

GEOLOGICA ULTRAIECTINA

Mededelingen van de
Faculteit Aardwetenschappen
Universiteit Utrecht

No. 180

**Statistical properties of travel time measurements
and
the structure of the Earth's mantle**

Axel H. E. Röhm

GEOLOGICA ULTRAIECTINA

Mededelingen van de
Faculteit Aardwetenschappen
Universiteit Utrecht

No. 180

**Statistical properties of travel time measurements
and
the structure of the Earth's mantle**

Axel H. E. Röhm

**Statistical properties of travel time measurements
and
the structure of the Earth's mantle**

Statistische eigenschappen van looptijdmetingen
en
de structuur van de aardmantel
(met samenvatting in het Nederlands)

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR AAN DE
UNIVERSITEIT UTRECHT, OP GEZAG VAN DE
RECTOR MAGNIFICUS, PROF. DR H.O. VOORMA,
INGEVOLGE HET BESLUIT VAN HET COLLEGE VOOR PROMOTIES IN
HET OPENBAAR TE VERDEDIGEN OP
DONDERDAG 18 NOVEMBER 1999 DES OCHTENDS TE 10.30 UUR

DOOR

Axel Hermann Erhard Röhm

Geboren op 19 december 1967, te Stuttgart - Bad Cannstatt

Promotor: Prof. Dr R.K. Snieder

The research described in this thesis was carried out at the Faculty of Earth Sciences, Utrecht University, Budapestlaan 4, 3584 CD Utrecht, The Netherlands. This work was supported by the Netherlands Organization for Scientific Research (NWO) through the Pionier project PGS 76-144.

ISBN 90-5744-038-5

The works of the Lord are great,
sought out of all them that have pleasure therein.

Psalm 111.2

Contents

Outline	1
1 Delay times	5
1.1 Introduction	5
1.2 The ISC Bulletin	5
1.3 How to measure arrival times?	7
1.3.1 Picking onsets versus cross-correlation techniques	8
1.3.2 On Fresnel zones and causality zones	8
1.4 Difficulty of picking onset times	9
1.5 Relation between arrival and delay times	11
1.6 Conclusions	14
2 Bias in reported seismic arrival times	17
2.1 Introduction	18
2.2 Identification of the variations	20
2.3 Results	24
2.3.1 Station location	26
2.3.2 Station networks	29
2.3.3 Changes in noise level	29
2.3.4 'Fake' arrival times	30
2.4 Discussion	31
2.5 Conclusions	36
3 Effects of arrival time errors on travel time tomography	39
3.1 Introduction	40
3.2 Random errors due to finite reading precision	40
3.2.1 Variance due to finite reading precision	42
3.2.2 Reading precision less than 1 second	44
3.3 Systematic errors in station time	45
3.3.1 Test results	46
3.4 Discussion	48
3.5 Conclusion	50

4	Correlation functions of mantle structure	51
4.1	Introduction	52
4.2	Method	53
4.3	Data	55
4.4	Variance	55
4.5	Correlation	60
4.6	Inversion for mantle variance and correlation functions	69
	4.6.1 Theoretical constraints on the autocorrelation function . . .	72
	4.6.2 Model parameterization and inversion algorithm	73
4.7	Inversion results	75
	4.7.1 Regional differences	76
	4.7.2 Comparison with tomographic models	77
4.8	Discussion	78
4.9	Conclusions	82
5	Thermal structure of continental upper mantle	83
5.1	Introduction	84
5.2	Seismic tomography and tectonic regionalization	87
5.3	Surface heat flow	91
5.4	Constraints on temperatures from S wave velocity	93
5.5	Comparison with heat flow derived geotherms	98
5.6	Discussion	101
5.7	Conclusions	102
A	Variance of delay times in a two-dimensional random medium	105
B	Constraints on the correlation function	107
	B.1 Fluctuations with zero mean	107
	B.2 Continuous fluctuations	108
	B.3 Relation with delay time covariance	109
	References	111
	Samenvatting (Summary in Dutch)	119
	Zusammenfassung (Summary in German)	125
	Acknowledgments	131
	Curriculum Vitae	135

Outline

The Earth's deep interior is sampled daily by waves originating from earthquakes. Different types of waves gather information along various paths of propagation. Each seismogram recorded by one of the several hundred seