impressive

lateral control (absence of aliasing) in figure 3.11 argue otherwise. The power total for the layer 2, $l = 2$ test inversion is only 30% of the power contained in sampled voxels, far short of the 90% and 75% found for the tests in which $l = 3$ and $l = 10$ were introduced to the first layer. The 30% for $l = 2$ and 39% for $l = 4$ found here are more characteristic of the totals found for subsequent tests and actually make more sense than values approaching 100%. Tomographic inversions typically deal with a poorly conditioned system of constraining equations and the model space of solutions is quite large. Additional prejudices must be added in order to find (or define) a unique solution. We employ a minimum norm criterion and also apply a smoothing operator which reduce the variability of our voxel values. Our dilemma requires choosing between a generally smooth model and a model which fits the data exactly. As a result, the model we pick from the set of acceptable solutions has both a small norm and low variability between adjacent voxels. It is reasonable to believe that this model will generally underestimate the power contained in the real Earth, or in a simulated Earth input to our inversion scheme. We could easily pick other models from the range of least squares solutions that overestimate actual power, but we do not currently know how to specify a unique model that estimates actual power accurately and is physically plausible.

We choose four single harmonics to test for the transition zone, 400-670 km. These are $l = 2$, due to its prominence in the results of previous studies [e.g., *Masters et al.*, 1982, *Romanowicz et al.*, 1987, *Inoue et al.*, 1990, *Vasco et al.*, 1990, and *Pulliam et al.*, 1991], $l = 4$, due to its intriguingly small value in figure 3.5, $l = 6$, because of its moderately large value in figure 3.5 and its constructive behavior when several layers in the upper mantle are averaged, and $l = 12$, in order to test the resolution of our model parametrization at higher degrees. Figure 3.13a shows the input pattern for the $l = 2$ pattern introduced to the transition zone and the re-expansions with all voxels contributing, first, and with only the sampled voxels contributing, second.

The pattern is returned almost exactly when all voxels are incorporated. Of course, this result is not unique to this layer since it does not depend on ray coverage. The power lost with unsampled voxels amounts to just 15% of the input. Figure 3.13b shows the power contained in the inversion performed with synthetic residuals. The pattern is indeed returned as $l = 2$, with some smearing to adjacent layers above and below but with more aliasing to other degrees within the same layer. Interestingly, $l = 1$, 3, 5, and 6 all show equivalent power, but $l = 4$, which also shows small values in studies of the real Earth, returns a relatively small component. The structure of the problem seems to discourage placing power in the $l = 4$ degree. Although the largest harmonic, $l = 2$, is returned as 0.04 $(km/sec)^2$ the total power contained in the entire model is 0.07 $(km/sec)^2$. This total represents about 30% of the power contained in sampled voxels.

Results for the $l = 2$ harmonic test suggest we perform a similar test for the $l = 4$ degree in the same layer. At $l = 4$ the parametrization interferes with the full recovery of the input harmonic, but only slightly (figure 3.14a). Likewise, the re-expansion performed with unsampled voxels set to zero reveals 80% of the input power is returned to the $l = 4$ component, but small amounts are returned to adjacent degrees, for a total of 88% of the input in the harmonic series. More surprising results turn up in the inversion of synthetic residuals (figure 3.14b). While significant power is aliased to other harmonics within the same layer and some power is smeared to the $l = 4$ harmonics above and below the anomalous layer, very little power appears elsewhere in the model. The total power in the output model is 43% of the power input to the sampled voxels, comparable to the 36% retrieved in the $l = 2$ case. The harmonic that shows the most power other than $l = 4$ is $l = 5$, which fits the tendency noted above to move power to the next higher harmonic. Here that tendency is observed even though the $l = 5$ component appears with a relatively small value in a model produced for the mantle with real data (ISC5_LSQR, presented in *Pulliam et al.*, 1991. In short,

the $l = 4$ component would be expected to show up noticeably if it exists in the transition zone of the real Earth, but would be preferentially aliased to the $l = 5$ harmonic. The fact that neither of these harmonics appears significantly in our mantle models, coupled with these experiments, leads us to believe the $l = 4$ harmonic is not significant in the mantle's pattern of heterogeneity.

Results for a test of $l = 6$ in the second depth interval (figures 3.15a,b) are nearly identical to the $l = 4$ results. 44% of the sampled voxels' input is returned to the inverted model. The next higher harmonic, $l = 7$, benefits most from the inversion's misplacement of power. Smearing upward and downward in depth is minimal and the pattern contained in the output model is unambiguously recognizable as the input pattern. Figures 3.5 and 3.6 indicate that $l = 6$ is a large component of heterogeneity in the upper mantle and the results shown here, combined with results described below for a more general test, suggest that an existing $l = 6$ pattern should be clearly revealed in an inversion.

Pushing our model parametrization to the limits of its resolution with a spherical harmonic representation, we test the $l = 12$ harmonic in the transition zone (figures 3.16a,b). The most drastic degradation of the input pattern occurs when the series is recombined to the model grid and re-expanded with all voxels' contributions (figure 3.16a). Only 76% of the input power is contained in the $l = 12$ component of this second series, though 93% is contained in the entire series. In contrast, the loss due to unsampled pixels is only an additional 8% of the input $l = 12$ power. The test inversion results (figure 3.16b) show that more power is moved to the next higher harmonic than is observed for any other component, though the $l = 12$ signal still achieves a value comparable to the $l = 2$ signal in its test for this layer. In this case, however, the signal-to-ghost ratio has dropped considerably, to 2.5. Though absent here, in the presence of systematic and random data errors and more complex patterns of heterogeneity, the $l = 12$ component may approach or exceed the resolvability limit of our

model parametrization and ray coverage. We explore this resolvability in the presence of random errors and complex heterogeneity in the next section.

3.3.2 A Higher-order $l = 2$ Pattern in the Upper Mantle

Modeling of heterogeneity in the upper mantle with free oscillations [e.g., *Masters et al.,* 1982] and surface waves [e.g., *Romanowicz et al.,* 1987] reveals an anomalously large $l = 2$ harmonic component. In addition, tomographic studies of body waves also point to an important degree two component [e.g., *Inoue et al.,* 1990; *Vasco et al.,* 1990; *Pulliam et al.,* 1991]; yet some disagreement remains about the depth range at which this pattern of heterogeneity is most dominant. *Masters et al.* [1982] localize the pattern to the transition zone, 400-670 km. *Romanowicz et al.* [1987] find a similar pattern at 300-400 km depth. The body wave studies all suggest the transition zone is the most likely location, but special considerations concerning ray coverage and the determination of raypaths in the upper mantle lead us to wonder how strong is our radial control of velocity anomalies in the upper mantle. We pursue this question with two experiments in which we place the best-fitting pattern found by *Masters et al.* [1982] in the 200-400 km layer, first, and in the 400-670 km layer, second, and perform separate synthetic tests for each case. Figures 3.17a and 3.17b show recombined Mollweide projections of the input pattern and the re-expansion with ideal coverage, respectively. Figure 3.18 shows the same information for the harmonic series before they are recombined for presentation in figures 3.17a and 3.17b. Here we show the amplitudes, rather than power, of the spherical harmonic series coefficients. Due to the confusing concentration of information in this figure, only power will be shown in figures for subsequent tests. This long-wavelength pattern is retrieved almost exactly when re-expanded with ideal ray coverage (figures 3.17b and 3.18). However, since the ray coverage in the upper mantle is so incomplete, the re-expansions with null ray coverage show some degradation of the input pattern. Figures 3.17c and 3.18 show the null coverage re-expansion for the 200-400 km layer; figures 3.17d and 3.18

show the 400-670 km layer. The highs and lows are recovered adequately even in the southern hemisphere. Only the portions with amplitudes near zero and values in oceanic regions are poorly recovered. The deeper layer, which has more complete ray coverage, shows a more faithful facsimile of the input, though both layers' results are acceptable.

Inversion results for the 200-400 km layer are far superior to the 400-670 km results, particularly with regard to depth control. For the 200-400 km layer (figures 3.19 and 3.20) the B_2^2 coefficient is well-recovered, as are all coefficients but A_2^2, which nearly disappears. A great deal of aliasing to higher harmonics within the same layer occurs and the power smeared to the 400-670 km layer reaches nearly the same level as the power recovered in the 200-400 km interval. In contrast, when the pattern is input to the 400-670 km depth layer (figures 3.21 and 3.22) the inversion returns both a more faithful reproduction better confined to the input layer. Higher harmonics in the same layer receive less of the prodigal power and the anomalies transferred to adjacent layers above and below are smaller in amplitude than are the smeared anomalies in the 200-400 km case.

3.3.3 Coherent Patterns in the Upper Mantle

To test whether a coherent pattern of heterogeneity for the upper or lower mantle is recoverable, we must first choose the pattern to test. Our guide to this choice is figure 3.6, in which the results for an actual inversion are averaged over the top three layers (0-670 km depth) for the upper mantle and the bottom eleven layers (670-2898 km depth) for the lower mantle. For the lower and upper mantle we choose two test patterns each, all of which are drawn from the results shown in figure 3.6. In each case, the first test deals with a pattern that appears prominently in the actual results and the second test considers a pattern that generally shows low power in the actual inversion. The amplitudes of these input coefficients are drawn from actual inversion

results. Figure 3.23a shows the input pattern for the first test of the upper mantle, the series averaged over the top three layers of the ideal coverage re-expansions, and the average formed from series re-expanded with actual ray coverage. Also shown are the averaged outputs from test inversions performed with and without random noise added to the synthetic travel time residuals. The added noise follows a two-sided exponential distribution with variance adjusted to match the variance of actual residuals calculated (approximately) through model ISC5_LSQR of *Pulliam et al.* [1991]. Figure 3.23b shows the patterns re-expanded for each layer with ideal and actual ray coverage. Note the increase in power contained in each series with increasing depth. This trend reflects the greater number of voxels sampled in each layer at greater depths, though it does not reflect the increasing homogeneity with depth of ray coverage among sampled voxels. The total power in the re-expansion with ideal ray coverage is 96% of the input; the re-expansion with actual coverage, representing the best inversion results we can hope to obtain, totals 72%. The inversion results contain 80% of the total power in sampled voxels, for the case without noise, and 96%, for the case with noise. Whereas the addition of exponentially-distributed noise to an inversion procedure that minimizes the l^2 norm of the residuals clearly influences the total power contained in output models, the spherical harmonic patterns generated in each case are nearly identical (figure 3.23c and 23d). The largest values are returned in the topmost layer, 0-200 km. Again, this is not a complete surprise since the epicentral distance range of our rays extends to 0°, which allows some portion of these rays to bottom in the uppermost layer. The greatest aliasing to adjacent harmonics occurs in the top layer. This may be due to the extremely uneven ray sampling in this layer; one voxel under Europe is visited more than 20,000 times by our set of rays while large portions of the oceans, particularly in the Pacific and the southern hemisphere, are not sampled at all and many voxels are sampled by just a few rays. Still, the input harmonics return with power two to three times that of the closest ghost harmonic, clearly dominant and

unambiguous when compared to the real anomalies' nearest neighbors. The second layer, 200-400 km, of figures 3.23c and 3.23d show the same disturbing effect with respect to the layers above and below. Power is consistently underestimated in this layer, although the power ratio of real to ghost harmonics is higher than for the top layer. This effect may be due to a defect in our one-dimensional starting model, a modified Jeffreys-Bullen P-velocity model *[Jeffreys,* 1960], which contains no discontinuities. Many studies have pointed to the need for sharp jumps in velocity at 400 km and 670 km depth; one effect of these discontinuities on our initial raytracing would be that first arrivals would commonly be from waves refracted at the discontinuities. Fewer raypaths would have their turning points in the 400-670 km depth layer in a model with discontinuities. With respect to the real Earth our rays probably oversample the 400-670 km layer at the expense of the 200-400 km layer. The 400-670 layer in figures 3.23c and 3.23d show the best ratios of actual to ghost harmonics of all three upper mantle layers. The amount of smearing to layers below the three perturbed layers is small. fully an order of magnitude below the values returned for actual perturbations. The tendency to alias power upward, rather than downward in the spectrum, is observed even for the $l = 2$ harmonic, despite the strong $l = 1$ component of the ray coverage expansions (figures 3.3 and 3.4).

The second test for the upper mantle, shown in figures 3.24a-d, considers a pattern that appears in actual results with small power. The ratios of maximum power of inputs to outputs and total power of inputs to outputs are quite similar to those found in the first upper mantle test (compare figures 3.23a and 3.24a). Similarly, the increase in power recovered when noise is added to the synthetic data mirrors the increase found previously and the portions of signal lost due to incomplete ray coverage follow the percentages established in the first test (figure 3.24b). The output patterns tell a slightly different story, however. Once again the second layer recovers the minimum power total, though the discrepancy is not nearly so drastic, and the third layer shows

the cleanest signal (figure 3.24c). Aliasing to shorter wavelengths occurs in all layers, though the dominant occurrence is again in the 0-200 km depth layer. The $l = 8$ harmonic is barely distinguishable from the noise in the top layer, and only reliably retrieved in the 400-670 km depth range. The smearing of power downward in the mantle appears to be a greater problem for these harmonics; though surprisingly the aliased power propagates more strongly than the perturbed harmonics. Particularly worrisome is the $l = 2$ component, which does not exist at all in the input pattern but appears significantly in the 670-870 km depth layer. There it reaches one-third the power of the largest harmonic recovered in the layer above. The addition of noise (figure 3.24d) generally degrades the signal to ghost power ratio for the perturbed layers, smears power from perturbed harmonics downward from the 400-670 km layer to the 670-870 km layer, and reduces the power of the ghost $l = 2$ component in this lower mantle layer. The addition of noise to the synthetic data apparently obscures a structural tendency of the inverse problem to funnel power to the $l = 2$ harmonic.

3.3.4 Lower Mantle

The first test for the lower mantle considers the $l = 2, 7$, and 12 harmonics (figures 3.25a,b,c and d), which appear with relatively large power when model ISC5_LSQR is averaged over the lower mantle (figure 3.6). The input pattern is taken directly from the model ISC5_LSQR and the same constructive pattern seen throughout the lower mantle model is introduced to each of the eleven layers spanning the depth range 670 km to 2998 km. The re-expansions of the input pattern for ideal and null ray coverage, shown averaged over all layers in figure 3.25a, and for individual layers in figure 3.25b, show only slight degradation of the input pattern. The total power contained in the eleven lower mantle layer expansions with ideal coverage comprises 96% of the power contained in the input series. The total power contained in the null coverage expansions (figure 3.25b) amounts to 95% of the input. Again tests with and without added noise were performed and the two inversions (figures

3.25c and d) show the noise to have the same effects demonstrated in the upper mantle tests. Amplitudes of recovered harmonics differ, but the redistribution of power, the effect relevant to our purpose here, is the same for each inversion. The residuals' variance is reduced by nearly 85% in the noise-free case, while the noisy data show only a 10% variance reduction. Figures 3.25c and d show good recovery of the $l = 2$ harmonic, though some power is smeared into the upper mantle layers and the lowermost layer, D", is very poorly imaged. Again it is surprising that almost no power is aliased to the $l = 1$ term even though the ray coverage shows a strong $l = 1$ component. Some power is aliased to the $l = 3$ term, though not very much. The $l = 7$ term aliases power in both directions in the power spectrum, with an apparently greater tendency to move power upward to $l = 8$. Recovery of the higher harmonics drops off with depth, simply as a result of the decreasing ratio of the small wavelengths of the higher harmonic patterns to depth. Still, the $l = 7$ harmonic is recoverable in the depth range 2470-2670 km but is indistinguishable from noise in D". This must be an effect of inadequate ray coverage rather than wavelength, in contrast to the $l = 12$ harmonic which drops into the noise at about 2270 km depth. As discussed earlier, this layer suffers from poor ray coverage because we restrict our epicentral distance range to less than 96°.

A second test for the lower mantle considers a pattern that appears at low power in model ISC5_LSQR. Harmonics $l = 1, 4, 8$, and 11 are introduced with the same procedure described for the first lower mantle test (figure 3.26a). Because of the fairly complete ray coverage of the lower mantle, 95% of the input power is contained in the re-expanded series of sampled voxels' contributions. Figure 3.26b shows the full recovery of power for all layers but the 670-870 km and the 870-1070 km layers. The inversion for noise-free data (figure 3.26c) returns 37% of the sampled voxels' input power to its spherical harmonic series. The inversion for noisy data (figure 3.26d) returns 52%. The inversion results for noise-free and noisy data show more aliasing of

power to other harmonics in the second lower mantle test than in the first lower mantle test (compare figures 3.26c,d to figures 3.25c,d). Though the pattern is clearly recognizable in each case, the spurious harmonics in the second test return with much greater power than do the spurious peaks in the first test. We infer from this result that the pattern that appears prominently in inversions of real data is probably retrieved reliably, but stands out with relatively high power because other patterns, such as the one introduced in our second lower mantle test, are not retrieved as reliably.

3.3.5 Lowermost Mantle

A last test considers the pattern of heterogeneity found in the lowermost mantle by *Dziewonski* [1984] and contained in the P-velocity model L02.56. We extract only the dominant harmonic components which correspond to our deepest two layers, 2470-2670 km and 2670-2898 km. Model L02.56 documents deviations from PREM, a different radial average than our starting model, and so our adoption of this pattern is not rigorously correct. However, these tests are intended as examples and the particular patterns tested and the amplitudes of these patterns are not critical to the effects we wish to investigate. Figures 3.27a and 3.27b show the input patterns to the 2470-2670 km and the 2670-2898 km layers, respectively. Figures 3.27c and 3.27d show the re-expansions with null ray coverage for the same layers. These re-expansions are nearly identical to the re-expansions with ideal coverage. While the ray sampling of the mantle's two deepest layers is nearly complete, in the sense that nearly every voxel is sampled, the quality of sampling in terms of strength of constraints placed on model values is quite poor. This is partly due to our restriction of rays to less than 96° and partly due to the fact that each ray that samples the lowermost mantle also samples many voxels in the remaining portions of the mantle. The two layers considered here comprise just 8% of the mantle's volume. Figures 3.3 and 3.4 document the uneven sampling for these two layers compared to the rest of the mantle. Figure 3.28a shows the input power contained in the $l = 2$ and $l = 6$ harmonics only, and the re-

expansions for ideal and null ray coverage. The re-expansions return 99% and 98% of the input power, respectively. The test inversion for noise-free data (figure 3.28b) returns an entirely inadequate image for the lowermost layer. Only 4.5% of the power input to sampled voxels is recovered. While the $l = 2$ component is recognizable above spurious signals, its power is fully two orders of magnitude below the input, and just a factor of three above the $l = 3$ ghost harmonic. The $l = 6$ component is not recognizable above the power of aliased harmonics. Recalling results for the $l = 7$ harmonic in the bottom layer of the first lower mantle test, we know that this result is due partly to the diminished resolvability of shorter wavelengths with depth. Another contributor is the poor ray sampling in this layer. The layer just above, 2470-2670 km, returns much more power to the $l = 2$ component, but this level is still an order of magnitude less than the input power. Again, the $l = 6$ component is not distinguishable from spurious peaks. Results for an inversion of noisy data (figure 3.28c) are similar to the noise-free results, although instead of drastically underestimating the input power the total power in the output model exceeds the input by 25%. Results of the noisy inversion for the bottom four layers are shown as recombined patterns in figure 3.29. Power in the southern hemisphere is particularly underestimated and the breadth and details of the input pattern are not recovered. Only general features of the $l = 2$ pattern are recognizable; details of the shorter-wavelength $l = 6$ pattern have been lost. Clearly, our inversion results for the lowermost mantle are not reliable.

3.4 Conclusions

We find that while the unique ray sampling of our top layer allows us to recover the power input to a particular pattern of heterogeneity, the pattern itself is poorly determined in this layer. This result confirms our intuitive expectation. Sparse sampling of the spherical harmonic pattern results in a poor fit of the series to the data. In contrast to the top layer, a pattern in the 200-400 km layer is more precisely determined, though the power contained in the pattern is consistently underestimated and

more leakage to the first and third layers occurs. The transition zone, 400-670 km, shows similarly strong control of lateral heterogeneity patterns but tests return a more accurate estimate of input power than for the second layer. Lateral control of heterogeneity patterns is nearly as strong as lateral control in the lower mantle. The l = 2, 4, and 6 components all are recovered accurately in the 400-670 km layer. This supports previous findings from inversions with real data that l = 2 is a significant pattern of heterogeneity in the mantle's transition zone, and that l = 4 is not a significant pattern. For the upper mantle, l = 6 appears to be a dominant, constructive pattern. Similarly, the clear recovery of an l = 1 pattern and absence of aliasing to this pattern from other harmonics confirms the absence of an l = 1 component in the 0-200 km and 200-400 km layers, and the reliability of a strong l = 1 component in the 400-670 km layer of model ISC5_LSQR.

The l = 1 component in the lower mantle is also recovered reliably. Model ISC5_LSQR contains a large l = 1 component through the transition zone, changes sign at about 1100 km depth, increases in amplitude through the mid-mantle, then decreases in amplitude and again changes sign at about 1900 km depth. An examination of l = 1 coefficients for each layer of our lower mantle test does not reveal the change of sign which appears in model ISC5_LSQR. Nor does such a change of sign occur for any other harmonic. The ISC5_LSQR result apparently cannot be explained as an artifact of our inversion procedure or ray coverage. These tests also demonstrate the inability of our inversion scheme to retrieve shorter-wavelength features in the lower mantle. Results of inversions with real data for our lowermost layer, D", must be considered suspect due to the inadequate constraints placed on model values by our ray coverage.

The approach to evaluating the resolvability of mantle features presented here is specific to large-scale, coherent patterns of heterogeneity. Short wavelength features such as rift and subduction zones are considered, in a general sense, because their

large velocity contrasts contribute significant power to the long wavelength patterns of mantle heterogeneity. However, there will be additional effects due to scattering of short wavelength seismic energy which will diminish our capacity to image these features by means of body-wave, travel time tomography. *Snieder et al.* [1991] conduct a detailed study of the effects these small-scale features have on the splitting of normal modes and the consequences of scattering for global inversions performed with normal mode data. In addition, complications due to poor data quality will further degrade our inversion results. Nevertheless, our approach can help to evaluate the usefulness of global tomographic results as modeling constraints in geodynamics, to confirm the patterns produced by inversions of real data, and to identify patterns that must be viewed with skepticism.

3.5 References

Dziewonski, A.M., Mapping the lower mantle, Determination of lateral heterogeneity in P velocity up to degree and order 6, *J. Geophys. Res., 89,* 5929–5952, 1984.

Humphreys, E., and R.W. Clayton, Adaptation of back projection tomography to seismic travel time problems, *J. Geophys. Res., 93,* 1073–1085, 1988.

Humphreys, E., and R.W. Clayton, Adaptation of back projection tomography to seismic travel time problems, *J. Geophys. Res., 93,* 1073–1085, 1988.

Inoue, H., Y. Fukao, K. Tanabe, and Y. Ogata, Whole mantle P–wave travel time tomography. *Phys. Earth and Planet. Int., 59,* 294–328, 1990.

Jeffreys, H., *The Earth,* Cambridge University Press, London, 1960.

Masters, G., T.H. Jordan, P.G. Silver, and F. Gilbert, Aspherical Earth structure from fundamental spheroidal-mode data, *Nature, 298* 609–613, 1982.

Nolet, G., and R. Snieder, Solving large linear inverse problems by projection, *Geophys. Jour. Int., 103,* 565–568, 1990.

Pulliam, R.J., D.W. Vasco, and L.R. Johnson, A Simultaneous, Tomographic Inversion of ISC Travel Time Residuals for Mantle P-velocity, Source Mislocations, and

Station Corrections, submitted to *J. Geophys. Res.*, March 1991.

Romanowicz, B., G. Roult, and T. Kohl, The upper mantle degree two pattern: constraints from GEOSCOPE fundamental spheroidal mode eigenfrequency and attenuation measurements, *Geophy. Res. Lett., 14,* 1219–1222, 1987.

Scales, J. A., Tomographic inversion via the conjugate gradient method. *Geophysics, 52,* 179–185, 1987.

Snieder, R., J. Beckers, and F. Neele, The effect of small-scale structure on normal mode frequencies and global inversions, *J. Geophys. Res., 96,* 501–515, 1991.

Spakman, W., and G. Nolet, Imaging algorithms, accuracy and resolution in delay time tomography, in *Mathematical Geophysics,* N.J. Vlaar, G. Nolet, M.J.R. Wortel, and S.A.P.L. Cloetingh (eds.), Reidel, Dordrecht, pp. 155–187, 1987.

Vasco, D. W., Bounding seismic velocities using a tomographic method, *Geophysics, 56,* 472–482, 1991.

Vasco, D.W., R.J. Pulliam, and L.R. Johnson, Tomographic inversion of ISC travel times for mantle P wave velocity structure using an l^1 norm criterion, submitted to *J. Geophys. Res.*, November 1990.

3.6 List of Tables

Table 3.1 Details of the model parametrization and the sampling provided by our data set. Include are the average number of hits for sampled voxels in each layer along with the the average sum of ray segments in a voxel at a given depth and the number of voxels sampled in each depth interval. These averages include only voxels which have non-zero sampling.

3.7 List of Figures

Figure 3.1 Flow chart outlining the steps performed in the series of experiments.

Figure 3.2 The model mantle is parametrized as approximately equal-area voxels, 5° × 5° at the equator and generally 200 km thick, for a total of 22,876 voxels.

Figure 3.3 Power in the spherical harmonic expansions of ray samples ("hits") of voxels for each depth interval as a function of angular degree. All values are normalized to the maximum value appearing in the figure. Numbers on the right refer to the maximum power for each layer.

Figure 3.4 Power in the spherical harmonic layer expansions of the column sums of our coefficient matrix, **A**, indicating the total sampling of each voxel by our data set.

Figure 3.5 Power in the spherical harmonic expansions of model ISC5_LSQR for each depth interval as a function of angular degree.

Figure 3.6 Power in spherical harmonic series generated by averaging ISC5_LSQR layer expansions through the whole mantle and through the upper and lower mantle separately.

Figure 3.7 a) The spherical harmonic $l = 3$, $m = 0$ pattern introduced to the 0-200 km layer of our model. The amplitude of the A_3^0 coefficient is 0.25 km/sec. (b) The pattern produced after the input series is recombined onto our model

grid and re-expanded using values from all voxels ("ideal" ray coverage). (c) The re-expansion using only values for voxels which are sampled by our data set ("null" ray coverage).

Figure 3.8 The results of an inversion performed with a set of synthetic travel time residuals produced by estimating the delays caused by the introduced $l = 3$ velocity anomaly are shown. The top three layers are shown sequentially from top to bottom: 0-200 km, 200-400 km, and 400-670 km.

Figure 3.9 a) The power contained in the spherical harmonic series presented in figure 3.7 on Mollweide projections is presented here in more detail. At the top is the power of the spherical harmonic $l = 3$ pattern introduced to our model's top layer. In the middle is the power contained in the spherical harmonic series produced after the input series is recombined and re-expanded with ideal ray coverage. At bottom is the re-expansion with null ray coverage. b) Power in the spherical harmonic expansions for selected layers of a model produced by inverting synthetic data generated with our input heterogeneity pattern.

Figure 3.10 a) The $l = 10$ power introduced to the 0-200 km layer (top), the re-expansion with ideal ray coverage (middle), and the re-expansion using only sampled voxels (bottom). b) Power contained in the spherical harmonic layer expansions of our layer 1, $l = 10$ test inversion results.

Figure 3.11 a) The $l = 2$ power introduced to the 200-400 km layer (top), the re-expansion with ideal ray coverage (middle), and the re-expansion using only sampled voxels (bottom). b) Power contained in the spherical harmonic layer expansions of our layer 2, $l = 2$ test inversion results.

Figure 3.12 a) The $l = 4$ power introduced to the 200-400 km layer (top), the re-expansion with ideal ray coverage (middle), and the re-expansion using only

sampled voxels (bottom). b) Power contained in the spherical harmonic layer expansions of our layer 2, $l = 4$ test inversion results.

Figure 3.13 a) The $l = 2$ power introduced to the 400-670 km layer (top), the re-expansion with ideal ray coverage (middle), and the re-expansion using only sampled voxels (bottom). b) Power contained in the spherical harmonic layer expansions of our layer 3, $l = 2$ test inversion results.

Figure 3.14 a) The $l = 4$ power introduced to the 400-670 km layer (top), the re-expansion with ideal ray coverage (middle), and the re-expansion using only sampled voxels (bottom). b) Power contained in the spherical harmonic layer expansions of our layer 3, $l = 4$ test inversion results.

Figure 3.15 a) The $l = 6$ power introduced to the 400-670 km layer (top), the re-expansion with ideal ray coverage (middle), and the re-expansion using only sampled voxels (bottom). b) Power contained in the spherical harmonic layer expansions of our layer 3, $l = 6$ test inversion results.

Figure 3.16 a) The $l = 12$ power introduced to the 400-670 km layer (top), the re-expansion with ideal ray coverage (middle), and the re-expansion using only sampled voxels (bottom). b) Power contained in the spherical harmonic layer expansions of our layer 3, $l = 12$ test inversion results.

Figure 3.17 Recombined Mollweide projections of (a) the higher-order degree 2 pattern of Masters et al. (1982) input to our experiment, (b) the re-expansion with ideal coverage, and the null coverage provided by (c) the 200-400 km depth layer and (d) the 400-670 km layer.

Figure 3.18 Amplitudes of individual series coefficients in the layer expansions of the higher-order $l = 2$ test pattern. Coefficients are grouped in pairs: $[A_l^m, B_l^m]$. Recall that the B_l^0 coefficient is always zero.

Figure 3.19 Recombined Mollweide projections of the output from a test inversion in which the higher-order l=2 pattern is introduced to the 200–400 km layer. Results shown are for the depth intervals (a) 0-200 km, (b) 200-400 km, and (c) 400-670 km. Note the change of scale from the input patterns (figure 3.17).

Figure 3.20 Series coefficients in the top five layers' expansions of the inversion results for the test in which the pattern is introduced to the 200-400 km layer.

Figure 3.21 Recombined Mollweide projections of the output from a test inversion in which the higher-order l=2 pattern is introduced to the 400–670 km layer. Results shown are for the depth intervals (a) 200-400 km, (b) 400-670 km, and (c) 670-870 km.

Figure 3.22 Series coefficients in the top five layers' expansions of the inversion results for the test in which the pattern is introduced to the 400-670 km layer. The output anomaly is more completely confined to the anomalous input layer than is the output of the 200-400 km test (figure 3.20).

Figure 3.23 a) At the top is the power of the spherical harmonic pattern input for our first upper mantle test to all three upper mantle layers of our model. Next is the power contained in the spherical harmonic series after the input series is recombined onto our model grid and re-expanded using values from all voxels ("ideal" ray coverage). Third is the average of the re-expanded series for each layer using only values for voxels which are sampled by our data set ("null" ray coverage). Fourth and fifth are the series expansions averaged over all three layers of the test inversions using synthetic data without noise and with noise, respectively. b) Power in the re-expansions for all three layers individually using ideal ray coverage (top) and null coverage (bottom). c) Power contained in the layer series expansions of our test inversion using noise-free synthetic data. d) Power contained in the layer series expansions of our test inversion using synthetic data with noise which follows a two-sided exponential

distribution.

Figure 3.24 Figure captions are the same as for figure 3.23 for the second upper mantle test.

Figure 3.25 a) At the top is the power of the spherical harmonic pattern input for our first lower mantle test to all 11 lower mantle layers of our model. Second is the power contained in the spherical harmonic series after the input series is recombined onto our model grid and re-expanded with ideal ray coverage. Third is the average of the re-expanded series for each layer using null ray coverage. Fourth and fifth are the series expansions averaged over all 11 layers of the test inversions using synthetic data without noise and with noise, respectively. b) Power in the re-expansions for all 11 layers individually using null coverage (bottom). c) Power contained in the layer series expansions of our test inversion using noise-free synthetic data. d) Power contained in the layer series expansions of our test inversion using synthetic data with added noise which follows a two-sided exponential distribution.

Figure 3.26 Figure captions are the same as for figure 3.25 for the second lower mantle test.

Figure 3.27 The pattern extracted from model L02.56 (Dziewonski, 1984) input to our experiment for (a) the 2470-2670 km layer, (b) the 2670-2898 km interval, and the re-expansions with null ray coverage for the same two layers: (c) 2470-2670 km, (d) 2670-2898 km.

Figure 3.28 a) The top two traces show the power of the spherical harmonic pattern input to the bottom two layers of our model. The middle two traces show the power contained in the respective layer expansions after the input series is recombined onto our model grid and re-expanded with ideal ray coverage. The bottom two traces indicate the power contained in the respective layer

expansions using null ray coverage. b) Power contained in the layer series expansions of our test inversion using noise-free synthetic data. c) Power contained in the layer series expansions of our test inversion using synthetic data with added noise which follows a two-sided exponential distribution.

Figure 3.29 The results of our test inversion with a pattern extracted from L02.56 (Dziewonski, 1984) and introduced to our two lowermost layers. Noise following a two-sided exponential distribution has been added to the synthetic travel time residuals. The bottom four layers are shown: (a) 2070-2270 km, (b) 2270-2470 km, (c) 2470-2670 km, and (d) 2670-2898 km.

Table 3.1: Model layer depths and average voxel sampling

Layer number	Depth range (km)	Average velocity (km/s)	Average voxel volume (10^7 km^3)	Number of voxels sampled	% of voxels sampled	Average number hits	Average column sum of A (10^4 km)
1	0-200	7.94	6.06	1195	73.1	1586	31.02
2	200-400	8.58	5.67	1333	81.6	1342	25.97
3	400-670	9.88	7.07	1494	91.4	1409	34.34
4	670-870	10.95	4.82	1577	96.5	1261	27.38
5	870-1070	11.40	4.49	1604	98.2	1149	25.21
6	1070-1270	11.68	4.16	1627	99.6	991	21.29
7	1270-1470	11.96	3.85	1631	99.8	881	18.38
8	1470-1670	12.24	3.55	1632	99.9	816	16.96
9	1670-1870	12.51	3.26	1634	100.0	753	15.62
10	1870-2070	12.76	2.98	1634	100.0	675	13.79
11	2070-2270	13.01	2.72	1634	100.0	604	12.45
12	2270-2470	13.24	2.46	1634	100.0	494	10.04
13	2470-2670	13.47	2.22	1634	100.0	395	8.11
14	2670-2898	13.64	2.26	1626	99.5	179	3.68

Figure 3.1

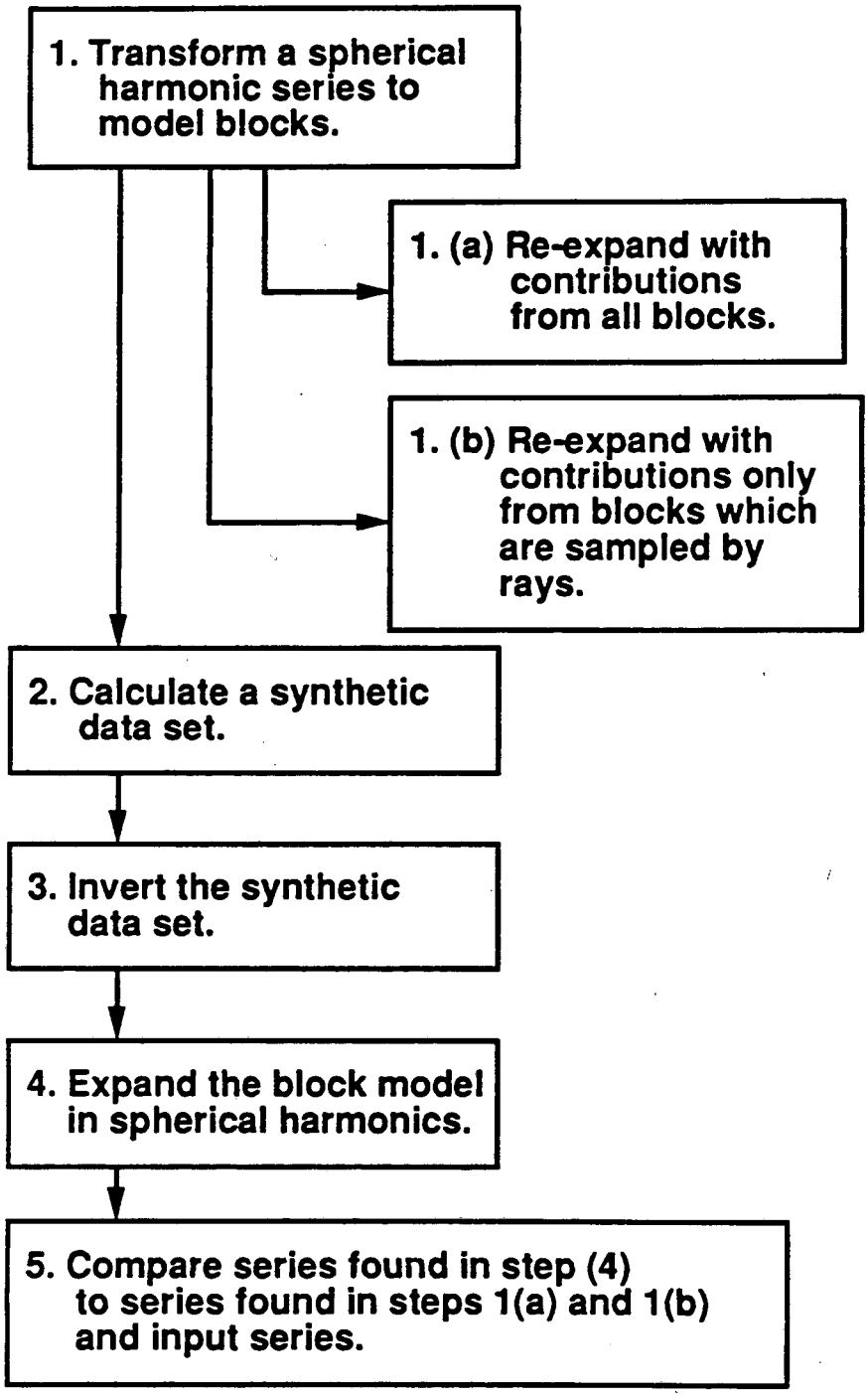
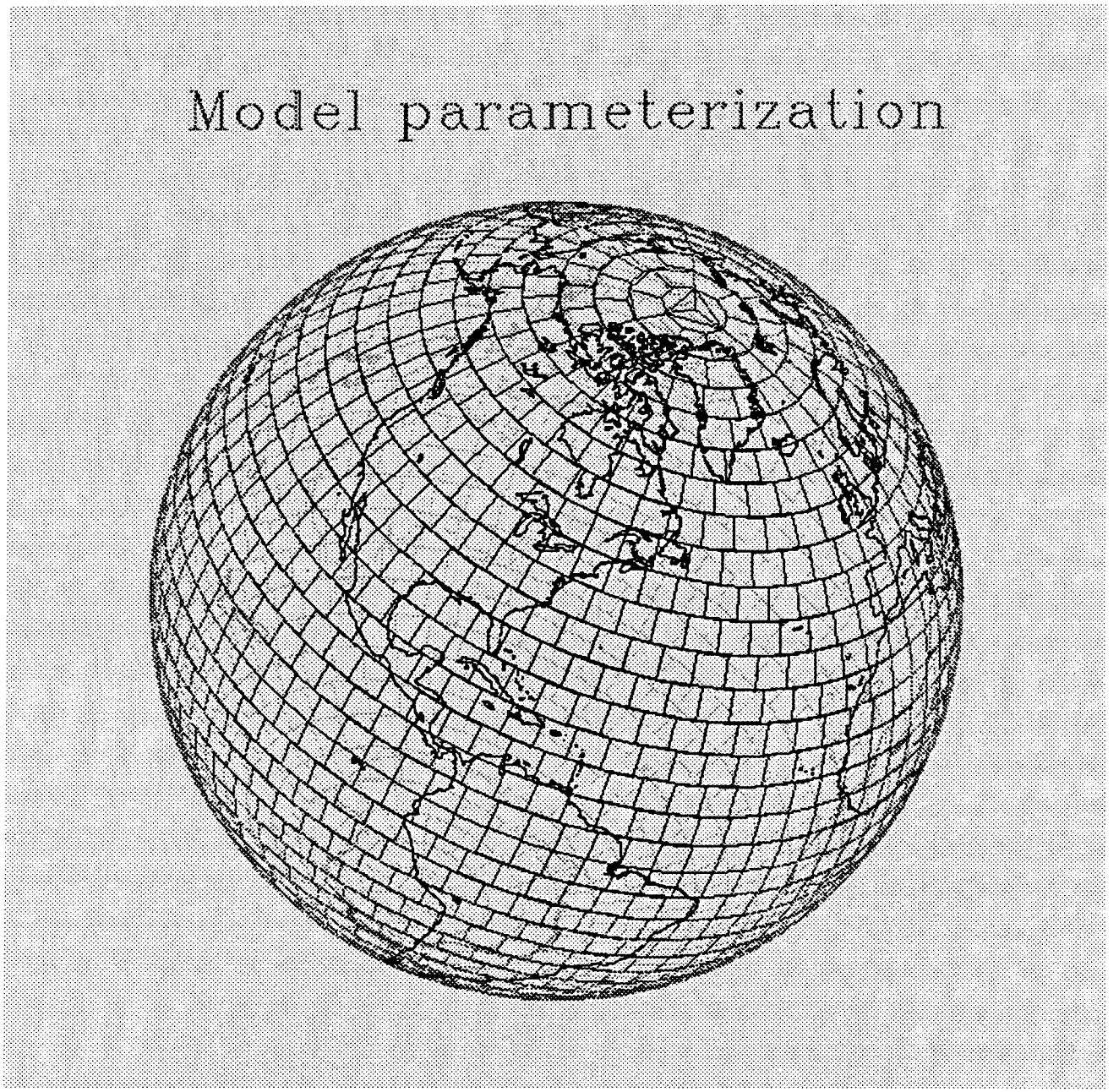

1. Transform a spherical harmonic series to model blocks.
1. (a) Re-expand with contributions from all blocks.
1. (b) Re-expand with contributions only from blocks which are sampled by rays.
2. Calculate a synthetic data set.
3. Invert the synthetic data set.
4. Expand the block model in spherical harmonics.
5. Compare series found in step (4) to series found in steps 1(a) and 1(b) and input series.

Figure 3.2



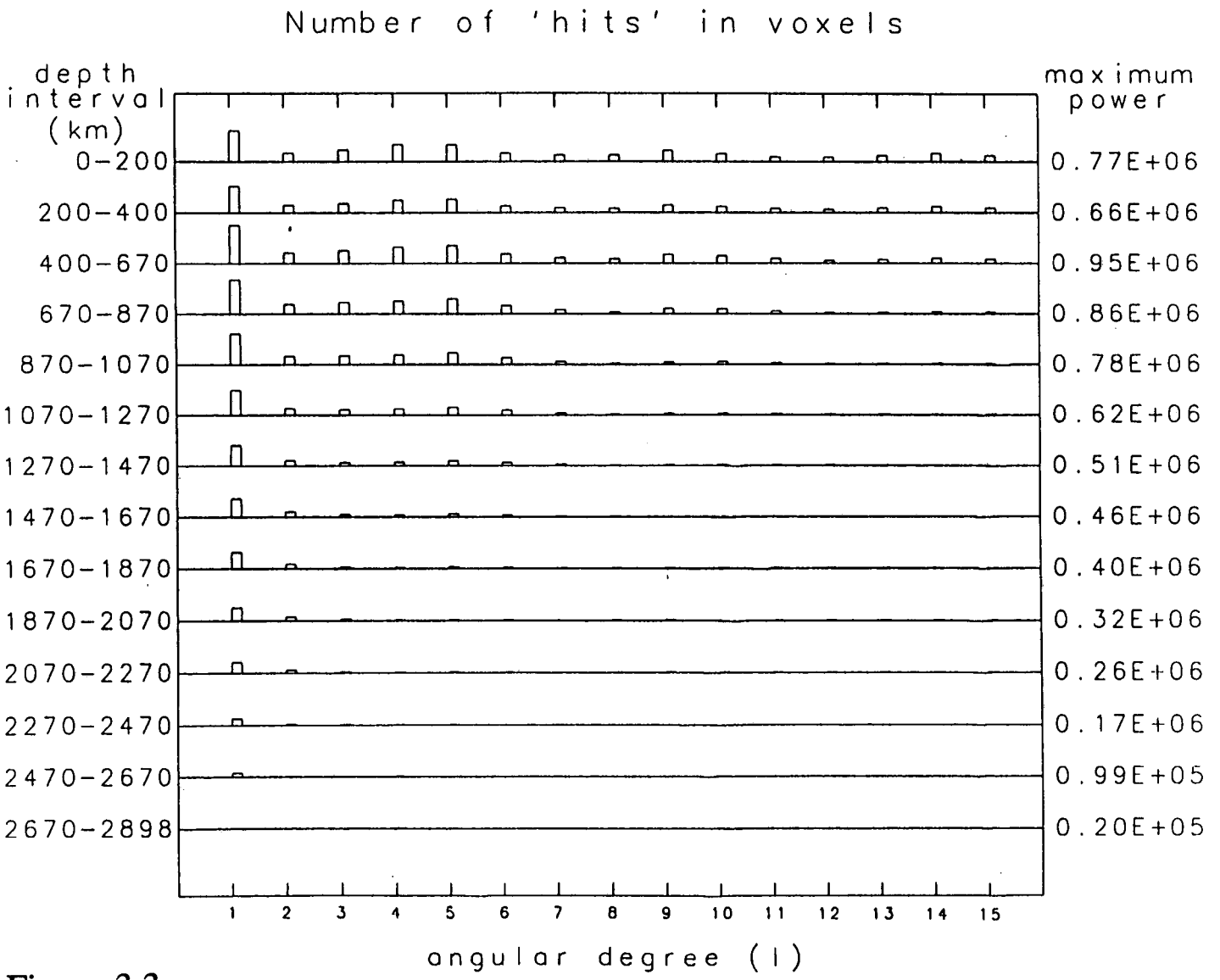

Number of 'hits' in voxels
depth interval (km)
maximum power
0-200
0.77E+06
200-400
0.66E+06
400-670
0.95E+06
670-870
0.86E+06
870-1070
0.78E+06
1070-1270
0.62E+06
1270-1470
0.51E+06
1470-1670
0.46E+06
1670-1870
0.40E+06
1870-2070
0.32E+06
2070-2270
0.26E+06
2270-2470
0.17E+06
2470-2670
0.99E+05
2670-2898
0.20E+05
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.3

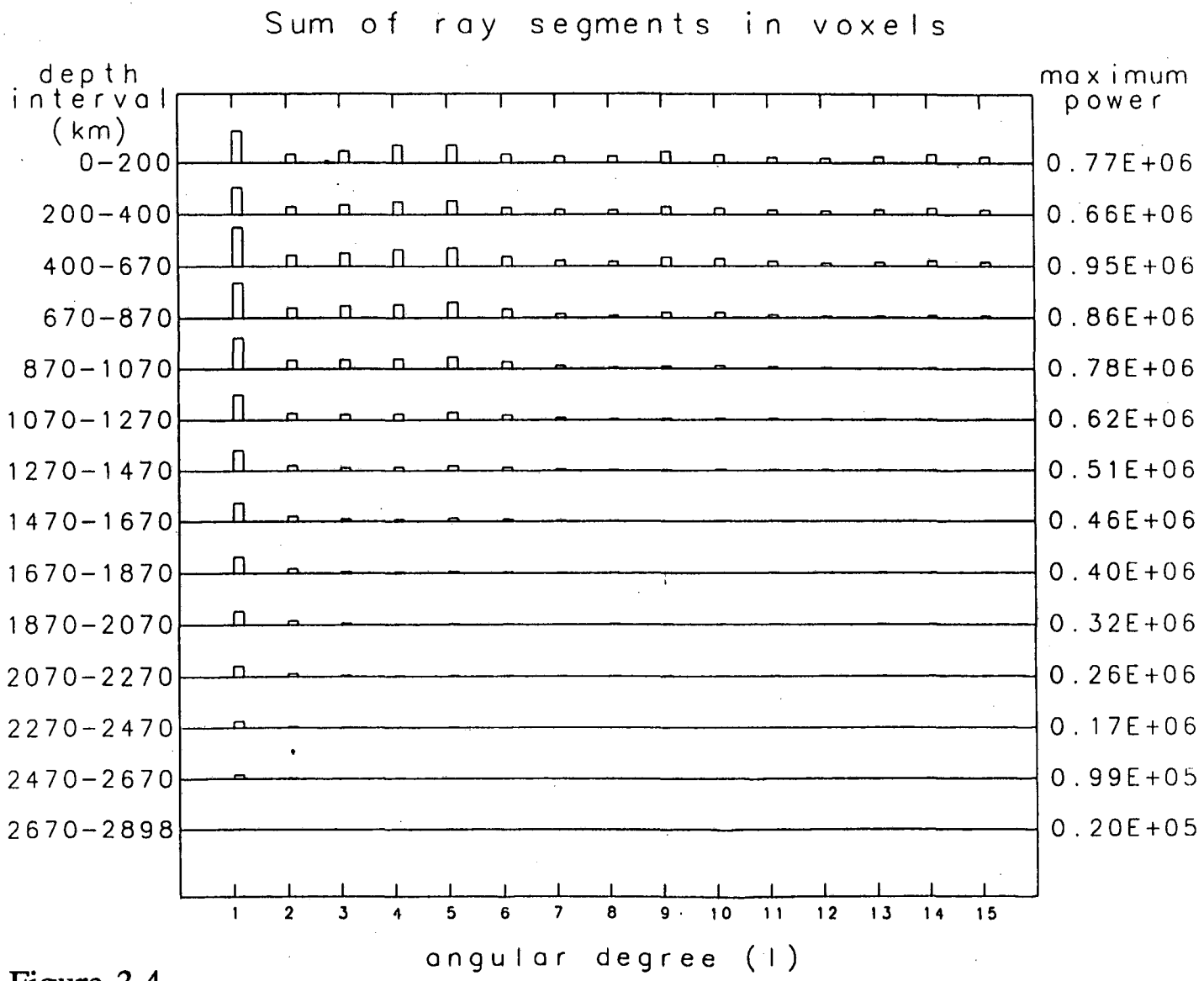

Sum of ray segments in voxels
depth interval (km)
maximum power
0-200
0.77E+06
200-400
0.66E+06
400-670
0.95E+06
670-870
0.86E+06
870-1070
0.78E+06
1070-1270
0.62E+06
1270-1470
0.51E+06
1470-1670
0.46E+06
1670-1870
0.40E+06
1870-2070
0.32E+06
2070-2270
0.26E+06
2270-2470
0.17E+06
2470-2670
0.99E+05
2670-2898
0.20E+05
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.4

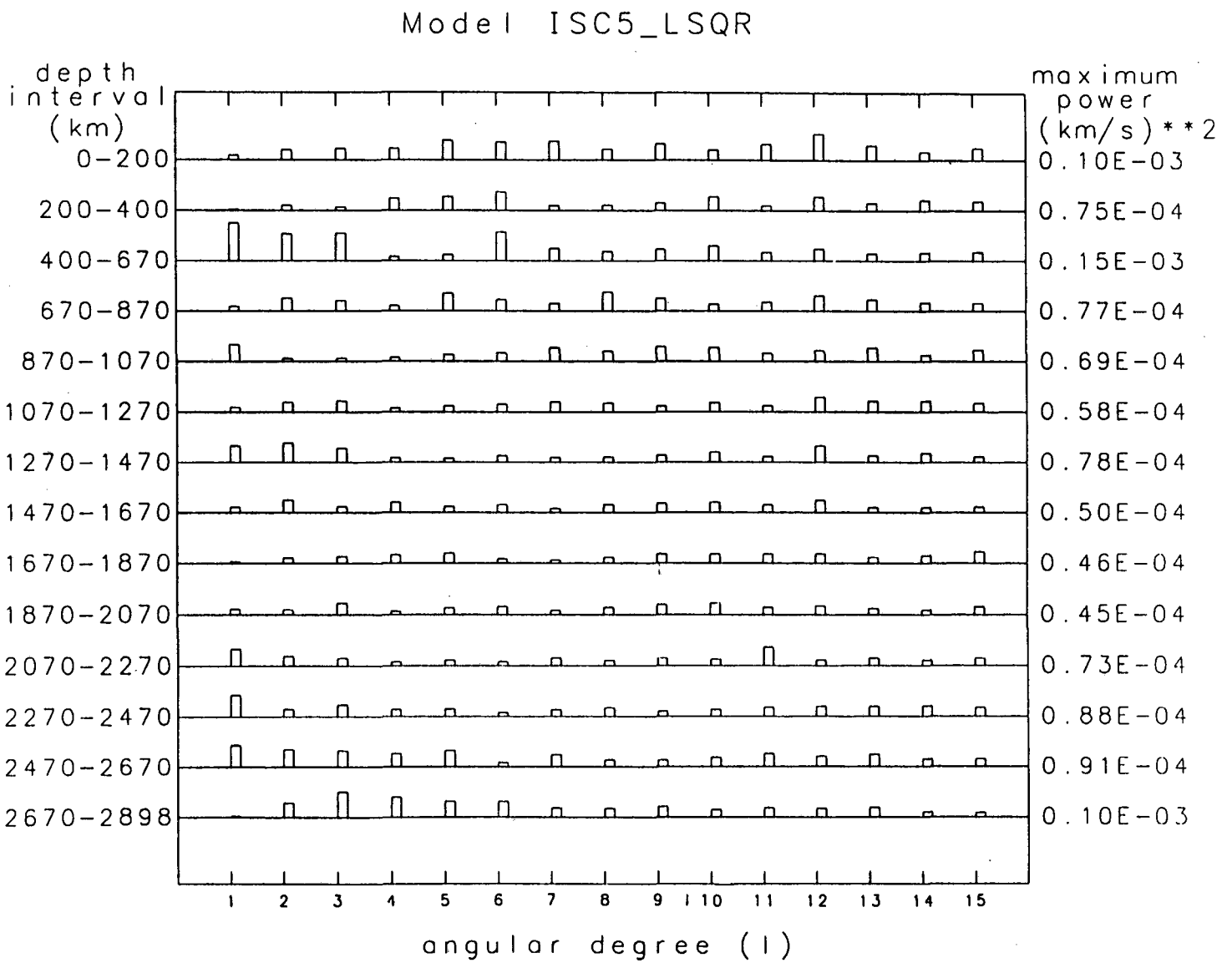

Model ISC5_LSQR
depth interval (km)
maximum power (km/s)**2
0-200
200-400
400-670
670-870
870-1070
1070-1270
1270-1470
1470-1670
1670-1870
1870-2070
2070-2270
2270-2470
2470-2670
2670-2898
0.10E-03
0.75E-04
0.15E-03
0.77E-04
0.69E-04
0.58E-04
0.78E-04
0.50E-04
0.46E-04
0.45E-04
0.73E-04
0.88E-04
0.91E-04
0.10E-03
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.5

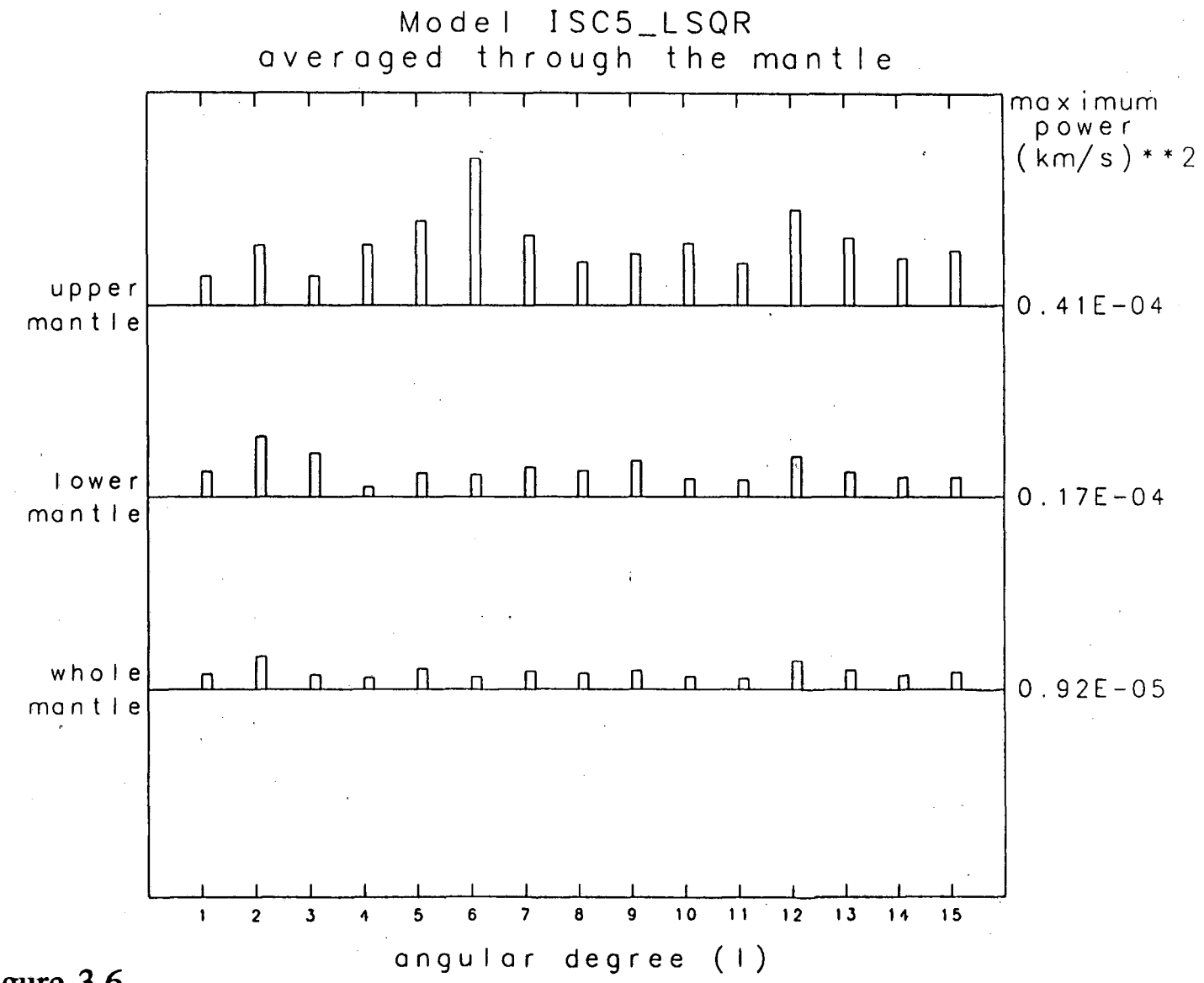

Model ISC5_LSQR
averaged through the mantle
maximum power (km/s)**2
upper mantle
0.41E-04
lower mantle
0.17E-04
whole mantle
0.92E-05
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.6

Figure 3.7

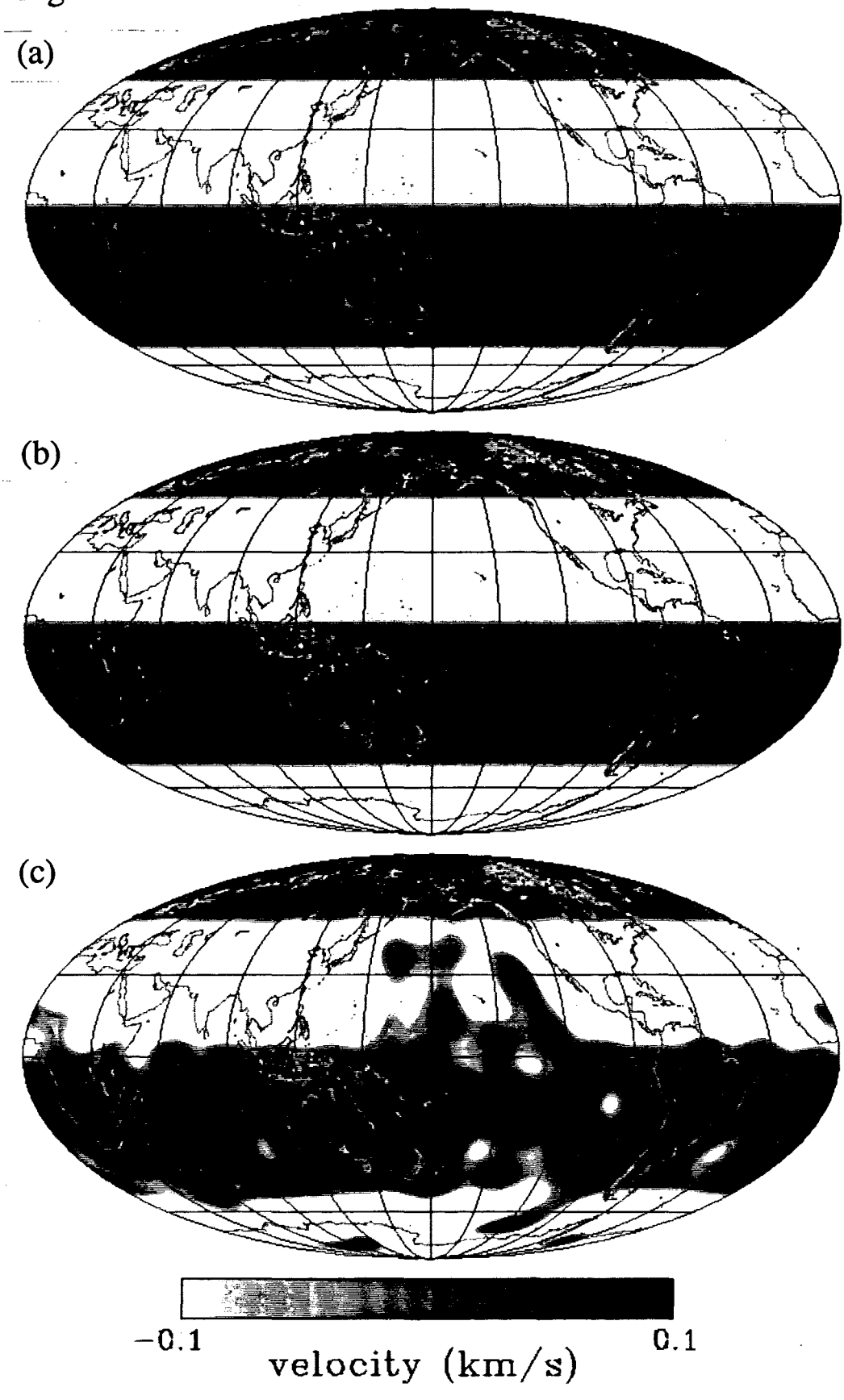
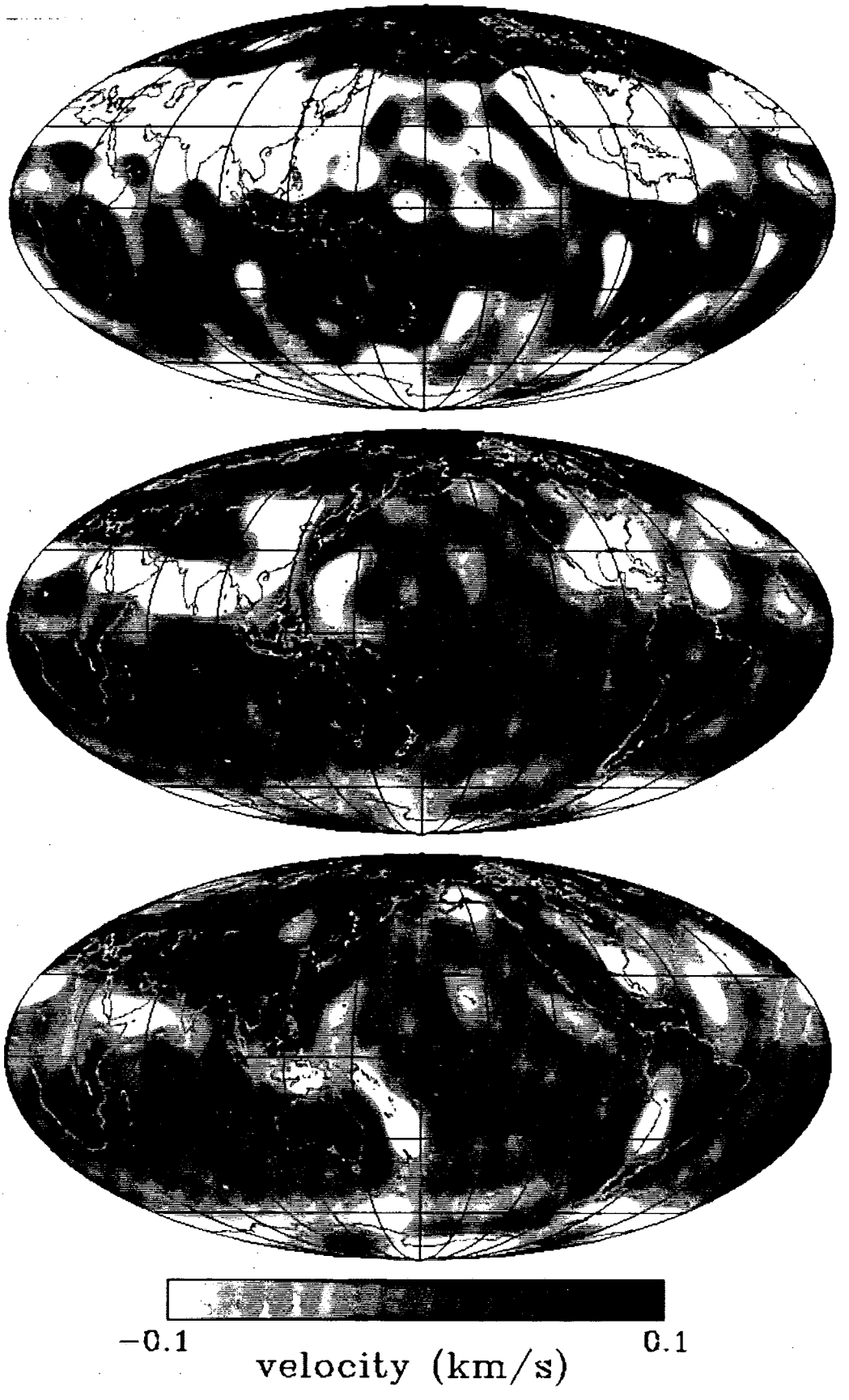


Figure 3.8



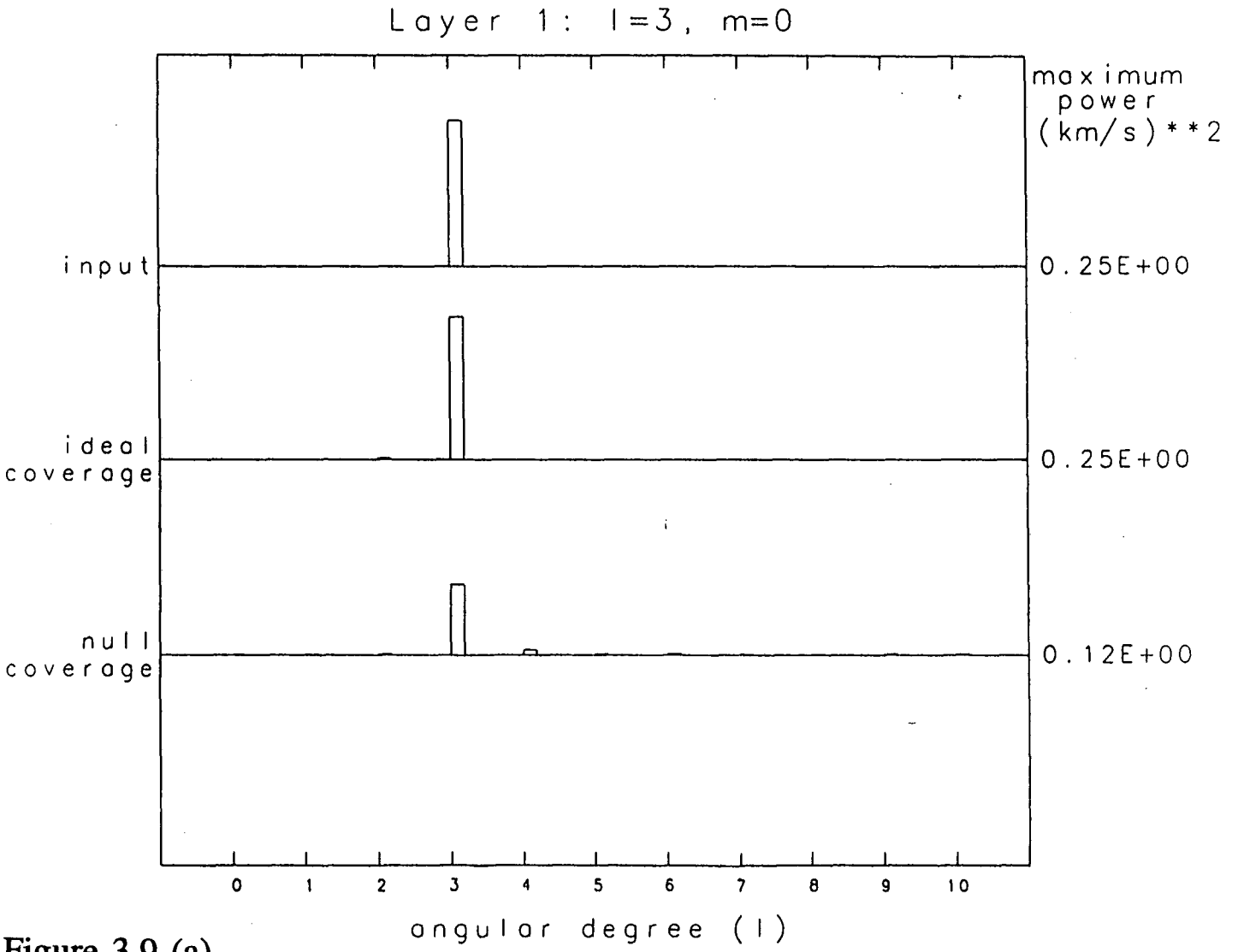

Layer 1: l=3, m=0
maximum power (km/s)**2
input
0.25E+00
ideal coverage
0.25E+00
null coverage
0.12E+00
0
1
2
3
4
5
6
7
8
9
10
angular degree (l)

Figure 3.9 (a)

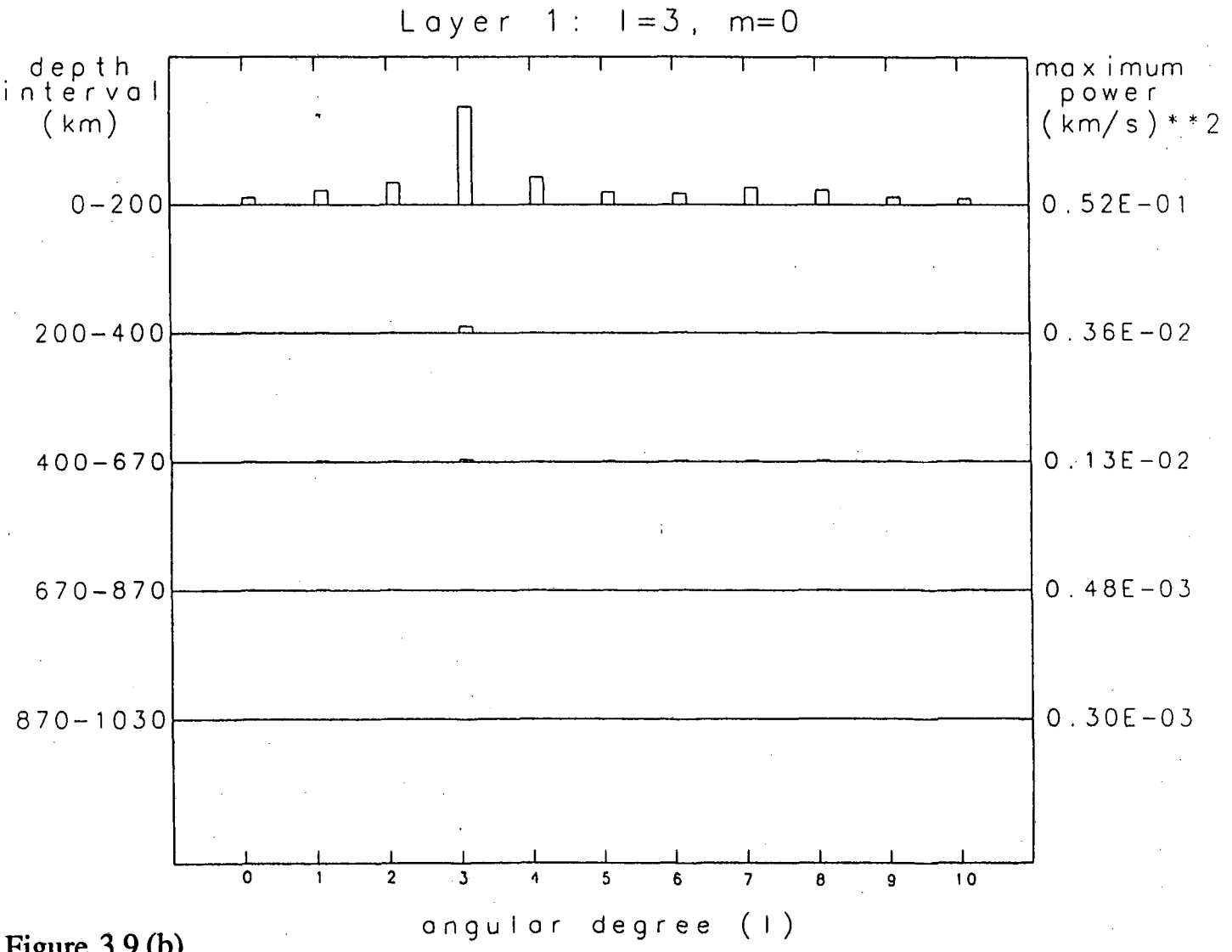

Layer 1: l=3, m=0
depth interval (km)
maximum power (km/s)**2
0-200
200-400
400-670
670-870
870-1030
0.52E-01
0.36E-02
0.13E-02
0.48E-03
0.30E-03
0
1
2
3
4
5
6
7
8
9
10
angular degree (l)

Figure 3.9 (b)

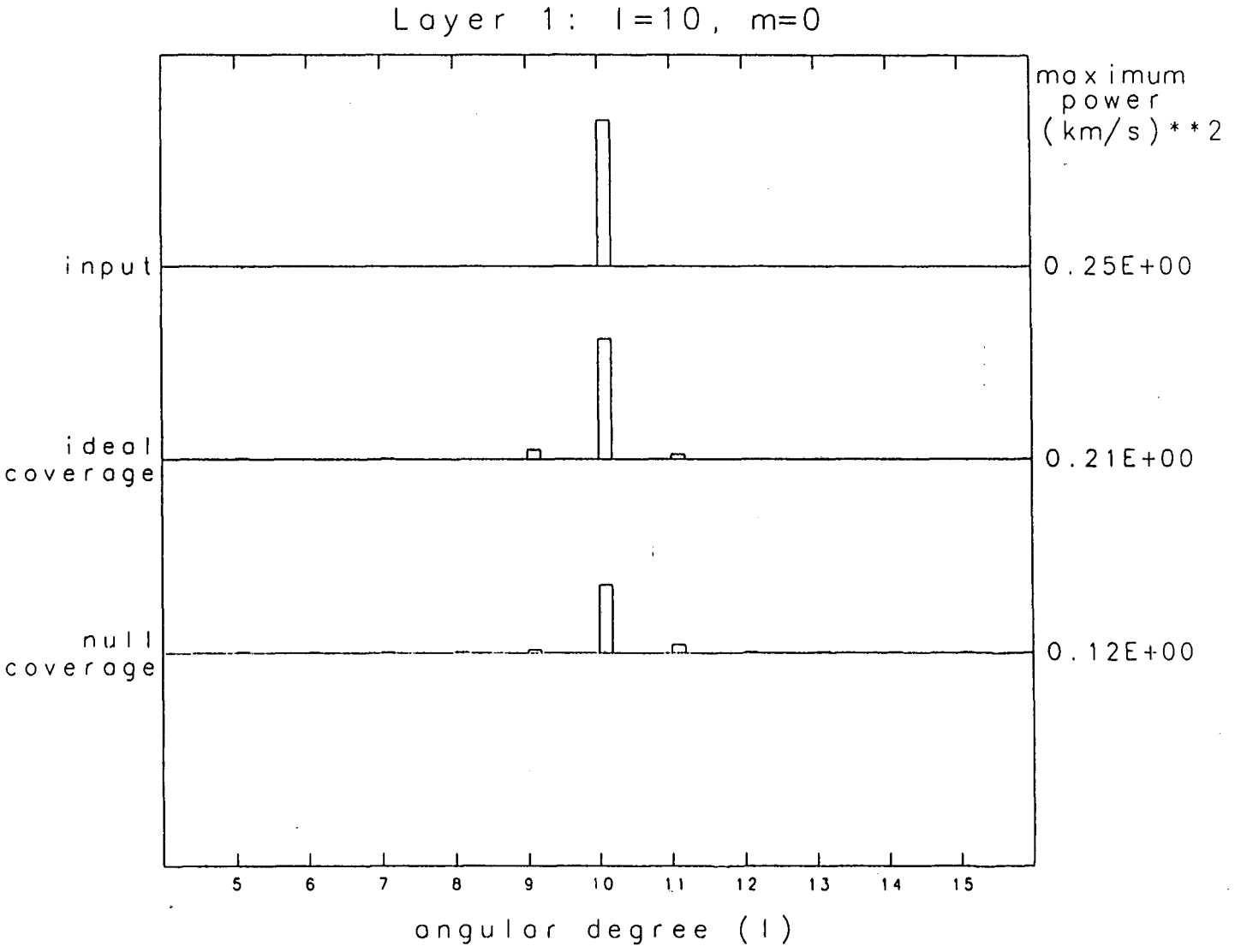

Layer 1: l=10, m=0
maximum power (km/s)**2
input
0.25E+00
ideal coverage
0.21E+00
null coverage
0.12E+00
5
6
7
8
9
10
11
12
13
14
15
angular degree (l)

Figure 3.10 (a)

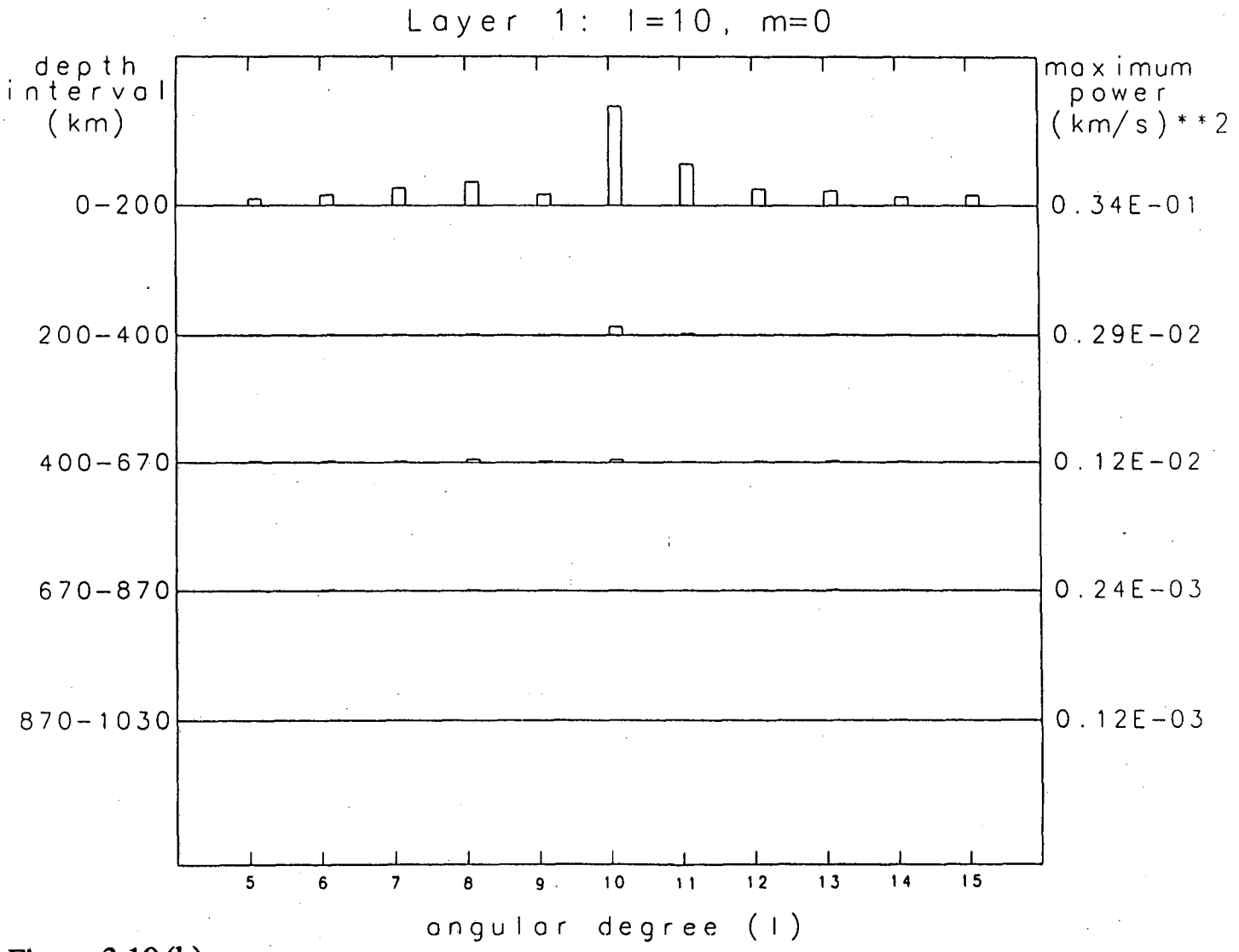

Layer 1: l=10, m=0
depth interval (km)
maximum power (km/s)**2
0-200
0.34E-01
200-400
0.29E-02
400-670
0.12E-02
670-870
0.24E-03
870-1030
0.12E-03
5
6
7
8
9
10
11
12
13
14
15
angular degree (l)

Figure 3.10 (b)

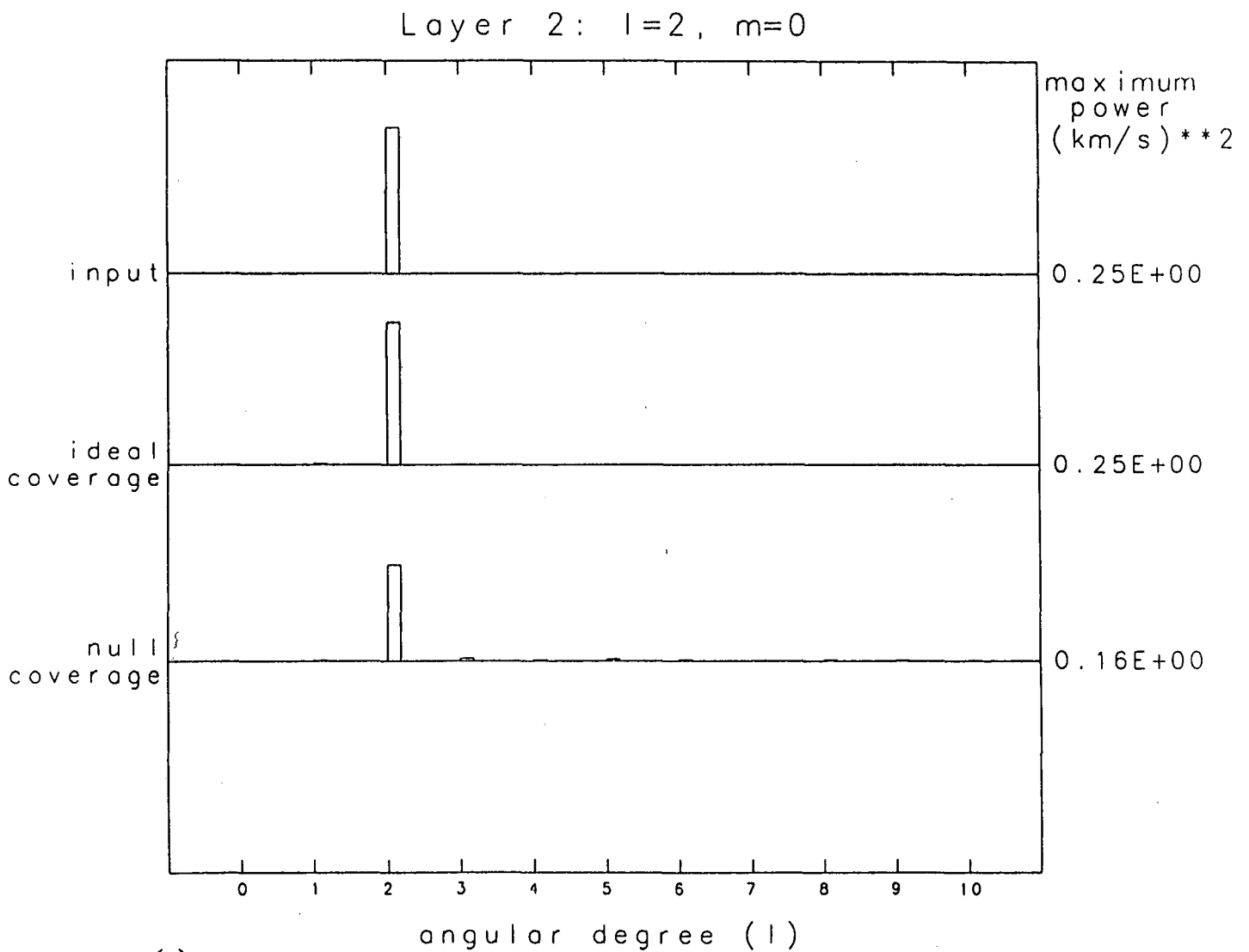

Layer 2: l=2, m=0
maximum power (km/s)**2
input
0.25E+00
ideal coverage
0.25E+00
null coverage
0.16E+00
0
1
2
3
4
5
6
7
8
9
10
angular degree (l)

Figure 3.11 (a)

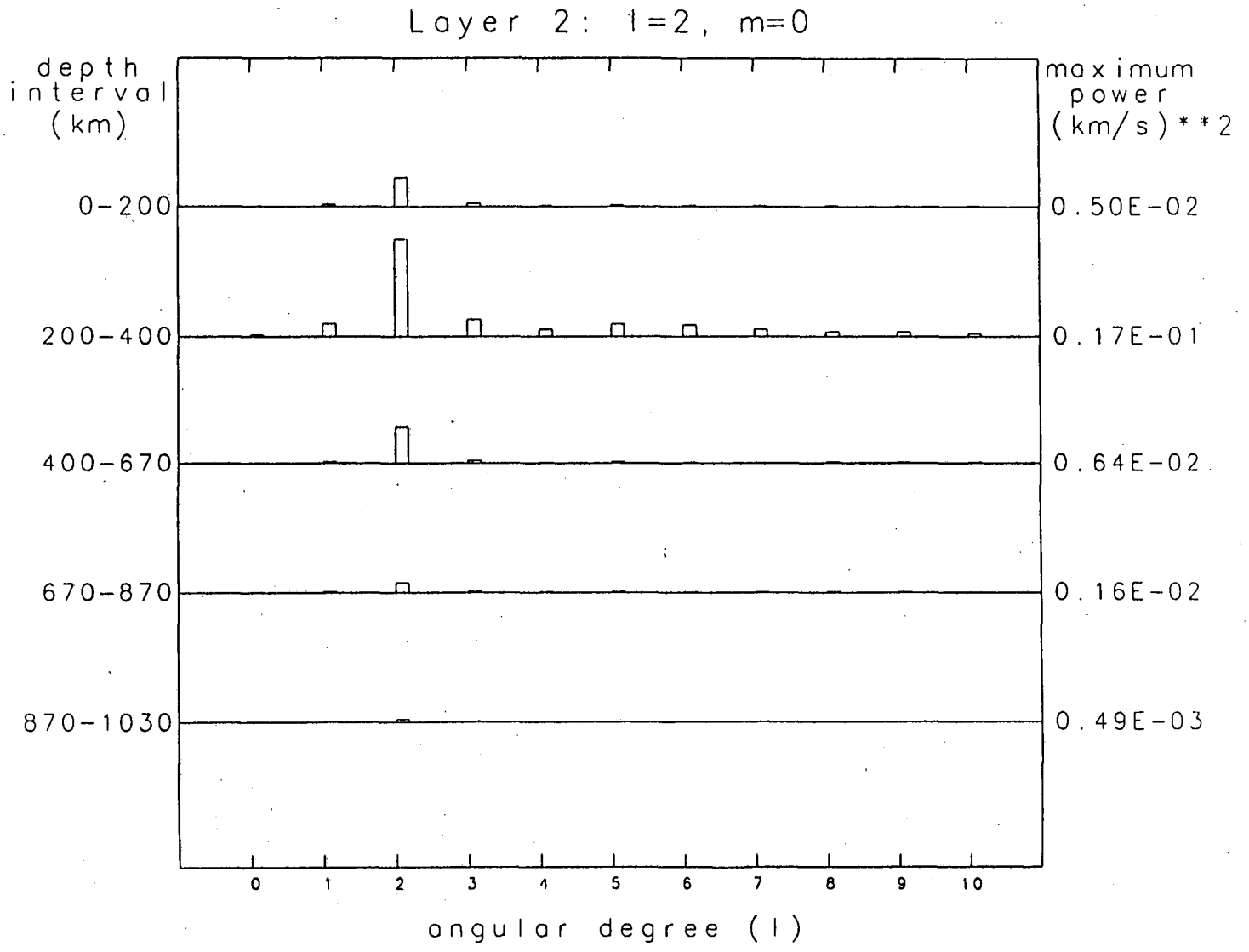

Layer 2: l=2, m=0
depth interval (km)
maximum power (km/s)**2
0-200
0.50E-02
200-400
0.17E-01
400-670
0.64E-02
670-870
0.16E-02
870-1030
0.49E-03
0
1
2
3
4
5
6
7
8
9
10
angular degree (l)

Figure 3.11 (b)

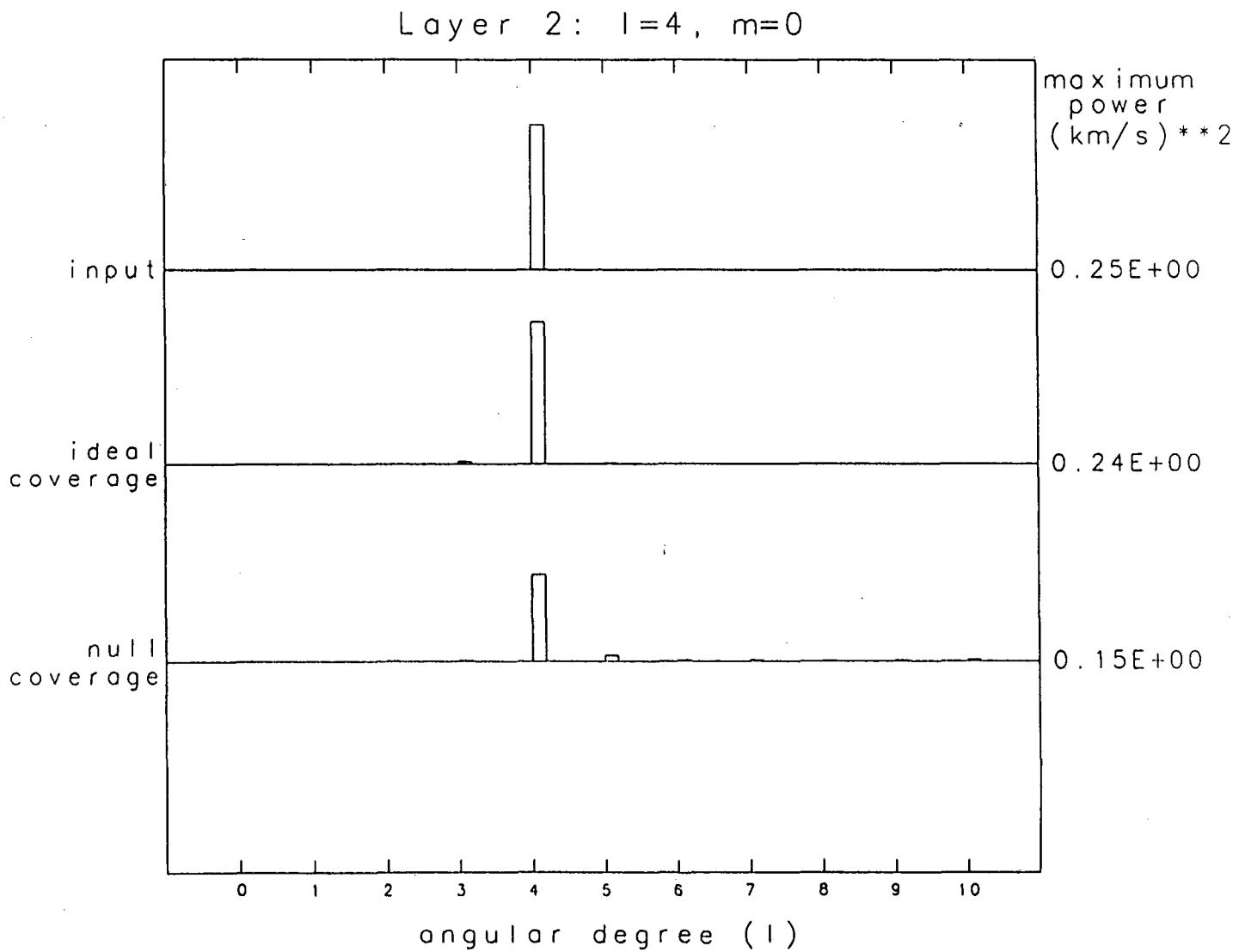

Layer 2: l=4, m=0
maximum power (km/s)**2
input
0.25E+00
ideal coverage
0.24E+00
null coverage
0.15E+00
0
1
2
3
4
5
6
7
8
9
10
angular degree (l)

Figure 3.12 (a)

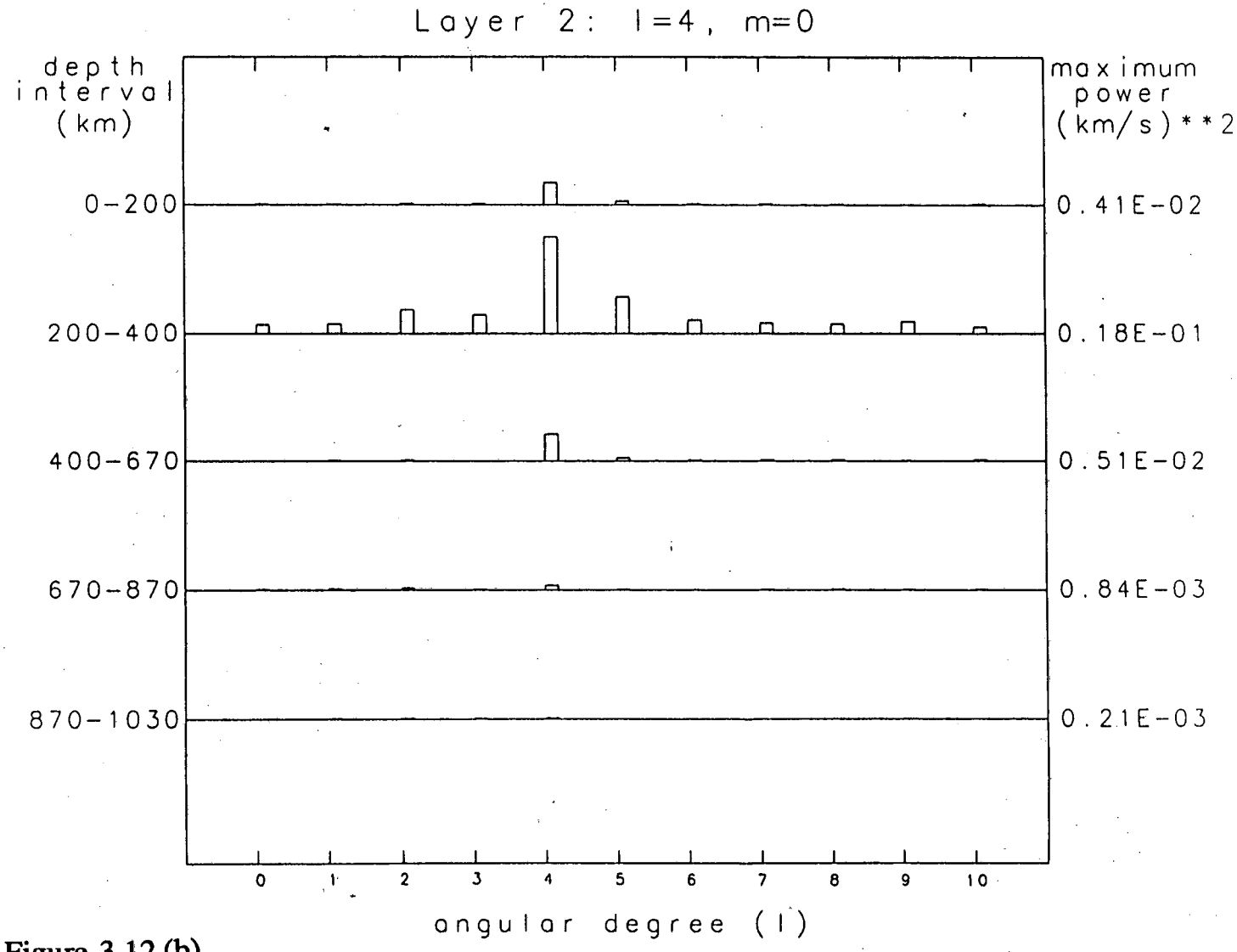

Layer 2: l=4, m=0
depth interval (km)
maximum power (km/s)**2
0-200
0.41E-02
200-400
0.18E-01
400-670
0.51E-02
670-870
0.84E-03
870-1030
0.21E-03
0
1
2
3
4
5
6
7
8
9
10
angular degree (l)

Figure 3.12 (b)

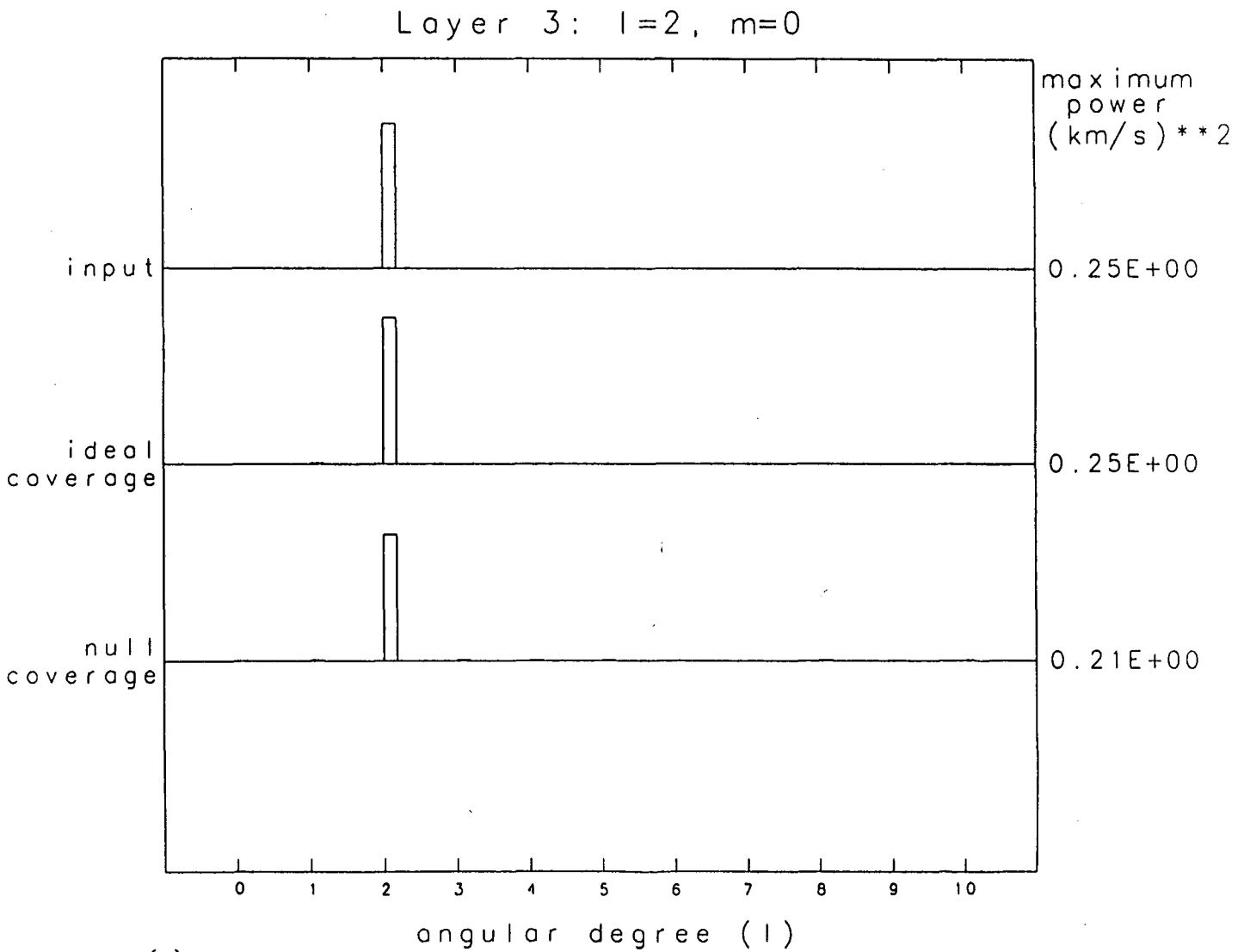

Layer 3: l=2, m=0
maximum power (km/s)**2
input
0.25E+00
ideal coverage
0.25E+00
null coverage
0.21E+00
0
1
2
3
4
5
6
7
8
9
10
angular degree (l)

Figure 3.13 (a)

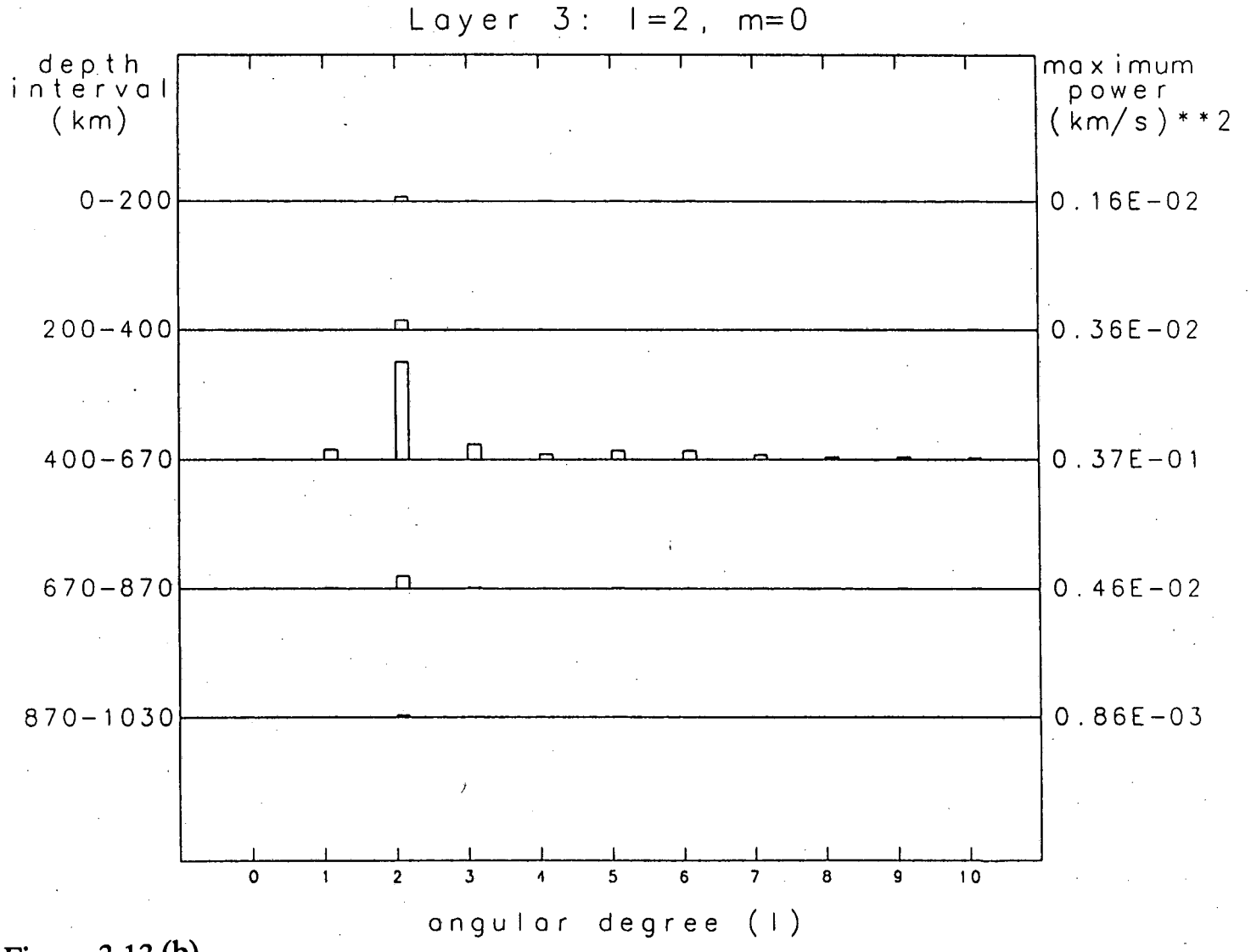

Layer 3: l=2, m=0
depth interval (km)
maximum power (km/s)**2
0-200
200-400
400-670
670-870
870-1030
0.16E-02
0.36E-02
0.37E-01
0.46E-02
0.86E-03
0
1
2
3
4
5
6
7
8
9
10
angular degree (l)

Figure 3.13 (b)

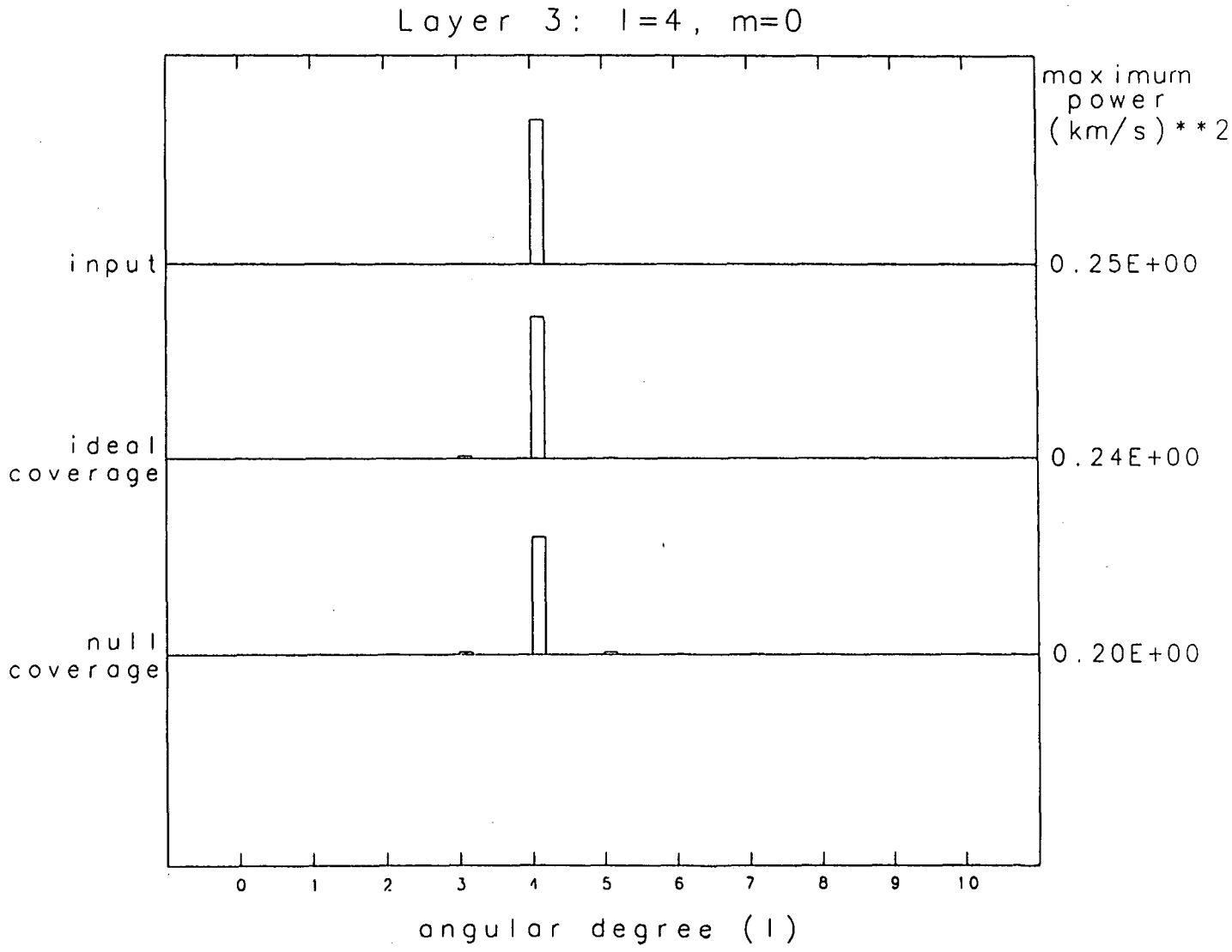

Layer 3: l=4, m=0
maximum power (km/s)**2
input
0.25E+00
ideal coverage
0.24E+00
null coverage
0.20E+00
0
1
2
3
4
5
6
7
8
9
10
angular degree (l)

Figure 3.14 (a)

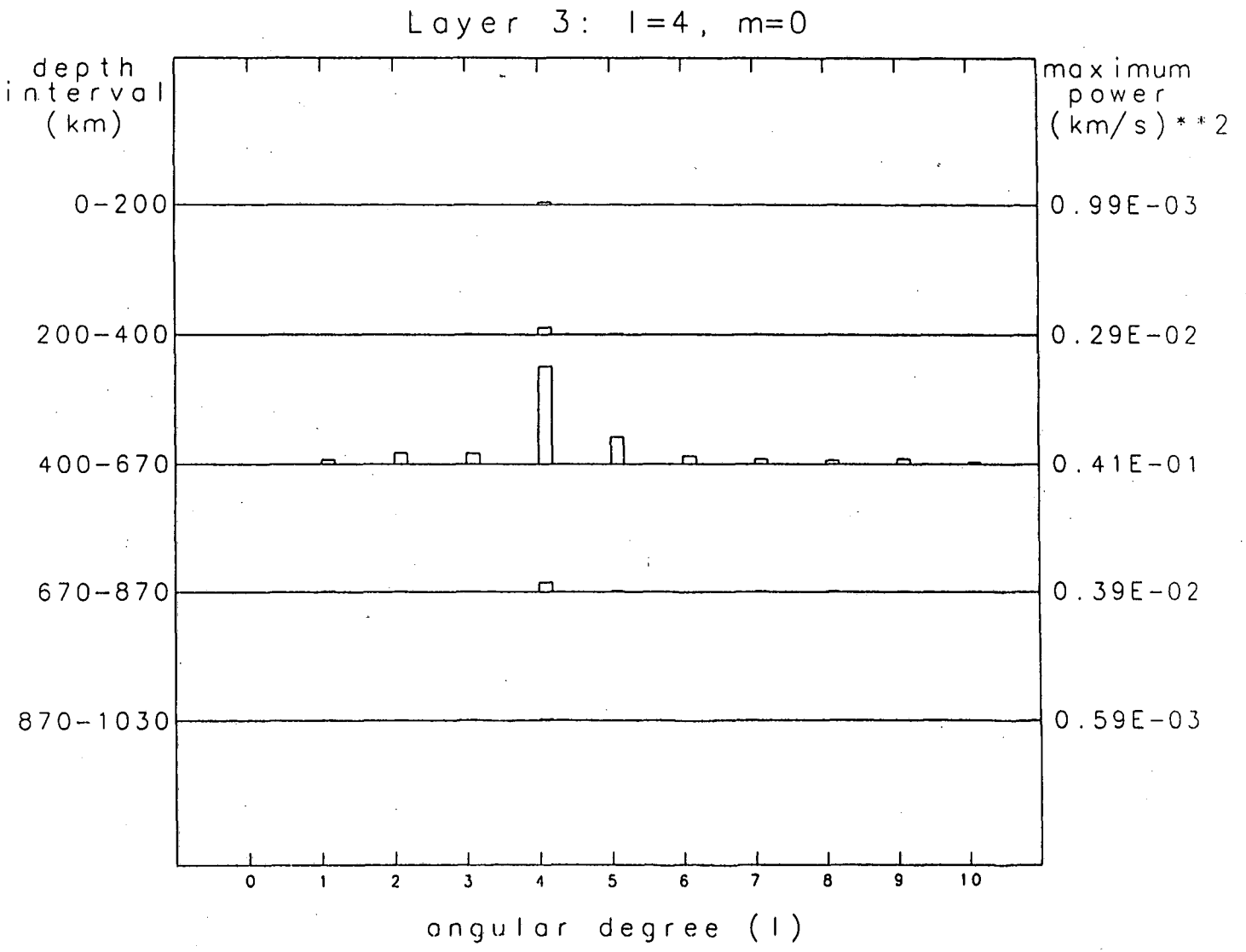

Layer 3: l=4, m=0
depth interval (km)
maximum power (km/s)**2
0-200
0.99E-03
200-400
0.29E-02
400-670
0.41E-01
670-870
0.39E-02
870-1030
0.59E-03
0
1
2
3
4
5
6
7
8
9
10
angular degree (l)

Figure 3.14 (b)

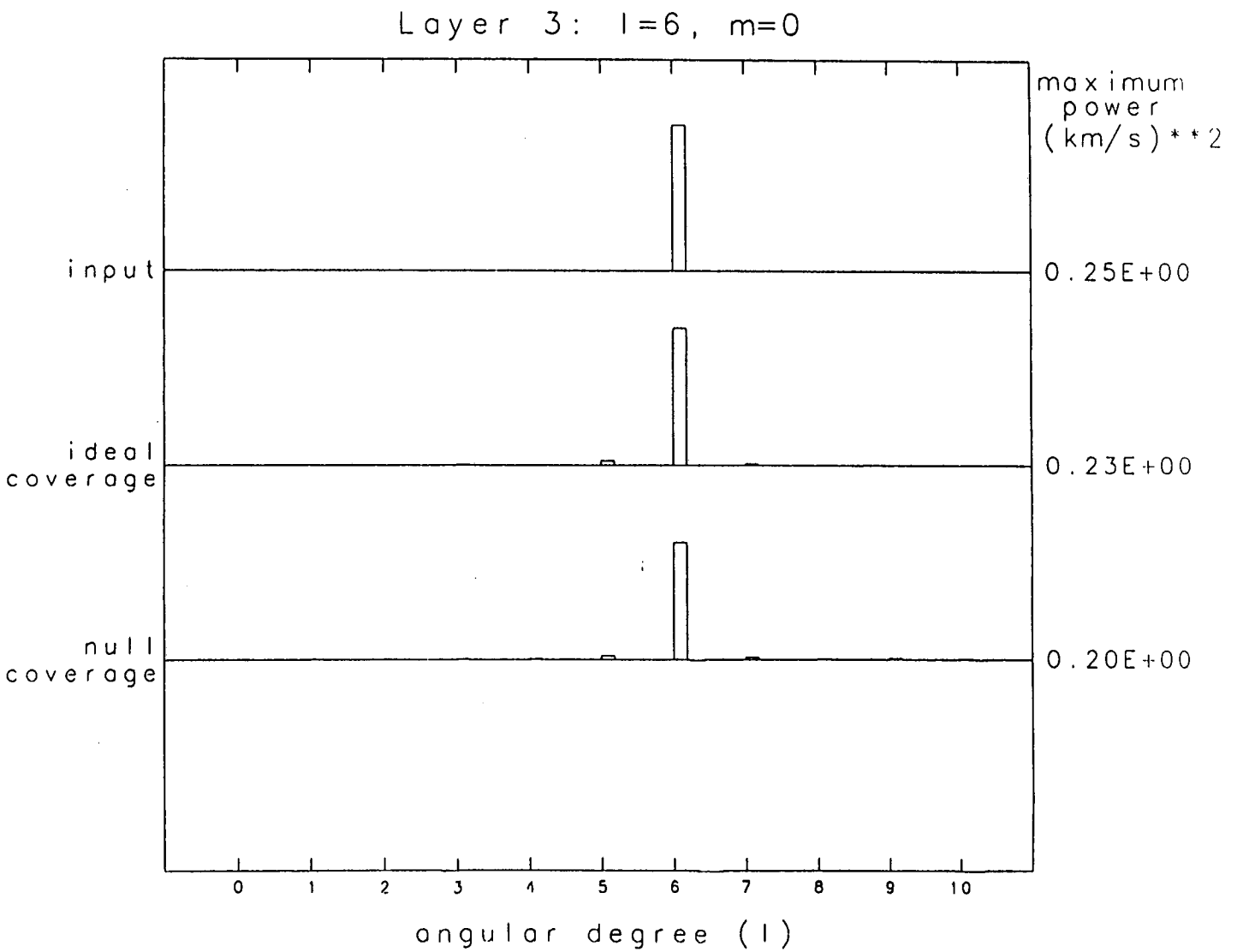

Layer 3: l=6, m=0
maximum power (km/s)**2
input
0.25E+00
ideal coverage
0.23E+00
null coverage
0.20E+00
0
1
2
3
4
5
6
7
8
9
10
angular degree (l)

Figure 3.15(a)

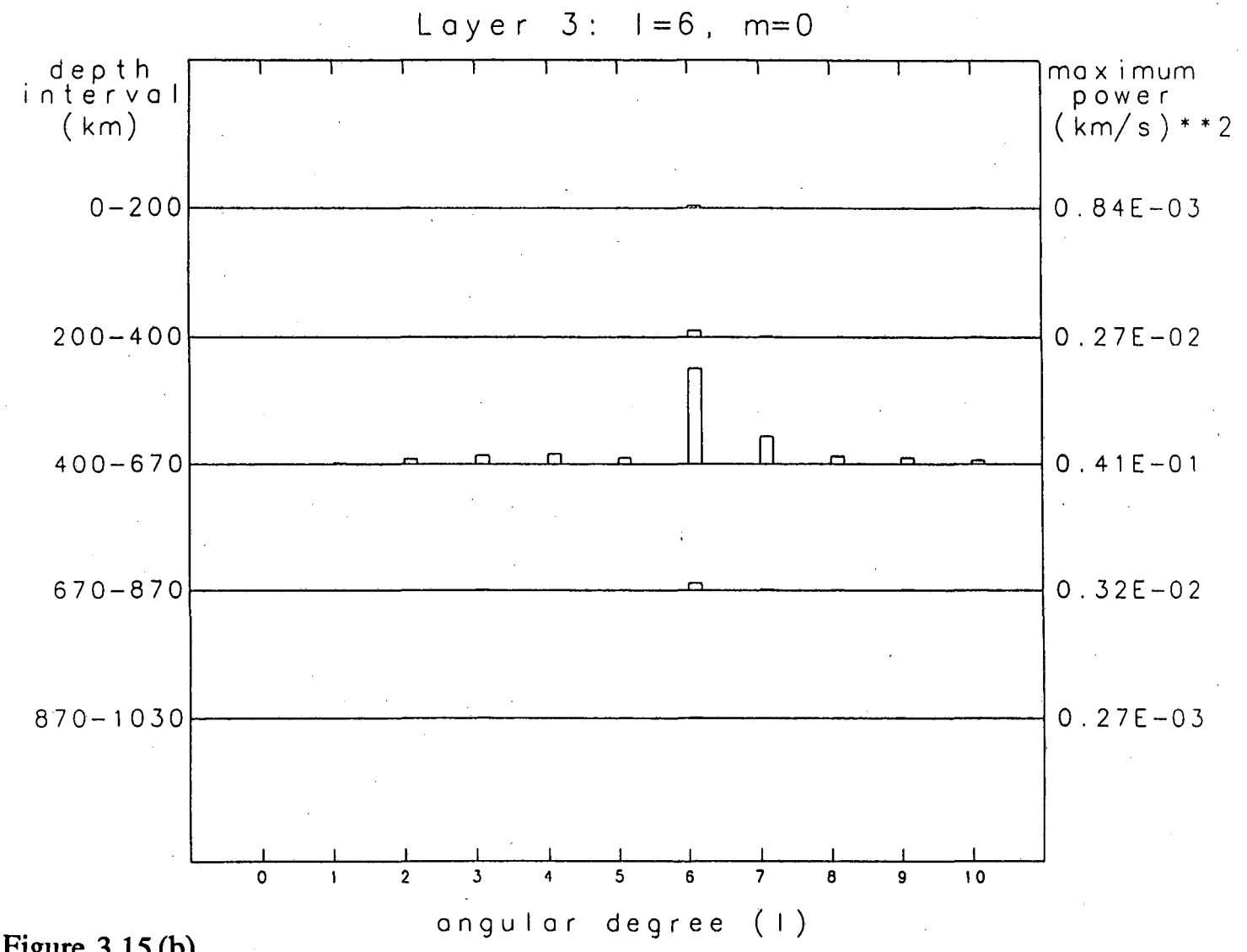

Layer 3: l=6, m=0
depth interval (km)
maximum power (km/s)**2
0-200
0.84E-03
200-400
0.27E-02
400-670
0.41E-01
670-870
0.32E-02
870-1030
0.27E-03
0
1
2
3
4
5
6
7
8
9
10
angular degree (l)

Figure 3.15 (b)

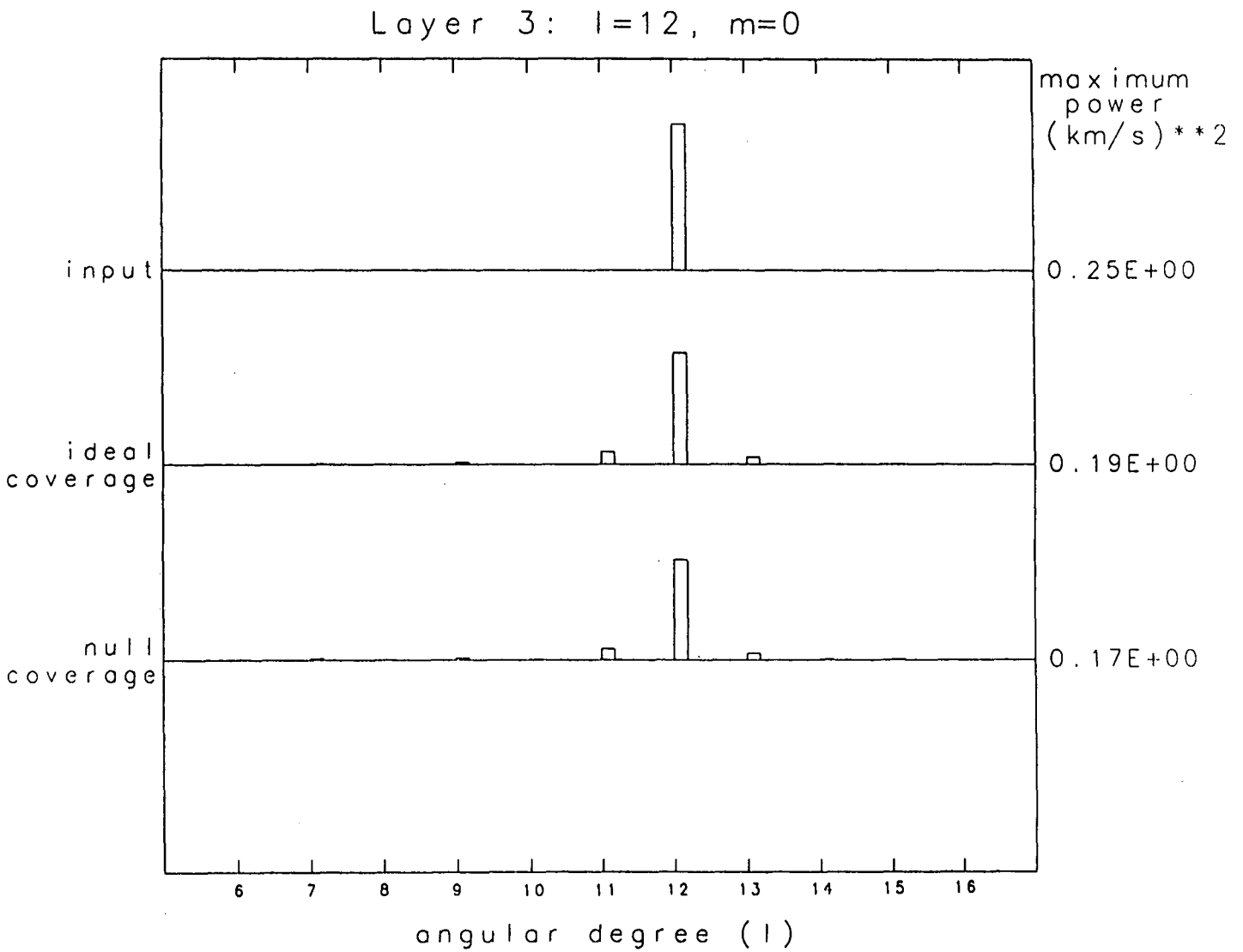

Layer 3: l=12, m=0
maximum power (km/s)**2
input
0.25E+00
ideal coverage
0.19E+00
null coverage
0.17E+00
6
7
8
9
10
11
12
13
14
15
16
angular degree (l)

Figure 3.16 (a)

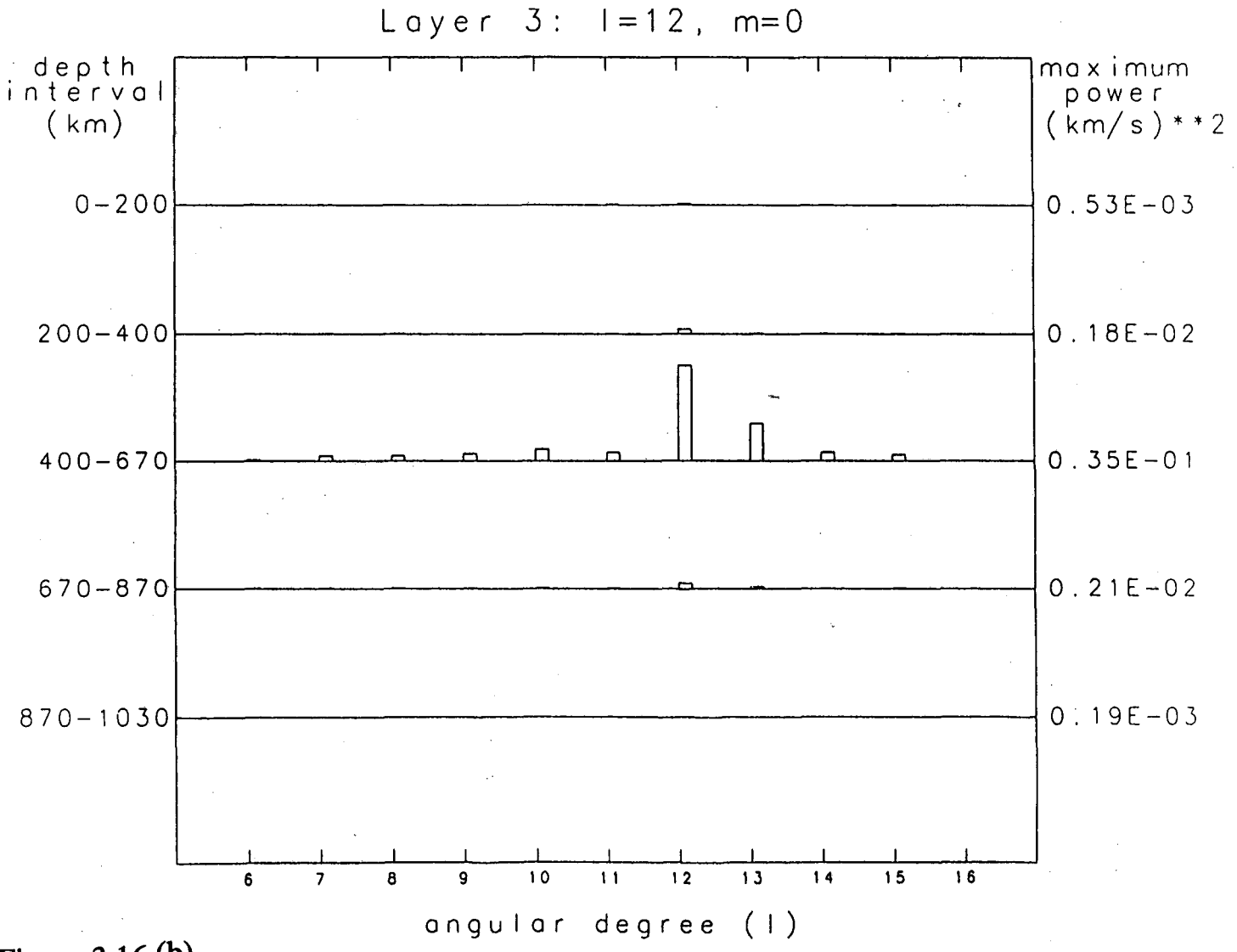
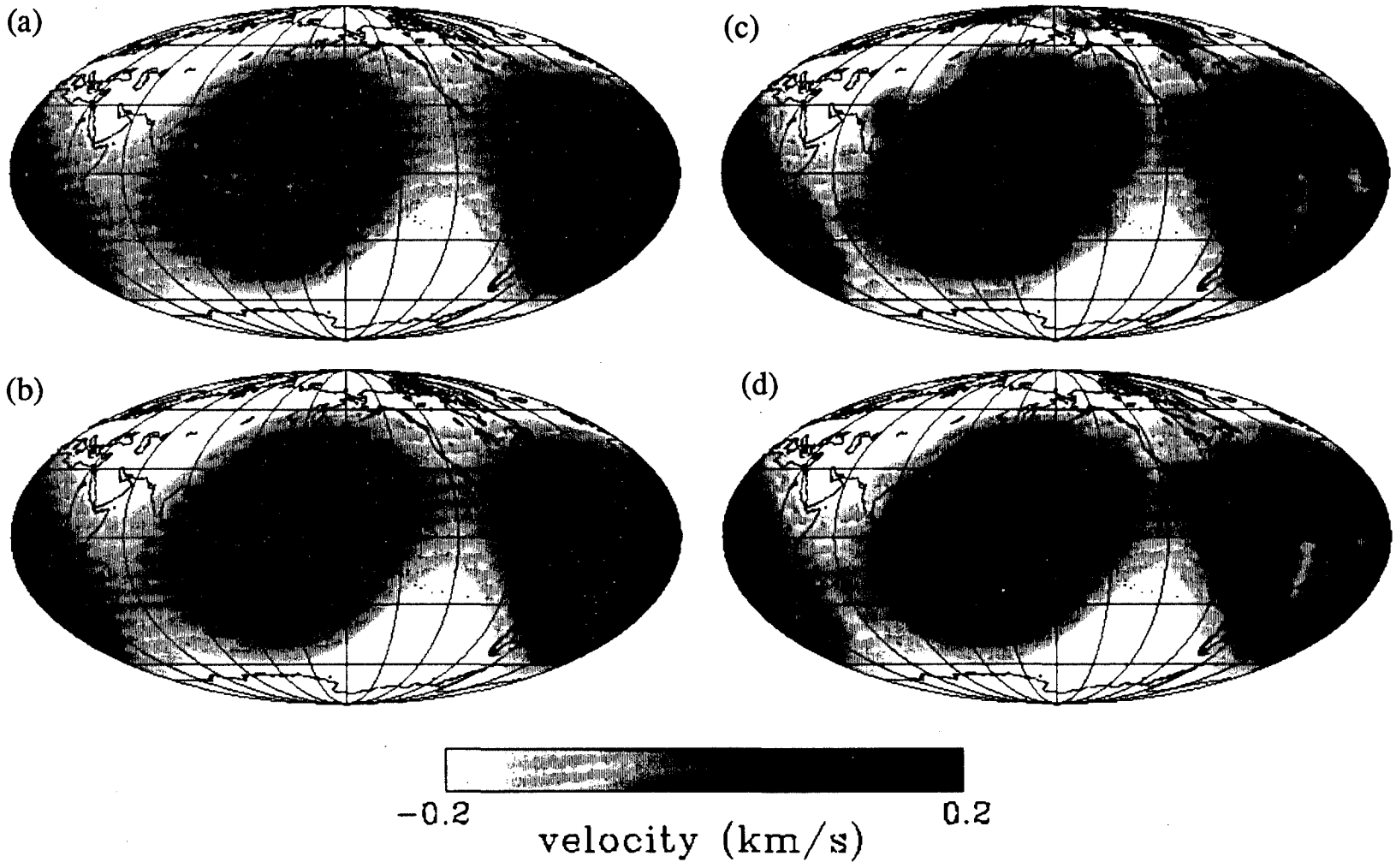

Layer 3: l=12, m=0
depth interval (km)
maximum power (km/s)**2
0-200
200-400
400-670
670-870
870-1030
0.53E-03
0.18E-02
0.35E-01
0.21E-02
0.19E-03
6
7
8
9
10
11
12
13
14
15
16
angular degree (l)

Figure 3.16 (b)

Figure 3.17



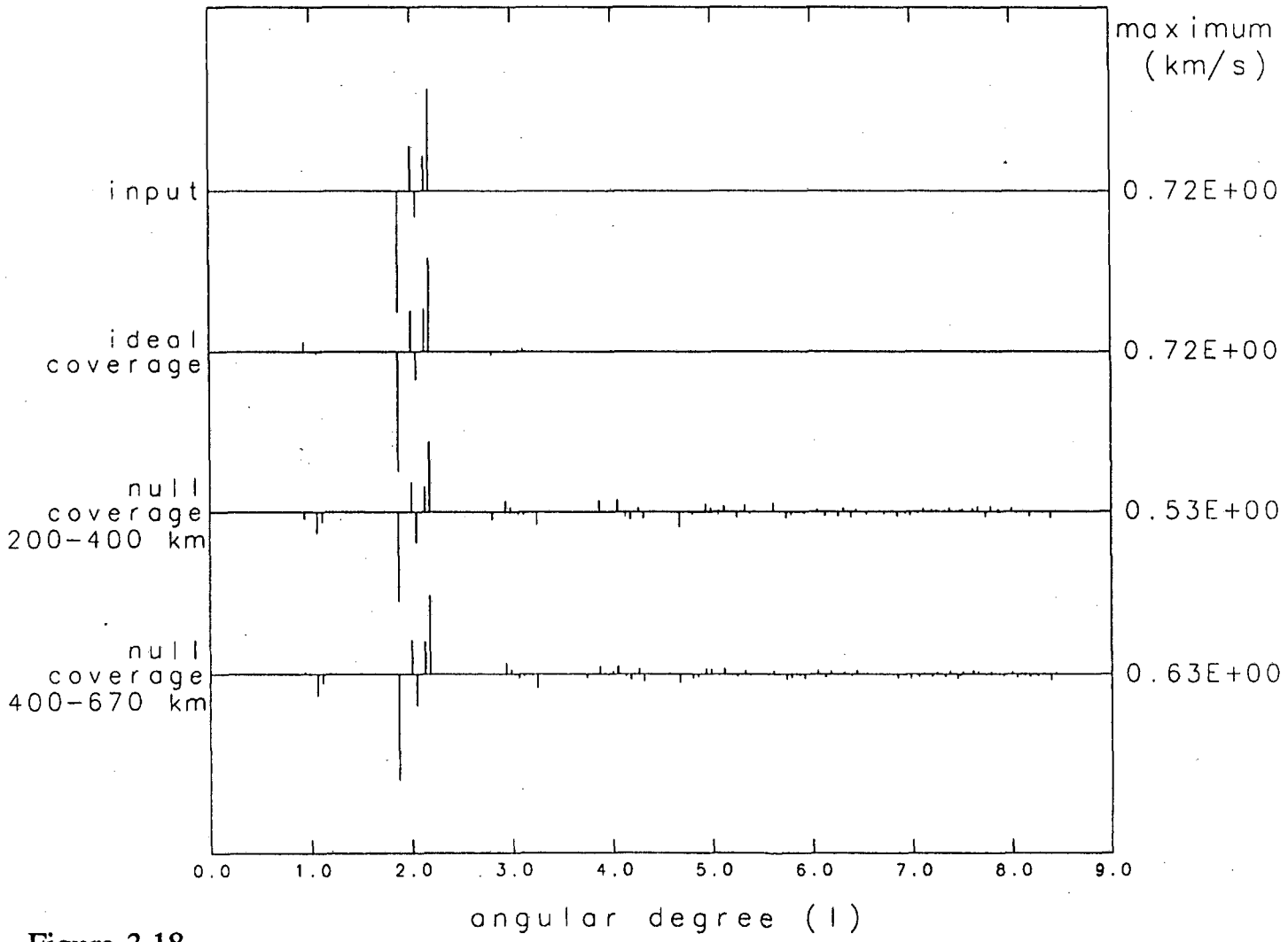
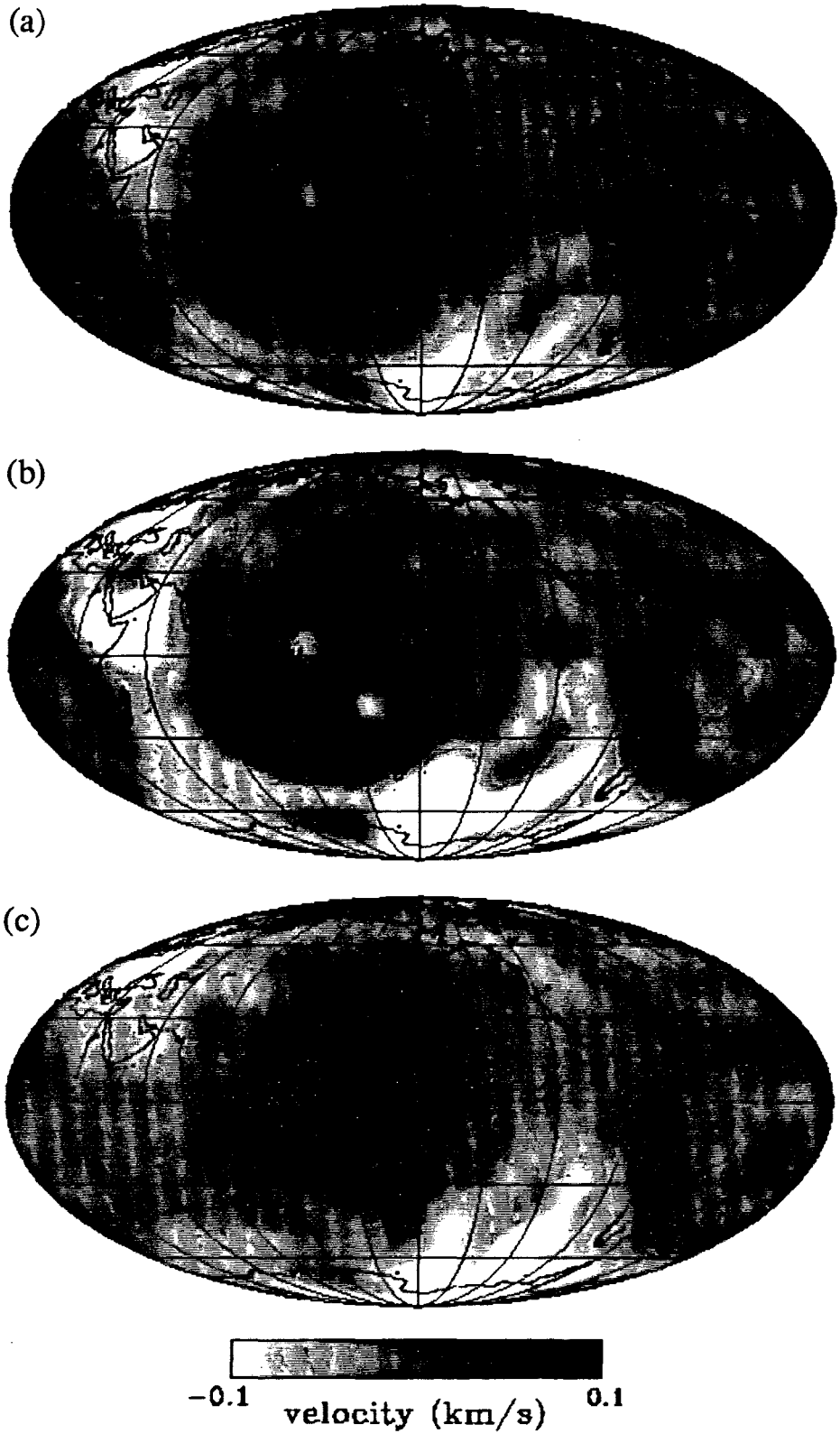

maximum
(km/s)
input
0.72E+00
ideal
coverage
0.72E+00
null
coverage
200-400 km
0.53E+00
null
coverage
400-670 km
0.63E+00
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
angular degree (l)

Figure 3.18

Figure 3.19



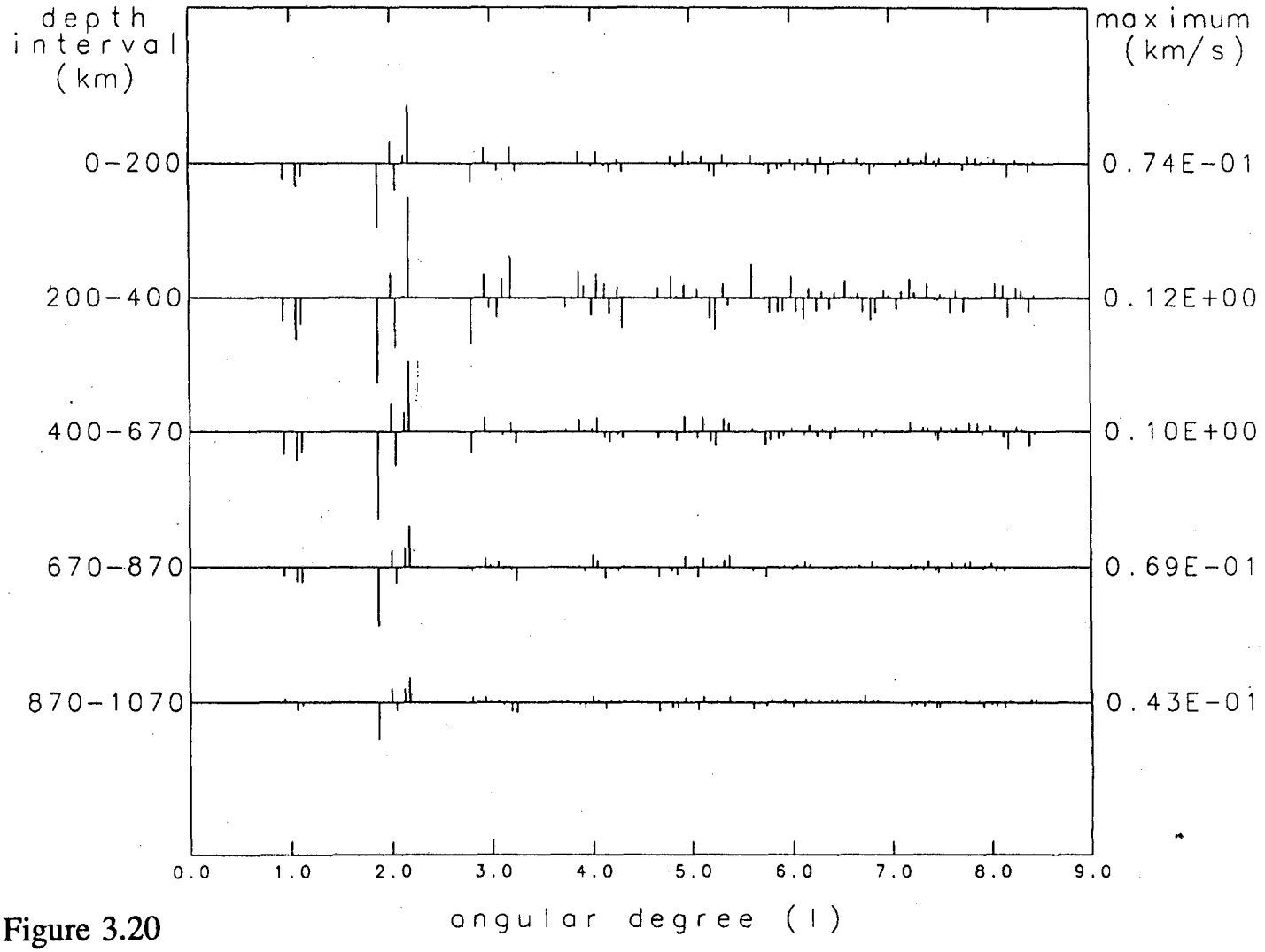
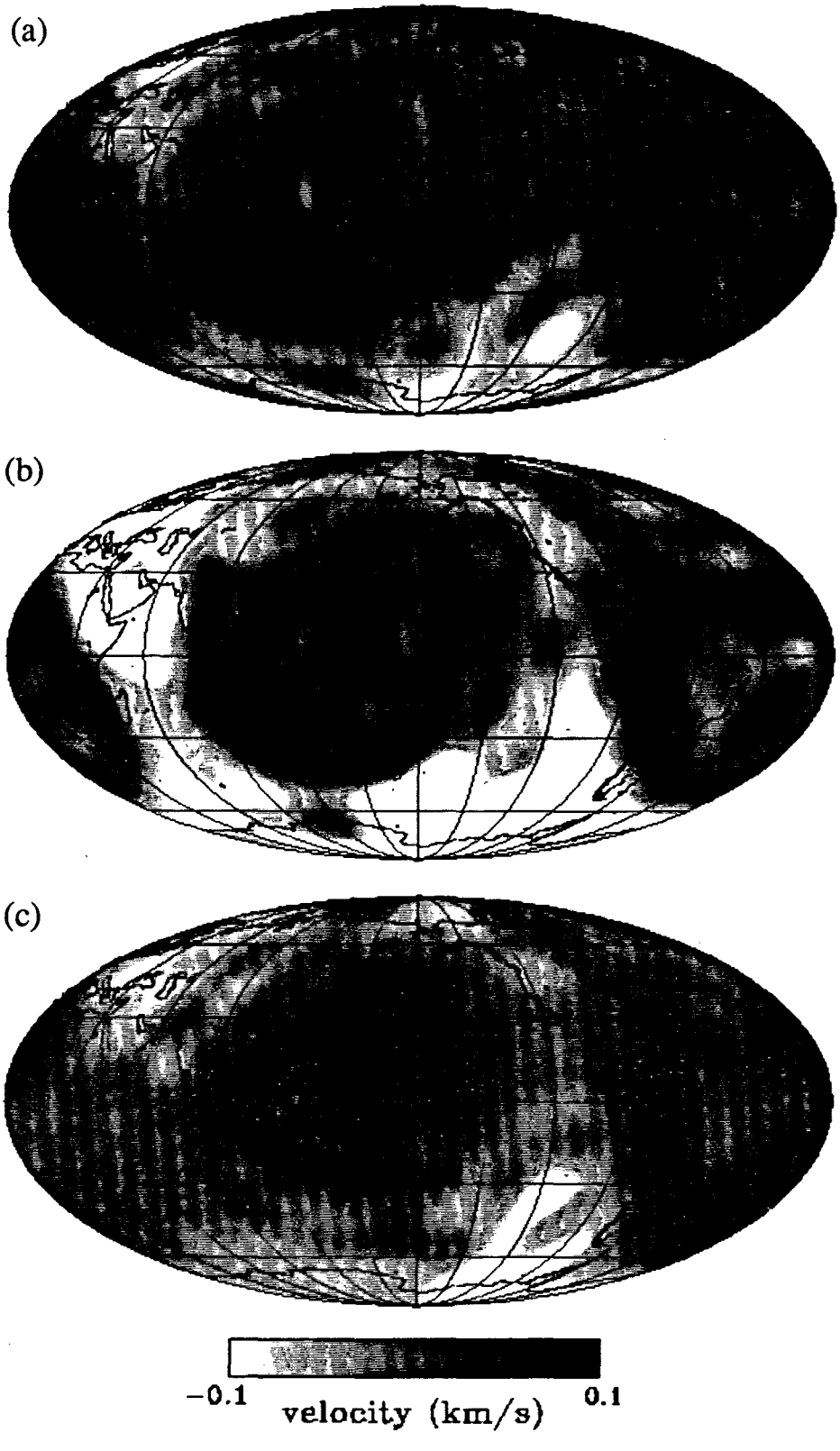

depth
interval
(km)
maximum
(km/s)
0-200
0.74E-01
200-400
0.12E+00
400-670
0.10E+00
670-870
0.69E-01
870-1070
0.43E-01
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
angular degree (l)

Figure 3.20

Figure 3.21



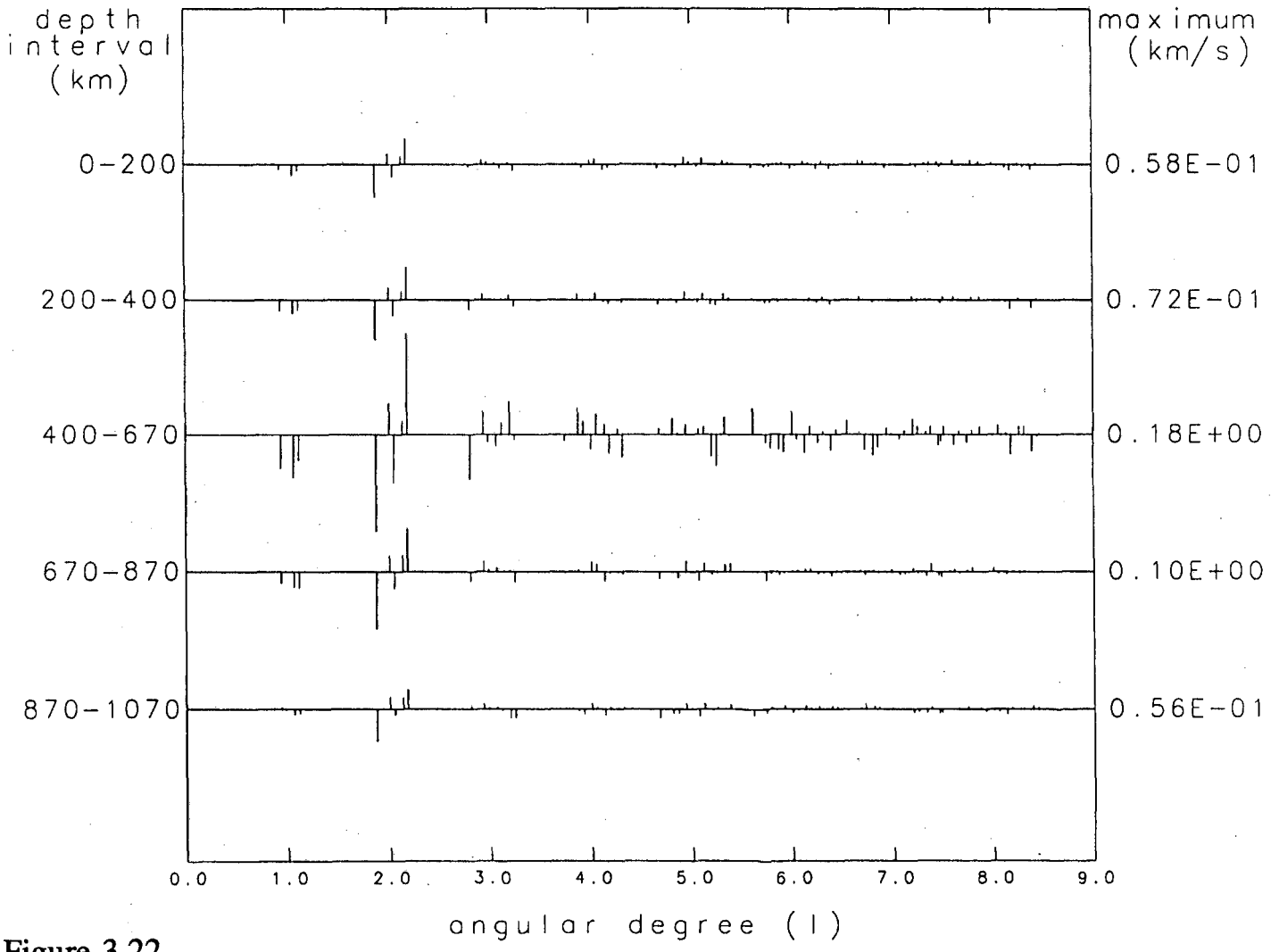

depth
interval
(km)
maximum
(km/s)
0-200
0.58E-01
200-400
0.72E-01
400-670
0.18E+00
670-870
0.10E+00
870-1070
0.56E-01
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
angular degree (l)

Figure 3.22

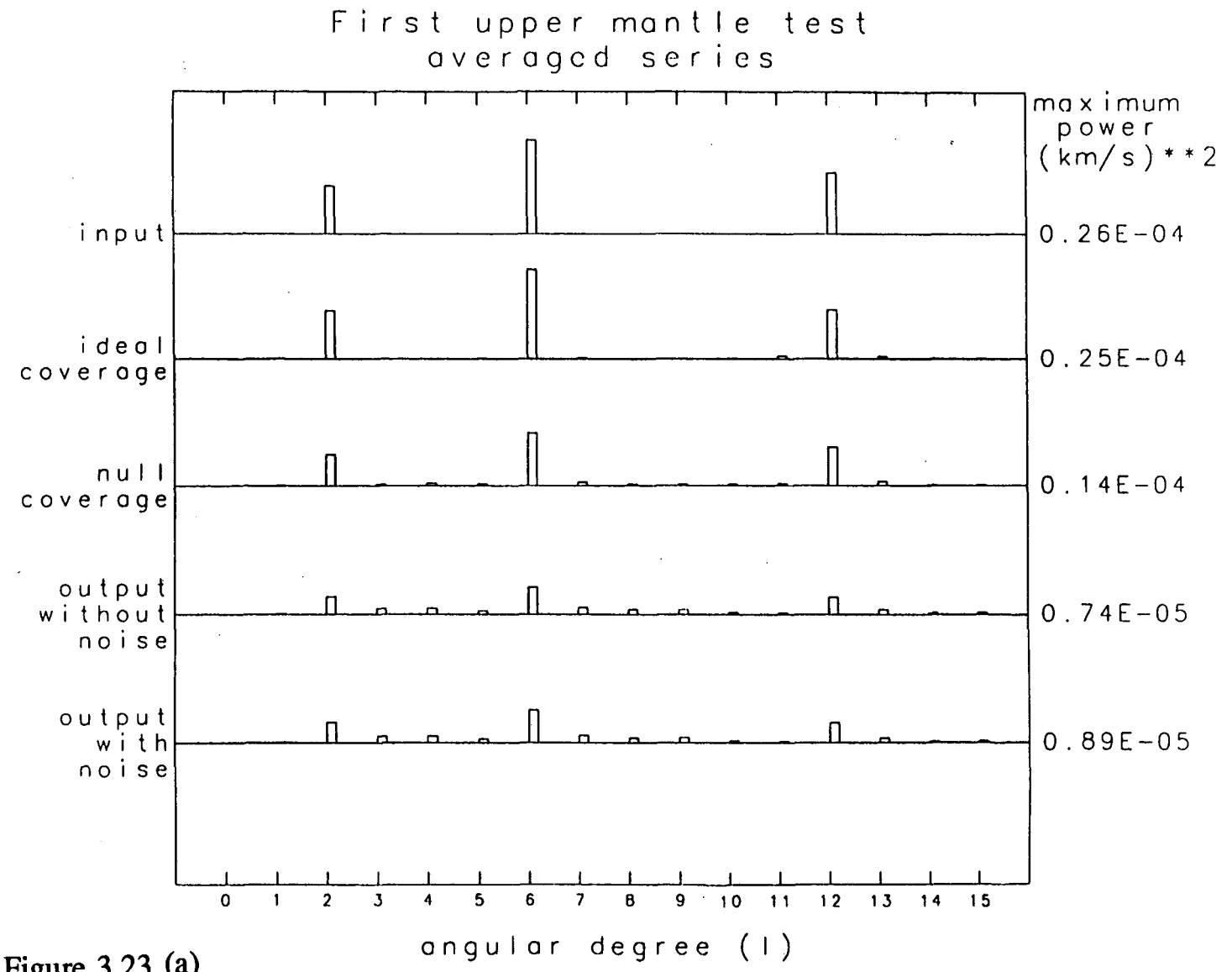

First upper mantle test
averaged series
maximum
power
(km/s)**2
input
0.26E-04
ideal
coverage
0.25E-04
null
coverage
0.14E-04
output
without
noise
0.74E-05
output
with
noise
0.89E-05
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.23 (a)

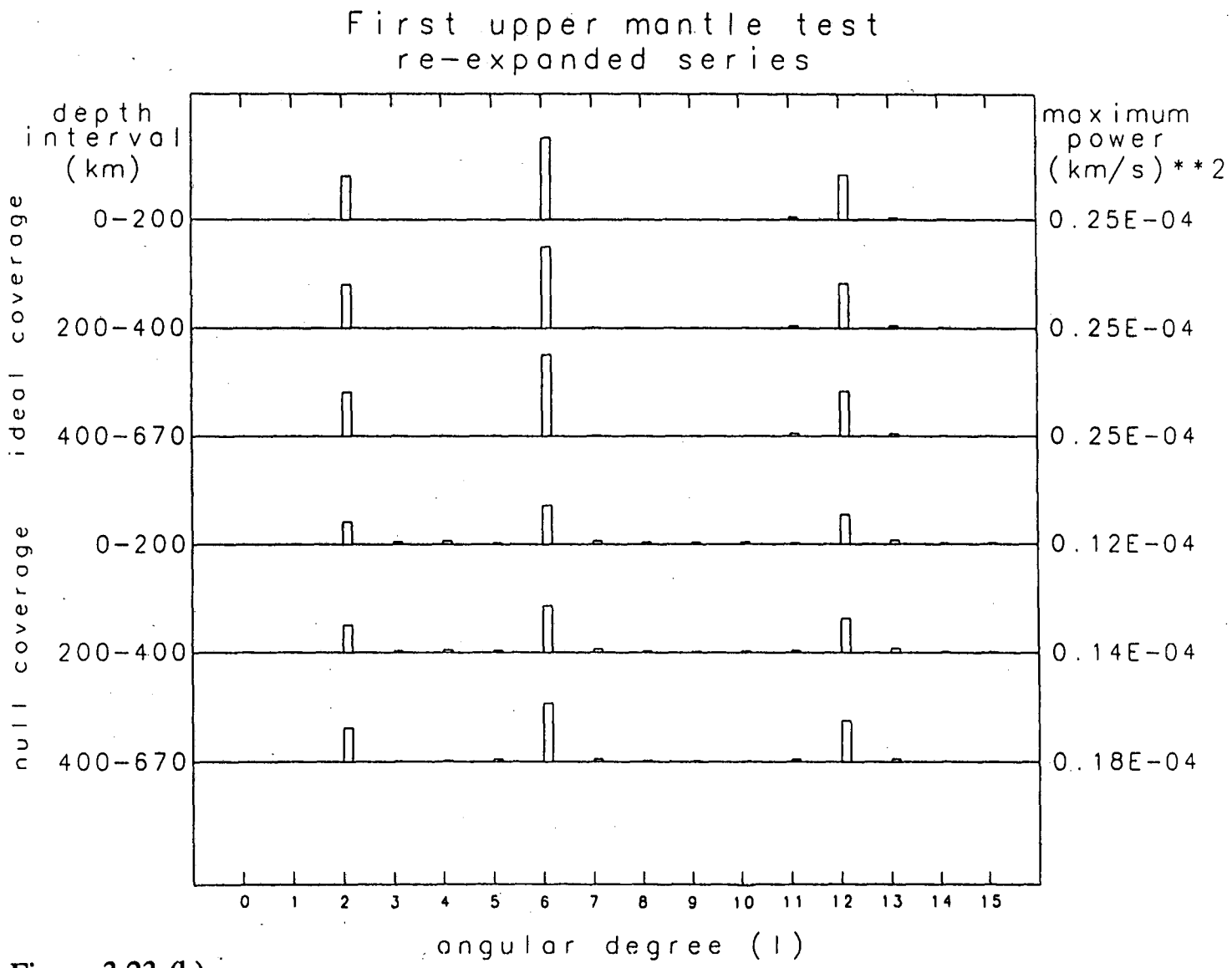

First upper mantle test
re-expanded series
depth interval (km)
maximum power (km/s)**2
ideal coverage
null coverage
0-200
200-400
400-670
0-200
200-400
400-670
0.25E-04
0.25E-04
0.25E-04
0.12E-04
0.14E-04
0.18E-04
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.23 (b)

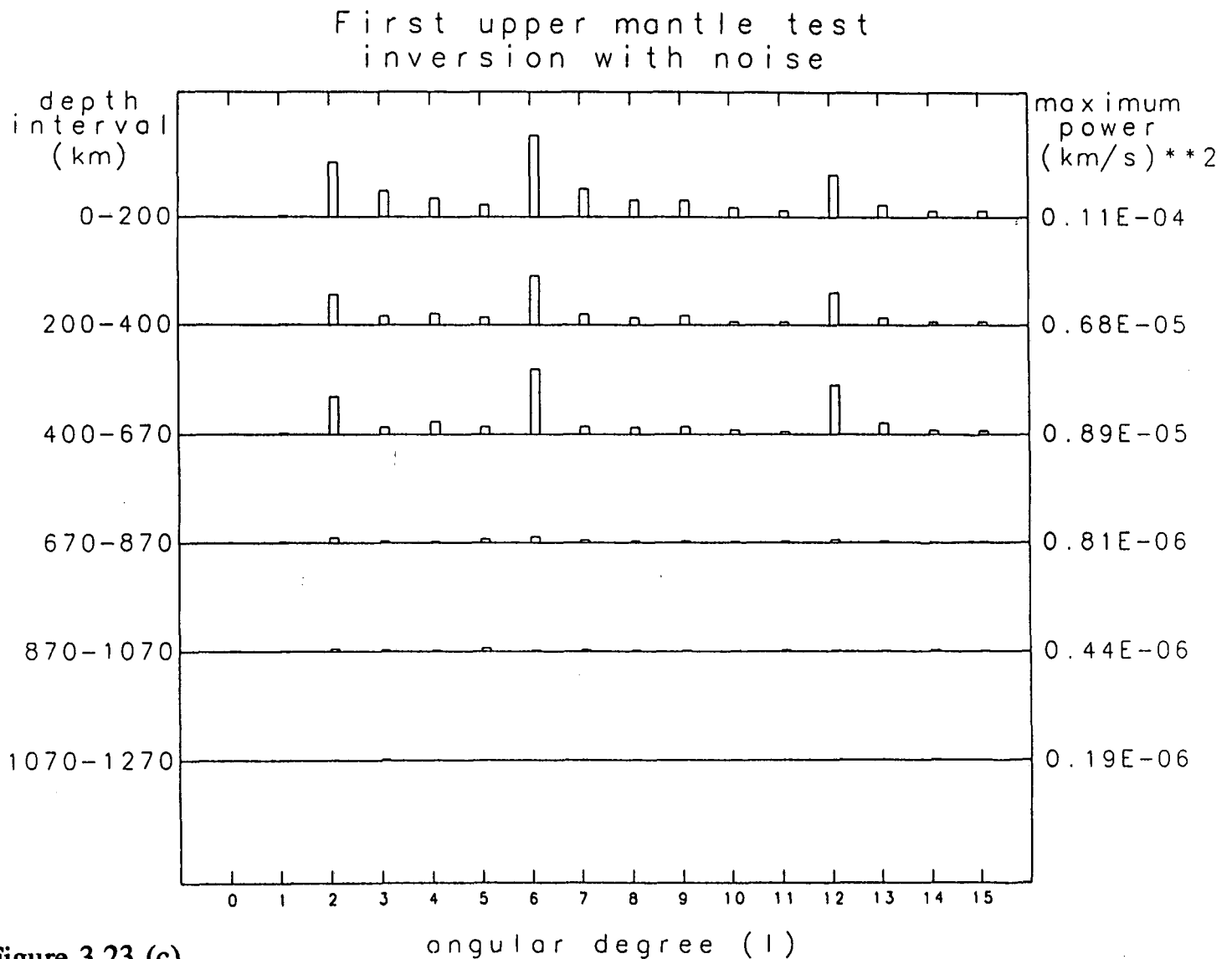

First upper mantle test
inversion with noise
depth
interval
(km)
maximum
power
(km/s)**2
0-200
0.11E-04
200-400
0.68E-05
400-670
0.89E-05
670-870
0.81E-06
870-1070
0.44E-06
1070-1270
0.19E-06
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.23 (c)

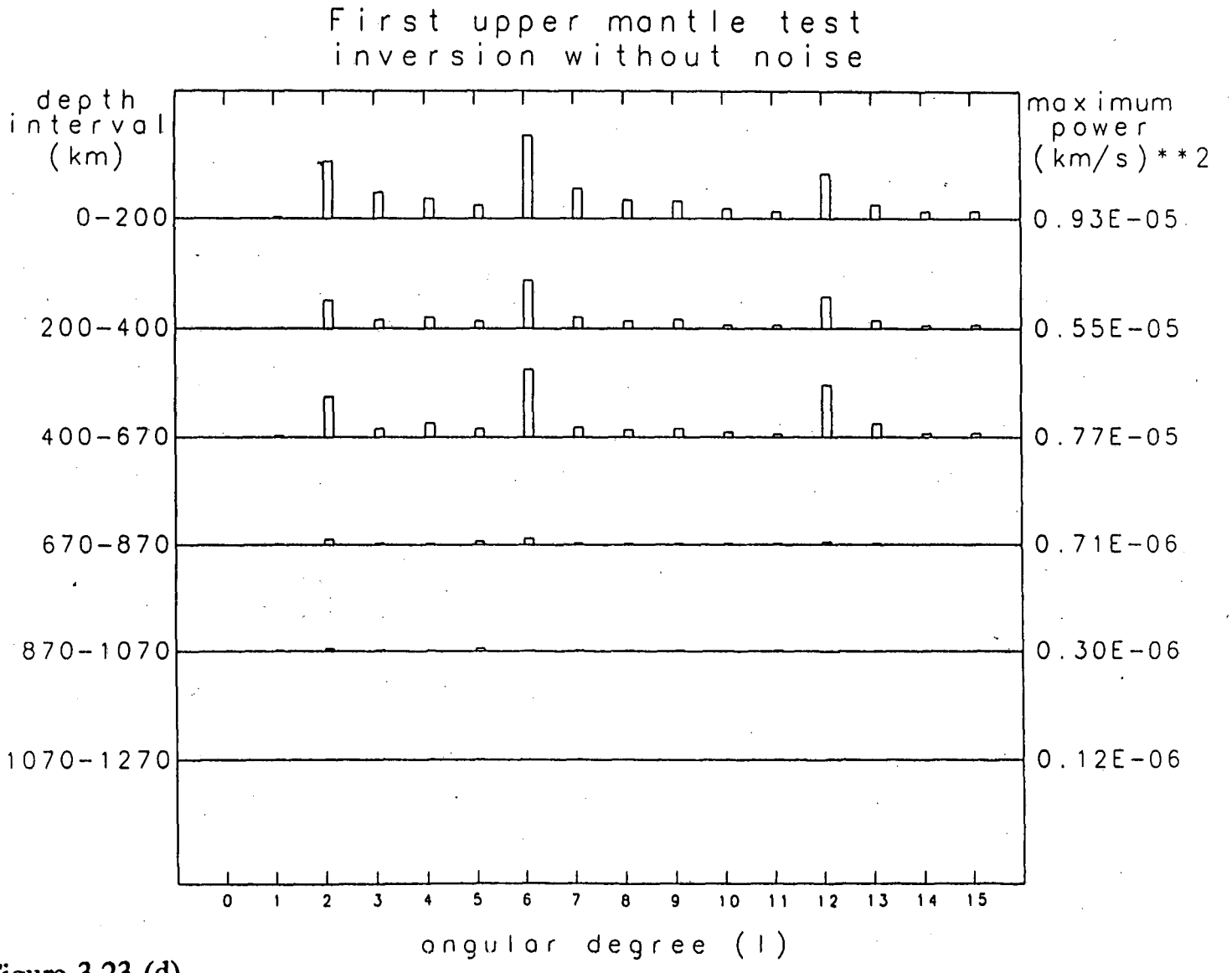

First upper mantle test
inversion without noise
depth
interval
(km)
maximum
power
(km/s)**2
0-200
0.93E-05
200-400
0.55E-05
400-670
0.77E-05
670-870
0.71E-06
870-1070
0.30E-06
1070-1270
0.12E-06
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.23 (d)

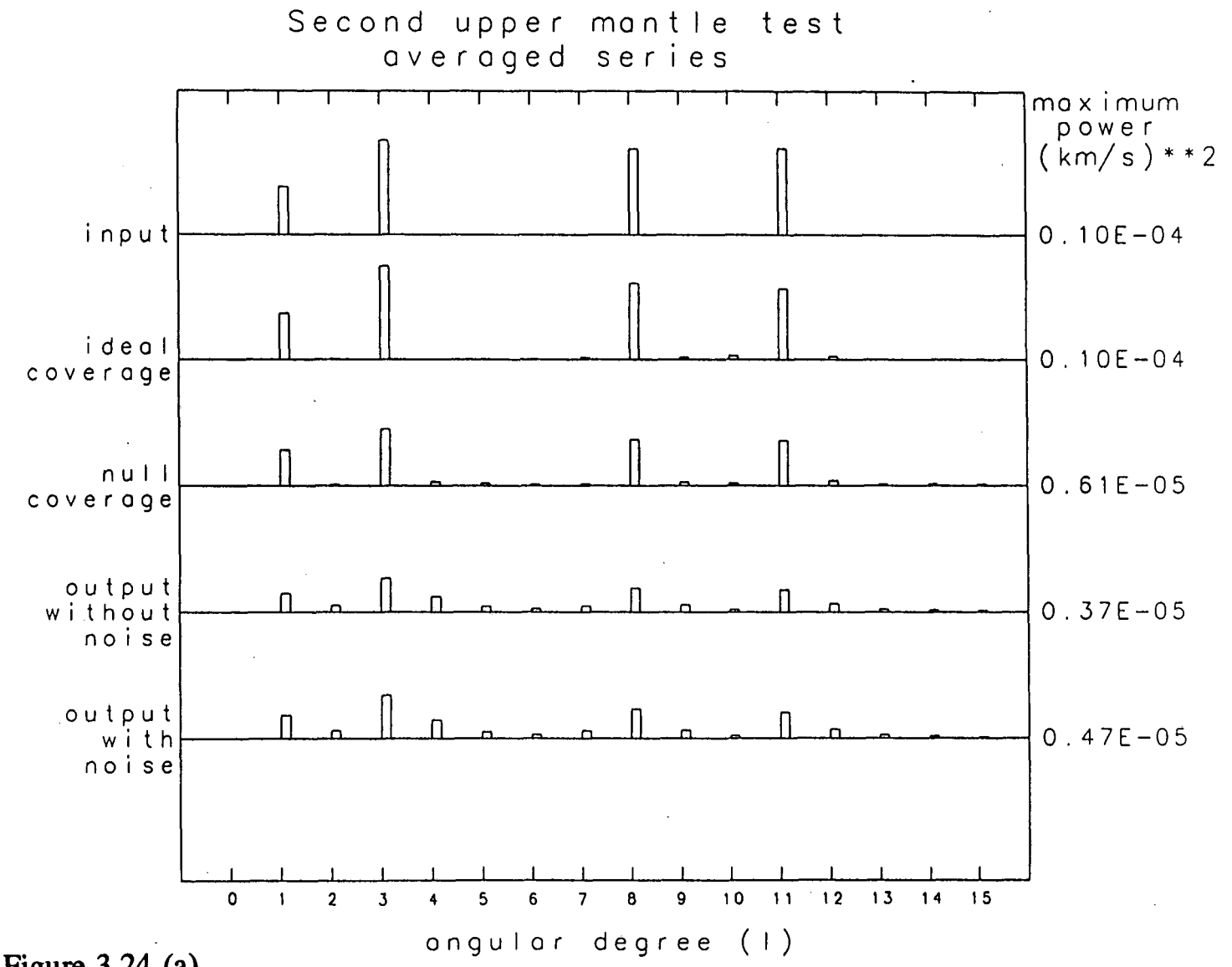

Second upper mantle test
averaged series
maximum
power
(km/s)**2
input
0.10E-04
ideal
coverage
0.10E-04
null
coverage
0.61E-05
output
without
noise
0.37E-05
output
with
noise
0.47E-05
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.24 (a)

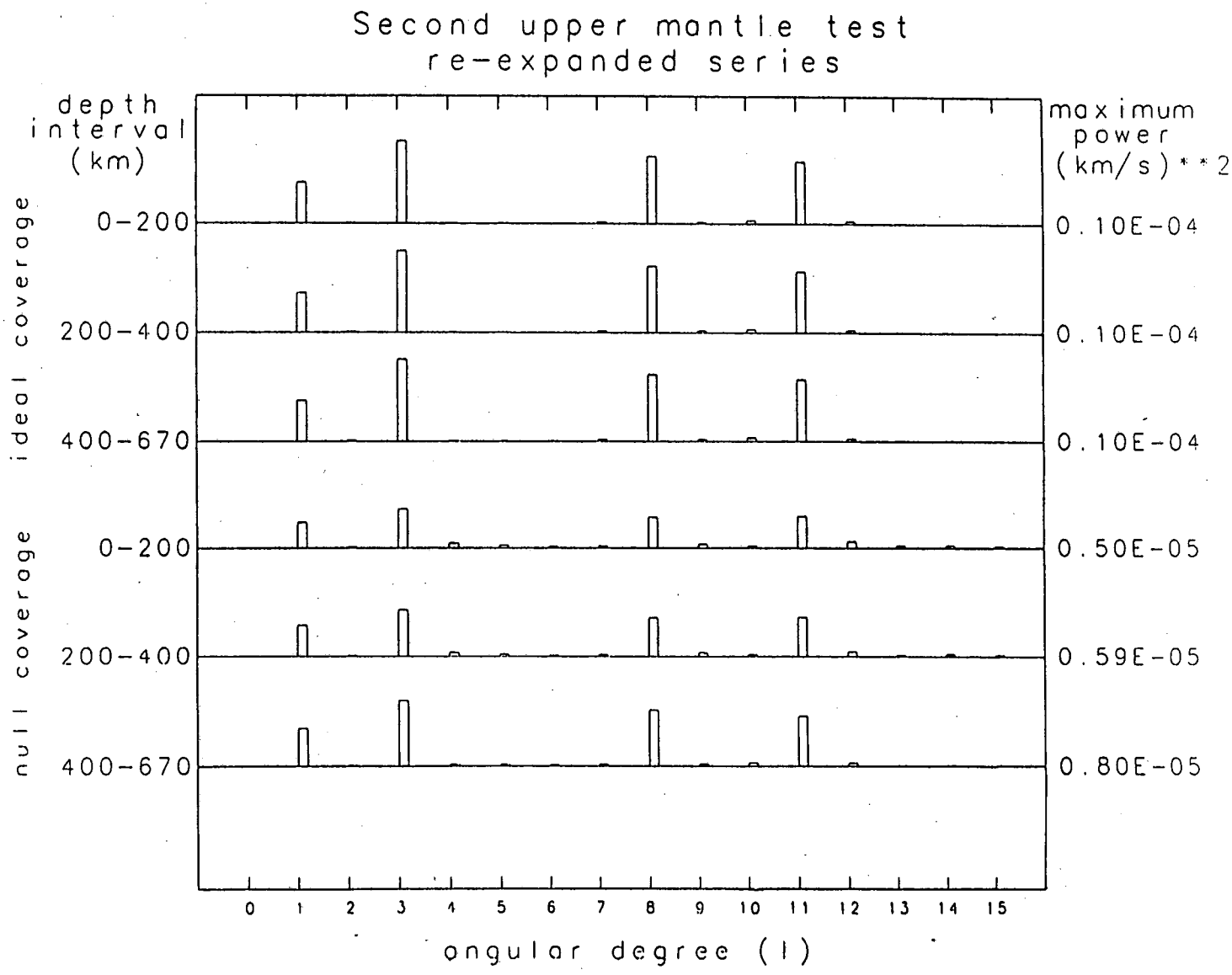

Second upper mantle test
re-expanded series
depth interval (km)
maximum power (km/s)**2
ideal coverage
null coverage
0-200
200-400
400-670
0-200
200-400
400-670
0.10E-04
0.10E-04
0.10E-04
0.50E-05
0.59E-05
0.80E-05
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.24 (b)

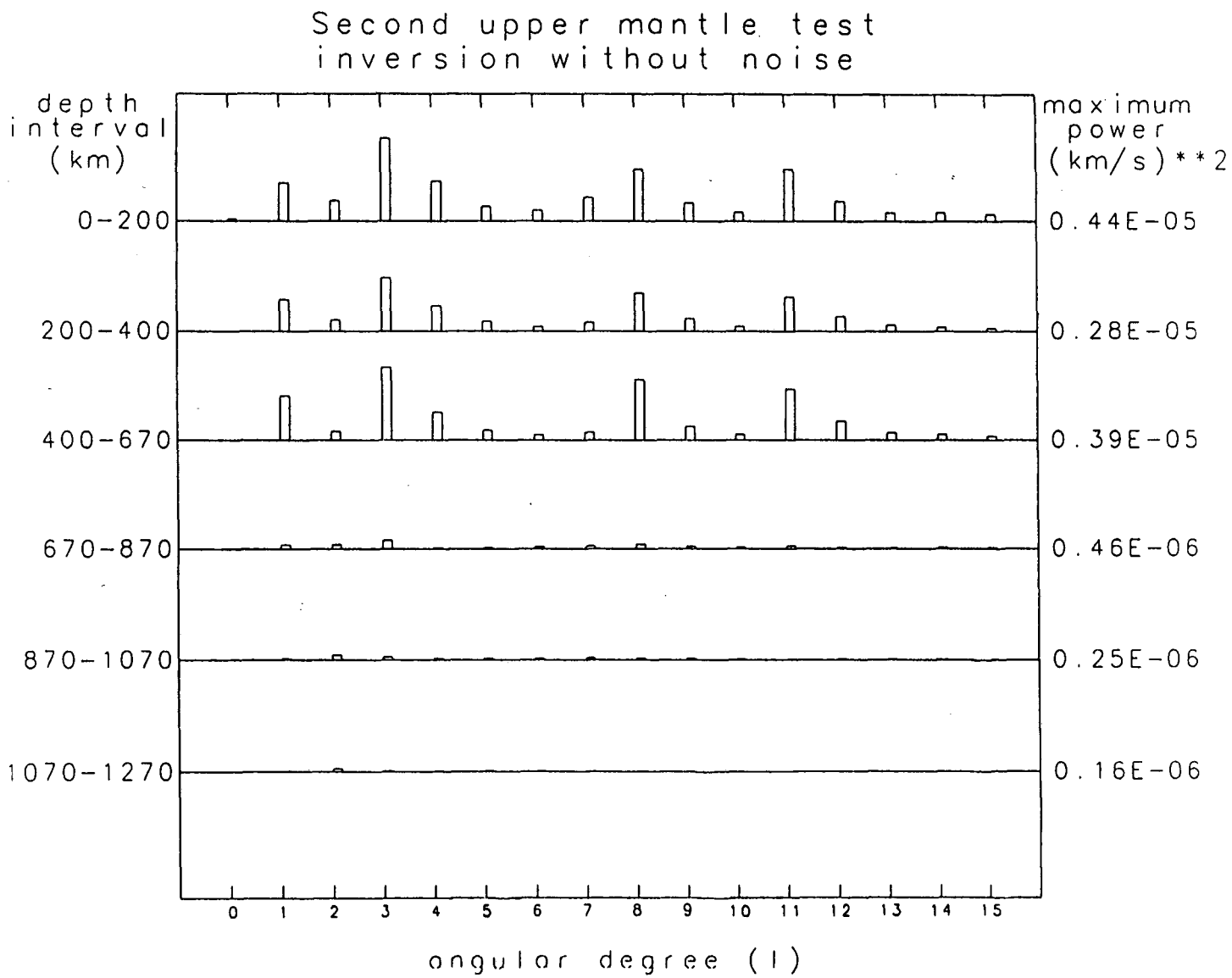

Second upper mantle test
inversion without noise
depth
interval
(km)
maximum
power
(km/s)**2
0-200
0.44E-05
200-400
0.28E-05
400-670
0.39E-05
670-870
0.46E-06
870-1070
0.25E-06
1070-1270
0.16E-06
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.24 (c)

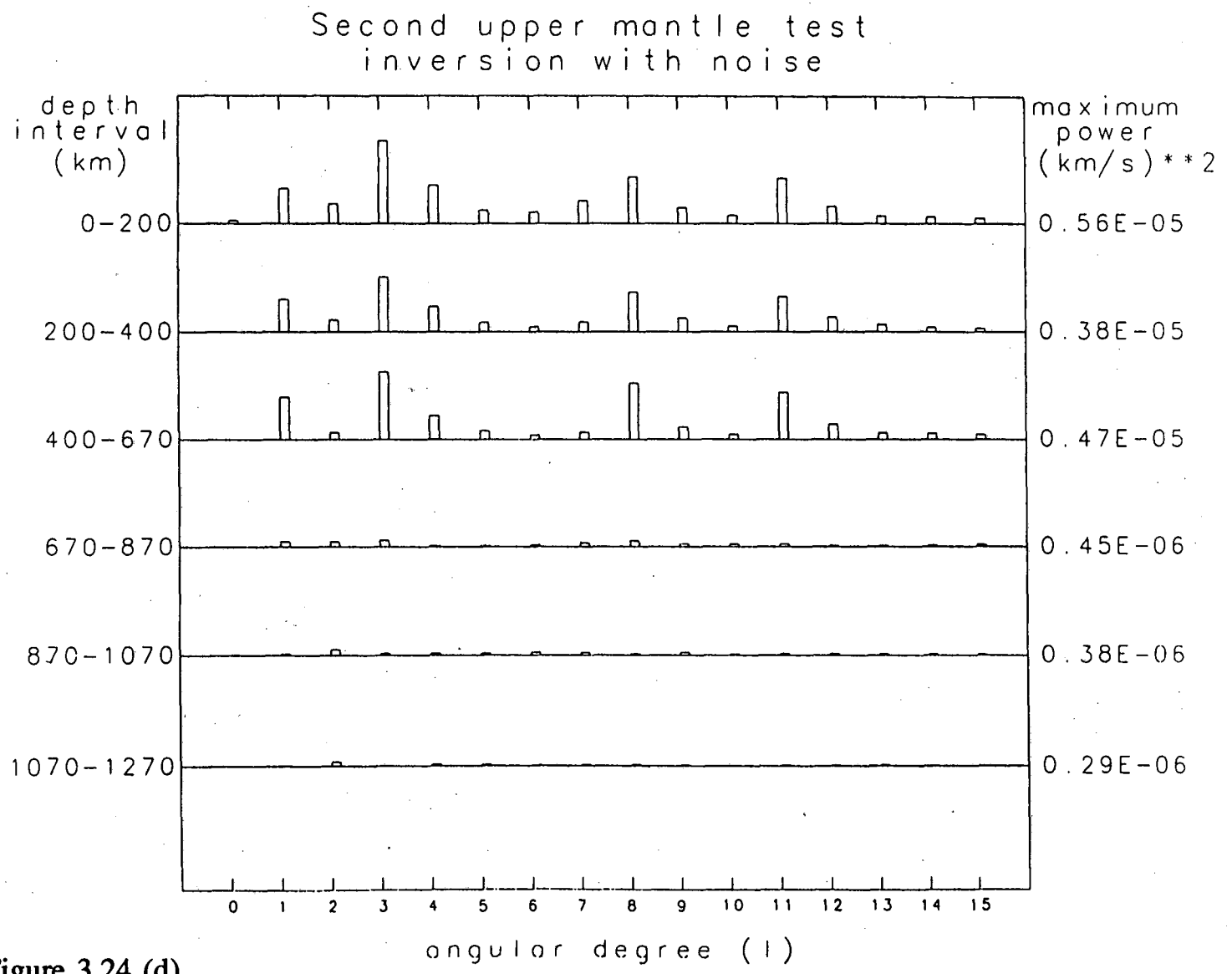

Second upper mantle test
inversion with noise
depth
interval
(km)
maximum
power
(km/s)**2
0-200
0.56E-05
200-400
0.38E-05
400-670
0.47E-05
670-870
0.45E-06
870-1070
0.38E-06
1070-1270
0.29E-06
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.24 (d)

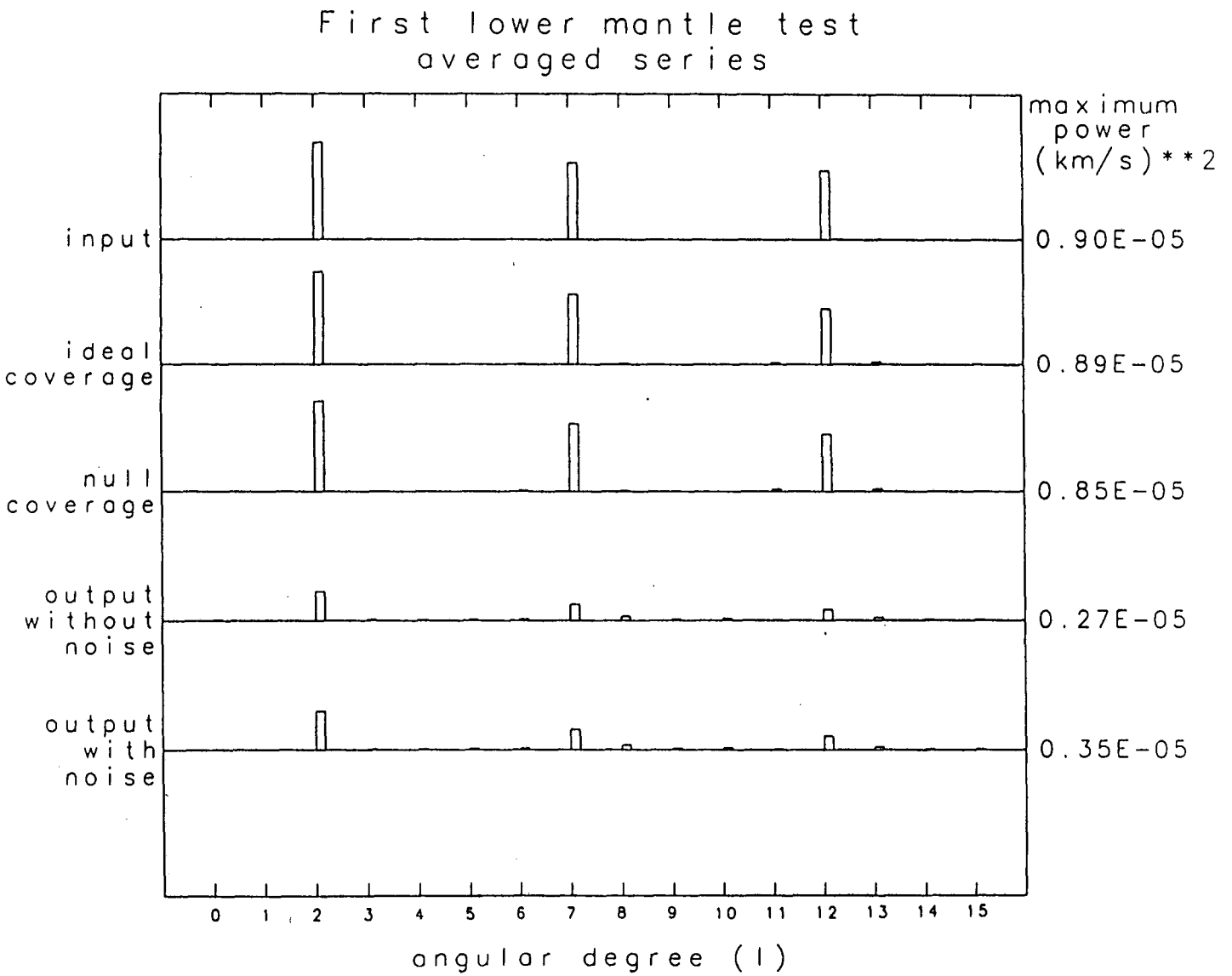

First lower mantle test
averaged series
maximum
power
(km/s)**2
input
0.90E-05
ideal
coverage
0.89E-05
null
coverage
0.85E-05
output
without
noise
0.27E-05
output
with
noise
0.35E-05
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.25 (a)

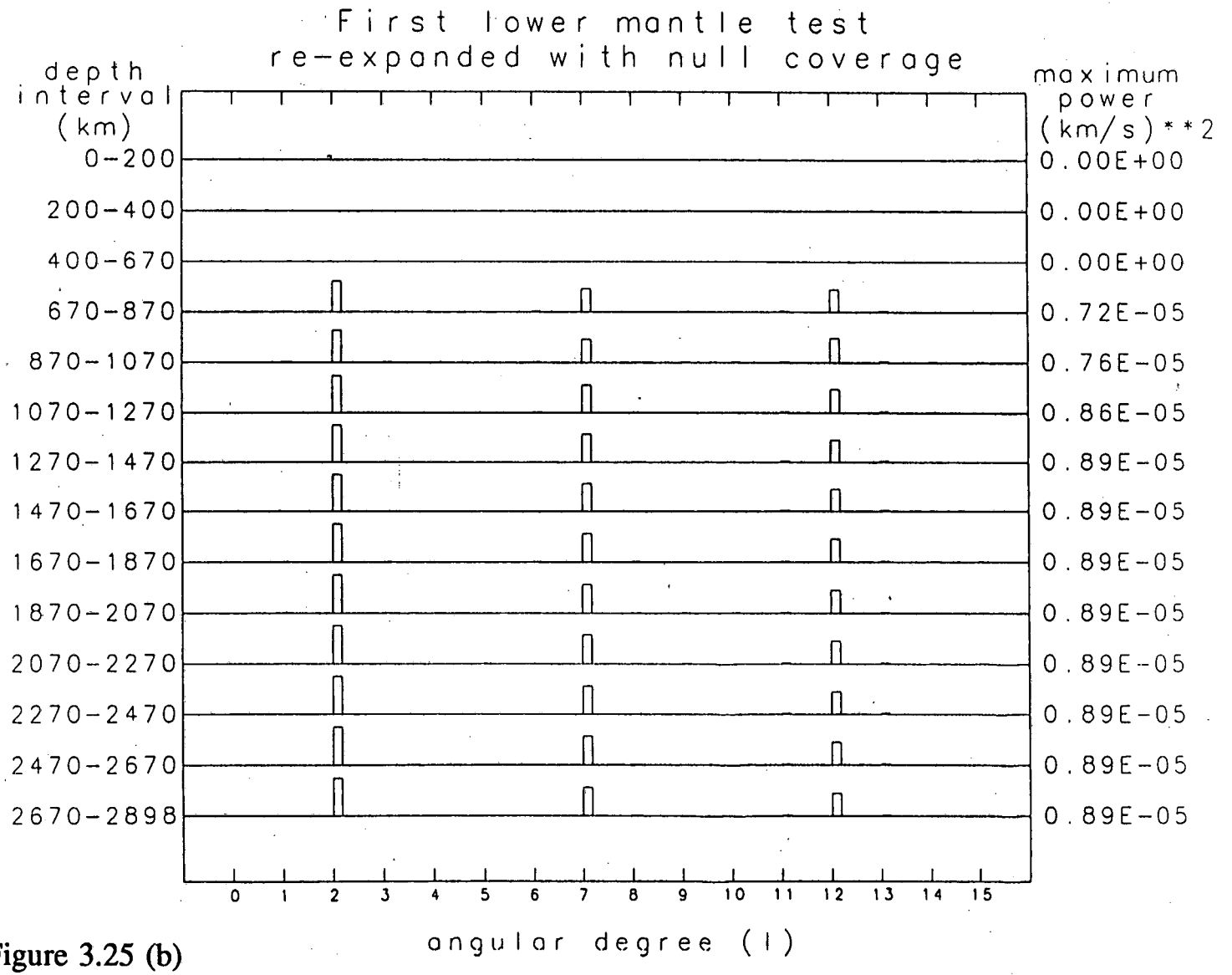

First lower mantle test
re-expanded with null coverage
depth interval (km)
maximum power (km/s)**2
0-200 0.00E+00
200-400 0.00E+00
400-670 0.00E+00
670-870 0.72E-05
870-1070 0.76E-05
1070-1270 0.86E-05
1270-1470 0.89E-05
1470-1670 0.89E-05
1670-1870 0.89E-05
1870-2070 0.89E-05
2070-2270 0.89E-05
2270-2470 0.89E-05
2470-2670 0.89E-05
2670-2898 0.89E-05
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.25 (b)

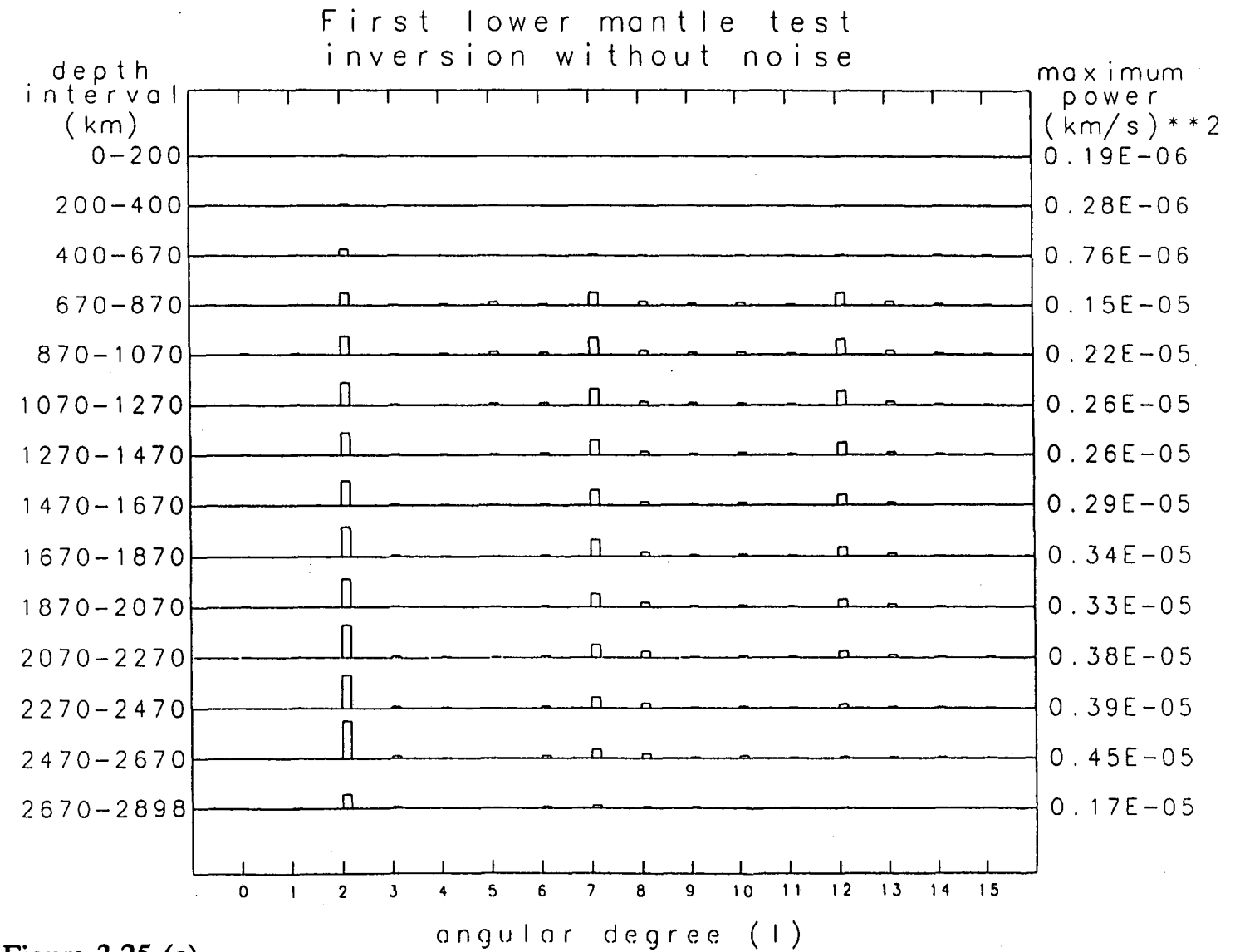

First lower mantle test
inversion without noise
depth
interval
(km)
0-200
200-400
400-670
670-870
870-1070
1070-1270
1270-1470
1470-1670
1670-1870
1870-2070
2070-2270
2270-2470
2470-2670
2670-2898
maximum
power
(km/s)**2
0.19E-06
0.28E-06
0.76E-06
0.15E-05
0.22E-05
0.26E-05
0.26E-05
0.29E-05
0.34E-05
0.33E-05
0.38E-05
0.39E-05
0.45E-05
0.17E-05
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.25 (c)

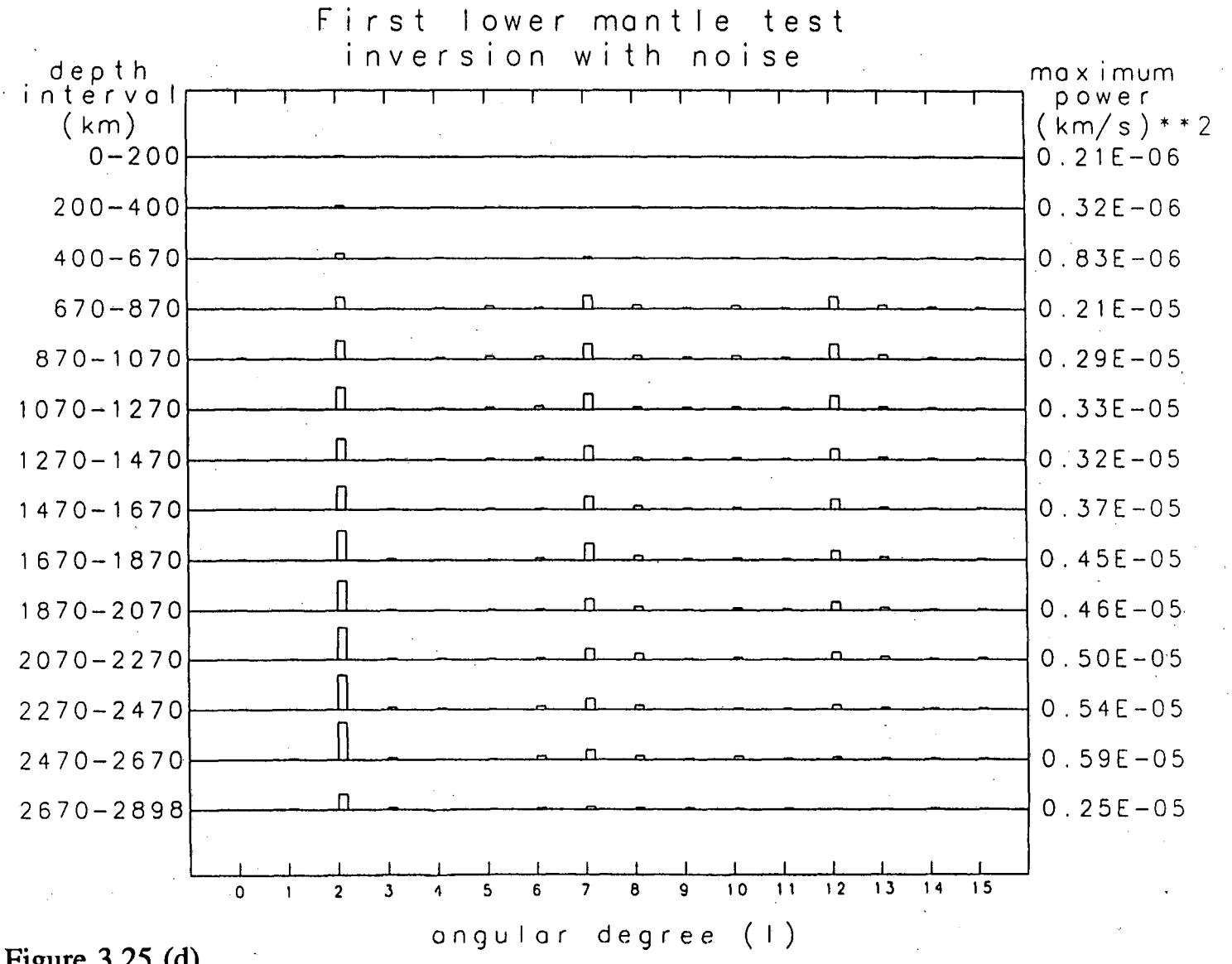

First lower mantle test
inversion with noise
depth
interval
(km)
0-200
200-400
400-670
670-870
870-1070
1070-1270
1270-1470
1470-1670
1670-1870
1870-2070
2070-2270
2270-2470
2470-2670
2670-2898
maximum
power
(km/s)**2
0.21E-06
0.32E-06
0.83E-06
0.21E-05
0.29E-05
0.33E-05
0.32E-05
0.37E-05
0.45E-05
0.46E-05
0.50E-05
0.54E-05
0.59E-05
0.25E-05
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.25 (d)

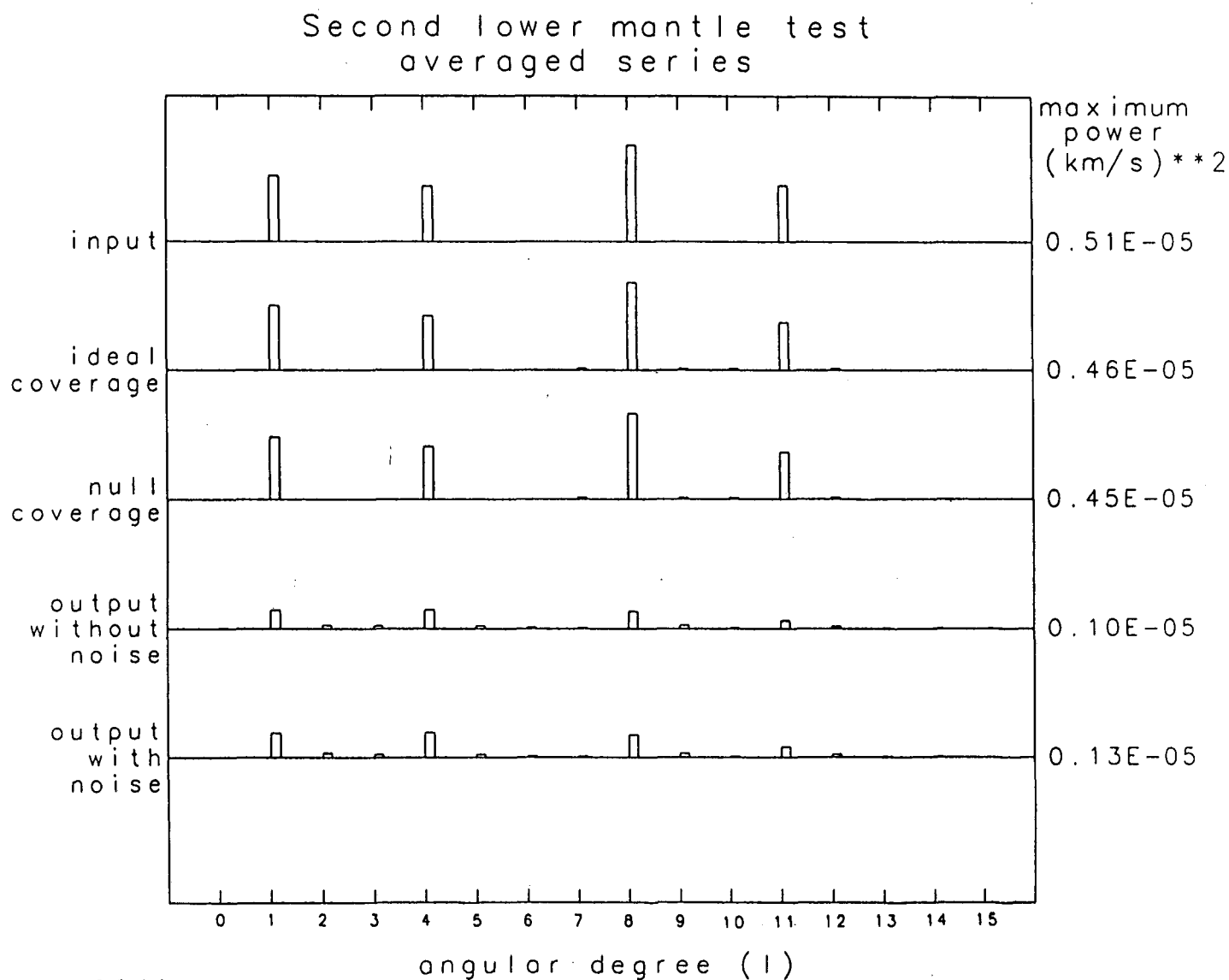

Second lower mantle test
averaged series
maximum
power
(km/s)**2
input
0.51E-05
ideal
coverage
0.46E-05
null
coverage
0.45E-05
output
without
noise
0.10E-05
output
with
noise
0.13E-05
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.26 (a)

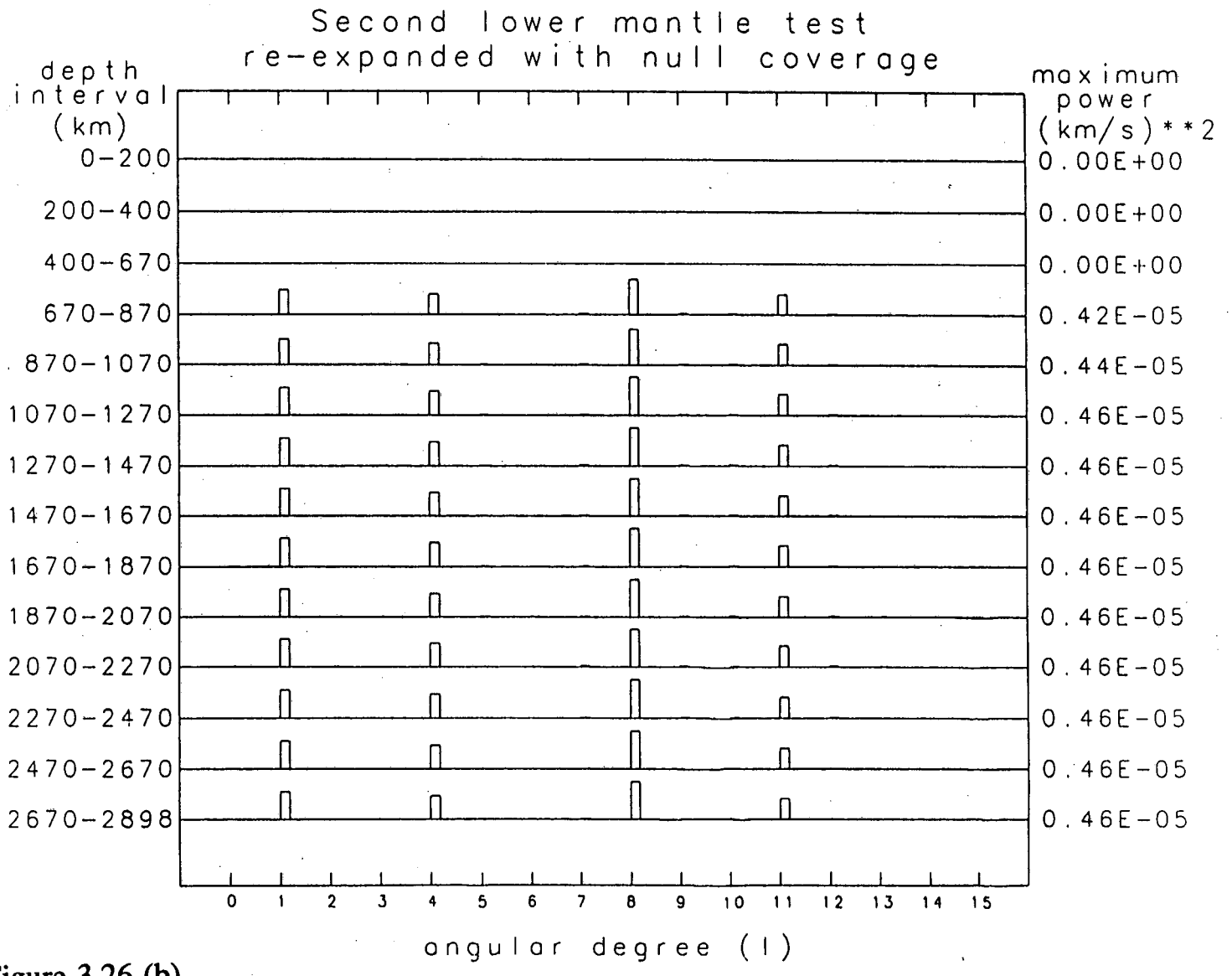

Second lower mantle test
re-expanded with null coverage
depth interval (km)
maximum power (km/s)**2
0-200 0.00E+00
200-400 0.00E+00
400-670 0.00E+00
670-870 0.42E-05
870-1070 0.44E-05
1070-1270 0.46E-05
1270-1470 0.46E-05
1470-1670 0.46E-05
1670-1870 0.46E-05
1870-2070 0.46E-05
2070-2270 0.46E-05
2270-2470 0.46E-05
2470-2670 0.46E-05
2670-2898 0.46E-05
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.26 (b)

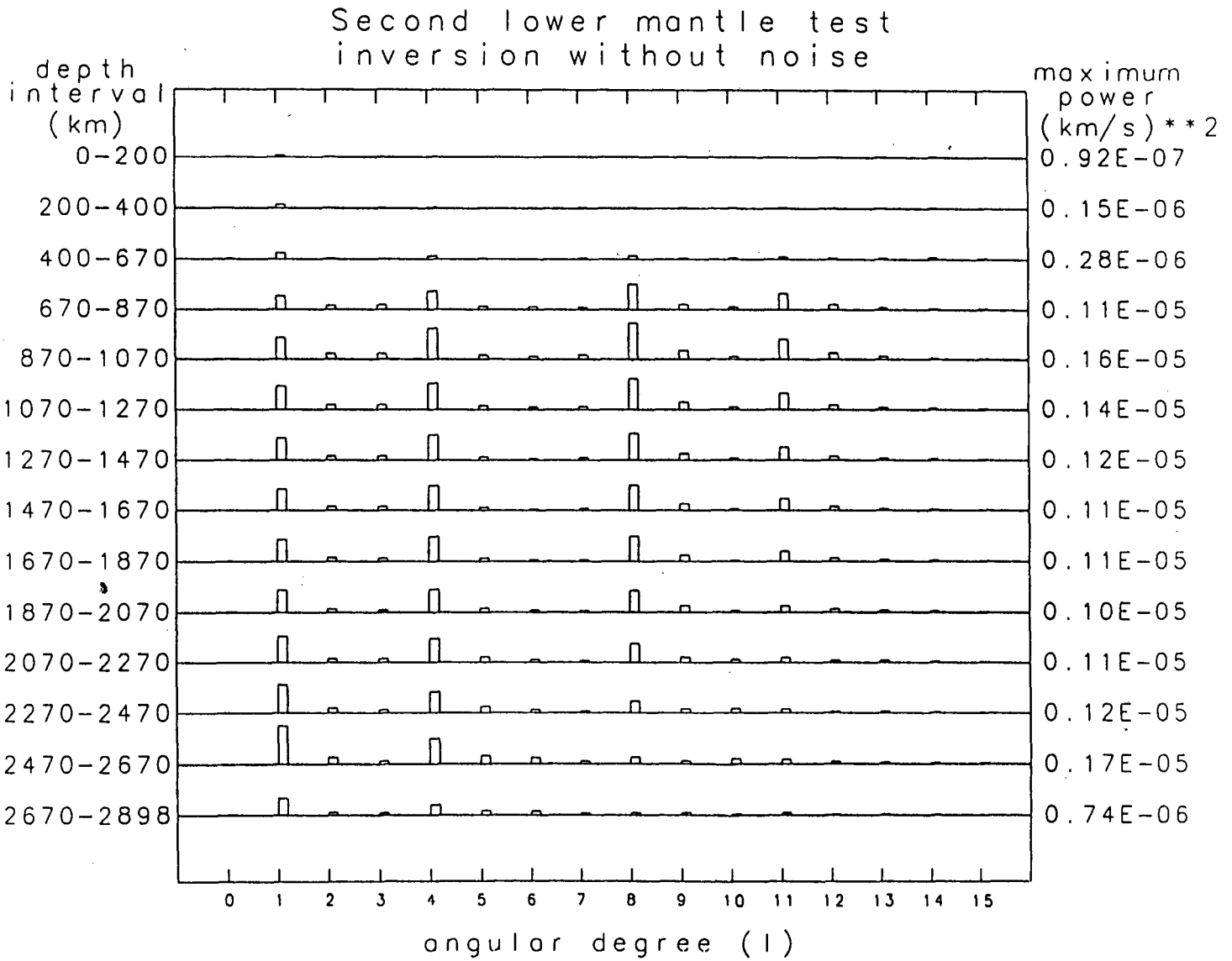

Second lower mantle test
inversion without noise
depth interval (km)
maximum power (km/s)**2
0-200
0.92E-07
200-400
0.15E-06
400-670
0.28E-06
670-870
0.11E-05
870-1070
0.16E-05
1070-1270
0.14E-05
1270-1470
0.12E-05
1470-1670
0.11E-05
1670-1870
0.11E-05
1870-2070
0.10E-05
2070-2270
0.11E-05
2270-2470
0.12E-05
2470-2670
0.17E-05
2670-2898
0.74E-06
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.26 (c)

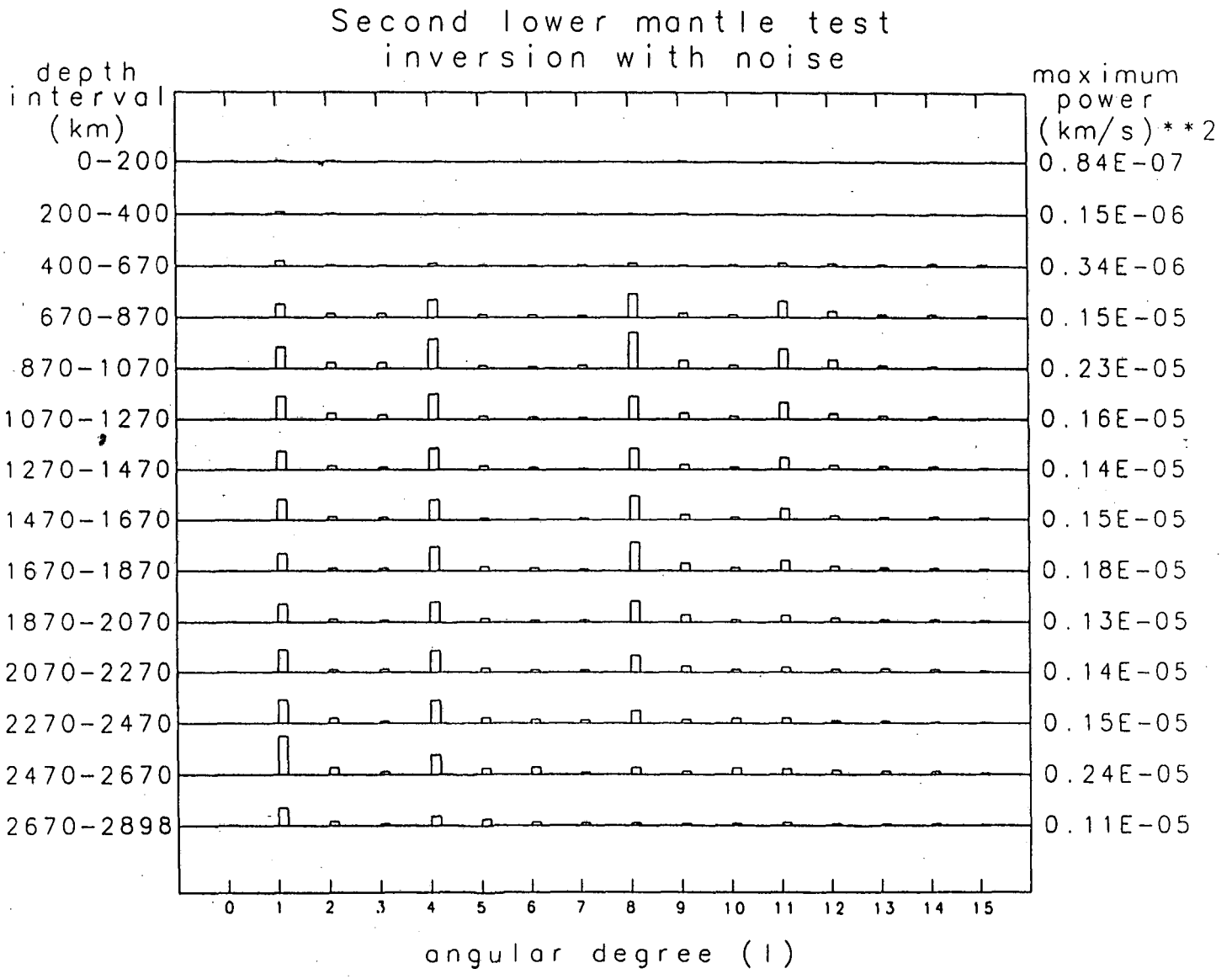
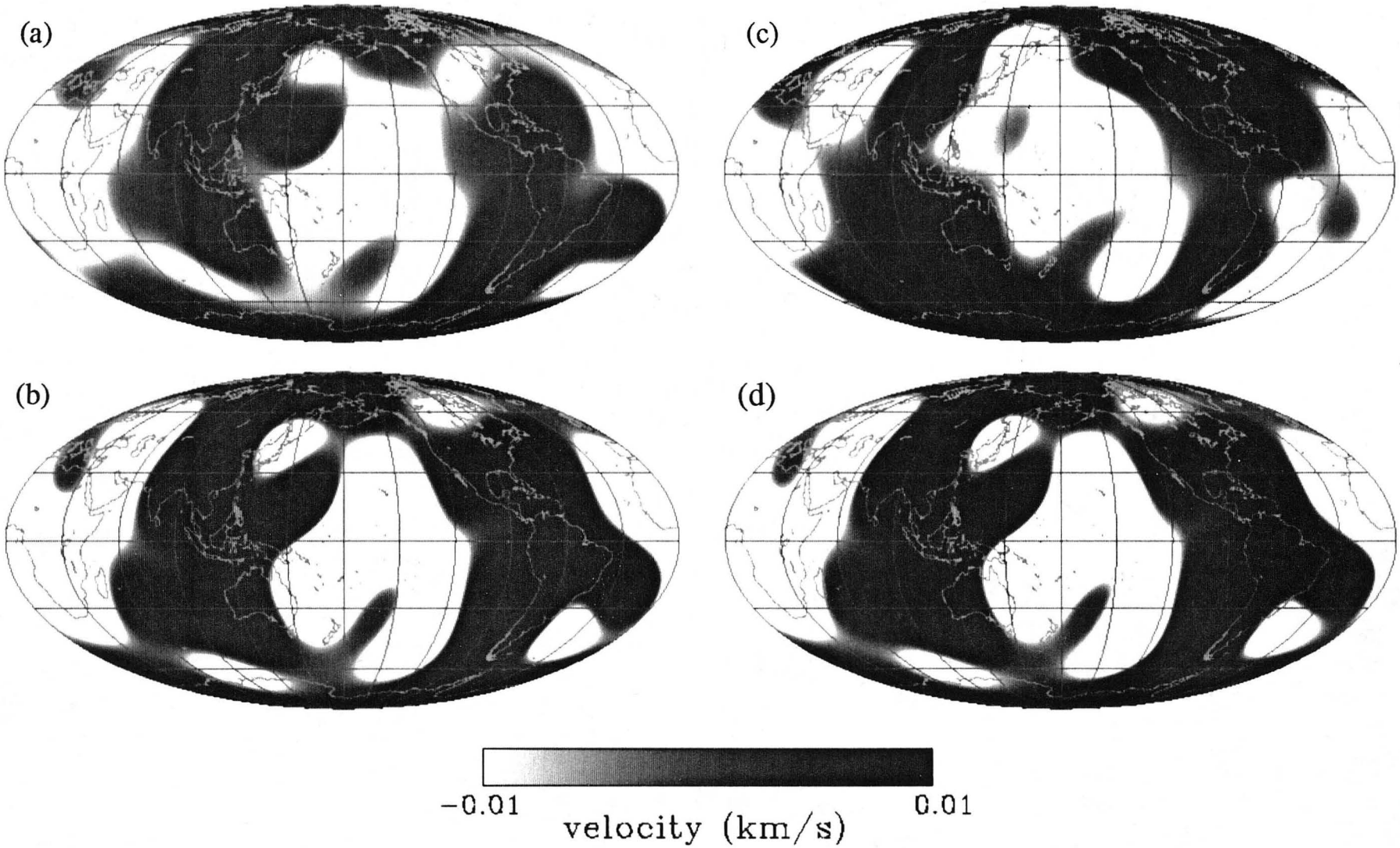

Second lower mantle test
inversion with noise
depth
interval
(km)
maximum
power
(km/s)**2
0-200
0.84E-07
200-400
0.15E-06
400-670
0.34E-06
670-870
0.15E-05
870-1070
0.23E-05
1070-1270
0.16E-05
1270-1470
0.14E-05
1470-1670
0.15E-05
1670-1870
0.18E-05
1870-2070
0.13E-05
2070-2270
0.14E-05
2270-2470
0.15E-05
2470-2670
0.24E-05
2670-2898
0.11E-05
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
angular degree (l)

Figure 3.26 (d)

Figure 3.27



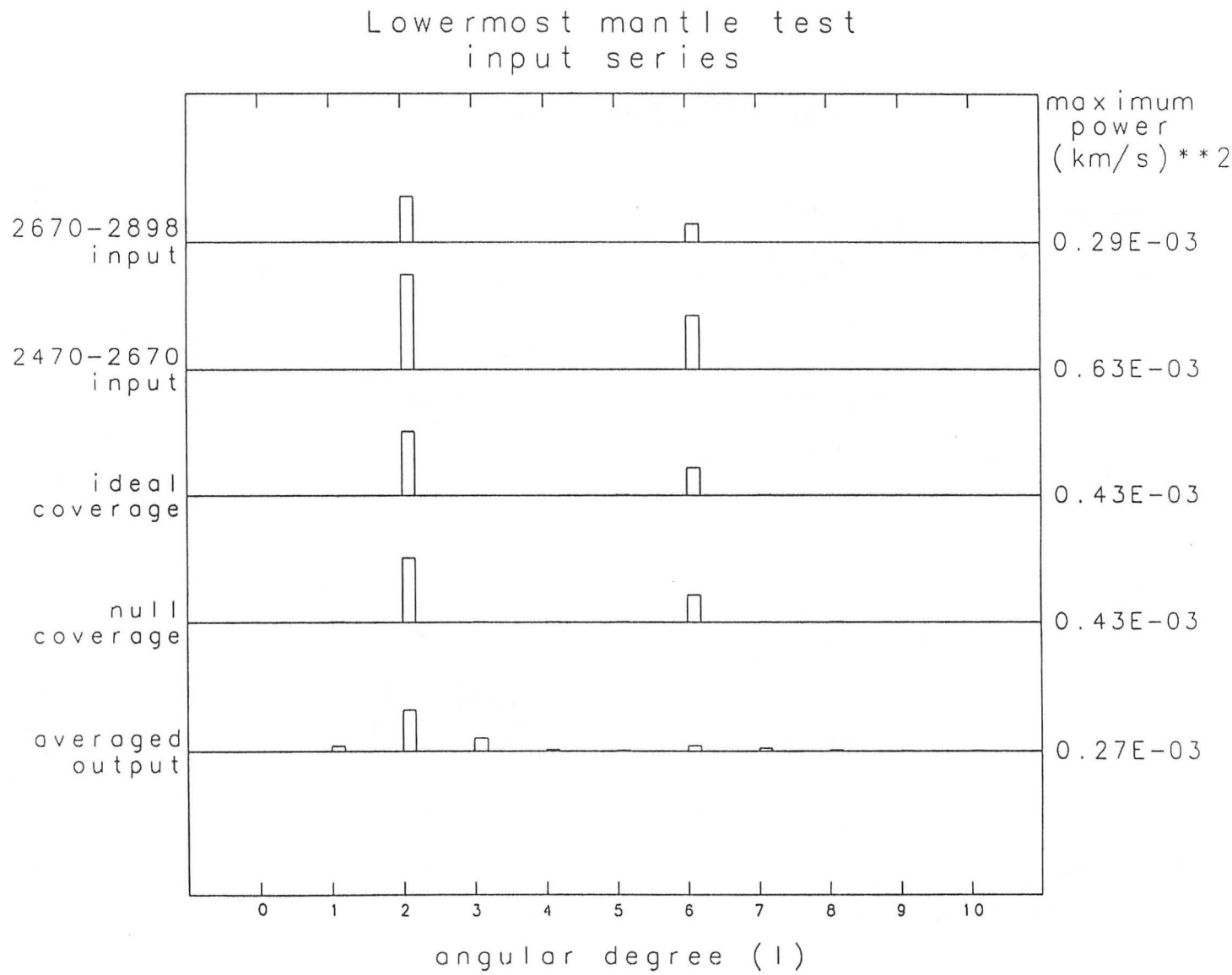

Lowermost mantle test
input series
maximum
power
(km/s)**2
2670-2898
input
0.29E-03
2470-2670
input
0.63E-03
ideal
coverage
0.43E-03
null
coverage
0.43E-03
averaged
output
0.27E-03
0
1
2
3
4
5
6
7
8
9
10
angular degree (l)

Figure 3.28 (a)

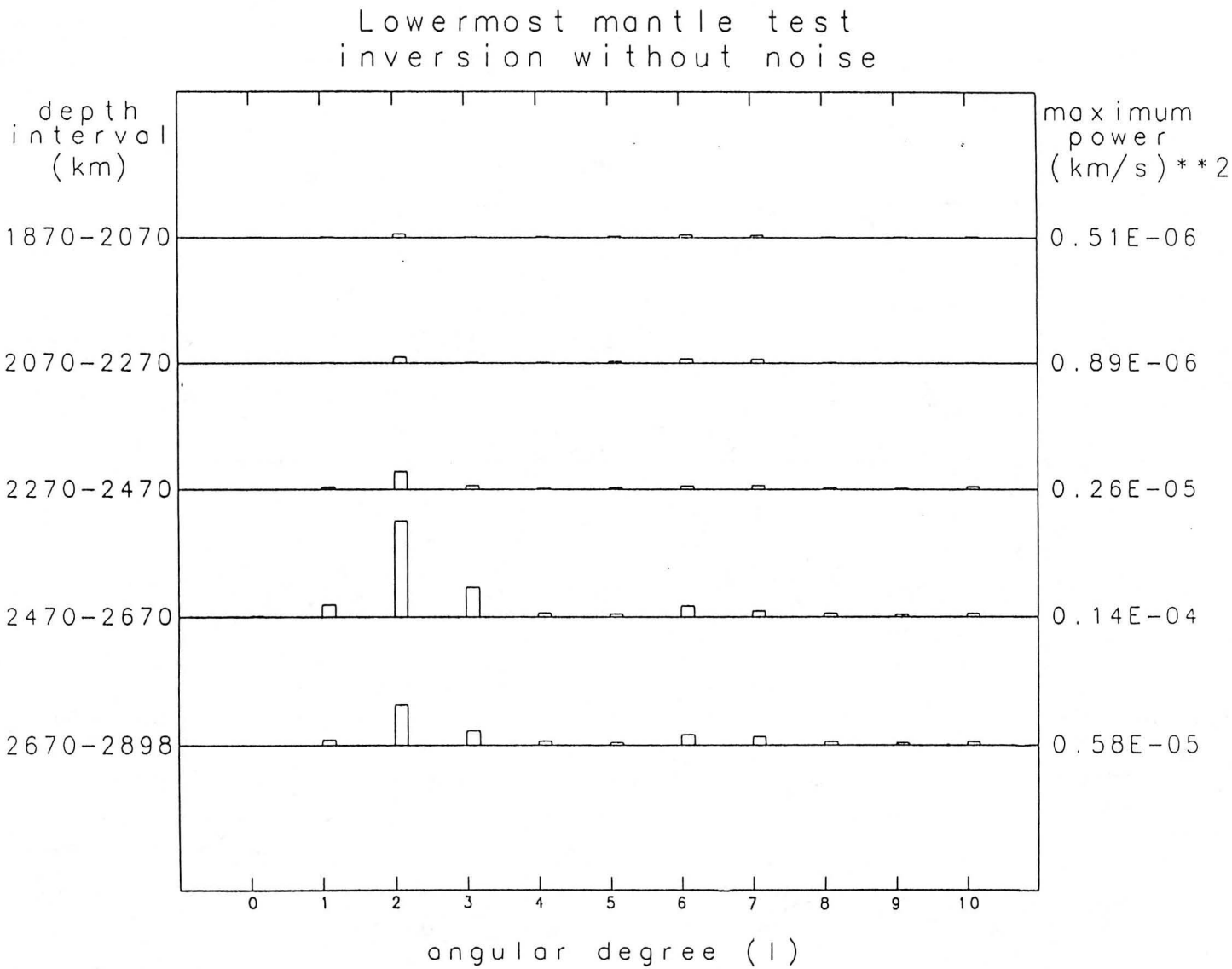

Lowermost mantle test
inversion without noise
depth
interval
(km)
maximum
power
(km/s)**2
1870-2070
0.51E-06
2070-2270
0.89E-06
2270-2470
0.26E-05
2470-2670
0.14E-04
2670-2898
0.58E-05
0
1
2
3
4
5
6
7
8
9
10
angular degree (l)

Figure 3.28 (b)

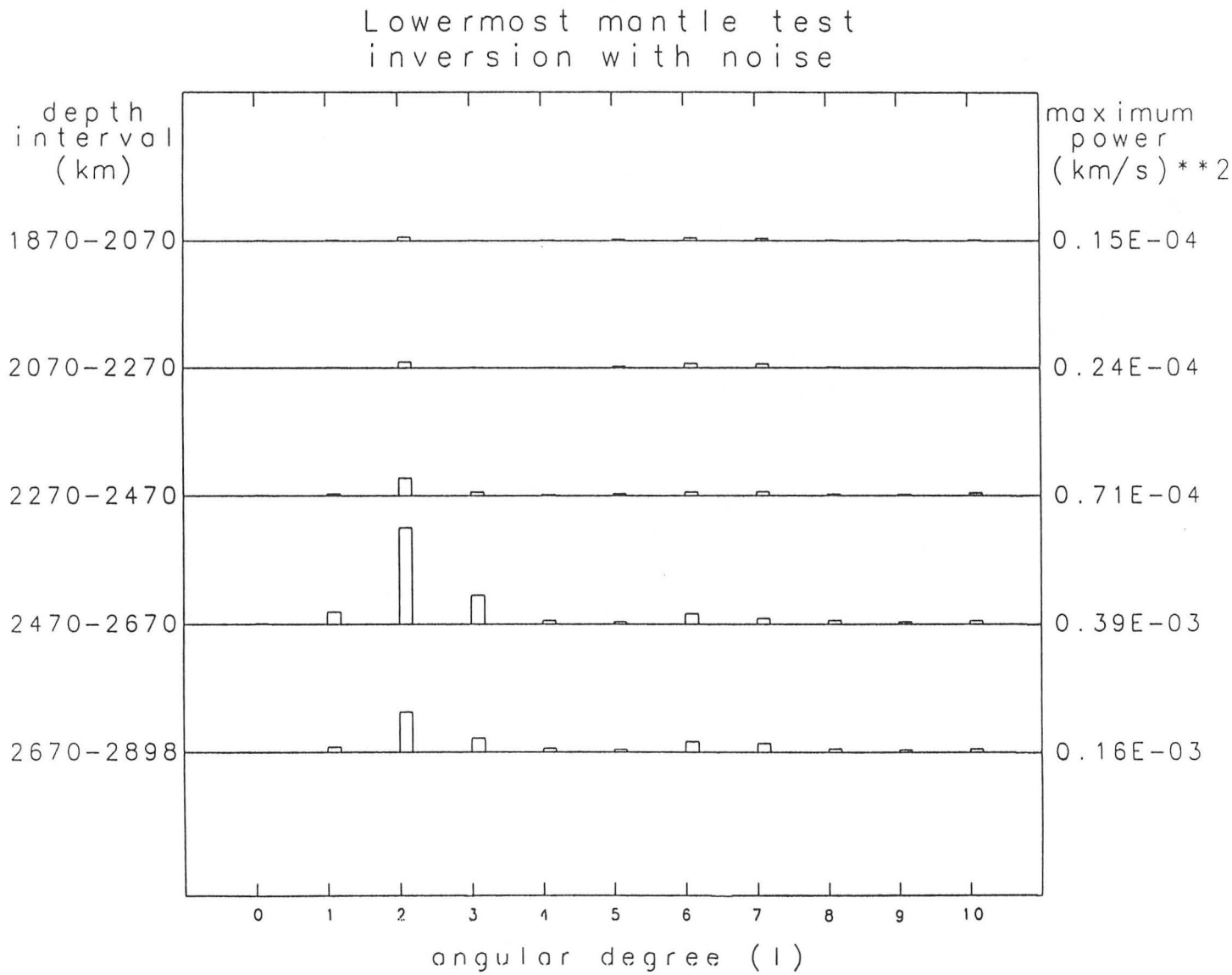
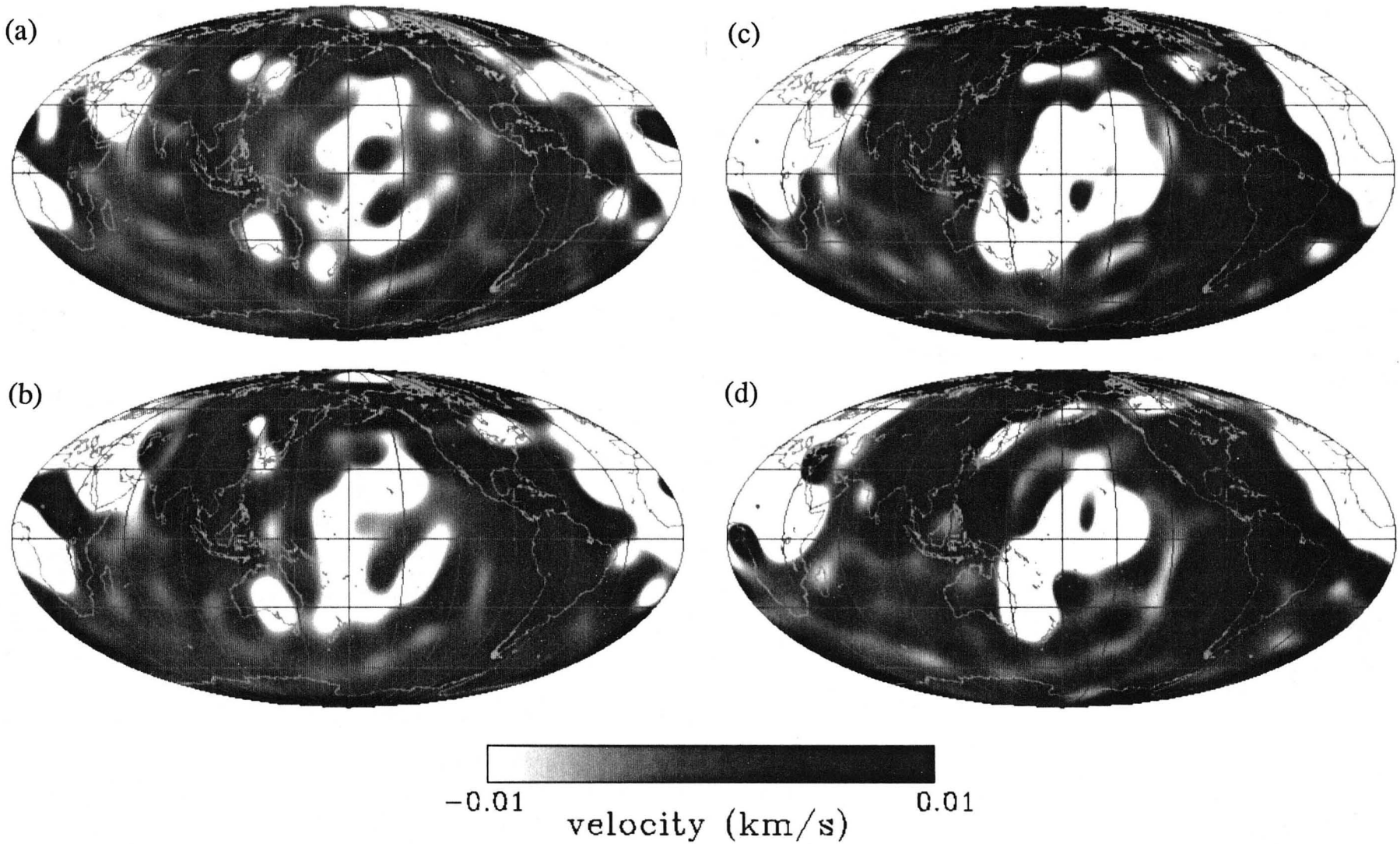

Lowermost mantle test
inversion with noise
depth
interval
(km)
maximum
power
(km/s)**2
1870-2070
0.15E-04
2070-2270
0.24E-04
2270-2470
0.71E-04
2470-2670
0.39E-03
2670-2898
0.16E-03
0
1
2
3
4
5
6
7
8
9
10
angular degree (l)

Figure 3.28 (c)

Figure 3.29



Chapter 4

A Simultaneous, Tomographic Inversion of ISC Travel-Time Residuals for Mantle P velocity, Source Mislocations, and Station Corrections

4.1 Introduction

Attempts to image the three-dimensional seismic heterogeneity of Earth's mantle differ in their approaches to parametrizing the model mantle and in their formulation and solution of the constraining equations. Popular model parametrizations include regionalization of the crust and mantle based on geographical association with surface tectonic processes, spherical harmonic series expansion of the anomalous velocity field, cubic splines, and division into a number of non-overlapping blocks. No one parametrization has been demonstrated to be clearly superior to the others. The bias inherent in a tectonic regionalization makes such a scheme inappropriate for studies of the lower mantle, uneven and incomplete ray coverage makes it useful for studies of the upper mantle. Spherical harmonic expansions and cubic splines require fewer terms to describe a model to the same level of detail as a block parametrization, but do not offer the blocks' geometrical simplicity. With independent block parameters, one may examine the ray sampling, resolution, and covariance of a geographical location more easily.

In addition to differences in model parametrization, studies differ in their construction of the system of equations to be solved and the numerical methods employed to solve them. Early efforts to map the three-dimensional velocity structure of Earth, all of which use a block parametrization, were limited in their structural detail by numerical methods that calculate the explicit inverse of the coefficient matrix. Those methods allow the formal calculation of covariance and resolution matrices to evaluate

the reliability of the model, but severely restrict the number of parameters available to describe the model due to limitations of existing computers. Increased detail became possible when iterative, approximate techniques were employed to solve the system of constraining equations. The cost of this increased detail was that resolution and covariance could not be calculated formally and presented alongside the mantle model. Several means to approximate resolution and covariance were developed in conjunction with SIRT backprojections and other methods. For example, some workers calculate a "point spread function", a column of the non-symmetric resolution matrix, and invert model error distributions to investigate the propagation of errors in the data through the algorithm to the solution. Others perform an inversion for a set of anomalies, each of which extends beyond the bounds of a single voxel, distributed throughout the model as a means to evaluate resolution.

While global studies performed to date vary in their methods for obtaining model solutions and in their approaches to parametrizing the model, the data involved in each study of P-velocity are nearly identical. Since 1964, the International Seismological Center has collected seismic arrival times from around the world. They employ these times in a sophisticated procedure in which times are associated into "events" and the events are located with P-arrival times and the one-dimensional, Jeffreys-Bullen (J-B) travel time tables and a standard least-squares technique [*Adams et al.*, 1982]. For the purposes of seismic imaging and accurate location of events, the geographical distribution of sources (earthquakes and large explosions) and seismographic stations, that so far are located almost exclusively on continents, is unfortunate. With the oceans nearly empty of stations, large portions of Earth, particularly in the southern hemisphere, are poorly sampled by recorded seismic energy, and event locations are inadequately constrained geometrically. Also, using the one-dimensional J-B model, which has known deficiencies in its representation of the spherically averaged structure of Earth in addition to its inability to account for lateral velocity variations, produces

inaccurate event locations. The problems surrounding these inaccuracies in source locations form a central concern of this thesis.

In this chapter we use ISC P-arrival data from January 1964 through January 1987 (frontispiece) and the LSQR algorithm to solve for a three-dimensional P-velocity model of the mantle, source mislocations, and station corrections. Reliability of our model is checked by approximate means, and the models produced by *Dziewonski* [1984], *Clayton and Comer* [1983; *Hager and Clayton*, 1988], and *Inoue et al.* [1990] provide valuable comparisons. Our study differs from *Clayton and Comer's* [1983; *Hager and Clayton*, 1988] in that we use the LSQR algorithm instead of the SIRT algorithm to solve the constraining equations, and that we solve for source and station terms simultaneously. In contrast to *Dziewonski* [1984] who parametrizes the model mantle with spherical harmonics, and solves the equations with the generalized inverse, we use a block parametrization and the LSQR algorithm. *Inoue et al.'s* [1990] procedure severely downweights outliers, while ours does not. The consequences of disregarding outliers are discussed more fully later in this chapter. Also, we solve for source and station terms differently than *Inoue et al.* [1990], use summary rays formed from the entire data set instead of using subsets consisting of actual rays, and perform the model smoothing differently. A companion study to this one, undertaken by *Vasco et al.* [1990] with the same data set, considers the case in which the l^1 norm of the travel time residuals is minimized, rather than the l^2 norm. The former case more closely approximates *Inoue et al.'s* [1990] treatment of outlying residuals than the l^2 minimization performed in this chapter.

4.2 Method

4.2.1 Model Parametrization

The starting model used in this study is a one-dimensional, spherically-symmetric P-velocity model modified from *Jeffreys* [1960]. Modifications to the Jeffreys

velocity model were necessary to obtain a model consistent with the J-B travel-time tables [*Jeffreys and Bullen*, 1940]. These modifications are small but important because they remove a systematically slow trend for the mid-mantle from the model published by Jeffreys and make the model more consistent with the tables, that are used by the the ISC to find source locations. We use the modified J-B model rather than PREM [*Dziewonski and Anderson*, 1981] or the Herrin model [*Herrin et al.*, 1968], because in tests it provides equally accurate locations for sources with known hypocenters, and because using a model consistent with the J-B tables, that are used by the ISC to locate the events, eliminates the need to relocate all events in a new velocity model.

Our three-dimensional model is parametrized with approximately equal-area volume elements (voxels), 5° by 5° laterally and 14 shells, generally 200 km thick, for a total of 22,876 voxels (see figure 4.1 and table 1). The exceptions to the 200 km thick shells occur in the upper mantle, in order to place a radial boundary at the 670 km discontinuity, resulting in a 270 km thick shell, and above the core-mantle boundary, where the lowermost shell is 228 km thick.

We write the vector of travel time residuals for a single event as the sum of three terms:

$$\delta \mathbf{t}_j = \mathbf{t}_j^{obs} - \mathbf{t}_j^{calc} = \delta \mathbf{t}_j^{hypo} + \delta \mathbf{t}_j^{model} + \delta \mathbf{t}_j^{station}, \tag{4.1}$$

where the length of all vectors, m_j, is equal to the number of arrival times reported for the jth event. The hypocentral term itself has four contributions: so

$$\delta \mathbf{t}_j^{hypo} = \frac{\partial t_i}{\partial h_1}\delta h_{1_j} + \frac{\partial t_i}{\partial h_2}\delta h_{2_j} + \frac{\partial t_i}{\partial h_3}\delta h_{3_j} + \frac{\partial t_i}{\partial h_4}\delta h_{4_j}, \tag{4.2}$$

where

t_i = the travel time for the ith ray,

δh_{1_j} = origin time correction for the jth event,

δh_{2_j} = latitude correction for the jth event,

δh_{3_j} = longitude correction for the jth event,

δh_{4_j} = depth correction for the jth event,

$i = 1,2,...,m_j;\ j = 1,2,...,n_e$,

m_j = the number of arrival times for the jth event,

n_e = the number of events in the data set.

In matrix notation,

$$\delta \mathbf{t}_j^{hypo} = \mathbf{H}_j \delta \mathbf{h}_j, \tag{4.3}$$

where

$(H_{il})_j = \dfrac{\partial t_i}{\partial h_l}$ = matrix of source mislocation partial derivatives,

$l = 1,2,3,4,$

$\delta \mathbf{h}_j$ = vector of hypocenter corrections for event j.

The "model" term in equation (4.1), $\delta \mathbf{t}_j^{model}$, represents the deviation of travel times predicted by our starting velocity model, $\hat{c}(\mathbf{r})$ from travel times through Earth's actual velocity structure, $c(\mathbf{r})$, for the jth event. For single travel time residual, δt_i^{model}, we assume that our starting velocity model is within a few percent of the true structure and seek to reconcile the discrepancy by solving for the perturbation term,

$$\delta t_i^{model} = \int_{S_i} \frac{ds}{c(\mathbf{r})} - \int_{\hat{S}_i} \frac{ds}{\hat{c}(\mathbf{r})} \tag{4.4}$$

$$\approx \int_{\hat{S}_i} \left(\frac{1}{c(\mathbf{r})} - \frac{1}{\hat{c}(\mathbf{r})}\right) ds$$

$$\approx -\int_{\hat{S}_i} \frac{\delta c(\mathbf{r})}{\hat{c}^2(\mathbf{r})} ds,$$

where $\hat{S}_i$ is the path of the ith ray through the starting velocity model, $\hat{c}(\mathbf{r})$.

We represent the velocity perturbations as a finite linear combination of "basis" functions,

$$\delta c(\mathbf{r}) = \sum_{k=1}^{n} \gamma_k f_k(r), \tag{4.5}$$

and choose a set of local basis functions in which the medium under investigation is divided into non-overlapping cells, or voxels. Following *Nolet* [1987], let

$$f_k(\mathbf{r}) = \begin{cases} 1 & \text{if } \mathbf{r} \text{ is in cell } k \\ 0 & \text{elsewhere} \end{cases} . \tag{4.6}$$

Our choice of a local basis is arbitrary in many respects. We prefer the block parametrization to a series representation because it allows a more accurate assessment of ray sampling of Earth and the resulting coefficient matrix is quite sparse. This sparseness may be exploited to solve the matrix problem efficiently. Fewer terms are required to describe the model to the same level of resolution with the global spherical harmonic basis (fewer by up to an order of magnitude), but the coefficient matrix in the spherical harmonic case is dense.

An expression for the travel time perturbations in terms of velocity perturbation basis functions results:

$$\delta t_i^{model} = \sum_{k=1}^{n} -\int_{\hat{S}_i} \frac{\gamma_k f_k(\mathbf{r})}{\hat{c}(\mathbf{r})^2} ds = \sum_{k=1}^{n} A_{ik} \gamma_k, \tag{4.7}$$

where

$$A_{ik} = -\int_{\hat{S}_i} \frac{f_k(\mathbf{r})}{\hat{c}(\mathbf{r})^2} ds.$$

In matrix form,

$$\delta \mathbf{t}^{model} = \mathbf{A}\boldsymbol{\gamma}. \tag{4.8}$$

The "station" term of equation (4.1) includes errors in observed travel time residuals resulting from incorrect observations, such as instrument errors and systematic phase mispicks or misidentifications at a particular station, or from inaccuracies in the starting model near the station that occur on a scale too small to be resolved by the 3-D model parametrization. We express these contributions to the travel time residual as

$$\delta \mathbf{t}^{station} = \mathbf{S}\boldsymbol{\mu}, \tag{4.9}$$

where

$$S_{ik} = \begin{cases} 1 & \text{if } k = \text{station number} \\ 0 & \text{if } k \neq \text{station number,} \end{cases}$$

$\mu_k =$ the station correction for the kth station.

Substituting equations (4.3), (4.8), and (4.9) into equation (4.1), we find the problem we wish to solve is

$$\delta \mathbf{t} = \mathbf{A}\,\boldsymbol{\gamma} + \mathbf{H}\,\delta \mathbf{h} + \mathbf{S}\,\delta\boldsymbol{\mu}, \tag{4.10}$$

where

$\delta \mathbf{t} \in R^{M\times 1}$ = vector of travel time residuals,

$\mathbf{A} \in R^{M\times n_p}$ = matrix of ray segments in voxels,

$\boldsymbol{\gamma} \in R^{n_p\times 1}$ = vector of coefficients in the expansion of perturbations to the starting model,

$\mathbf{H} \in R^{M\times 4n_e}$ = matrix of partial derivatives for all events,

$\delta \mathbf{h} \in R^{4n_e\times 1}$ = vector of perturbations to the hypocenters,

$\mathbf{S} \in R^{M\times n_s}$ = matrix of partial derivatives for stations,

$\delta\boldsymbol{\mu} \in R^{n_s\times 1}$ = vector of station corrections,

M = number of data (reported arrivals),

n_e = number of events,

n_p = number of model blocks,

n_s = number of reporting stations.

4.2.2 Simultaneous vs. Progressive Inversion

At this point we may combine matrices into a single, partitioned matrix and solve for all parameters simultaneously, or take advantage of the problem's natural separation into three distinct classes of parameters and solve for each class progressively. Each approach has its appeal. Simultaneous inversion is simpler conceptually and requires fewer computational operations, but deals with a much larger matrix ($O(M \times n_p + 4n_e + n_s)$), so it demands more core memory, even in row-active implementations. Further, due to the different nature of the parameters to be estimated and their differing scales, results are very sensitive to the scaling applied to the coefficient matrix. Progressive inversion, on the other hand, enables us to exploit the natural separation of the matrix problem to solve for each set of parameters in a step-wise fashion, reducing demands for core memory and eliminating the need for careful scaling of parameter classes. This approach follows *Pavlis and Booker* [1980], *Spencer and Gubbins* [1980], *Jordán and Sverdrup* [1981], *Kennett and Williamson* [1988] and *O'Connell and Johnson* [1991], among others, and allows a more detailed analysis of resolution and uncertainty in the determination of mislocation terms than would be practical otherwise. Unfortunately, the projections involved for each source mislocation matrix tend to fill in elements in originally sparse coefficient matrices, which increases the computation time required to solve the complete problem and increases disk-access time for row-active algorithms. This last point is critical for global-scale problems involving large data sets and large numbers of model parameters. Still, each method is feasible for problems involving matrices of the order $10^6 \times 10^4$. Preliminary results of a progressive inversion algorithm have been presented [*Pulliam and*

Johnson, 1989b], as well as synthetic tests of the algorithm [*Pulliam and Johnson*, 1989a] (see chapter 2).

4.2.3 Simultaneous Inversion

Here we choose to combine the three coefficient matrices and solve for all parameters simultaneously, i.e.

$$\left[\mathbf{A}|\mathbf{H}|\mathbf{S}\right]\begin{bmatrix}\boldsymbol{\gamma}\\ \hline \boldsymbol{\delta}\mathbf{h}\\ \hline \boldsymbol{\delta}\boldsymbol{\mu}\end{bmatrix} = \boldsymbol{\delta}\mathbf{t}, \tag{4.11}$$

or

$$\mathbf{G}\mathbf{x} = \boldsymbol{\delta}\mathbf{t}, \tag{4.12}$$

where

$$\mathbf{G} = \left[\mathbf{A}|\mathbf{H}|\mathbf{S}\right]$$

and

$$\mathbf{x} = \begin{bmatrix}\boldsymbol{\gamma}\\ \hline \boldsymbol{\delta}\mathbf{h}\\ \hline \boldsymbol{\delta}\boldsymbol{\mu}\end{bmatrix}.$$

H and **S** are first scaled so that each row has the same euclidean norm as the same row of **A**.

Equation (4.12) presents us with a classical linear inverse problem. Typically, the $M \times N$ coefficient matrix, **G**, will have many more data than parameters ($M \gg N$) and, given that errors are contained in the data, the equations will almost surely be inconsistent. As a criterion for minimizing the misfit of parameters to data we choose the euclidean (l^2) norm, resulting in the least squares problem:

$$Min\,\|\mathbf{G}\mathbf{x} - \boldsymbol{\delta}\mathbf{t}\|^2 = Min\,(\mathbf{G}\mathbf{x} - \boldsymbol{\delta}\mathbf{t})^T(\mathbf{G}\mathbf{x} - \boldsymbol{\delta}\mathbf{t}). \tag{4.13}$$

This is not a simple choice. Strictly, least-squares is most appropriate for problems involving a Gaussian distribution of errors. When applied to such a distribution, least-squares produces the maximum likelihood solution to the linear matrix equation. However, residuals contained in ISC travel time data are not clearly Gaussian [*Buland*, 1984; *Vasco et al.*, 1990]. More observations are found in the distribution's tails than one would expect in a Gaussian distribution. We may transform the entire matrix problem to one involving a Gaussian distribution of residuals by applying a set of weights to rows of the problem (4.12). These weights may be obtained through uniform variance reduction analysis of the original vector of residuals, $\delta \mathbf{t}$ [*Jeffreys*, 1939; *Buland*, 1986]. Alternatively, we can remove the bulk of the blunders and gross random errors by truncating the distribution. We choose to truncate the summary residual distribution (figure 4.7) at ±7 seconds. This truncation value assures us we will not discard too many reliable observations and, as a test of whether we are keeping too many unreliable observations, we apply the uniform variance reduction method to our truncated set of residuals. After inverting the sets of modified and unmodified summary residuals, we find differences in velocities for individual voxels on the order of 0.01%. Apparently the outliers are sufficiently few in number, relative to the central portion of the distribution, that the influence they exert on the final model is minor.

Analysis of a set of travel time residuals to which corrections for a three-dimensional model are applied reveals that their distribution may be more similar to a two-sided exponential than a Gaussian. If this is true, minimizing the l^1 norm would be more appropriate than l^2 minimization. Since the l^1 norm is less sensitive to outliers in the distribution of residuals, gross errors in the dataset -- due to mis-identification of phases, mis-readings of time codes, faulty instruments, and source mislocation, for example -- would be less likely to propagate through the inversion to the model. On the other hand, some of the outlying residuals in the tails of the distribution constitute real and significant data, indicating relatively large velocity

differences between the real earth and our starting model. Raytracing from "calibration" events, sources with known locations, produce travel time residuals amounting to 5 seconds or more at some stations. Providing the residuals resulting from the more extreme velocity anomalies in the real earth do not violate the assumptions under which we linearized the originally nonlinear travel time problem, these data are the ones we wish to emphasize in the inversion, not the extreme errors or the minor deviations clustered around zero. Both l^2- and l^1-minimization approaches warrant our attention in order to compare resulting models. The l^1-minimization is pursued in a companion study by *Vasco et al.* [1990] using the same dataset.

From here on our development parallels the development of the inverse problem in chapter 2. Ultimately we arrive at the modified set of equations

$$\begin{bmatrix} \tilde{\mathbf{b}} \\ 0 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{G}} \\ \lambda \mathbf{B} \end{bmatrix} \tilde{\mathbf{x}}, \tag{4.14}$$

where

$$\tilde{\mathbf{G}} = \mathbf{W}_d \mathbf{G} \mathbf{W}_x,$$

$$\tilde{\mathbf{b}} = \mathbf{W}_d \delta \mathbf{t},$$

$$\mathbf{x} = \mathbf{W}_x \tilde{\mathbf{x}}$$

subject to the minimization of $\|\tilde{\mathbf{x}}\|^2$. Our estimate of $\mathbf{W}_d$ comes from the standard errors of ISC travel time residuals as a function of epicentral distance. Assuming the data are independent, we form the diagonal matrix

$$\mathbf{W}_d = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2} & \dots & 0 \\ . & 0 & \dots & . \\ . & . & \dots & . \\ . & . & \dots & . \\ 0 & 0 & \dots & \frac{1}{\sigma_M} \end{bmatrix}$$

M = number of rays in the data set,

σ_i = standard errors of ISC travel time residuals,

$i = 1,2,\dots,M$.

Our estimate of $\mathbf{W}_x$ is

$$W_{x_{kk}} = \frac{\sum_i^{n_k} l_i}{n_k v_k^{\frac{1}{3}}},$$

where

l_i = the length of the ith ray segment in voxel k,

v_k = the volume of the kth voxel,

n_k = the number of ray segments in the kth voxel.

We solve these equations with the LSQR algorithm.

4.3 Data

4.3.1 Data Selection

The data inverted in this study were obtained from the catalog of the International Seismological Centre (ISC) for the period January 1964 through January 1987 (frontispiece). To avoid contamination of our mantle phases by Earth's core we limit the

range of our coverage to epicentral distances between 0° and 96°. The confusing scatter caused by refractions from the 400 km and 670 km discontinuities at about 15° to 25° is dealt with in the inversion process by weighting each summary ray by the inverse of the standard deviation of travel time residuals as a function of delta. A single event must have a minimum of twenty-five reporting stations, and source depths as reported by the ISC must be greater than 0 km and less than 70 km. We discard all events located by the ISC at Earth's surface, but retain events located at the other default depths. All observations are corrected for ellipticity by integration along the raypath. Travel time residuals are formed by subtracting the time calculated by tracing rays through the spherically-symmetric starting velocity model from the observed time corrected for ellipticity. Raytracing is performed by a shooting method involving the direct numerical integration of the eikonal equations with an integration scheme that checks the local error at each integration step. Lengths of ray segments in voxels are found by integrating distance along the curved raypath and finding the intersections of rays with voxel boundaries. Rays associated with residuals greater than seven seconds are discarded. Approximately 3.02 million rays satisfy these criteria. Figures 4.2 and 4.3 show the locations of the selected events and seismographic stations, respectively.

We examined histograms of travel time residuals associated with events assigned by the ISC to five different default depths, 0, 5, 10, 15, and 33 km, and compared the residual distributions for these events to the residual distribution of remaining events. Due to a problem with our FORTRAN subroutine, ISC records in which the source depth was left blank defaulted to zero source depth. Although the residual distribution shows a clear bimodal pattern, we were unable to distinguish reliably between true, zero-depth locations and blank-depth locations after the data had been extracted from the ISC master set. For this study we discarded all events with source depth equal to zero. However, histograms for the four remaining default depths are indistinguishable from the histogram for all remaining events, so we cannot justify culling events with

source depth = 5, 10, 15, or 33 km. A histogram of the culled data set is shown in figure 4.4 along with the first four moments of the travel time residual distribution

4.3.2 Summary Rays

Summary rays were formed as composites of rays that sample nearly the same portion of Earth. Bins consist of 2° by 2° voxels divided into seven depth intervals of 10 km each, for a total of 71,904 bins. Residuals of rays emanating from and ending in the same two bins are averaged into a single, summary ray. These bins are quite small, even compared to our model voxels. The result is a set of summary rays in which most are composed of very few actual rays, typically two or three. We apply no minimum cutoff, so nearly half of the resulting rays consist of just a single ray, not a composite at all. By constructing summary rays we seek to reduce the redundancy of the data set, in order to mitigate the effects of nonuniform sampling of Earth on our final model, and remove variations in travel times due to heterogeneity on a scale smaller than our velocity model parametrization. At the same time we would like to preserve as much of the original variation of residuals as possible. We wish to allow the inversion algorithm to reconcile the discrepancies in travel times, rather than remove these discrepancies in a pre-processing step. This approach allows us to evaluate the performance of our algorithm using test cases that better represent the true case of inconsistent and erratic travel times in Earth. Figures 4.5 and 4.6 show the locations of summary events and summary stations.

Our final data set consists of a total of 725,993 summary rays emanating from 5,986 summary events. Figure 4.7 (compare to figure 4.4) shows the mean, variance, and skewness of the data are reduced significantly by constructing summary rays. However, the statistics of the two distributions are not directly comparable because we truncate the summary residuals at ±7 seconds. Compared to the distribution of actual residuals, the distribution of summary residuals is slightly more like a Gaussian

distribution, as indicated by the sizes of its tails, and slightly less like a two-sided exponential, as indicated by a comparison of cumulative distributions, than is the distribution of actual residuals. Ultimately, the residuals remaining after three-dimensional structure is accounted for will resemble a two-sided exponential distribution more closely than a Gaussian.

4.3.3 Ray Coverage

Figures 4.8a-f show the distribution of ray segment lengths that make up our coefficient matrix, **A**. These values consist of column sums, indicating the total sampling of each individual voxel by the data set used in this study. In the absence of a weighting matrix, $\mathbf{W}_x$, that balances the column norms, results of an inversion would be expected to follow this pattern quite closely. Figure 4.8a shows the clear demarcation of plate boundary source regions that are well-sampled. Asia, North America, Europe, and Australia are also well-sampled. In contrast, other regions tend to be quite poorly-sampled. There are also large oceanic areas that are completely unsampled by our data set. These voxels do not enter into the inversion. The next depth layer, figure 4.8b, shows a broadening of the well-sampled regions and a slight reduction of the unsampled oceanic areas. At 400-670 km and 670-870 (figures 4.8c and 4.8d) these trends continue, and by the mid-mantle (figure 4.8e) virtually all voxels are sampled. In general, sampling becomes more homogeneous with depth, and at the bottom of the mantle (figure 4.8f) the sampling is much more uniform than in the first layer. However, in absolute numbers the sums of ray segments in voxels decrease with depth, even as more voxels are sampled in each layer. Table 4.1 details the average number of hits for sampled voxels in each layer along with the the average sum of ray segments in a voxel at a given depth and the number of voxels sampled in each depth interval. These averages include only voxels that have non-zero sampling. The trends in Table 4.1 show that while homogeneity of sampling increases with depth, voxels tend to be less frequently and less heavily traversed by recorded seismic rays. Note in

all six figures the strong bias toward the northern hemisphere and toward continents. As we will discover in our treatment of resolution and covariance, the uneven mantle sampling translates directly into uneven constraints and resolution for our final model.

3.4 Stochastic vs. Deterministic Analysis

Gudmundsson et al. [1990] show that there exists a minimum level of stochastic noise in the ISC data set below which we cannot expect to resolve structure and which casts doubt upon the reliability of schemes such as ours to resolve the apparently small velocity anomalies in the lower mantle. This noise might arise from reading or instrument errors or very small-scale structure in the upper mantle that causes multi-pathing, or a breakdown of the ray approximation in general. However, *Gudmundsson et al.'s* [1990] analysis indicates the level at which the behavior of travel time residuals becomes non-systematic is well below the starting level of our data. For the scale of our model blocks, 5° × 5°, their figures show maximum extrapolated variances of somewhat greater than 2 sec^2 at the distance range contaminated by reflections from discontinuities, 15° to 25°, and averaging slightly less than 1 sec^2 outside this range. In contrast, our residuals average variance is 3.1 sec^2 for the raw travel time residuals, which is reduced to 2.6 sec^2 after the formation of summary rays and 2.1 sec^2 after inversion. Our variance values are not consistent with those of *Gudmundsson et al.* [1990] because they choose to truncate the residual distribution at 4 sec, whereas we truncate at 7 sec. Clearly, even if half the original variance in travel time residuals cannot be accounted for by our inversion method, the half that can be accounted for is substantial and significant. *Gudmundsson et al.* [1990] note as well that the signal to random noise ratio in the teleseismic ISC P-wave data is about $S/N \approx 2$. Their other results elegantly confirm previous indications that upper mantle structure is significantly more heterogeneous than mid- and lower-mantle structure. These results do not stand in the way of imaging lower mantle structure, where that structure is comparable in amplitude to upper mantle heterogeneity, with the portions of the travel

time residuals that do vary systematically.

4.4 Resolution and Uncertainty

Equally important to producing a seismic velocity model is a thorough investigation of the reliability of that model. We need to evaluate the "resolution," the image of an input model as seen through the "filter" of the inverse method, and the "uncertainty," the errors contained in our output resulting from errors in the input data that propagate through the inversion. For a discrete problem of the form (4.14) in which the data contain Gaussian errors, the estimate of uncertainty takes the form of an a posteriori covariance matrix [*Tarantola*, 1987],

$$\mathbf{C}_{\hat{m}} = (\tilde{\mathbf{G}}^T \mathbf{C}_d^{-1} \tilde{\mathbf{G}} + \mathbf{B}^T \mathbf{C}_m^{-1} \mathbf{B})^{-1}. \tag{4.15}$$

The resolution matrix is then

$$\mathbf{R} = \mathbf{C}_{\hat{m}} \tilde{\mathbf{G}}^T \mathbf{C}_d^{-1} \tilde{\mathbf{G}}. \tag{4.16}$$

However, due to the large numbers of data and model parameters required to image Earth's interior to a useful level of detail, formal calculation of covariance and resolution matrices has been beyond our computational capacity. Calculation of resolution and covariance matrices in tomographic inversions have been necessarily approximate and incomplete, and a number of methods have been developed to evaluate a tomographic model's reliability.

4.4.1 Approximating Resolution

Humphreys and Clayton [1988] explore the resolution of an inversion by means of a synthetic test in which a velocity perturbation is introduced to one or more voxels in a region of interest. Using Fermat's principle, which holds that travel times calculated through the three-dimensional Earth are insensitive to changes in raypath, one may calculate the travel time residuals that would be produced by the synthetic anomalies without tracing rays in the 3-D model. Synthetic data is constructed with

the same raypaths as the original data set and the same model parametrization. The synthetic data is inverted and the voxels adjacent to the perturbed voxels examined for smeared and "ghost" images that are artifacts of the inversion. The result may be thought of as the response of the algorithm to an impulse introduced to the system, and forms one column of the resolution matrix, which is non-symmetric. *Humphreys and Clayton* [1988] call this vector the "point spread function," distinguishing it from the "resolving kernel" which is the corresponding row of the resolution matrix. One limitation of this approach is that the anomalies are introduced in the span of the parametrization, and the resulting estimate of "resolution" may therefore be misleadingly high.

Figures 4.9-4.14 show point spread functions for six voxels. A 5% velocity anomaly is introduced at (17°S, 178°W, 600 km depth) beneath the Tonga subduction zone (figure 4.9a), and an inversion returns a value of 3.7% for the anomalous voxel (figure 4.9b). Note the relatively small leakage to adjacent voxels, indicating the region is well-resolved by our method. The maximum smeared anomalies occur in voxels directly above and below the perturbed voxel, at 0.7% and 0.6%, respectively. No other smeared anomaly exceeds 0.2%. A second impulse of 3%, shown in figure 4.10a, was placed beneath the New Hebrides subduction complex at (16°S, 166°W, 900 km depth). The inversion returns a value of 2.1% and neighboring voxels returned 0.4%, 0.1%, 0.2%, and 0.1% in the same layer (figure 4.10b). Some streaking along raypaths to the south and north appears at low amplitudes. Both images of anomalies placed in subduction zones (figures 4.9b and 4.10b) display the broadening and smearing with depth noted by *Spakman et al.* [1989] in simulated tests with a much finer model grid. Although the amplitudes of the smeared anomalies are quite small, the smearing is indeed systematic and apparently oriented along the dip of the subduction zone features we will discover later in our model. The voxels above and below return 0.2% and 0.6%, respectively. The voxel two layers below the introduced

anomaly returns 0.3%.

The third and fourth tests examine negative anomalies. The first, shown in figure 4.11a, is introduced to the first layer beneath the Hawaiian islands at (21.4°N, 158°W, 70 km depth). The -3% anomaly is returned as -2.1%, and the maximum spurious perturbation, -0.66%, occurs in the voxel immediately below and to the west of the perturbed voxel (figure 4.11b). Note the smearing in this case into voxels to the west and east, which lie along common raypaths to sources in the northwest Pacific and to North America, respectively. This smearing results from poor geometrical constraint of the voxels beneath Hawaii; most arriving rays travel along parallel paths. An anomaly beneath Iceland produces a similar smearing pattern to the east and west. A -4% anomaly at (65°N, 18°W, 70 km depth) (figure 4.12a) returns as -2.5% with a smeared value in the voxel just below reaching -1.5% and the voxel two layers below reaching -0.3% (figure 4.12b). Figure 4.13a shows a 3% anomaly introduced beneath central South America at (0°, 65°W, 800 km depth). A value of 1.8% is returned by our inversion, with smeared values above reaching 0.4% and below reaching 0.3% (figure 4.13b). A systematic smearing feature dips to the east, possibly along rays emanating from subduction zone events, but at very low amplitudes. A 4% velocity anomaly is introduced at (40°N, 85°W, 800 km depth) beneath the eastern North America (figure 4.14a), and an inversion returns a value of 2.5% for the anomalous voxel (figure 4.14b). Here the leakage to adjacent pixels displays an interesting pattern. As we will see later, a significant anomaly appears in our model through the Carribbean and into central South America that corresponds to the smeared velocities shown here. As in other tests, the largest values appear in voxels directly above and below the perturbed voxel, both at 0.5%. No other value exceeds 0.2%. While the amplitudes of smeared values are small, the systematic error suggests we would not have strong control on the spatial extent of an actual anomaly and warns us to be cautious in the interpretation of results in this region. In addition to the values returned to perturbed voxels

and their immediate neighbors, the test inversion produces several "ghost" images away from introduced anomalies. The largest of these amounts to 0.1%, or one-tenth of the smallest value returned by the inversion for an introduced anomaly.

Inoue et al. [1990] show a way to approximate the resolving kernel for one model parameter, as well as the corresponding row of the covariance matrix. The idea is to use the LSQR algorithm to solve

$$\tilde{\mathbf{G}}^T \tilde{\mathbf{G}} \mathbf{x}_j = \mathbf{e}_j,$$

where $\mathbf{e}_j$ is a vector whose j^{th} element is 1, while all other elements are 0. After $\mathbf{x}_j$ is found, the jth row vector, $\mathbf{y}_j$ of $\mathbf{R}$ is given by

$$\mathbf{y}_j = \mathbf{x}_j \tilde{\mathbf{G}}^T \mathbf{C}_d^{-1} \tilde{\mathbf{G}}.$$

The drawback of calculating a single row or column of the resolution matrix is that each interesting feature must be examined individually with a separate inversion. The result is a visual representation of resolution presented in a number of figures that must be compared simultaneously to the model produced with real data.

A similar, though more complete, approach is to introduce a full model, so that a value is specified for every voxel, and invert the synthetic data generated through this model. *Inoue et al.* [1990] advocate a checkerboard pattern in which adjacent voxels alternate between two extreme values. Let the model checkerboard pattern be

$$\mathbf{m}_{cb} = \beta \sum_j c_j \mathbf{e}_j,$$

where $c_j =$ 1 or -1 and β is a scaling factor. Calculate the synthetic travel times from real sources to the real stations included in the data set (e.g., using the data kernel $\tilde{\mathbf{G}}$), so

$$\mathbf{d}_{cb} = \tilde{\mathbf{G}} \mathbf{m}_{cb}.$$

Use the LSQR algorithm to solve for an estimate of $\mathbf{m}_{cb}$:

$$\hat{\mathbf{m}}_{cb} = \beta \mathbf{R} \sum_j c_j \mathbf{e}_j = \beta \sum_j c_j \hat{\mathbf{e}}_j,$$

that is simply a superposition of point spread functions. The estimated model image may then be compared to the starting model to identify regions with poorly recovered values. *Spakman and Nolet* [1987] use a harmonic function instead of a checkerboard pattern. Both test patterns share the advantage that a more complete sense of resolution for the model may be presented with just a few figures. Unfortunately, they also share the disadvantage that separate inversions must be performed for various input models with different wavelengths and amplitudes. Both patterns are parametric in the sense that we assume before inversion that we have some idea of the spatial scale and amplitudes of interesting features of the real earth. Impulse tests assume we know the location and amplitude of interesting features. Both methods require the same computation time as generating the best-fitting solution itself.

Figure 4.15a shows our input checkerboard model. Values alternate between +0.3 km/s and -0.3 km/s for adjacent voxels. Figure 4.15b shows the results for the top layer, 0-200 km. Well-resolved regions correlate strongly with good ray coverage (figure 4.8a), though continents are generally imaged more clearly than mid-ocean ridges. This discrepancy is probably due to better geometrical constraints on continental voxels. At mid-ocean ridges, rays connecting to seismographic stations propagate more vertically than horizontally, thus rays sampling ridges tend to be parallel. Continental voxels are generally sampled by a more complex set of criss-crossing rays that provide stronger constraints on the continental voxel's velocity. Checkerboard results for subsequent layers are incorporated into maps of our velocity model, which are shown in the next section.

Nolet and Snieder [1990] suggest a less time-consuming means of producing a resolving kernel with a reduced basis, produced by the LSQR algorithm. It is common for coefficient matrices in tomographic problems to be numerically singular, so the

matrix is rank-deficient and the range space may be spanned by a basis that is considerably smaller in dimension than the original matrix. Each iteration of the LSQR algorithm produces a single basis vector of the range space. One must decide when a sufficient number of basis vectors has been produced to represent the solution to the desired degree of accuracy. In the absence of the singular value spectrum, this decision presents a serious problem. *Scales* [1989] offers a way to obtain the singular values from the tri-diagonal matrix also produced by the LSQR algorithm, but he points out that numerical round-off errors can produce artifact entries in the set of singular values. We performed tests on a real, cross-hole tomographic problem in which the results of Scales' method were compared to singular values obtained via SVD and confirmed the deleterious effect of these errors, which do not allow the singular value spectrum to be produced reliably.

Vasco [1991] presents an extremal bound approach to evaluating resolution and uncertainty in a tomographic inversion. Instead of finding a single model that is "best-fitting" in some sense, he finds properties of the range of models that are consistent with the data. While this method is much different in its approach, it shares the computational drawbacks of methods mentioned previously. Each parameter must be considered individually and the computation time required for each block is comparable to the time required to find the entire "best-fitting" model. A subset of the model parameters could be examined, but calculating bounds for every block is not feasible.

4.4.2 Approximating Covariance

Attempting to evaluate covariance, we examine the distribution of travel time residuals and find that it more closely resembles a two-sided exponential distribution of deviates than a Gaussian, though it falls somewhere between the two [*Vasco et al.*, 1990]. To investigate the propagation of these errors through the inversion procedure to the final model, we replace the vector of travel time residuals with a vector of

synthetic residuals that are distributed, in the first case, as a two-sided exponential and, in the second case, as a Gaussian. In each case, the variance of the random distribution after truncation is adjusted to the same level as that of the actual data. We perform 25 inversions with different residual vectors for each case and find corresponding "model" vectors, ε^n, where $n = 1, 25$. An estimate of the model covariance may be obtained as

$$c_{ij} = \frac{1}{N}\sum_{n=1}^{N}\varepsilon_i^n\varepsilon_j^n$$

where

$\varepsilon_i =$ value for voxel i produced by inversion of errors.

Figures 4.16a and 4.16b show the covariance estimates for our model's top layer, 0-200 km, for the Gaussian and exponential cases, respectively. Results for both types of distribution share the same general patterns. Locations of large and small errors are quite similar; only the amplitudes of the errors vary significantly. This difference follows immediately from the main difference between Gaussian and exponential distributions: the exponential distribution has much longer tails than does the Gaussian. Given our inversion procedure, which minimizes the l^2 residual norm and therefore weights large deviates more heavily than small ones, the longer-tailed distribution naturally translates into larger "model" values. The most striking feature of these covariance estimates is their strong correlation with ray coverage. Areas sampled by many rays have large errors associated with them, while sparsely-sampled areas tend to have smaller errors. This reflects the tendency of the LSQR algorithm to image anomalies where they are well-constrained, but to bias model values toward zero where there is inadequate coverage. This tendency is desirable when dealing with real data and a real model; regions of large uncertainty receive small values rather than large values. However, the same tendency renders the values produced by this method poor estimates of

model covariance. In fact, these figures do indicate where random errors in our data are mapped in our model, but with the noise maps alone we cannot assess the reliability of individual model parameters. Perhaps the maps do serve as reliable estimates of model uncertainty in regions that are constrained reasonably well, but distinguishing regions that have reliable estimates from regions which must be dismissed from consideration is not possible.

A second approach to estimating model covariance, the jackknife, is outlined by *Efron* [1982] and applied to seismic data by *Lees and Crosson* [1989]. Here we need not assume a particular distribution for the errors, since model variability is assessed directly from the variability of the data. Unfortunately, our algorithm still comes into play and we must be again be wary of the influence of uneven ray coverage and artifacts of our parametrization. To form the jackknife estimate we perform N inversions of the real data, leaving out a subset of the data without replacement. For each inversion we produce a model

$$\hat{\mathbf{m}}^n = N\hat{\mathbf{m}} - (N-1)\hat{\mathbf{m}}^{\mathbf{n}}$$

$$\hat{\mathbf{m}} = \frac{1}{N}\sum_{n=1}^{N}\hat{\mathbf{m}}^{\mathbf{n}}$$

which has the variance

$$c_{ii} = \sum_{n=1}^{k}\frac{(m_i^n - \hat{m}_i)^2}{N(N-1)}$$

where

N = number of data subsets.

Figure 4.15 shows a map of jackknife standard errors for the 0-200 km depth layer. The general tendency of the jackknife estimate is also to place anomalies in

heavily-sampled regions, but some features distinguish the jackknife estimate of variance from the model error distributions estimates. The jackknife estimates show a weaker correlation with the number of rays sampling a voxel and the sum of ray lengths in a voxel than do the model error distributions' estimates, and shows more variability due to the geometrical distribution of rays sampling a given voxel. If leaving out a few rays produces a velocity estimate for a given voxel much different from previous estimates, the voxel is poorly constrained. However, the known effect of our algorithm again causes us to distrust the results of the jackknife procedure. Poorly constrained voxels are biased towards zero, exactly the opposite tendency we desire for an investigation of covariance.

Keeping in mind our algorithm's tendency, we should simply disregard results in poorly sampled regions and concentrate on portions of Earth that are well-sampled. Distinguishing well- from poorly-sampled regions is easily done with a block model parametrization. We adopt this approach in our presentation of model maps in the next section.

4.5 Results and Discussion

4.5.1 Source Mislocations and Station Corrections

The averaging procedure employed to construct summary rays should reduce both the mislocation of events and the station errors, at least in the cases where several actual stations are averaged into one summary station, and so we expect to see only general features of source mislocations and station errors in our results. These terms are included for the sake of a realistic formulation of the problem, and for stability in the inversion. When an inversion is performed without source mislocation or station terms, variance is reduced by 18%. Solving for velocity and source terms, but not for station terms, the total variance of summary travel time residuals is reduced by 19%. Solving for all velocity, source, and station terms reduces residual variance by 22%.

The amount of variance explained by our model depends critically on the weights given to the roughness penalty and damping coefficient; further reduction may be achieved by relaxing the minimum norm criterion and at the expense of a reasonably smooth and physically plausible model. However, the relative increase in variance reduction as velocity, source mislocation, and station terms are progressively included is a consistent feature of the inversions. While not orthogonal, the three parameter classes clearly account for distinct parts of the travel time residuals. Leaving out one set of terms from the inversion leaves a corresponding portion of the residual variance unexplained.

Due to the averaging procedure employed to construct summary events, source location corrections are generally small. Source location corrections in subduction zones generally move the source toward the positive velocity anomaly. Sources in continental regions have the smallest, nearly insignificant, corrections. Station corrections range from -0.94 to +1.07 sec, with no obvious correlation to tectonics or elevation of the station region.

4.5.2 Velocity Model

Figures 4.18a-c illustrate the technique we use to show maps of individual layers of our P velocity model for the 670-870 km depth range. For each layer we take the results from our checkerboard test (figure 4.18a), find the absolute value for each pixel, and normalize so the checkerboard output represents the portion of the recovered input value. The checkerboard results now range from 0 to 1. For the layer velocity map (figure 4.18b) we remove the mean from each layer and show velocity in gradations of red to blue. Red indicates slower-than-average velocity and blue marks a faster-than-average region. Next we apply the normalized checkerboard results, for a given layer, to that layer's velocities (figure 4.18c). Colors range from full saturation to white as each pixel's velocity is modified by its checkerboard resolution value. Full recovery

of the input checkerboard value is indicated by full color saturation and no recovery produces zero color saturation, in which case the pixel is white. Areas not sampled by our data set are left black. Figures 4.19a-f show six additional depth layers of model ISC5_LSQR.

4.5.3 Correlations with Surface Tectonics

4.5.3.1 Rift Zones

A significant feature of model ISC5_LSQR is slow anomalies associated with mid-ocean spreading centers (figure 4.19a). The mid-Atlantic rift, tracing the middle of the Atlantic Ocean, appears quite clearly, although not continuously. Anomalies reach -1.5%, but average closer to -0.25%. Resolution for parts of the rift is poor but anomalies that do appear are consistently slow, as expected. Resolution is particularly poor for the discontinuous parts of the anomaly in the South Atlantic. The rift extending to the South Sandwich subduction zone appears faintly, as do the Chile and East Pacific rifts. These areas, along with most of the southern hemisphere, suffer from poor resolution in our checkerboard tests.

A reasonably continuous, slow anomaly emerges from the Red Sea along the Carlsberg and mid-Indian rifts and diverges into two anomalies where the mid-Indian rift splits into the Southeast and Southwest Indian rifts. The anomaly associated with the Southeast Indian rift extends through the southern Indian Ocean and south of Australia, after which it disappears. The checkerboard tests indicate this anomaly is poorly resolved. The anomaly marking the Southwest Indian rift continues around the tip of South Africa to join the mid-Atlantic anomaly. The perturbations associated with Indian Ocean spreading centers reach -1%, but again average closer to -0.25%. The East African rift shows quite clearly in our model, with anomalies reaching -2.0%. The only major rift zone that finds no reliable expression in our model is the ridge that extends across the southern Pacific. The ISC catalog contains very few events located

along this ridge (see figure2), which results in poor resolution.

Many of the rift zone anomalies that appear in the 0-200 km range of model ISC5_LSQR also appear in the 200-400 km layer (figure 4.19b), including the mid-Atlantic, East African, and mid-Indian rifts in particular. All of these are diminished in their magnitudes, spatial extent, and continuity. Many anomalies associated with rifts do not appear at all, despite generally increased resolution in the second layer. By the third layer, 400-670 km depth, only the anomalies associated with the Carlsberg, mid-Indian, and East African rifts are clearly visible (figure 4.19c). In the fourth layer (figure 4.18c) the strong correlation between rift zones and slow anomalies is gone, although diffuse, slow anomalies beneath the mid-Indian and East African rifts persist.

4.5.3.2 Subduction Zones

The backarc basins in the western Pacific are clearly marked by slow anomalies that average around -1.0% (figure 4.19a). At 200-400 km the pattern is still clear (figure 4.19b), but in several cases, such as beneath the Aleutians and the Japanese island arc, the slow anomaly has been pinched out by an adjacent fast anomaly. This may be due to the dominance of the subducted slab over the excess volatiles released by slabs in mantle material below 200 km. However, in the Tonga-Kermadec and New Hebrides subduction zones the fast anomaly shows the opposite effect, having been partially displaced by a slow anomaly. Still, the Tonga anomaly is fast and clearly continuous through the 400-670 km layer (figure 4.19c). Another fast anomaly, much larger in lateral extent, appears in the 670-870 km layer and extends to more than 1670 km depth, but the continuity between this deep anomaly and the shallower anomaly associated with the subduction beneath Tonga is questionable. The shallower fast anomaly, 0-670 km, is consistent with the results of the regional study of *Zhou* [1990], though our model grid is not able to show the finer detail of the slow backarc

basin directly above the dipping slab. The large fast anomaly to the west of the Tonga trench does not appear in Zhou's cross-sections. Beneath Japan and the Kurile Islands another fast anomaly protrudes into the lower mantle at a much shallower angle than the Tonga feature. Again the continuity between a fast upper mantle anomaly dipping to the northeast, and a broadened, diffuse, but similarly fast anomaly below 670 km depth is not clear. The deeper anomaly appears to be continuous all the way to the lowermost layer in the mantle beneath northeastern Asia. The shallower features of our model correspond to anomalies shown by *Zhou and Clayton* [1990] but the deeper anomalies of model ISC5_LSQR fall in regions not included in their cross-sections. A fast anomaly beneath the Andes is abruptly pinched out in the second layer by slow anomalies. The fast anomaly reappears in the third layer below South America and continues into the 670-870 km layer. The Bering Sea appears as a slow anomaly, consistent with other back-arc basins around the Pacific. In the second layer a fast anomaly extends along the Aleutian trench and displaces the southernmost extension of the slow, Bering Sea anomaly. In the 400-670 km and 670-870 km layers the fast anomaly becomes progressively more dominant, but diminishes in the 870-1070 km depth interval. A significant fast anomaly appears in the 400-870 km depth range beneath the Mariana subduction zone, and a broad, fast region occurs in the 670-870 km layer underneath the Philippines.

In general, the broadening and flattening to sub-horizontal noted by *Zhou* [1990] and *Zhou and Clayton* [1990] also appear in our model at the 670-870 km layer (figure 4.18). *Van der Hilst and Spakman* [1989] demonstrate that inaccuracies of the J-B model in the upper mantle and the geometry of ray coverage provided by the P phase can lead to a similar sort of flattening and horizontal extension purely as artifacts of the imaging procedure. They caution that the interpretation of such images must be approached with caution, though a more accurate one-dimensional starting model can help limit these flattening artifacts. Our starting model differs from the

starting model used by *van der Hilst and Spakman* [1989] in that we do not include discontinuities in the upper mantle, but is similar in that we do not include a low-velocity zone in the uppermost mantle either. Without discontinuities our one-dimensional model will generate raypaths that appear to sample the transition zone well, while in fact the travel time residuals associated with those rays actually correspond to rays that were refracted at the discontinuities. One effect may be the inaccurate mapping of anomalies to locations in the transition zone, and perhaps to locations below the 670 km discontinuity. The extent of this mis-mapping is unclear. A second problem is associated with the apparent existence in some regions of the real earth of a low-velocity zone, which is not present in our starting model. A low-velocity zone in Earth's mantle would produce a shadow zone at the surface in which few rays would emerge and those that did emerge would have small amplitude. Readers of seismograms might overlook the actual first arrival and instead pick a larger-amplitude arrival that was refracted at the 400 km discontinuity or turned upward by a steep gradient below the low-velocity zone. Unless we can re-identify the picks supplied to the ISC and discard or use appropriately the late arrivals (which we cannot with a J-B model) we will introduce a set of systematically slow residuals which will propagate through the inversion to produce slow anomalies in our model. The distribution of travel time residuals as a function of epicentral distance contains an anomalously sparse section clustered about zero seconds between 10° and 17° but indeed shows a systematic trend toward slow (positive) residuals. Rays emerging at these distances follow paths that bottom between 100 and 250 km in the J-B model. Rays are most sensitive to velocity perturbations near their sources, receivers, and bottoming points, so unexpected slow anomalies found in these depths in models produced with J-B starting model should be considered suspect. If low-velocity zones were distributed according to some pattern, preferentially under continents rather than oceans, for example, rather than appearing consistently worldwide, the effect could be

even more difficult to uncover. Rather than appearing as a slow mean to the three-dimensional model, which may be removed and used to update the one-dimensional model, the effect would be specific to particular regions. This may provide an alternative explanation for images that suggest a detached slab feature due to a lack of continuity between fast features above and below about 200 km depth. We believe our results are relatively, though not entirely, uncontaminated by this effect since the inverse weighting by residuals' standard errors applied to rows of the constraining matrix equation reduces the influence of these data.

4.5.3.3 Continental Shields

Fast anomalies show quite clearly in continental shield regions (figure 4.19a). Alaska, Canada, Greenland, Fennoscandia, Siberia, and northern Australia all show fast anomalies on the order of +1.5%. All of these fast anomalies persist through the 200-400 km layer (figure 4.19b), but appear broken and discontinuous in the transition zone. Northern Africa and eastern South America are poorly resolved.

4.5.3.4 Hotspots

A number of hotspots correlate well visually with strong slow anomalies in our model. The Azores, Cape Verde, Canary Islands, Afar/Ethiopia, Lake Victoria/East Africa, Comores Islands, Kerguelen, Christmas Island, Tasmania, Caroline Islands, Hawaii, Galapagos, Vema Seamount, and Mt. Erebus hotspots all appear as isolated, negative velocity perturbations (figure 4.19a). In addition, the Yellowstone and Raton, New Mexico hotspots appear subsumed into the general slow anomaly covering the western United States. The Mehetia/Society Islands/Tahiti hotspot appears as one member of a complex set of four hotspots, including the MacDonald Seamount, and Marquesas Islands and Pitcairn Island/Gambier Islands hotspots. Between these four hotspots, marked by two diffuse slow anomalies, lies an apparent change to a fast anomaly. The Jan Mayen and Iceland hotspots appear as parts of the mid-Atlantic rift, as

does the St. Helena hotspot just to the west of Africa. It is interesting to compare the St. Helena hotspot anomaly to the Ascension hotspot, located about 8° closer to Africa, which finds no expression in our model. St. Helena is located near the rift and is thus illuminated by seismic events associated with the formation of new oceanic crust. In all, about half the set of hotspots compiled by *Richards et al.* [1988] are marked by slow anomalies in at least the first layer of our model.

None of these anomalies changes sign through the first three layers. Below 670 km none clearly changes, but the Yellowstone and Iceland anomalies appear displaced. In the 870-1070 km range, several more hotspot anomalies are displaced from their surface locations, but each anomaly persists in some form nearby. Beneath Hawaii a slow anomaly, reaching -1.5%, extends deep into the mantle, trailing off to the northwest with depth. Both *Vasco et al.* [1990] and *Inoue et al.* [1990] show this area to have a weak fast perturbation, contrary to our expectation. This difference can probably be explained by the differences in the way we treat the data. *Inoue et al.* [1990], though they use an l^2 residual norm minimization, severely downweight outliers. *Vasco et al.* [1990] minimize an l^1 norm of the residuals. Experience with calibration events that have known locations shows that some of the most extreme, slow travel times are recorded at the Hawaiian stations. Downweighting these extreme residuals in the inversion procedure causes the algorithm to overlook the anomalies that give rise to these slow travel times.

4.5.4 Transition zone

In the transition zone, 400-670 km, the pattern of anomalies changes completely (figure 4.19c). The correlations between anomalous velocities and surface tectonics observed in the top two layers do not exist here. Shield regions are not generally fast and backarc basins are not generally slow, though the region extending northward and westward from New Zealand, a complex subduction zone, is quite slow. A striking

slow feature, amounting to -2%, appears under southern Eurasia and India. The fast feature beneath Tonga is not clearly continuous through this layer, though neighboring voxels show a similar anomaly to the west in the next lower layer.

Several hotspot anomalies including those at Hawaii, Kerguelen, Iceland, Lake Victoria, Yellowstone, Raton, Afar, Mt. Erebus, Galapagos, Canary Islands, and Mehetia persist through this layer. In the 670-870 km layer, the features at Tonga spread laterally to the west. The slab-related anomaly under Japan has migrated a similar distance westward but without similar lateral extension.

Overall, our topmost layer shows a similar pattern to *Inoue et al.'s* [1990] 78-148 km layer, though our layer is not so heavily smoothed. Our 400-670 km layer is also quite similar to their 478-629 km layer in regions for which we have ray coverage. Though they do not show their hitcount map for this layer, the long slow feature extending northeast-southwest across the Pacific in their model may be an artifact of their smoothing procedure. Our data set shows the central Pacific to be largely unsampled in this depth range.

4.5.5 Mid Mantle

Confirming results of previous studies, our model shows diminished amplitudes of velocity anomalies in the mid-mantle. There is no obvious large-scale radial continuity throughout the mid-mantle. On a smaller scale, several hotspot anomalies persist. Most striking are fast anomalies beneath eastern North America, the Caribbean, and central South America, and the features, mentioned earlier, beneath Tonga and Japan/eastern Asia. The fast anomaly beneath eastern North America and the Caribbean appears in the same location as a large S-velocity anomaly reported by *Grand* [1987] and the anomaly's apparent continuation beneath South America also appears in Grand's recent results *(personal communication)* (see figures 4.18c and 4.20b). Similar features for P-velocity appear in the inversions performed by *van der Hilst* [1990].

The fast feature beneath Tonga broadens and continues to dip to the west to a depth of 1670 km. Beneath Japan and eastern Asia the fast anomaly is diffuse but extends all the way to the core-mantle boundary.

4.5.6 Lower Mantle

The lowermost layers, 2470-2670 km (figure 4.19f) and 2670 km-CMB, show a significantly different pattern from the mid-mantle. In general, our results for the 2470-2670 km layer agree with both Dziewonski's L02.56 model and *Hager and Clayton's* [1988] smoothed version of *Clayton and Comer's* [1983] model. When expanded in spherical harmonics and recombined using only $1 \leq l \leq 6$, our model also shows a large slow anomaly over southern Africa, though this anomaly is displaced relative to L02.56 and in a way that is more consistent with Hager and Clayton's result. Other slow anomalies appear beneath the southern Pacific Ocean, beneath the Bering Sea and toward the North Pole, in the northern Atlantic Ocean, and beneath Papua New Guinea. Fast anomalies appear beneath Asia, South America, and north of New Zealand. Some oscillation appears to occur between the lowermost layer and the 2470-2670 km layer just above. In the South Pacific, a large-amplitude slow anomaly just above the CMB trades off with a fast anomaly above it. A ring of slow material surrounds the fast anomaly in the second-to-bottom layer. The checkerboard tests show reasonable resolution in this area, but such oscillations may be just the sort of problem checkerboard tests cannot reveal. Results for the bottom layer, D", are suspect because our restriction to rays with epicentral distance less than 96° results in poor coverage of this layer. This restriction is intended to avoid contamination of our data set by arrivals diffracted at the core. Reduced resolution in the bottom layer is the price we pay for avoiding this contamination.

4.7 Continuity of Features

Figure 4.20a shows several of the features described above in a set of cross-sections through model ISC5_LSQR. A constant-latitude slice at 24°S through the Tonga-New Hebrides subduction complex shows the associated fast anomaly dipping to the west. A fast continental root appears further to the west at this latitude under northern Australia. To the east the Pitcairn Island/Gambier Islands hotspot is associated with a slow anomaly. Note the pinching out of the dipping Tonga anomaly between the 200-400 km and the 400-670 km layers. The appearance of such a broad and deep fast anomaly to the west of the subduction zone, extending to 1670 km depth in the dip direction of the downgoing slab, offers tantalizing circumstantial evidence for slab penetration into the lower mantle. Further evidence is shown in the other constant-latitude cross-section of figure 4.20a, at 52°N, where a fast anomaly appears under the Kuriles and dips to the northwest. This anomaly is much broader, extending continuously through the Japanese Island arc and under eastern Asia (see cross-section at 124°E), but is less distinct than the Tonga anomaly. A slight fast anomaly beneath the Aleutians is also shown.

A constant-longitude slice at 84°W, under eastern North America, the Caribbean, and western South America shows a distinct fast anomaly beginning at about 1070 km depth and extending to 2070 km. Under North America and the Caribbean, the anomaly is consistent in both size and location with an S-velocity anomaly reported by *Grand* [1987]. The fast anomaly veers to the east south of the Caribbean (see figure 4.20b). Above the fast anomaly in the Caribbean and through Central America the model is slow. A constant-depth section shows the 2470-2670 km layer with longer-wavelength anomalies than the upper mantle. A slow anomaly appears under the western Pacific and a region beneath the eastern Pacific and North America is fast.

Figure 4.20b shows the long fast anomaly beneath eastern North America, the Caribbean, and South America in a depth section at 670-870 km. A constant-longitude

slice at 88°W shows that this depth constitutes the top of the anomaly at this longitude, though the fast continental shield in the northern United States and Canada appears at the top of the section. To the east of the fast anomaly in the northern hemisphere lies a broad slow anomaly under the Atlantic Ocean. A slice at 29°S latitude shows the fast anomaly under central South America extends continuously to the surface at the Chilean subduction zone. A section at 64°N shows fast shields beneath Canada, Greenland, and Fennoscandia and a slow anomaly beneath Iceland. A broad slow region appears at 32°E under the East African rift zone.

4.5.8 Spherical Harmonic Expansion

Surface spherical harmonic series expansions to degree 15 were calculated by integration around the globe for each coefficient, rather than by fitting coefficients to model values by least-squares. In this way, coefficients are independent of each other and coefficient values are independent of the point of truncation of the harmonic series. That is, coefficients do not change if the series expansion is calculated a second time with a different number of terms. The associated Legendre polynomials are fully normalized, i.e.,

$$p_l^m(\theta) = \left[(2-\delta_{m,0})(2l+1)\frac{(l-m)!}{(l+m)!}\right]^{1/2} P_l^m(\cos\theta).$$

Figure 4.21 shows the total power in the series expansion for each layer plotted as a function of depth. The anomalously low power in the 200-400 km layer probably is due to the fact that rays that bottom in this layer, which emerge at the epicentral distance range $15° \leq \Delta \leq 20°$, have the largest variance of all the travel time residuals. These rays are the most sensitive to velocity perturbations in the 200-400 km layer, but in our inversion their influence on the final model is (severely) downweighted by the inverse of the residuals' standard errors. The first layer (0-200 km) and the transition zone (400-670 km) have the highest power, indicating the greatest heterogeneity

in our model occurs at these depths. Again, because our starting model does not contain discontinuities at 400 and 670 km depth we are probably mapping more power into the transition zone than is justified. Power decreases in the mid-mantle and increases again as we approach the core. The small increase in power at the 1270-1470 km depth layer is the result of an unusually large $l = 1$ component. The drop in power from the 2470-2670 km layer to the lowermost layer, 2670-CMB, is probably due to our poor ray coverage. Figure 4.22 shows the power in series expansions of each layer as a function of angular degree. The top two layers appear relatively devoid of power at the lower degrees, despite the strong concentration of ray coverage in $l = 1$–5 patterns. In the 400-670 km layer $l = 1$, 2, 3, and 6 dominate. The large $l = 2$ component confirms previous reports, but to our knowledge, no other study has shown the equally prominent $l = 1$ and 3 components. In the mid-mantle power is more or less evenly distributed across the harmonic terms. The exception is in the layer 1270-1470 km, where the $l = 1$ harmonic is strong and the $l = 2$ and 3 components rise above the higher-degree harmonics. These components clearly are responsible for the high power total of this layer.

Since sign information is not included in power calculations, figure 4.22 does not show how the distribution patterns for all layers combine constructively or destructively to form a pattern for the whole mantle. Figure 4.23 shows the power in the spherical harmonic expansions averaged through the whole mantle and through the upper and lower mantle separately. The averaging is performed on the the individual harmonic coefficients, weighted at each layer by the square of the layer mid-point's radius, which normalizes the power in each layer to the layer's surface area. For the upper mantle the power spectrum shows a dominant $l = 6$ component, along with prominent $l = 2$, 5, 12, and 13 terms. The $l = 1$ power for the upper mantle is low, simply reflecting the results in figure 4.22 which show the small $l = 1$ components in the first two layers. However, in the lower mantle the same component is

unexpectedly low, given the large values in several of the individual layers. This may mean the $l = 1$ component is poorly resolved in the lower mantle and trades off between layers in our model. In general, high power in harmonic degrees of the upper mantle coincides with high power in the same degrees of the lower mantle, though relative amplitudes once again suggest that heterogeneity is concentrated in the upper mantle. Degrees 2, 5, and 12 dominate the expansion averaged through the whole mantle. Surprisingly, degree 6 has the lowest average power even though its power dominates the top layers and contributes significantly to the lower mantle's total. From the relatively low power of the whole-mantle integration compared to the separate lower- and upper-mantle averaged series, it is clear that either one part of the real mantle is compensating for anomalies in the other part or our inversion scheme is trading off power between the lower and upper portions of our model mantle.

4.6 Conclusions

We present a three-dimensional P-velocity model for the Earth's mantle found by inverting ISC travel time data for the time period January 1964 - January 1987. Our inversion minimizes the l^2 norm of the travel time residuals, by means of the conjugate gradient variant LSQR algorithm. Model maps show values only for sampled voxels, with a weighting scheme based on our estimate of model resolution. This allows the most realistic presentation of what is known about mantle velocity structure.

Model values and patterns for the top two layers correlate well visually with surface tectonics. Backarc basins, rift zones, and some known hotspots all find expression as slow anomalies in our model while continental shields and some subduction zones are marked by unusually fast velocities. Comparison of model ISC5_LSQR to the model of *Inoue et al.* [1990] reveals great visual similarities in patterns of fast and slow velocities. Overall, the mid- and lower mantle show distinctly less heterogeneity than the upper mantle. Fast anomalies appear in both the mid- and lower mantle

beneath the Tonga subduction zone, eastern North America, the Caribbean, central South America, and Japan/eastern Asia. The lower mantle correlates well at low orders with the models reported by *Dziewonski* [1984], *Morelli and Dziewonski* [1985, 1986], and *Clayton and Comer* [1983; *Hager and Clayton*, 1988], though we cannot be confident of our results for the lowermost layer (D").

Seismological studies of lateral heterogeneity in the mantle differ in their general approaches and in specific decisions made along the way. Our model parametrization differs significantly from those employed by *Dziewonski* [1984] and *Inoue et al.* [1990], and differs slightly from *Clayton and Comer* [1983; *Hager and Clayton*, 1988]. We use summary rays while *Dziewonski* [1984] and *Inoue et al.* [1990] do not. We weight rows of the matrix problem by the inverse standard errors of the travel time residuals as a function of delta, a measure of data quality, while *Inoue et al.* [1990] weight preferentially according to the size of the residual. We weight columns of the coefficient matrix by a measure of the quality of each voxel's sampling, while *Clayton and Comer* [1983; *Hager and Clayton*, 1988] weight by the number of hits, and *Inoue et al.* [1990] do not weight columns at all. We use the LSQR algorithm to solve the constraining matrix equation while *Clayton and Comer* [1983; *Hager and Clayton*, 1988] and *Dziewonski* [1984] do not. While each of these differences has consequences that may be traced to differences in our final models, our results are quite similar in general. All these studies depend on traditional ray theory and Fermat's principle in the construction of the tomographic equations and, perhaps most importantly, each study makes use of the same set of global travel time data.

4.7 References

Adams, R.D., A.A. Hughes, and D.M. McGregor, Analysis procedures at the International Seismological Centre, *Phys. Earth. Planet. Int., 30,* 85–93, 1982.

Buland, R., Residual statistics, *Terra Cognita, 4,* 268, 1984.

Buland, R., Uniform reduction error analysis, *Bull. Seis. Soc. Am., 76,* 217–230, 1986.

Clayton, R.W., and R.P. Comer, A tomographic analysis of mantle heterogeneities from body wave travel times, *EOS Transactions AGU, 64,* 776, 1983.

Dziewonski, A.M., Mapping the lower mantle, Determination of lateral heterogeneity in P velocity up to degree and order 6, *J. Geophys. Res., 89,* 5929–5952, 1984.

Dziewonski, A.M., and D.L. Anderson, Preliminary reference Earth model, *Phys. Earth and Planet. Int., 25,* 297–356, 1981.

Dziewonski, A.M., B.H. Hager, and R.J. O'Connell, Large–scale heterogeneities in the lower mantle, *J. Geophys. Res., 82,* 239–255, 1977.

Efron, B., *The Jackknife, the Bootstrap and other Resampling Plans,* Soc. for Ind. and Appl. Math., Philadelphia, Pa, 1982.

Grand, S., Tomographic inversion for shear velocity beneath the North American Plate, *J. Geophys. Res., 92,* 14065–14090, 1987.

Gudmundsson, O., J.H. Davies, and R.W. Clayton, Stochastic analysis of global travel-time data, mantle heterogeneity and random errors in the ISC data, *Geophys. J. Int., 102,* 25–43, 1990.

Hager, B.H., and R.W. Clayton, Constraints on the structure of mantle convection using seismic observations, flow models, and the geoid, in *Mantle Convection,* W. R. Peltier (Ed.), Gordon and Breach, New York, 657–763, 1989.

Herrin, E., W. Tucker, J.N. Taggert, D.W. Gordon, and J.L. Lobdell, Estimation of surface focus P–travel times, *Bull. Seism. Soc. Am., 58,* 1273–1291, 1968.

Humphreys, E., and R.W. Clayton, Adaptation of back projection tomography to seismic travel time problems, *J. Geophys. Res., 93,* 1073–1085, 1988.

Inoue, H., Y. Fukao, K. Tanabe, and Y. Ogata, Whole mantle P–wave travel time tomography. *Phys. Earth and Planet. Int., 59,* 294–328, 1990.

Jeffreys, H., and K.E. Bullen, *Seismological Tables.* British Association for the Advancement of Science, London, 1940.

Jeffreys, H., *Theory of Probability,* Clarendon Press, Oxford, 1939.

Jeffreys, H., *The Earth,* Cambridge University Press, London, 1960.

Jordan, T.H., and K.A. Sverdrup, Teleseismic location techniques and their application to earthquake clusters in the South–Central Pacific, *Bull. Seis. Soc. Am., 71,* 1105–1130, 1981.

Kennett, B.L.N., and P.R. Williamson, Subspace methods for large–scale inversion, in *Mathematical Geophysics,* N.J. Vlaar, G. Nolet, M.J.R. Wortel, and S.A.P.L. Cloetingh (eds.), Reidel, Dordrecht, pp. 139–154, 1988.

Lawson, C.L., and R.J. Hanson, *Solving Least Squares Problems.* Prentice–Hall, Englewood Cliffs, N.J., 340 pp., 1974.

Lees, J.M., and R.S. Crosson, Tomographic inversion for three–dimensional velocity structure at Mount St. Helens using earthquake data, *J. Geophys. Res., 94,* 5716–5728, 1989.

Morelli, A., and A.M. Dziewonski, Stability of aspherical models of the lower mantle, *EOS Transactions AGU, 66,* 975, 1985.

Morelli, A., and A.M. Dziewonski, 3D structure of the Earth's core inferred from travel–time residuals, *EOS Transactions AGU, 67,* 311, 1986.

Nolet, G., Solving or resolving inadequate and noisy tomographic systems, *J. Comp. Physics, 61,* 463–482, 1985.

Nolet, G., Seismic wave propagation and seismic tomography, in *Seismic Tomography,* G. Nolet (ed.), Reidel, Dordrecht, pp. 1–23, 1987.

Nolet, G., and R. Snieder, Solving large linear inverse problems by projection, *Geophys. Jour. Int., 103,* 565–568, 1990.

O'Connell, D.R.H., Seismic velocity structure and microearthquake source properties at the Geysers, California, geothermal area, *Ph.D. Dissertation,* University of California, Berkeley, 1986.

O'Connell, D.R.H., and L.R. Johnson, Progressive Inversion for Hypocenters and P–wave and S–wave Velocity Structure, Application to the Geysers, California,

Geothermal Field, submitted to *J. Geophys. Res.*,

Pavlis, G.L., and J.R. Booker, The mixed discrete–continuous inverse problem, application to the simultaneous determination of earthquake hypocenters and velocity structure, *J. Geophys. Res., 88,* 4801–4810, 1980.

Pulliam, R.J., and L.R. Johnson, Effects of source mislocation in mantle delay time tomography, *Seismological Research Letters, 60,* 10, 1989a.

Pulliam, R.J., and L.R. Johnson, A Tomographic, Progressive Inversion of ISC P–wave Travel Times for 3–D Mantle Slowness Variations and Source Mislocations, *EOS Transactions, 70,* 1213, 1989b.

Richards, M.A., B.H. Hager, and N.H. Sleep, Dynamically supported geoid highs over hotspots, Observation and theory, *J. Geophys. Res., 93,* 7690–7780, 1988.

Scales, J. A., Tomographic inversion via the conjugate gradient method. *Geophysics, 52,*

Spakman, W., and G. Nolet, Imaging algorithms, accuracy and resolution in delay time tomography, in *Mathematical Geophysics,* N.J. Vlaar, G. Nolet, M.J.R. Wortel, and S.A.P.L. Cloetingh (eds.), Reidel, Dordrecht, pp. 155–187, 1987.

Spakman, W., S. Stein, R. van der Hilst, and R. Wortel, Resolution experiments for NW Pacific subduction zone tomography, *Geophys. Res. Letters, 16,* 1097–1100, 1989.

Spencer, C., and D. Gubbins, Travel–time inversion for simultaneous earthquake location and velocity structure determination in laterally varying media, *Geophys. J. R. Astron. Soc., 63,* 95–116, 1980.

Tarantola, B., *Inverse Problem Theory,* Elsevier, Amsterdam, 1987.

van der Hilst, R.D., Tomography with P, PP, pP delay-time data and the three-dimensional mantle structure below the Caribbean region. *Ph.D. Dissertation,* University of Utrecht, The Netherlands, 1990.

van der Hilst, R.D., and W. Spakman, Importance of the reference model in linearized

tomography and images of subduction below the Caribbean Plate, *Geophys. Res. Letters, 16,* 1093–1096, 1989.

Vasco, D. W., Bounding seismic velocities using a tomographic method, *Geophysics, 56,* 472–482, 1991.

Vasco, D.W., R. Jay Pulliam, and Lane R. Johnson, Tomographic inversion of ISC travel times for mantle P wave velocity structure using an l^1 norm criterion, submitted to *J. Geophys. Res.,* November 1990.

Zhou, H.W., Mapping of P–wave slab anomalies beneath the Tonga, Kermadec and New Hebrides arcs. *Phys. Earth and Planet. Interiors, 61,* 199–229, 1990.

Zhou, H.W., and R.W. Clayton, P and S wave travel time inversions for subducting slab under the island arcs of the northwest Pacific. *J. Geophys. Res., 95,* 6829–6851, 1990.

4.8 List of Tables

Table 4.1 Details of the model parametrization and the sampling provided by our data set. Include are the average number of hits for sampled voxels in each layer along with the the average sum of ray segments in a voxel at a given depth and the number of voxels sampled in each depth interval. These averages include only voxels which have non-zero sampling.

4.9 List of Figures

Figure 4.1 The model mantle is parametrized as voxels, 5° × 5° at the equator and generally 200 km thick, for a total of 22,876 model parameters. Voxels in a given layer have approximately equal surface area.

Figure 4.2 Locations of sources used in this study. The data set consists of about 46,000 shallow events located by the ISC for the time period January 1964 - January 1987.

Figure 4.3 Locations of seismographic stations reporting to the ISC in January 1987.

Figure 4.4 Histogram of the 3.02 million travel time residuals associated with the events included in this study. Also shown are the first four moments of the distribution.

Figure 4.5 Locations of summary sources. The averaging procedure, based on a 2° × 2° grid, reduces the number of sources to about 6,000.

Figure 4.6 Locations of summary stations. The number of stations is reduced to 979 by the ray averaging procedure.

Figure 4.7 Histogram of the summary residuals with the first four moments of the distribution.

Figure 4.8 The distribution of column sums of our coefficient matrix, **A**, indicating the total sampling of each voxel by our summary data set. Shown in each panel is a different depth layer: (a) 0-200 km, (b) 200-400 km, (c) 400-670 km, (d) 670-870 km, (e) 1270-1470 km, and (f) 2470-2670 km.

Figure 4.9 (a) A 5% velocity anomaly introduced at (17°S, 178°W, 600 km depth), beneath the Tonga subduction zone and (b) the resulting point spread function. Cross-sections are at 20°S, 174°W, 6°E, and 1700 km depth.

Figure 4.10 (a) A 3% anomaly placed beneath the New Hebrides subduction zone at (16°S, 166°E, 900 km depth) and (b) the resulting point spread function. Cross-sections are at 20°S, 168°E, 12°W, and 2470 km depth.

Figure 4.11 (a) A -3% anomaly introduced to the first layer beneath the Hawaiian islands at (21.4°N, 158°W, 70 km depth) and (b) the resulting point spread function. Cross-sections are at 20°N, 156°W, 24°E, and 1700 km depth.

Figure 4.12 (a) A -4% anomaly introduced to the first layer beneath Iceland at (65°N, 18°W, 70 km depth) and (b) the resulting point spread function. Cross-sections are at 64°N and 1300 km depth.

Figure 4.13 (a) A 3% anomaly introduced to the fourth layer beneath central South America at (0°, 65°W, 870 km depth) and (b) the resulting point spread function. Cross-sections are at 0°, 62°W, 118°E, and 1400 km depth.

Figure 4.14 (a) A 4% anomaly placed beneath eastern North America at (40°N, 85°W, 870 km depth) and (b) the resulting point spread function. Cross-sections are at 42°N, 58°W, 122°E, and 1900 km depth.

Figure 4.15 Map of the top layer, 0-200 km, of the resolution test using a synthetic "checkerboard" model. (a) Input values alternate between ±0.3 km/s. (b) Output of the test inversion.

Figure 4.16 Model standard error estimates for the 0-200 km layer using (a) a Gaussian distribution of deviates in place of the real data as input to the inversion and (b) a two-sided exponential distribution as input.

Figure 4.17 Model standard error estimates for the 0-200 km layer obtained with a jackknife procedure.

Figure 4.18 A set of layer maps for the 670-870 km depth range illustrating the technique used to plot velocity maps. The checkerboard resolution test results (a) are used to modify the velocity values for each block (b) based on the percentage of input value recovered by the test inversion. Final values (c) range from red (slow) to blue (fast) and from full color saturation, indicating full recovery of the input checkerboard value, to white, indicating no recovery.

Figure 4.19 (a-f) Six depth layers of model ISC5_LSQR: (a) 0-200 km, (b) 200-400 km, (c) 400-670 km, (d) 1270-1470 km, and (e) 1470-1670 km, and (f) 2470-2670 km. Each layer's mean has been removed. Velocity perturbations grade from red (slow) to blue (fast). In addition, color values are modified from full saturation, indicating the voxel is well-resolved as determined by the checkerboard test, to white, which indicates no recovery of the checkerboard value.

Figure 4.20 Cross-sections of model ISC5_LSQR. Shown in (a) are two constant-latitude slices at 24°S and 52°N, constant-longitude slices at 124°E and 84°W, and a constant-depth section showing the 2470-2670 km layer. Shown in (b) are two constant-latitude slices at 24°S and 64°N, two constant-longitude slices at 34°E and 98°W, and a constant-depth section at 670-870 km.

Figure 4.21 Power contained in surface spherical harmonic series expansions of model ISC5_LSQR for each depth layer.

Figure 4.22 Power in the spherical harmonic expansions for each depth interval as a function of angular degree. All values are normalized to the maximum value

appearing in the figure. Numbers on the right refer to the maximum power for each layer.

Figure 4.23 Power in spherical harmonic series generated by averaging ISC5_LSQR layer expansions through the whole mantle and through the upper and lower mantle separately.

Table 4.1: Model layer depths and average voxel sampling

Layer number	Depth range (km)	Average velocity (km/s)	Average voxel volume (10^7 km^3)	Number of voxels sampled	% of voxels sampled	Average number hits	Average column sum of A (10^4 km)
1	0-200	7.94	6.06	1195	73.1	1586	31.02
2	200-400	8.58	5.67	1333	81.6	1342	25.97
3	400-670	9.88	7.07	1494	91.4	1409	34.34
4	670-870	10.95	4.82	1577	96.5	1261	27.38
5	870-1070	11.40	4.49	1604	98.2	1149	25.21
6	1070-1270	11.68	4.16	1627	99.6	991	21.29
7	1270-1470	11.96	3.85	1631	99.8	881	18.38
8	1470-1670	12.24	3.55	1632	99.9	816	16.96
9	1670-1870	12.51	3.26	1634	100.0	753	15.62
10	1870-2070	12.76	2.98	1634	100.0	675	13.79
11	2070-2270	13.01	2.72	1634	100.0	604	12.45
12	2270-2470	13.24	2.46	1634	100.0	494	10.04
13	2470-2670	13.47	2.22	1634	100.0	395	8.11
14	2670-2898	13.64	2.26	1626	99.5	179	3.68

Figure 4.1

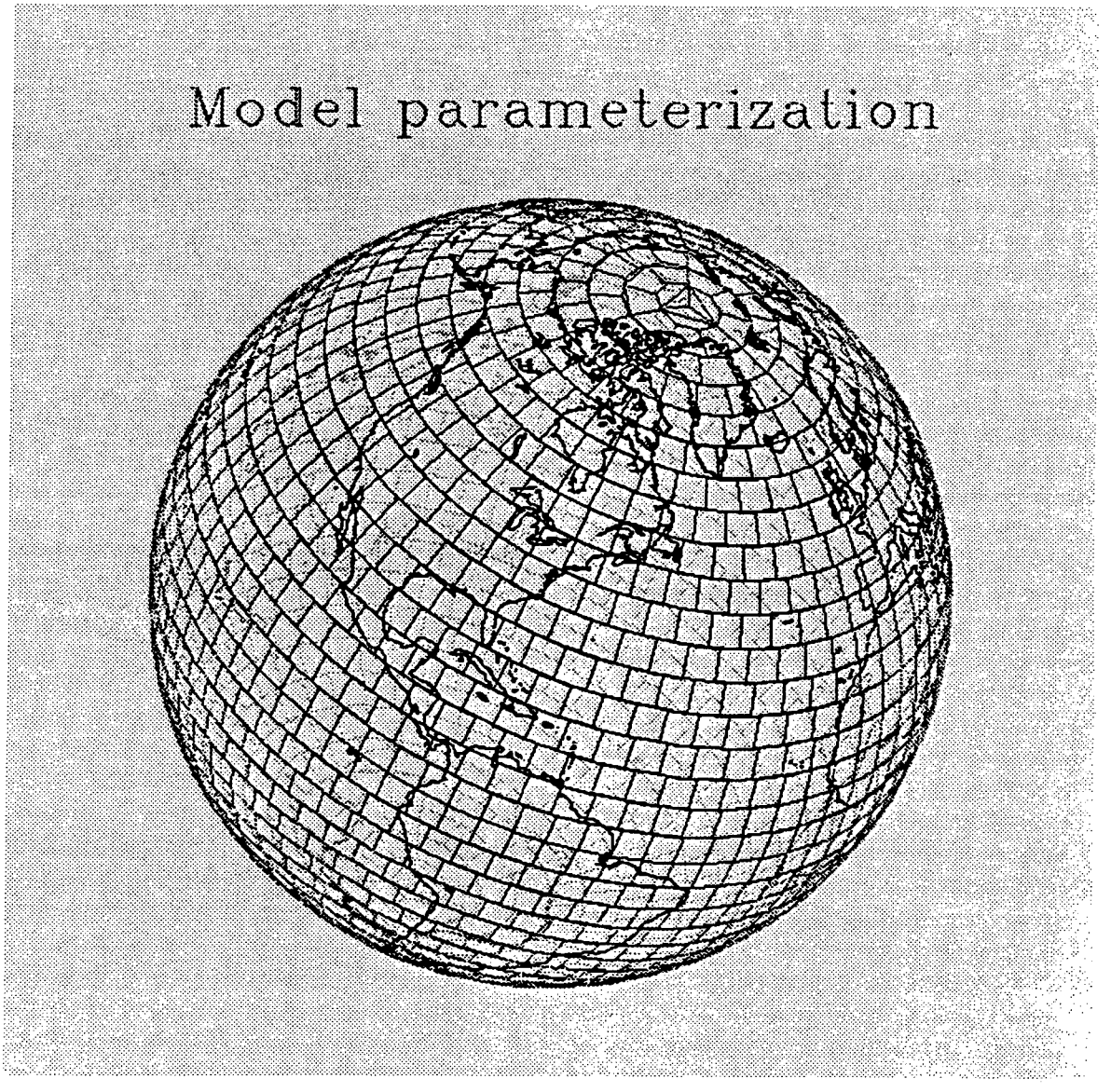


Figure 4.2

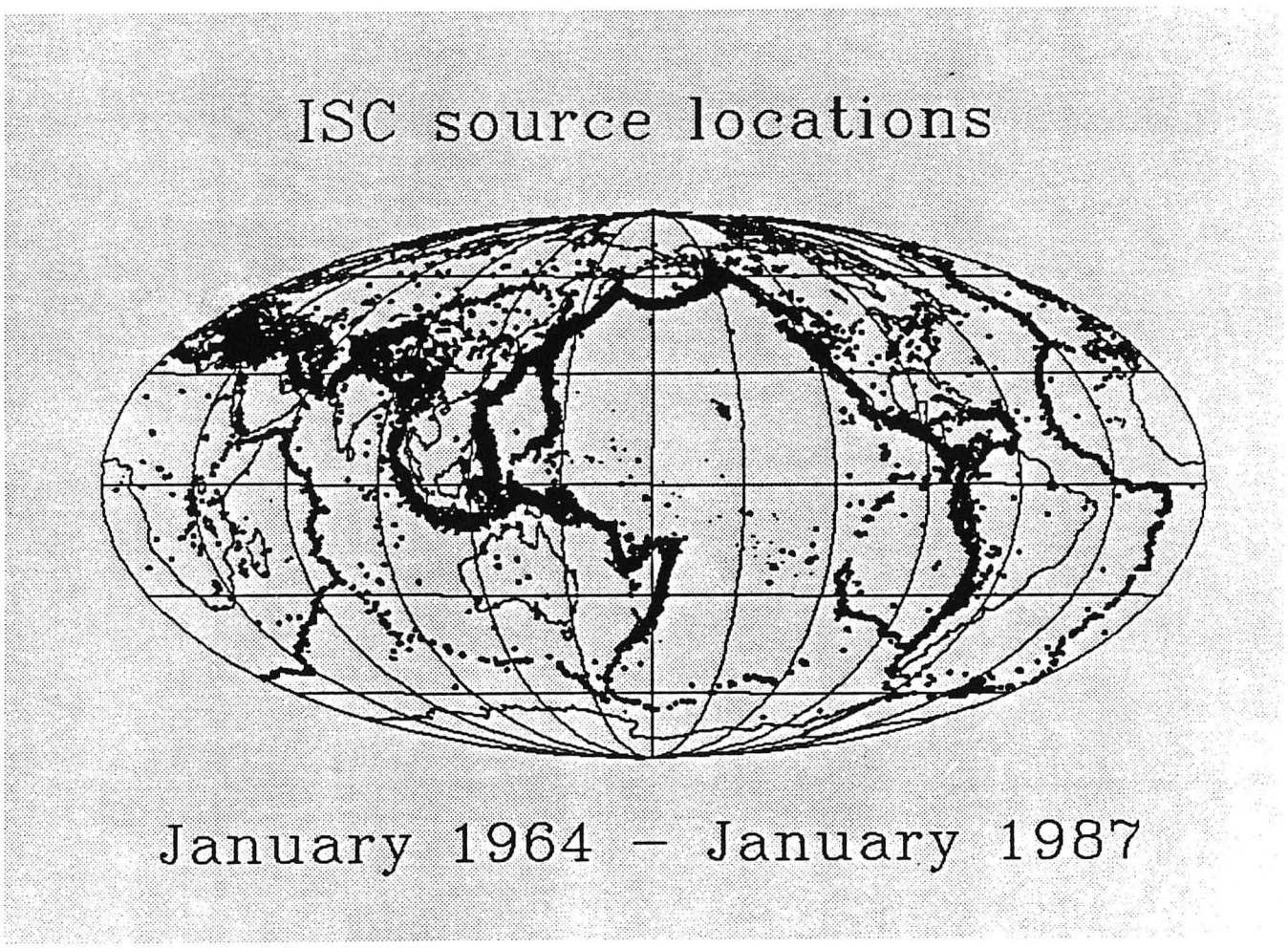


Figure 4.3

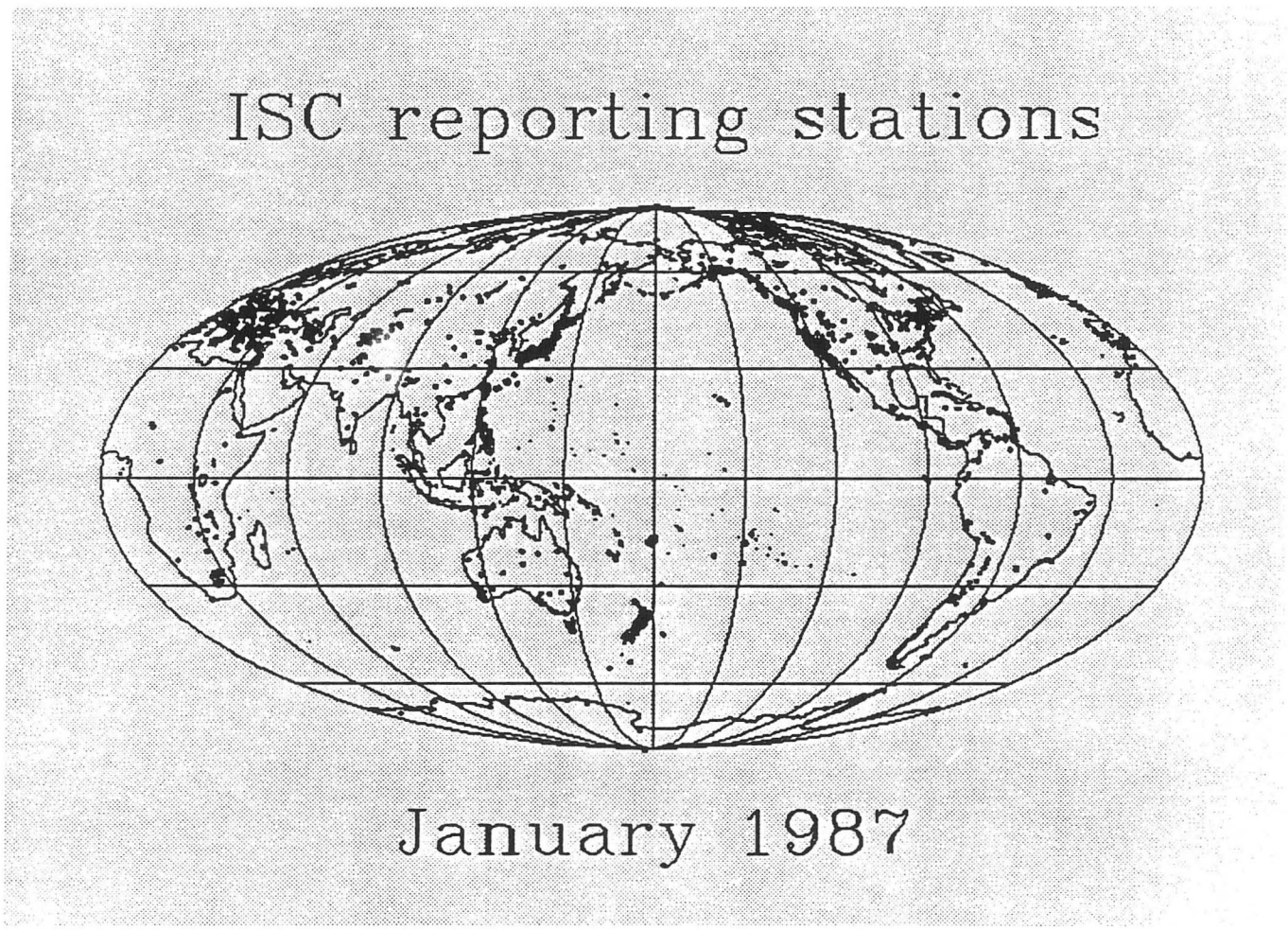


Figure 4.4

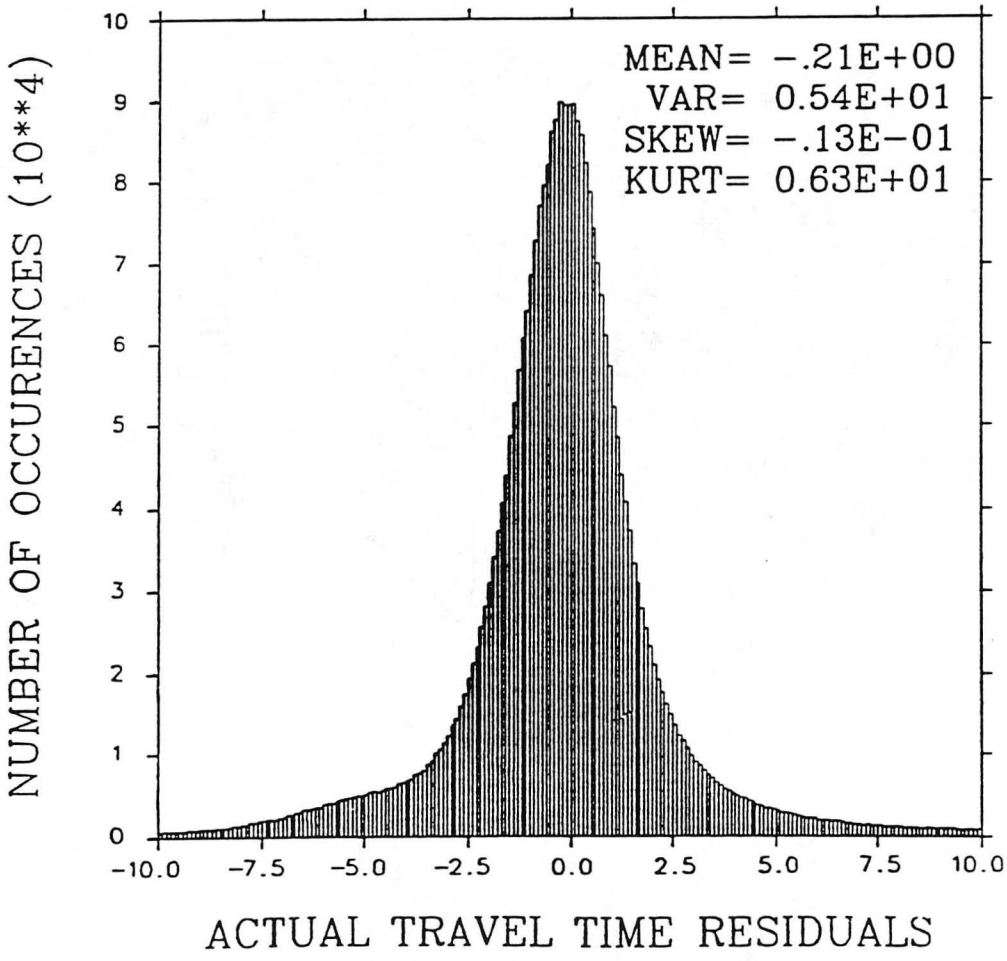

MEAN= -.21E+00
VAR= 0.54E+01
SKEW= -.13E-01
KURT= 0.63E+01
NUMBER OF OCCURENCES (10**4)
0
1
2
3
4
5
6
7
8
9
10
-10.0
-7.5
-5.0
-2.5
0.0
2.5
5.0
7.5
10.0
ACTUAL TRAVEL TIME RESIDUALS

Figure 4.5

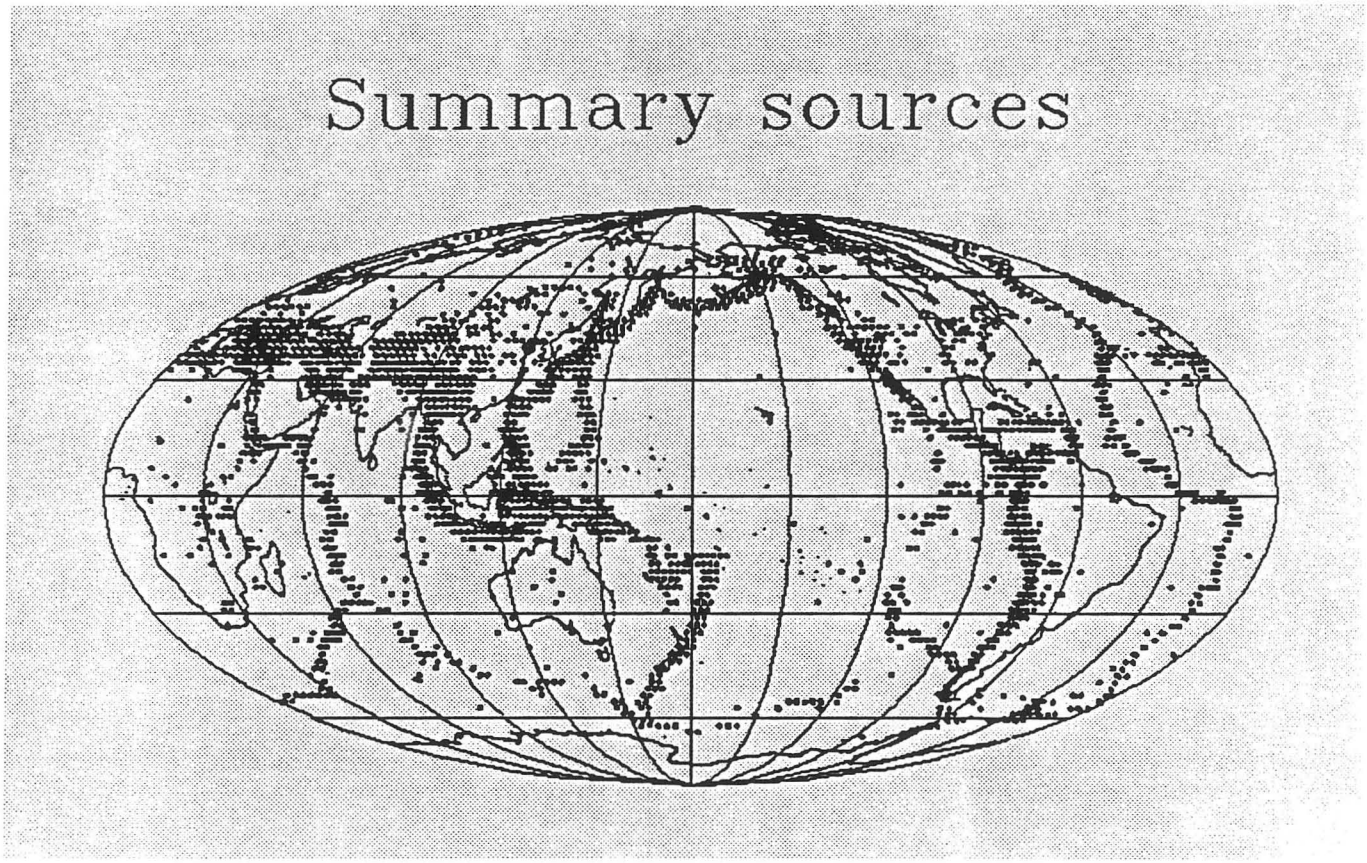


Figure 4.6

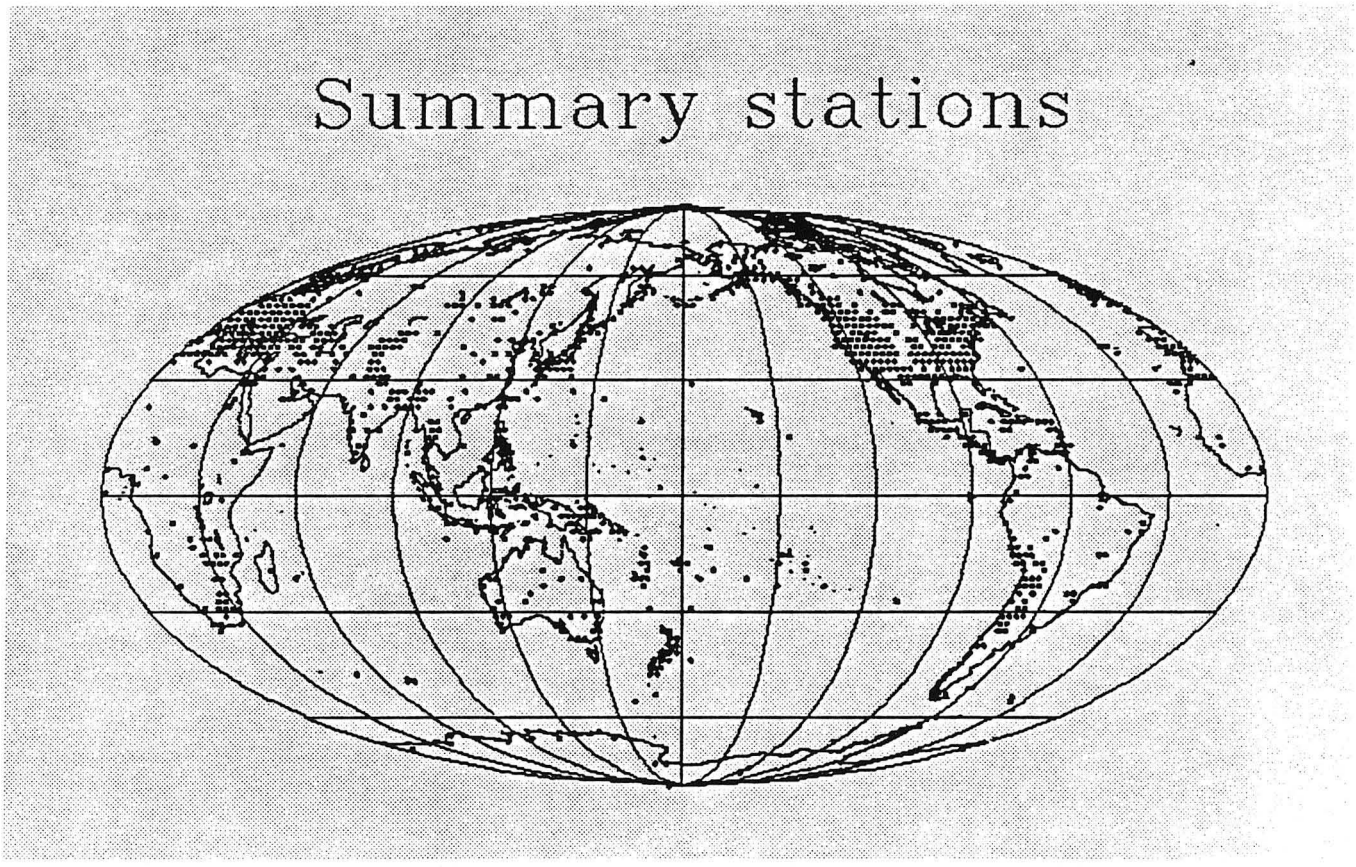


Figure 4.7

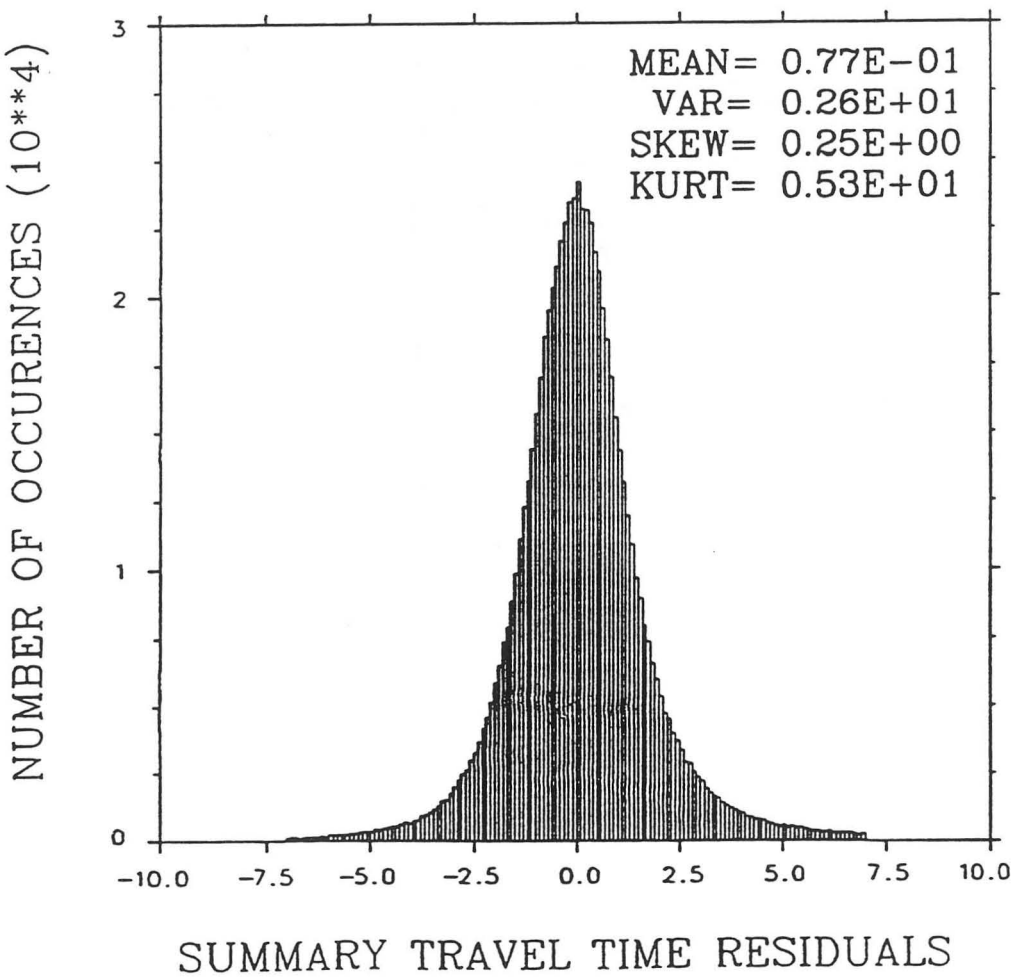

MEAN= 0.77E-01
VAR= 0.26E+01
SKEW= 0.25E+00
KURT= 0.53E+01
NUMBER OF OCCURENCES (10**4)
0
1
2
3
-10.0
-7.5
-5.0
-2.5
0.0
2.5
5.0
7.5
10.0
SUMMARY TRAVEL TIME RESIDUALS

Figure 4.8

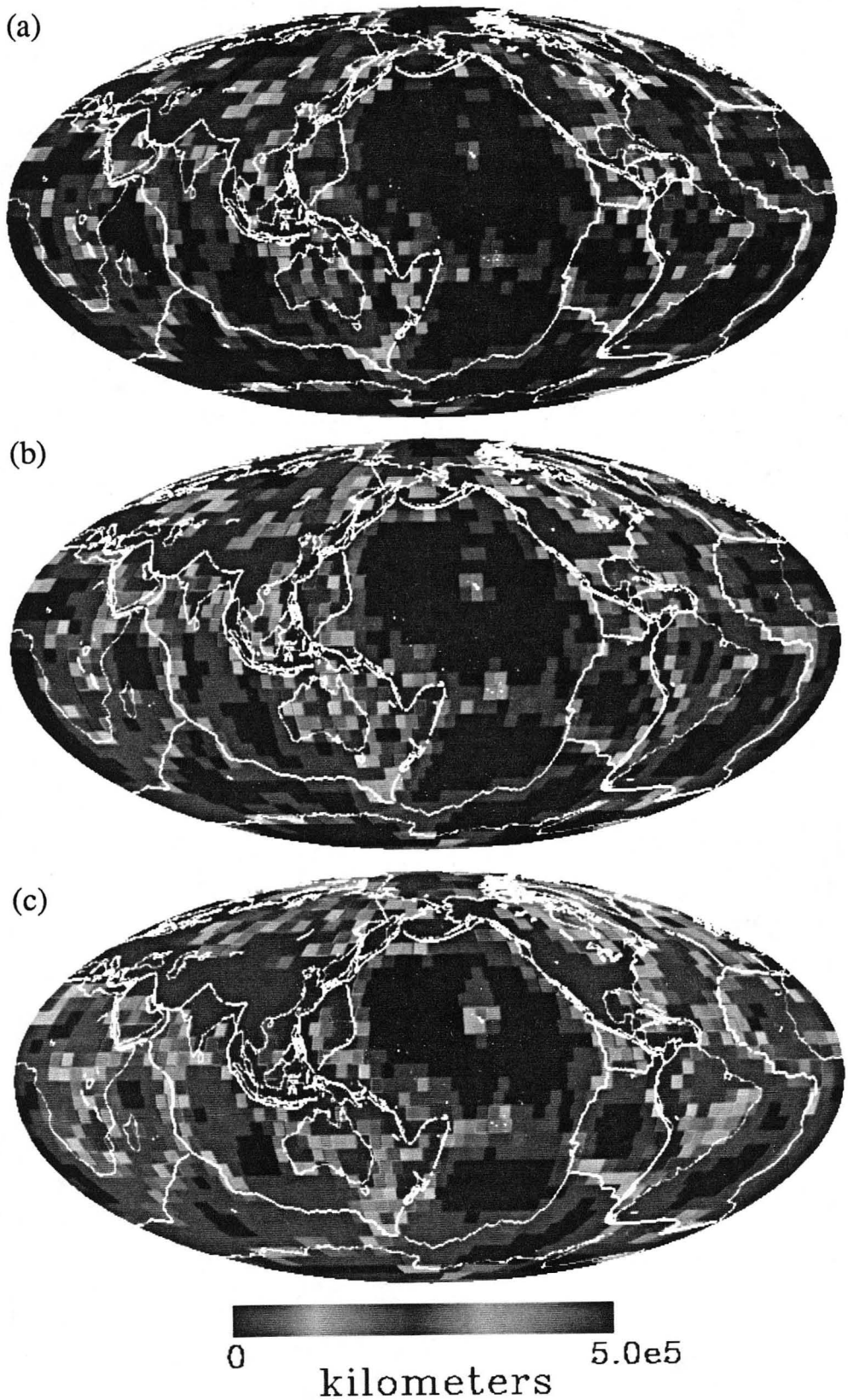
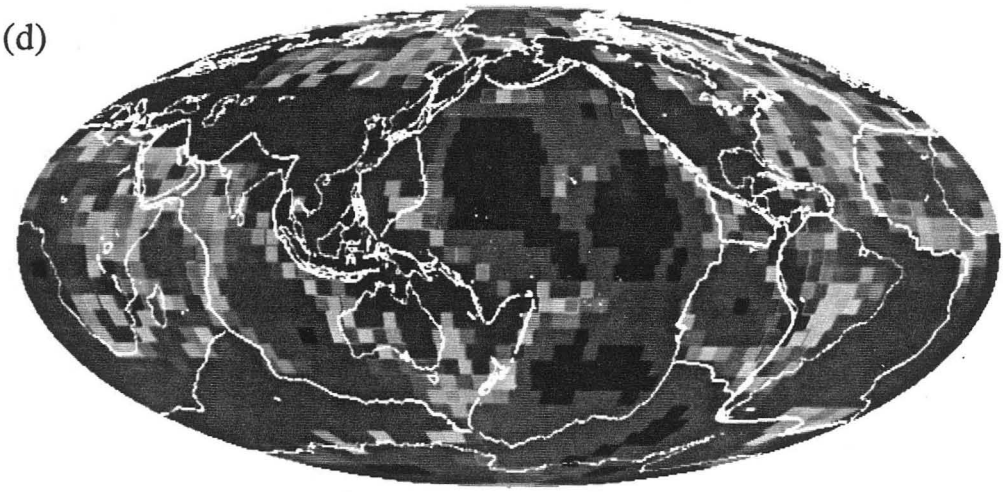
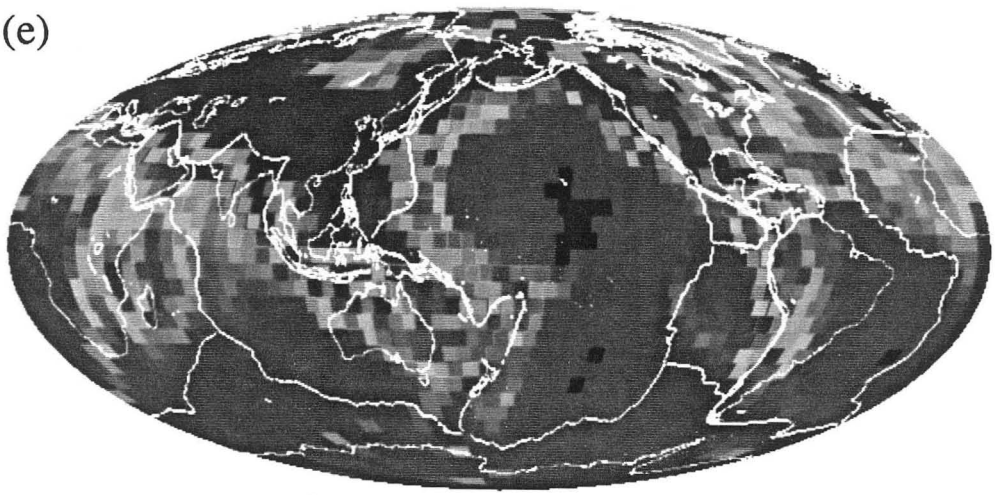


Figure 4.8

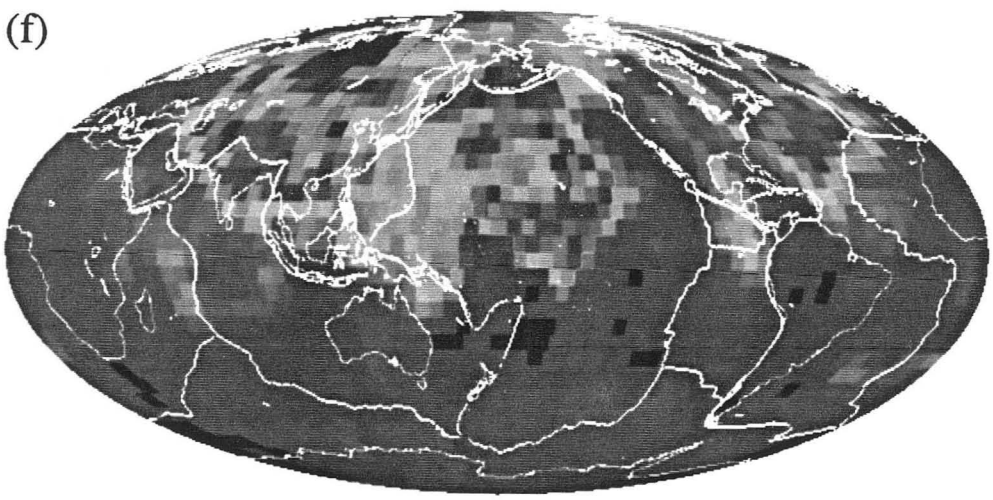
(d)



(e)



(f)



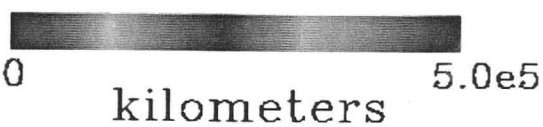


Figure 4.9 (a) (b)

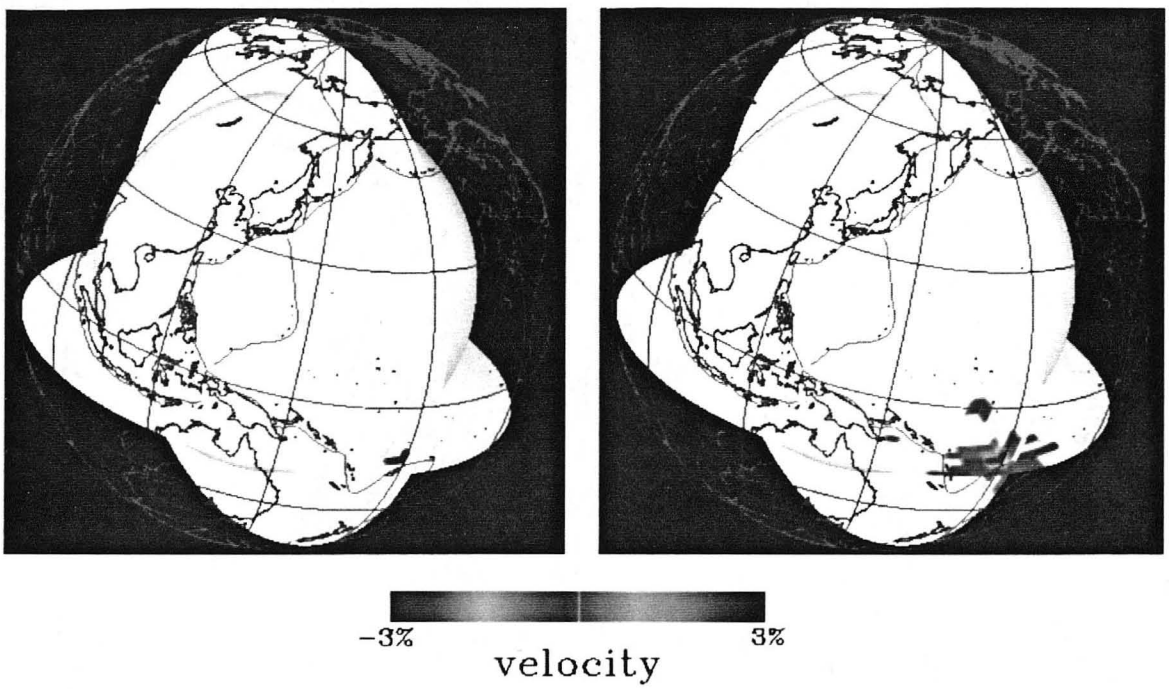


Figure 4.10 (a) (b)

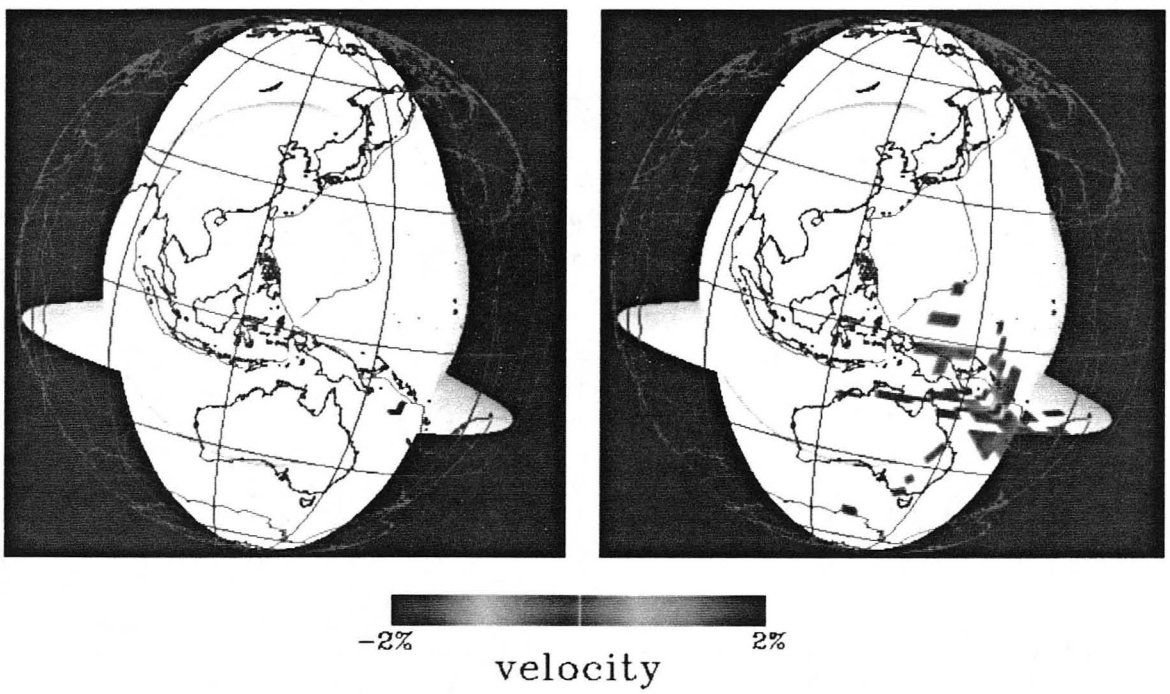


Figure 4.11 (a) (b)

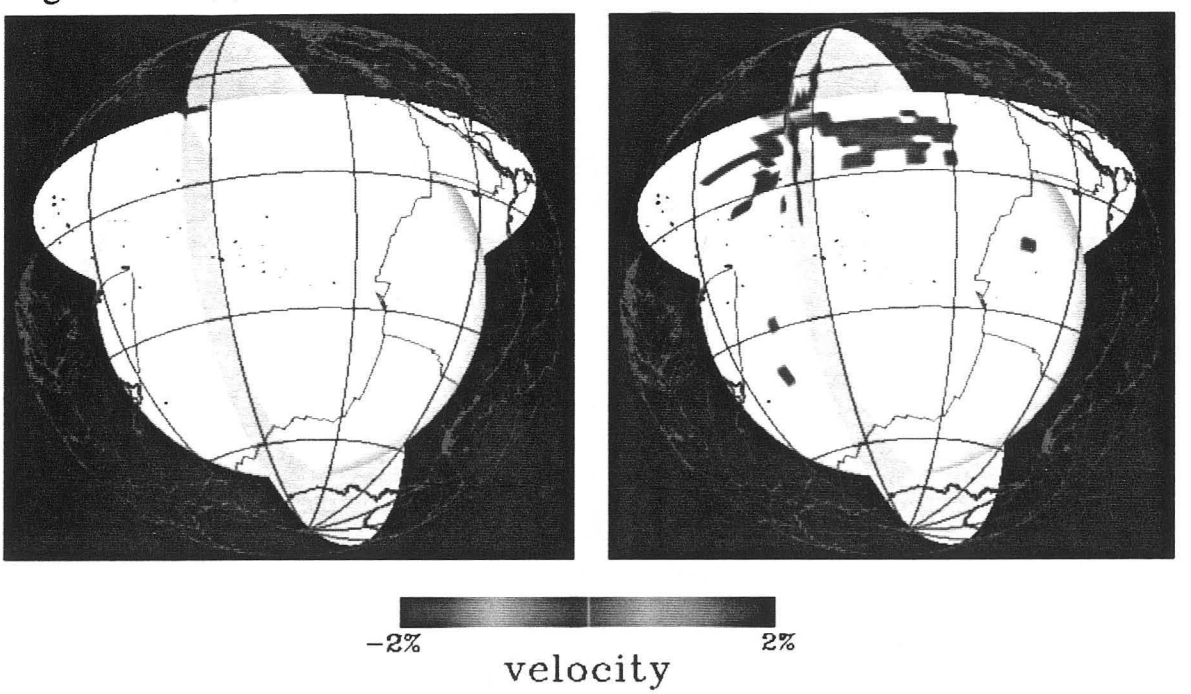


Figure 4.12 (a) (b)

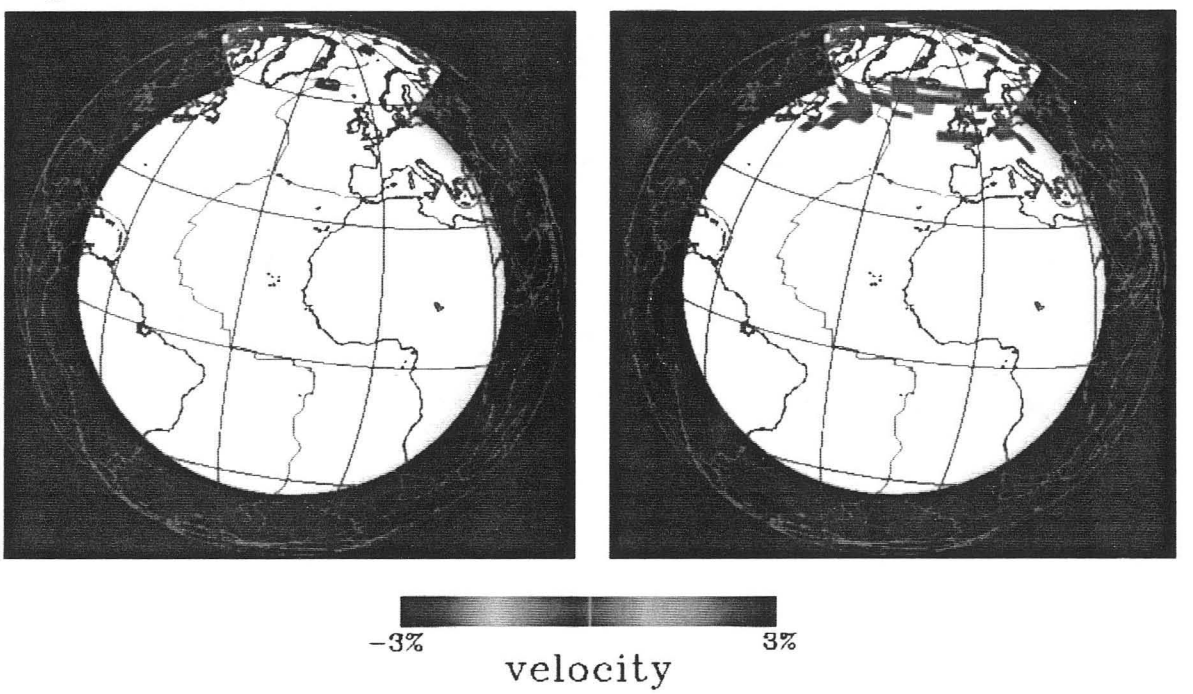


Figure 4.13 (a) (b)

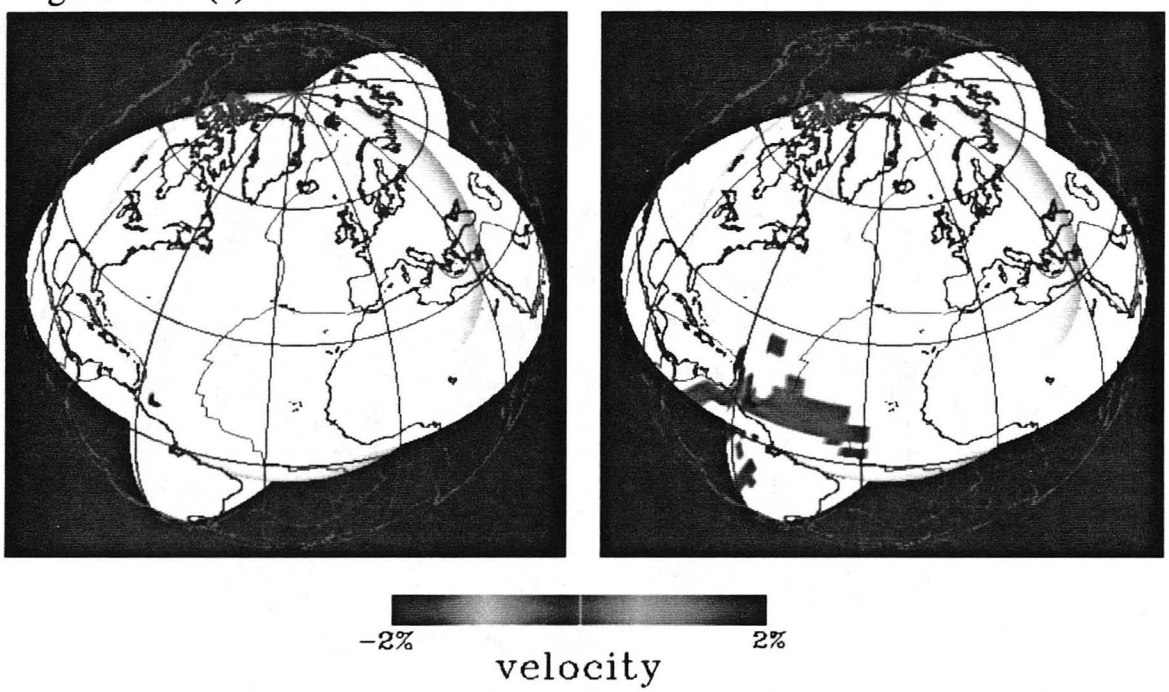


Figure 4.14 (a) (b)

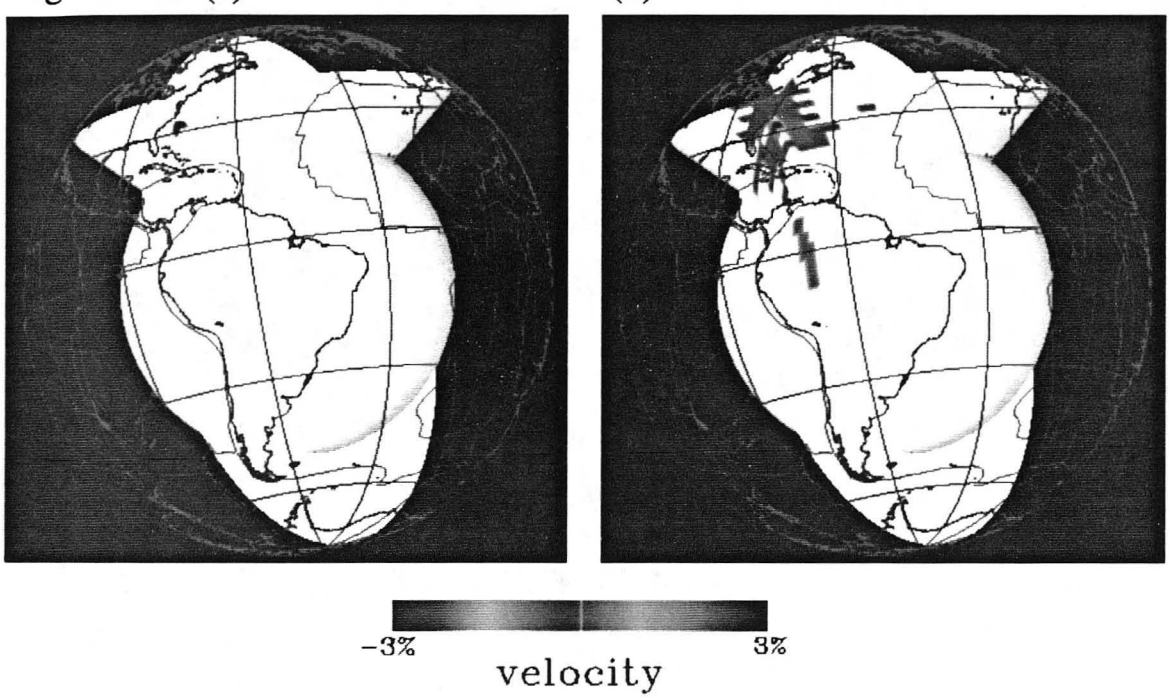


Figure 4.15

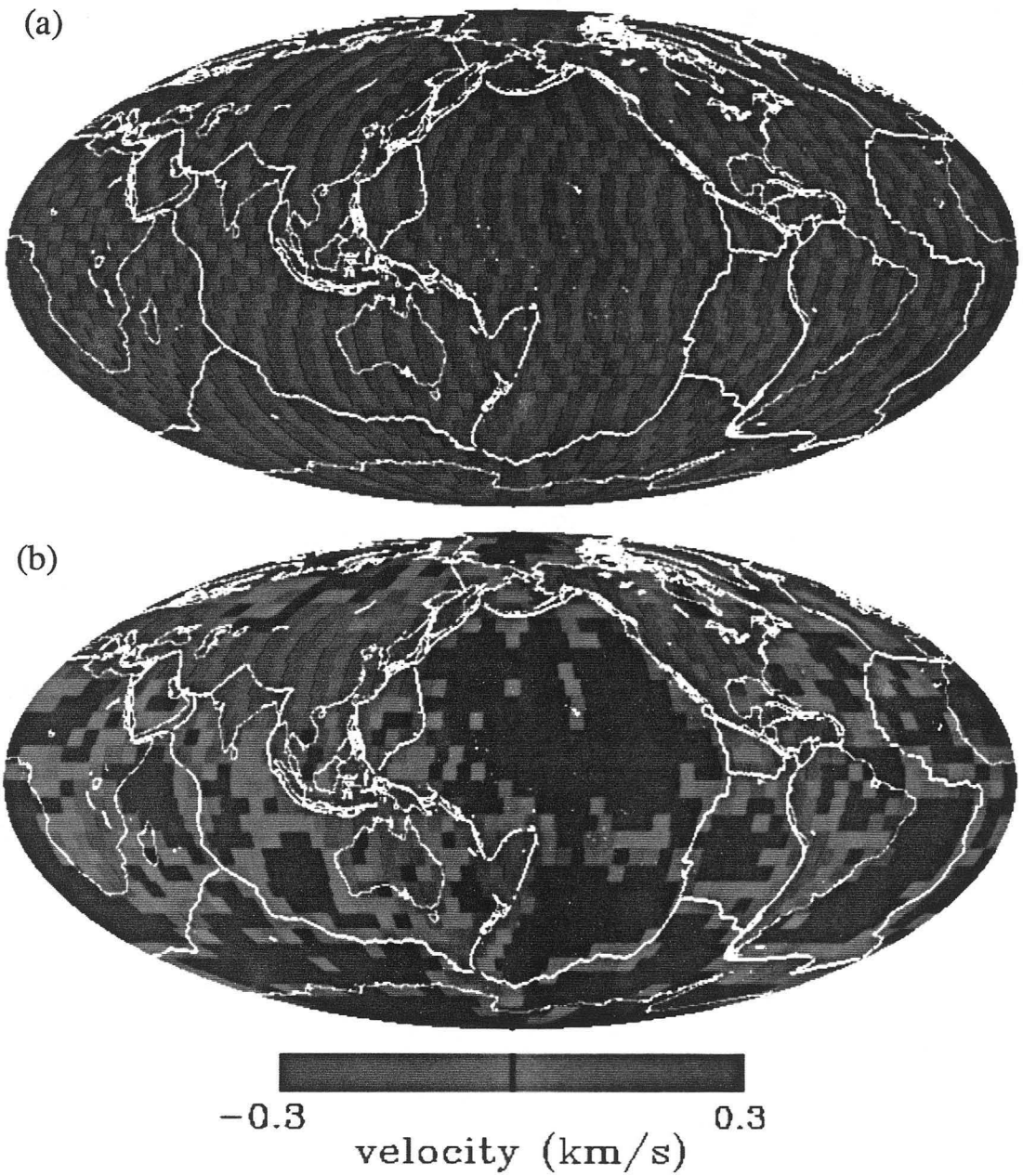
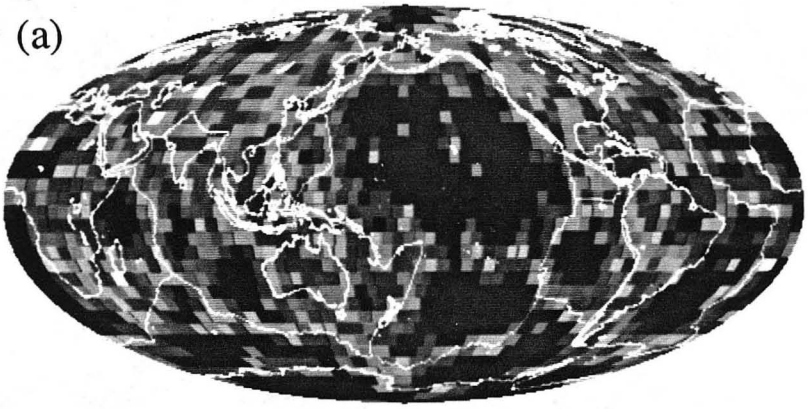


Figure 4.16

(a)



(b)

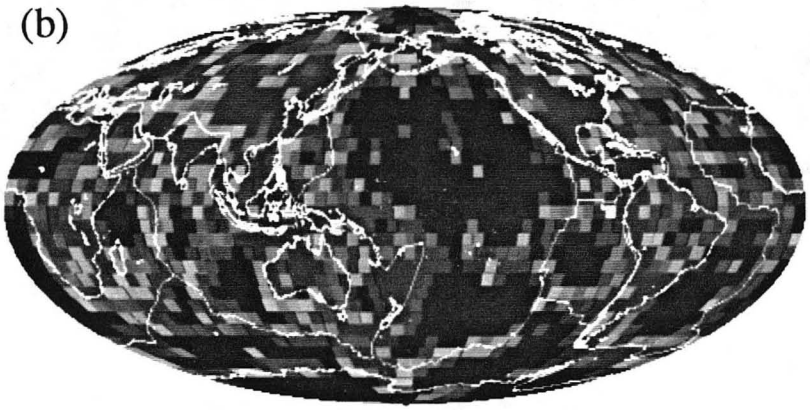
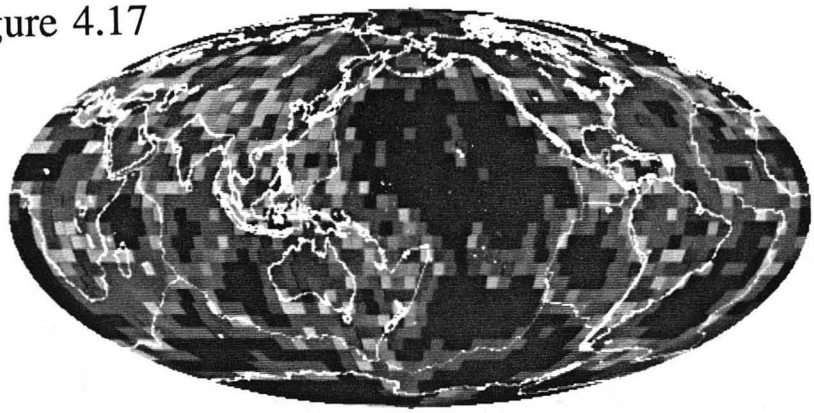


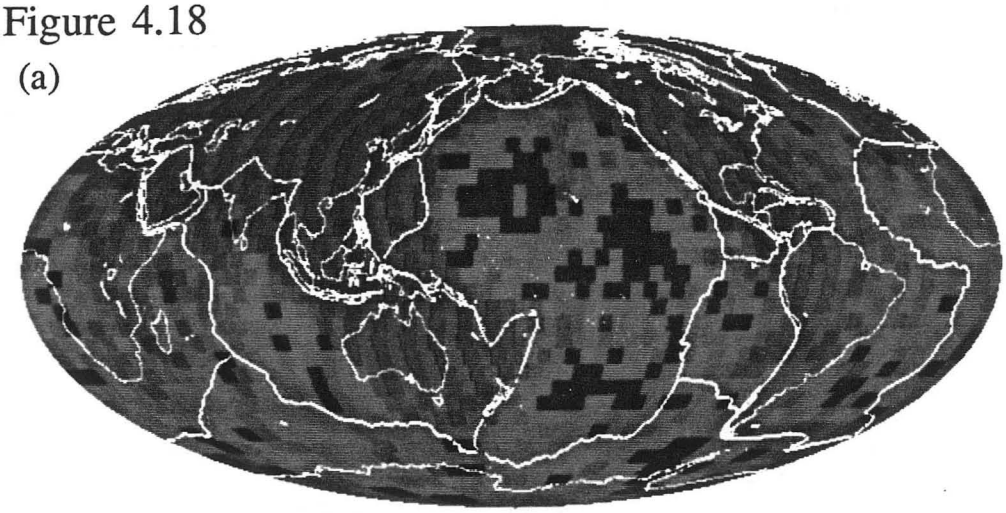
Figure 4.17



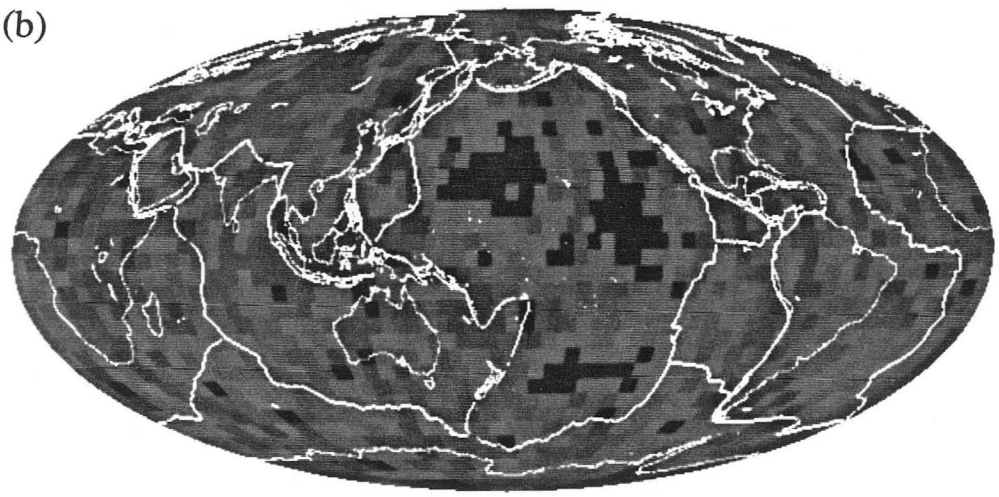
% of average velocity

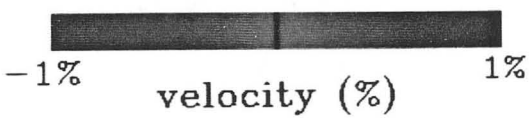
Figure 4.18

(a)

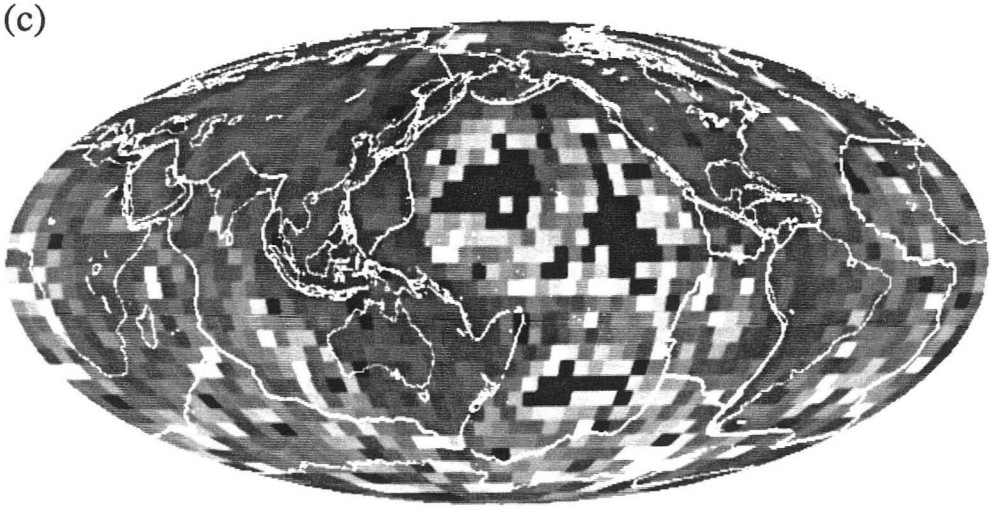


(b)





(c)



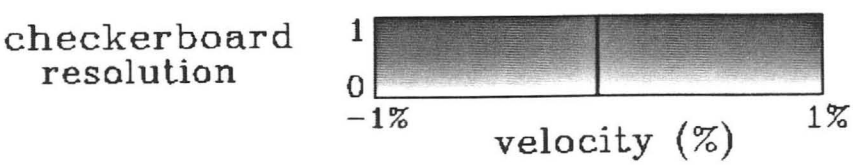
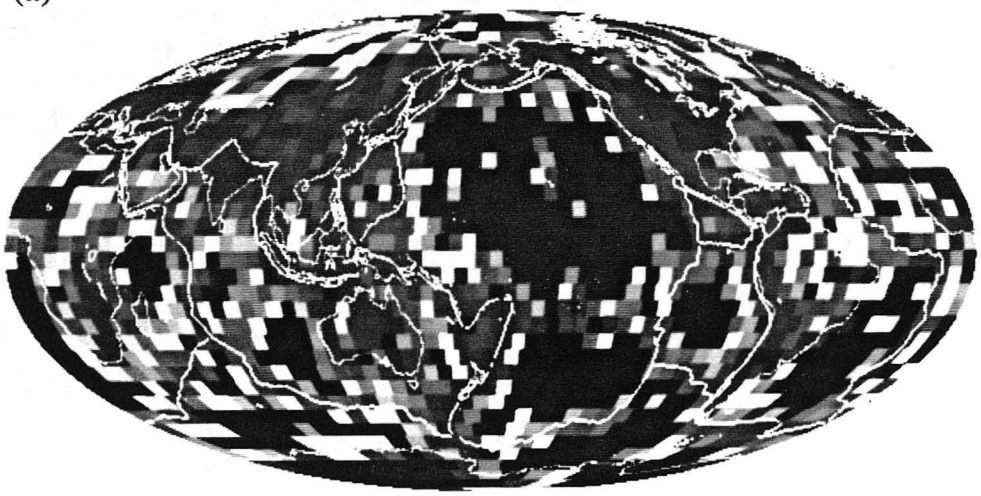
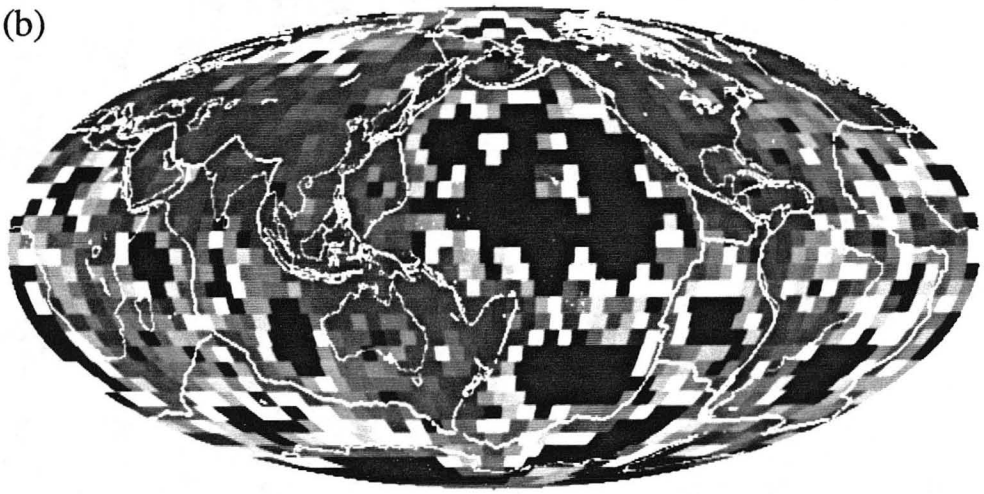


Figure 4.19

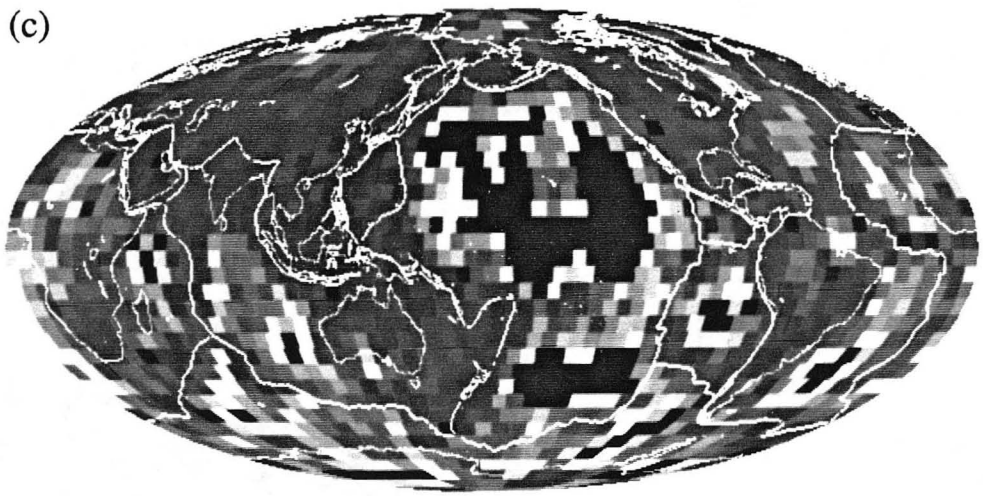
(a)



(b)



(c)



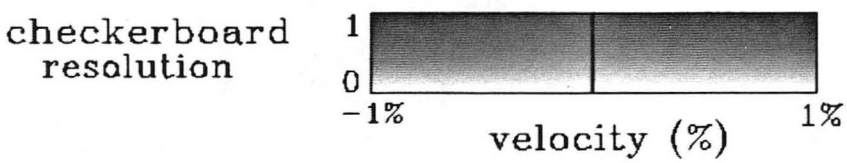
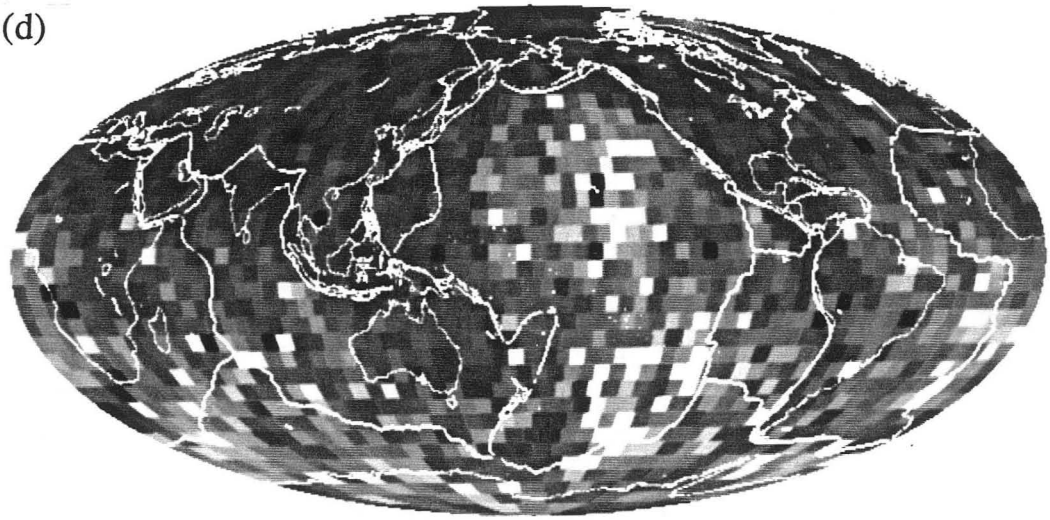
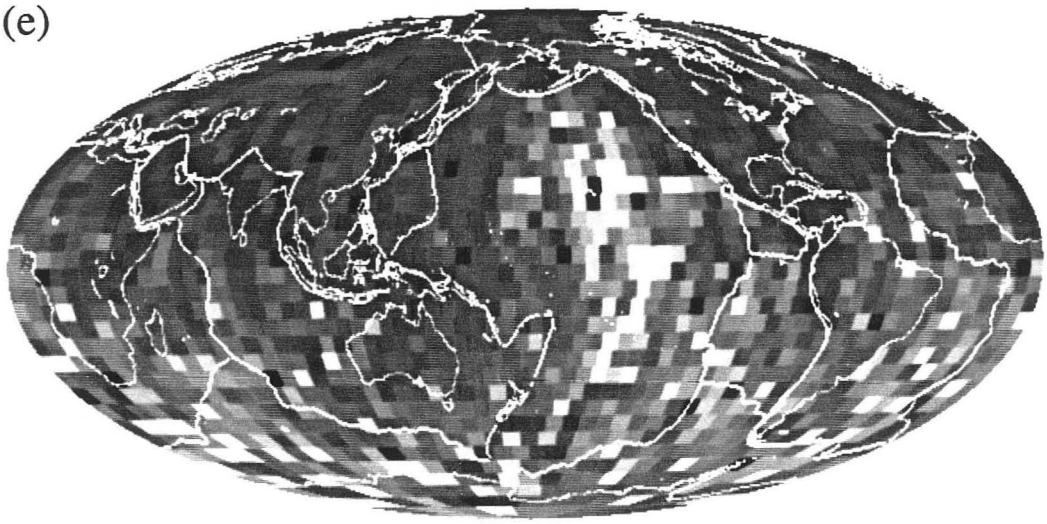


Figure 4.19

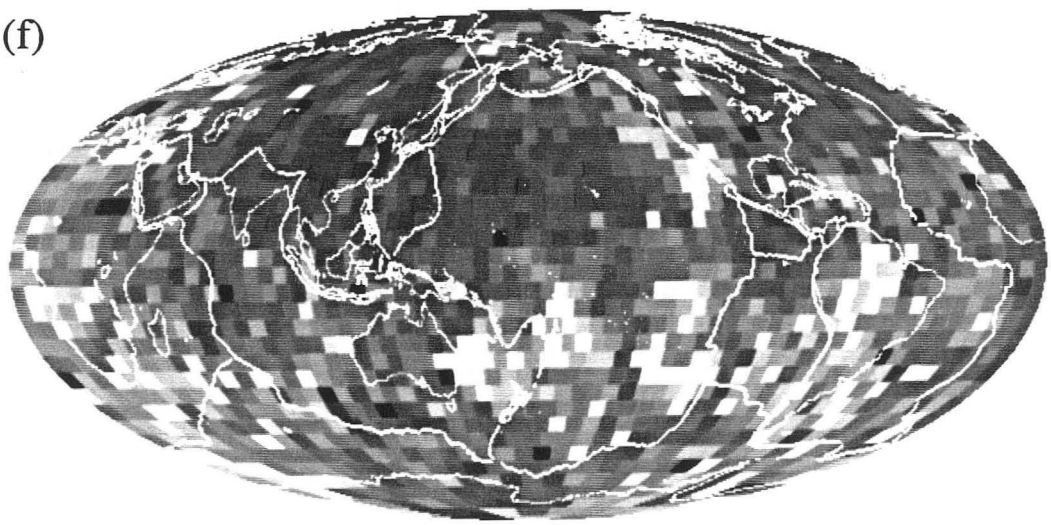
(d)



(e)



(f)



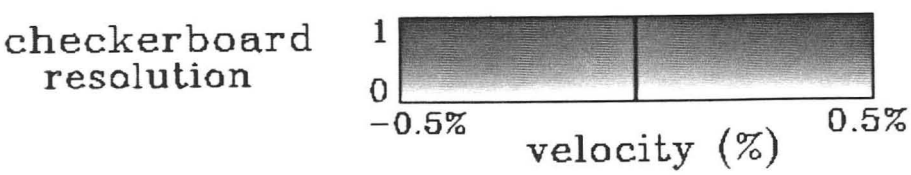


Figure 4.20

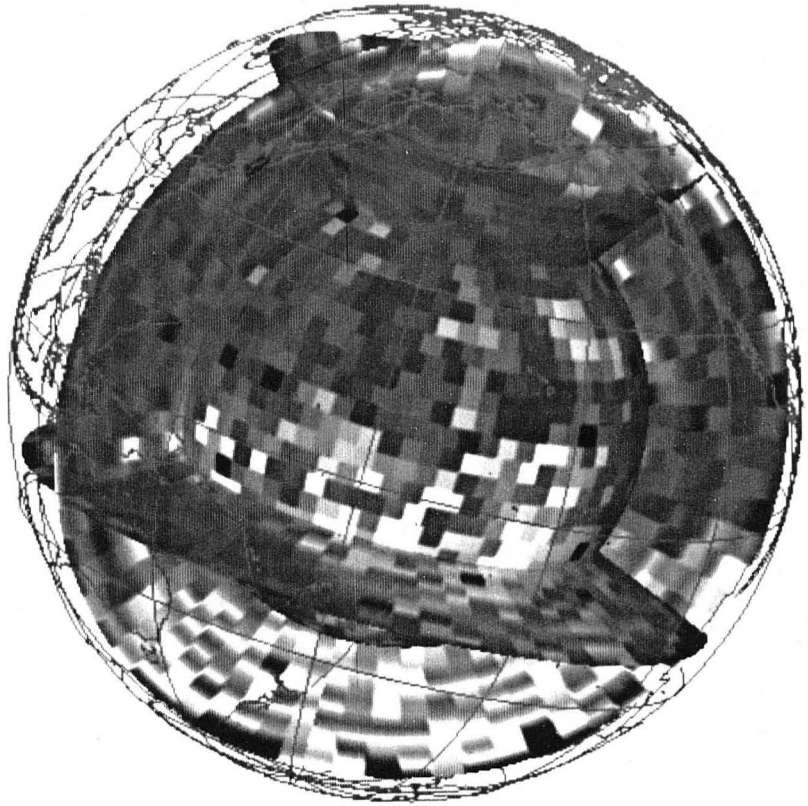
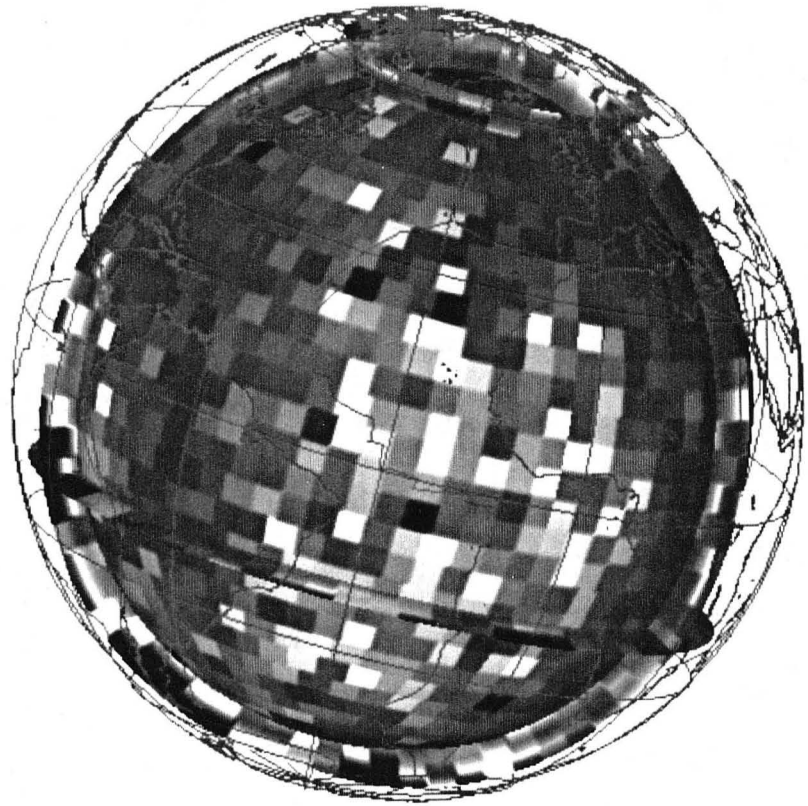


Figure 4.21



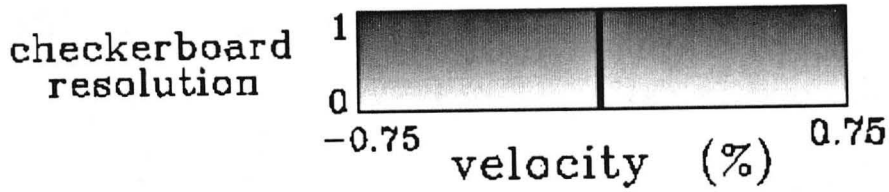


Figure 4.22

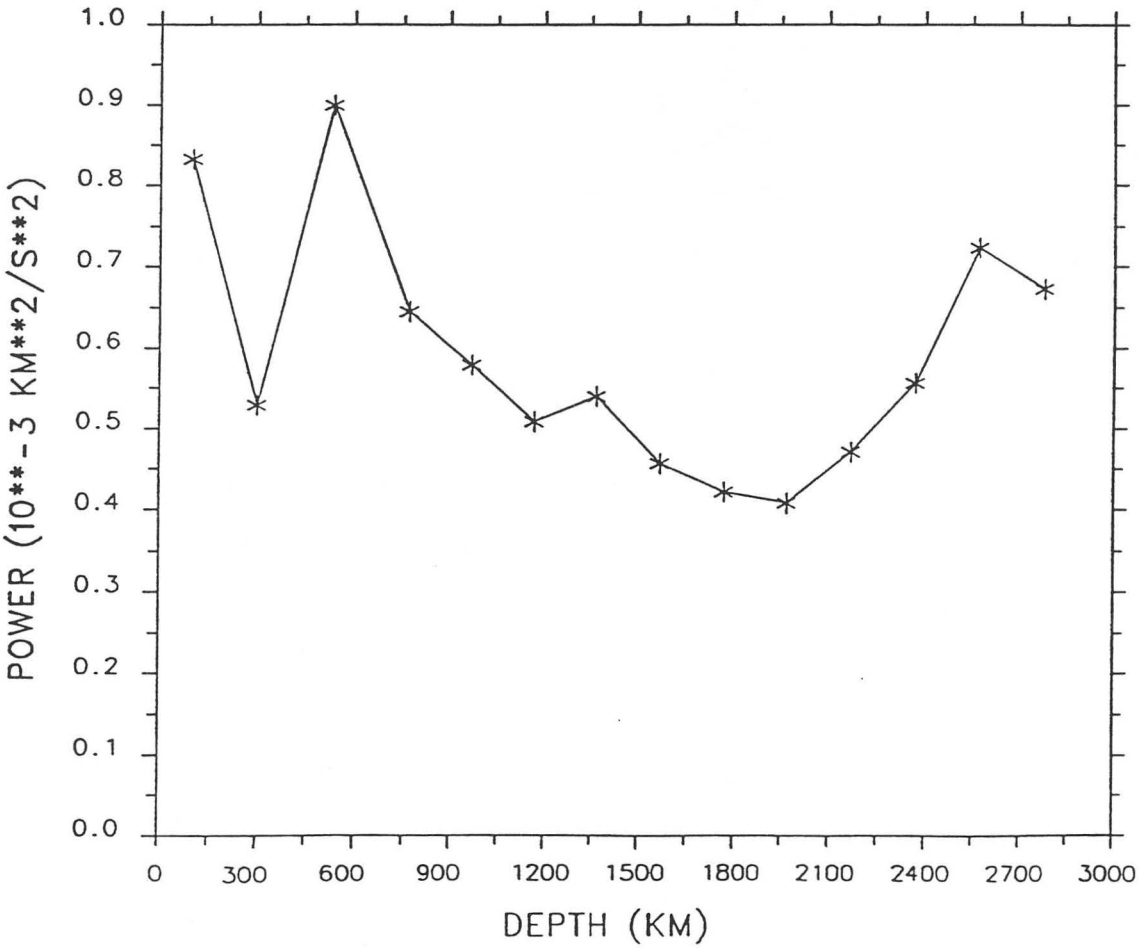

1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
POWER (10**-3 KM**2/S**2)
0
300
600
900
1200
1500
1800
2100
2400
2700
3000
DEPTH (KM)

Figure 4.23

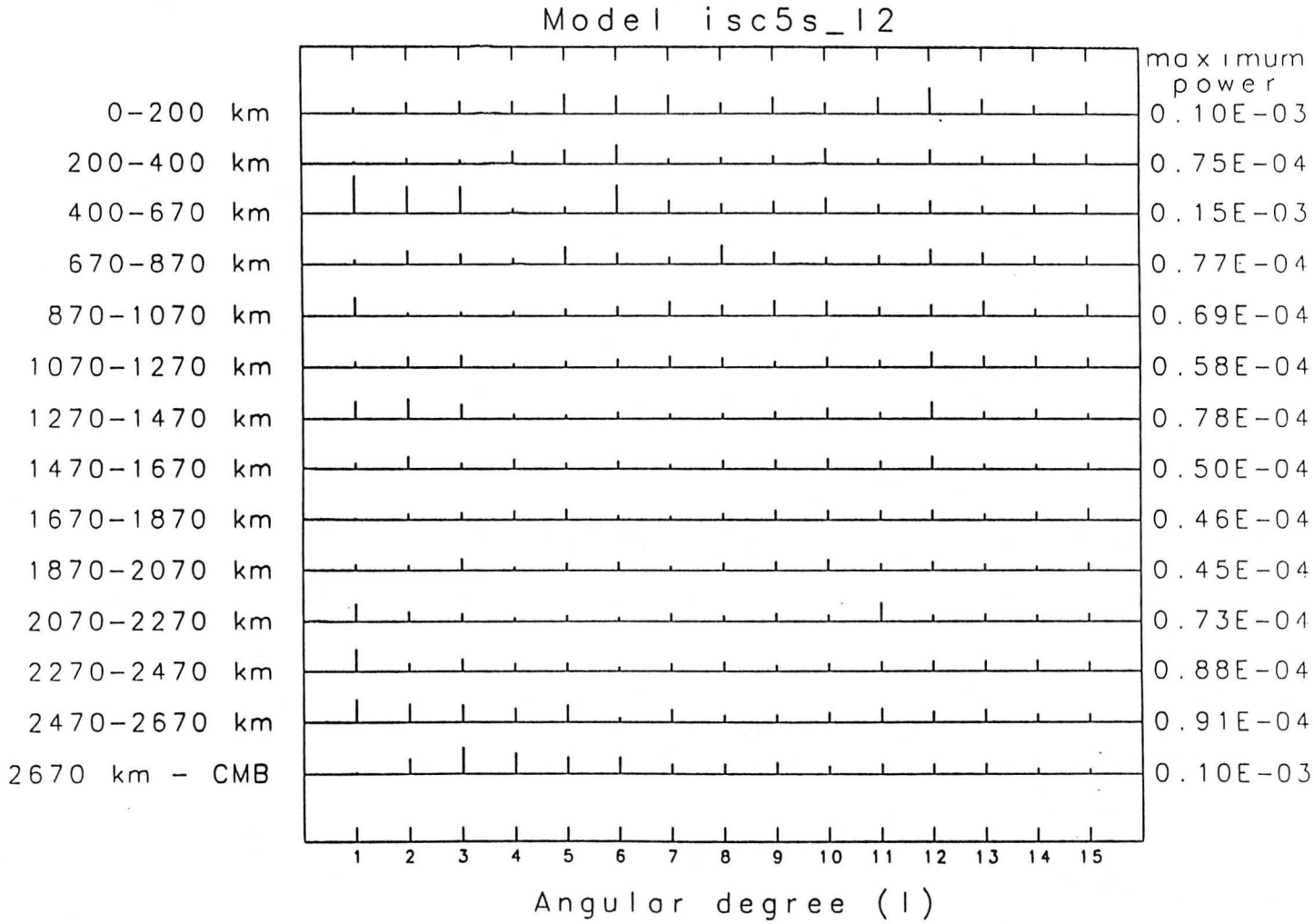
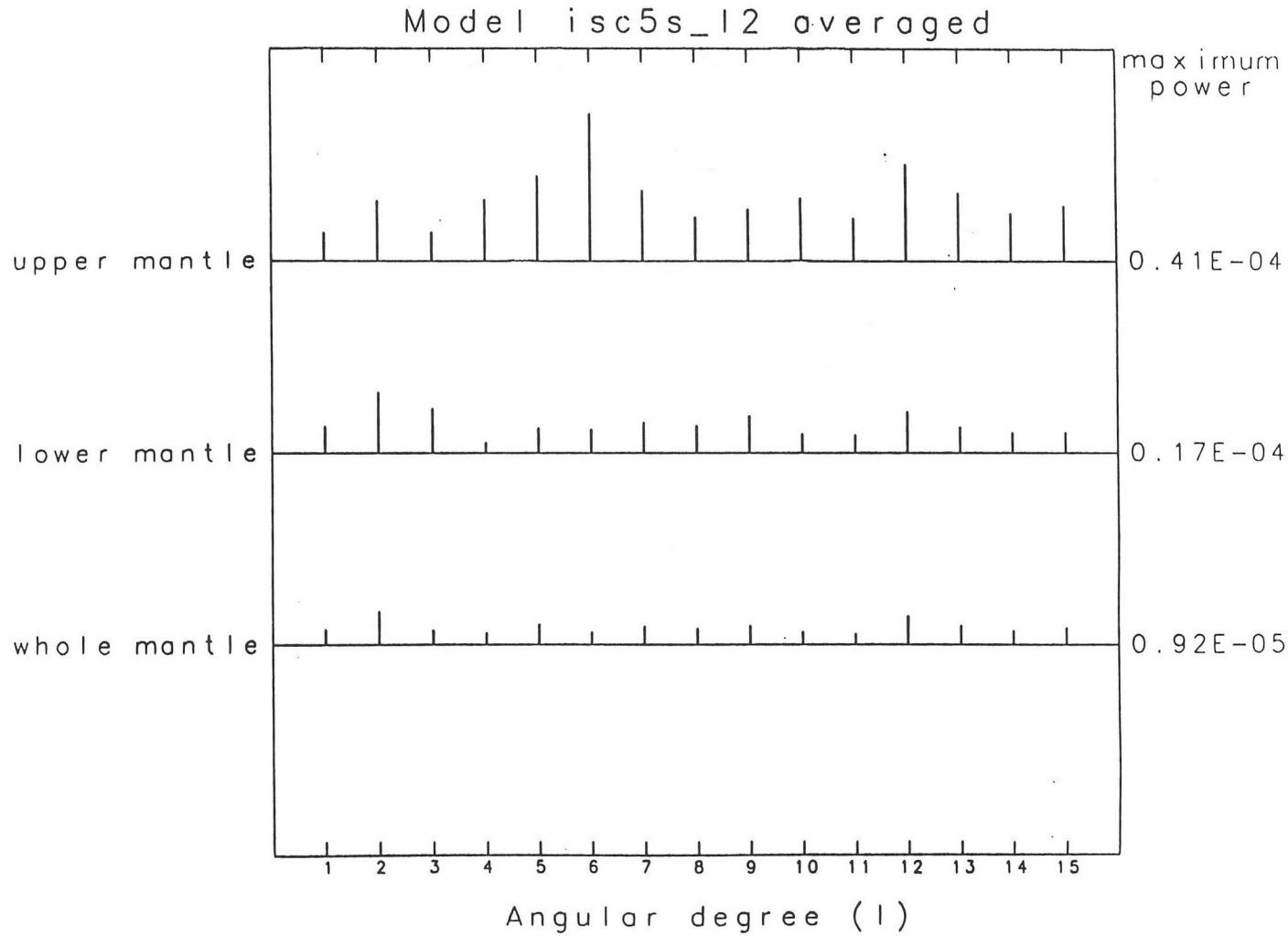

Model isc5s_l2
maximum power
0-200 km
0.10E-03
200-400 km
0.75E-04
400-670 km
0.15E-03
670-870 km
0.77E-04
870-1070 km
0.69E-04
1070-1270 km
0.58E-04
1270-1470 km
0.78E-04
1470-1670 km
0.50E-04
1670-1870 km
0.46E-04
1870-2070 km
0.45E-04
2070-2270 km
0.73E-04
2270-2470 km
0.88E-04
2470-2670 km
0.91E-04
2670 km - CMB
0.10E-03
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Angular degree (l)

Figure 4.24


Model isc5s_l2 averaged
maximum power
upper mantle
0.41E-04
lower mantle
0.17E-04
whole mantle
0.92E-05
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Angular degree (l)

Chapter 5

Conclusions

5.1 Summary and Discussion

We develop a formalism for the tomographic inversion of seismic travel-time residuals and solve the resulting system of travel-time equations in three ways: (a) directly, neglecting source terms, (b) simultaneously for both velocity model terms and corrections to the source locations, and (c) progressively, for each set of terms in succession. Both algorithms perform least-squares inversions which minimize the l^2 norm of the residuals. The methods differ primarily in their treatment of source mislocation terms. We explore the algorithms' performance in conjunction with synthetic data through simulations of the general procedure used to produce tomographic images of Earth's mantle from global earthquake data. Specifically, we investigate the effects of mislocated earthquakes on the velocity model obtained in an inversion and the ability of our simultaneous and progressive inversion techniques to correct mislocated earthquakes and produce an accurate velocity model.

Simulations of a global tomographic inversion scheme demonstrate that ignoring the effects of source mislocation results in underestimating velocity anomalies by up to 50%, creates smeared anomalies in adjacent voxels with values up to 50% of the retrieved velocity of its neighbor, and creates anomalies elsewhere in the mantle with values greater than those estimated for true anomalies. Clearly, careful treatment of the source location problem is critical to the accurate retrieval of three-dimensional velocity variations. The progressive inversion developed here generally produces more accurate source corrections and velocity anomaly estimates than does an inversion scheme in which both source corrections and velocity terms are found simultaneously. These results are superior particularly with respect to the suppression of artificial

anomalies in the velocity estimation.

We also invert real data supplied by the ISC. We use P arrival data from January 1964 through January 1987 and our inversion algorithms to solve for three-dimensional P velocity models of the mantle and source mislocations. The model mantle is parametrized by approximately equal-area blocks: $10° \times 10°$ and generally 200 km in depth. Nearly 345,000 rays from more than 3,000 shallow events satisfying selection criteria are included in the inversions.

Our visual comparison of the models found by neglecting source terms (ISC10_direct), by solving simultaneously for source and velocity terms (ISC10_sim), and progressively (ISC10_pro) reveals that upper mantle differences are located overwhelmingly in source regions, implying that the model differences result from the differences in our treatment of the source terms. Despite the concentration of model differences in source regions, the effects of different treatments of the source appear in the velocity models at both long and short wavelengths in the upper mantle. Differences between the models diminish with depth.

Our comparison of the source corrections produced by simultaneous and progressive inversion supports the interpretation that model differences are concentrated in source regions. Source corrections emerging from the progressive inversion are generally two to four times greater than the simultaneous inversion's corrections. Some events get extremely large corrections, in the progressive inversion, but only moderate corrections in the simultaneous inversion.

We investigate which patterns of seismic heterogeneity in the mantle would be returned reliably by a tomographic inversion in which the model mantle is parametrized by a set of discrete, non-overlapping voxels. We find that while the power input to a particular pattern of heterogeneity in the 0-200 km layer is generally recovered accurately, the pattern itself is poorly determined in this layer. A pattern in the 200-400 km layer is more precisely determined, though the power contained in the

pattern is consistently underestimated and more leakage occurs to the layers above and below. The transition zone, 400-670 km, shows similarly strong control of lateral heterogeneity patterns, but tests return a more accurate estimate of input power than for the second layer. The l = 2, 4, and 6 components are all recovered accurately in the 400-670 km layer. This result supports previous findings from inversions with real data that $l = 2$ is a significant pattern of heterogeneity in the mantle's transition zone and that $l = 4$ is not a significant pattern. For the entire upper mantle, $l = 6$ would be retrieved reliably and its constructive behavior in upper mantle models derived with real data is confirmed. These tests also demonstrate the inability of our inversion procedure to retrieve shorter-wavelength features in the lower mantle. Results for our lowermost layer, D", must be considered suspect due to the inadequate constraints placed on model values by our ray coverage and the sensitivity of these results to noise in the data.

In an effort to find the most accurate and smallest-scale model possible for the mantle, we use ISC P-arrival data (1964-1987) and the LSQR algorithm to solve for a three-dimensional P-velocity model, source mislocations, and station corrections simultaneously. The model mantle is parametrized by approximately equal-area blocks: 5° × 5° and generally 200 km in depth. More than 3 million rays from 46,000 shallow events satisfying selection criteria are averaged according to 2° × 2° × 10 km deep bins to construct nearly 726,000 summary rays for the inversion.

Due to the averaging procedure employed to construct summary events, source location corrections are generally small. Source location corrections in subduction zones generally move the source toward the positive velocity anomaly. Sources in continental regions have the smallest, nearly insignificant, corrections. Station corrections range from -0.94 to +1.07 sec, with no obvious correlation to tectonics or elevation of the station region.

Resolution and covariance are evaluated by approximate methods. Resolution is estimated by the inversion of a synthetic checkerboard test pattern, and calculating point spread functions for selected voxels. Covariance is estimated by averaging results from inversions of realistic errors and by a jackknife procedure.

We present our three-dimensional velocity model in conjunction with the resolution estimates produced by our checkerboard test. Normalized checkerboard output values, ranging from 0 to 1, are used to modify each voxel's red-blue velocity value from full color saturation, indicating good resolution, to white, which indicates no resolution. The velocity model shows a fast anomaly in the lower mantle beneath the Tonga-New Hebrides subduction zone to a depth of 1670 km, and another fast anomaly beneath the Japanese Island arc and eastern Asia reaching nearly to the core-mantle boundary. Continuity between these anomalies and shallower fast anomalies is not clear. A fast anomaly extending from 670 km to 2070 km depth appears beneath the eastern United States, Caribbean Sea, and Central South America. In addition, a number of slow anomalies associated with hotspots extend through the upper mantle but are extinguished in the lower mantle by our resolution weighting. Mid-ocean ridges are associated with moderately slow anomalies in the top 400 km of our model. The transition zone shows large l=1, 2, and 3 spherical harmonic components. Diminished heterogeneity in the lower mantle, reported by other authors, is confirmed by our study.

5.2 Recommendations for Further Studies

The checkerboard resolution tests and noise maps presented in chapter 4 indicate that the quality of our mantle images leave much to be desired in the southern hemisphere and in oceanic regions throughout the depths of the mantle. These tests offer visual images of the strength of the constraints placed on model values by P waves propagating directly through the mantle. Undersampling of a given voxel, in the sense

that the sampling is inadequate to provide a good average of the voxel's velocity variations, and parallel orientation of rays in a voxel will provide relatively weak constraints on the voxel's value. Completely unsampled regions cannot be imaged and thus decrease the value of our model as an independent reference for modeling of other physical parameters. One way to increase both the extent and quality of mantle ray coverage is to introduce additional seismographic stations in sparsely sampled or unsampled regions. While long recognized as a priority by seismologists, technical and cost considerations have slowed the realization of such a venture. Unsampled regions are generally in the deep oceans. Still, plans are underway to place a few prototype seismographs in the open ocean, if only temporarily and on relatively shallow ridges or seamounts. The problem of determining the ideal locations for the next few seismic instruments is not trivial. With cost and technical issues so daunting, instrument placement decisions cannot be made solely in pursuit of specific seismological goals. A formalism should be developed through which the value of additional instruments in various locations to different types of studies may be assessed. Unfortunately for our purposes here, placement decisions would probably be driven by the needs of global or regional surface wave studies and local studies of various types, rather than by body-wave studies of the deep Earth. Body-wave studies such as ours require data collected at many stations over quite a few years, so the marginal improvement to our models provided by an additional few stations operating for a few years would be small. Surface wave inversions, in contrast, use much more information from each seismogram and would benefit far more from a few new stations.

In the meantime, more extensive ray coverage with better geometrical orientation may be obtained by including additional phases from the ISC catalog in our inversions. Pioneering efforts by *vanderHilst and Engdahl* [1991] demonstrate the utility of reflected P phases, pP and PP, to inversions. The price we pay for these additional constraints is the larger errors inherent in data that arise from an indeterminate bounce

point. Reflections at the Mohorovicic discontinuity, the ocean-crust interface, and the ocean's free surface are likely to arrive within about 10 to 15 seconds of each other and would therefore all be included in a residual distribution truncated at ±7 seconds or greater. Corrections for bathymetry or topography would be approximate as well. The increased time spent by a bouncing phase in the upper mantle makes it relatively more sensitive to upper mantle, rather than lower mantle, structure. This is an advantage in that it allows us better to constrain upper mantle velocity anomalies and control the downward mapping of upper mantle anomalies. An additional advantage, particularly with the inclusion of pP data, is the improved constraint of earthquake hypocenter locations.

With respect to a three-dimensional model's veracity, every effort should be made to start the linearization required by the inversion about an accurate average (i.e., one-dimensional) model. Our starting model, the Jeffreys-Bullen P model, has known deficiencies. The upper mantle contains no discontinuities and no low velocity zone. The lower mantle was corrected from the model provided by *Jeffreys* [1960] to produce an accurate match to the J-B travel-time tables [*Jeffreys and Bullen*, 1940] which removed a systematic slow trend in the mid- to lower mantle. The ultimate effects of the lack of discontinuities and low velocity zone in our starting model is discussed extensively in chapter 4. Repeating our imaging procedure with a more accurate model, such as the *iasp*91 P model [*Kennett and Engdahl*, 1991] would surely produce a more accurate result for the upper mantle, though significant improvement in the lower mantle is not guaranteed. The cost of such a project is that the entire ISC catalog of events must first be relocated in the new one-dimensional model.

In the imaging procedure itself, one might correct the one-dimensional model after a three-dimensional model is produced, relocate the events in the updated 1-D model, and again produce a 3-D model, as do *Inoue et al.* [1990]. This represents a closer approximation to a nonlinear optimization for the solution to the travel-time

equations. However, this scheme will almost certainly produce a solution to the linearized problem that is associated with a local, rather than global, minimum to the nonlinear problem. Assessment of the estimated solution's deviation from the true solution is not possible and the improvement of the solution which emerges from several imaging and relocation steps over the solution found after one step is also problematic.

Complete relocation of the events in a 3-D model is not yet feasible, but the strongest three-dimensional effects on subduction zone events may be taken into account by a regionalized model, such as the ones produced by *Tralli and Johnson* [1986a] and used to relocate events by *Tralli and Johnson* [1986b] or by a subduction zone model only, such as the procedure used by *Engdahl and vanderHilst* [1991]. A regionalized relocation scheme which accounts for subducting lithosphere may provide significantly superior locations for the great majority of the world's earthquakes. From the point of view of tomographic imaging, a more sophisticated relocation procedure introduces additional possibilities for bias into the inversion for velocity structure. This is unfortunate but does not necessarily detract from the value of resulting models if the event relocations are demonstrably superior to the original locations.

The choice to minimize the l^2 norm of the travel-time residuals in the inversion is not a simple one. Strictly, least-squares is most appropriate for problems involving a Gaussian distribution of errors. When applied to such a distribution, least-squares produces the maximum likelihood solution to the linear matrix equation. But residuals contained in ISC travel-time data are not clearly Gaussian. More observations are found in the distribution's tails than one would expect in a Gaussian distribution. A precise characterization of the distribution of errors contained in the travel-time residuals, minimization of the appropriate residual norm, and an analysis of the consequences for the resulting velocity models would be an important contribution. A companion study to the one undertaken in chapter 4, by *Vasco et al.* [1990], considers the case in which the l^1 norm of the travel-time residuals is minimized, rather than the l^2 norm.

Further analysis of the implications this choice holds for the velocity models is in order.

Regional models of similar spatial resolution and global models of somewhat larger scale-length resolution have been produced with surface waves by researchers with data from various instruments [e.g., *Woodhouse and Dziewonski*, 1986; *Nataf et al.*, 1986; *Tanimoto*, 1990; *Montagner and Tanimoto*, 1990] These data and inversion schemes can constrain more physical variables, such as density and attenuation, than can studies such as ours that use travel-time picks for individual phase arrivals. However, global surface wave and free oscillation studies will likely not be able to reproduce the small scale-length resolution of body-wave studies in the upper mantle for some time to come, and perhaps may never be able to offer small scale-length models of the lower mantle. These global surface wave models for S velocity do provide valuable comparisons for our P velocity models and would provide comparisons for an S model produced with body waves as well. Such models have been produced on both a regional scale [e.g., *Grand and Helmberger*, 1984; *Grand*, 1987] and a global scale [*Davies*, 1986], but a procedure involving the relocation of events in an accurate one-dimensional model and perhaps solving simultaneously for both P and S velocity would provide an important tool for geophysicists. For geodynamical modeling, in particular, a reliable map of three-dimensional density variations in the mantle is more valuable than a map of variations in seismic velocity. A density map produced with both P and S models would be more accurate than a map produced by assuming a simple proportional relation between density and either P or S velocity alone.

Finally, the issue of resolution and uncertainty remains unsettled. The sheer size of the matrices involved in the linearized inversions renders both the calculation and evaluation of resolution and covariance matrices problematic. The capacity to calculate these matrices is within our reach, due to recent advances in massively-parallel computing, and initial attempts to evaluate the resolution and uncertainty in mantle P-

wave tomography have already proven informative with respect to tradeoffs between velocity parameters and spatial averaging [*Vasco et al.* 1990]. Resolution and covariance ideally would be calculated at the same time for all model parameters (velocity, source, and station terms) and the tradeoffs between different classes of parameters could be evaluated.

5.3 References

Davies, J.H., and R.W. Clayton, Lower-mantle S-wave tomography, *EOS Transactions AGU, 67,* 1099, 1986.

Engdahl, E.R., and R.D. van der Hilst, Reprocessing of ISC phase data for northwest Pacific earthquakes, *EOS Transactions AGU, 72,* 191, 1991.

Grand, S.P., Tomographic inversion for shear velocity beneath the North American Plate, *J. Geophys. Res., 92,* 14,065–14,090, 1987.

Grand, S.P., and D.V. Helmberger, Upper mantle shear structure of North America, *Geophys. J. R. Astron. Soc., 76,* 399–438, 1984.

Inoue, H., Y. Fukao, K. Tanabe, and Y. Ogata, Whole mantle P–wave travel time tomography. *Phys. Earth and Planet. Int., 59,* 294–328, 1990.

Jeffreys, H., and K.E. Bullen, *Seismological Tables.* British Association for the Advancement of Science, London, 1940.

Jeffreys, H., *The Earth,* Cambridge University Press, London, 1960.

Kennett, B.L.N. and E.R. Engdahl, Traveltimes for global earthquake location and phase identification, *Geophys. J. Int., 105,* 429–465, 1991.

Montagner, J.P., and T. Tanimoto, Global upper mantle tomography of seismic velocities and anisotropies, submitted to *J. Geophys. Res.,* September 1990.

Nataf, H.-C., I. Nakanishi, and D.L. Anderson, Measurements of mantle wave velocities and inversion for lateral heterogeneities and anisotropy. III. Inversion. *J. Geophys. Res., 91,* 7,261–7,303, 1986.

Tanimoto, T., Long-wavelength S-wave velocity structure throughout the mantle,

Geophys. J. Int., 100, 327–336, 1990.

Tralli, D.M., and L.R. Johnson, Lateral variations in mantle P velocity from tectonically regionalized tau estimates, *Geophys. J. R. Astr. Soc., 86,* 475–489, 1986a.

Tralli, D.M., and L.R. Johnson, Estimation of travel times for source location in a laterally heterogeneous Earth, *Phys. Earth and Planet. Interiors, 44,* 242–256, 1986b.

van der Hilst, R.D., and E.R. Engdahl, On the use of PP and pP data in delay time tomography, submitted to *Geophys. J. Int.,* February 1990.

Vasco, D.W., R. Jay Pulliam, and Lane R. Johnson, Tomographic inversion of ISC travel times for mantle P wave velocity structure using an l^1 norm criterion, submitted to *J. Geophys. Res.,* November 1990.

Vasco, D.W., R. Jay Pulliam, and Lane R. Johnson, Formal calculation of resolution and covariance matrices associated with mantle tomography, submitted to *Geophys. J. Int.,* December 1990.

Woodhouse, J.H., and A.M. Dziewonski, Mapping the upper mantle: three-dimensional modeling of earth structure by inversion of seismic waveforms, *J. Geophys. Res., 89,* 5,953–5,986, 1984.

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
INFORMATION RESOURCES DEPARTMENT
BERKELEY, CALIFORNIA 94720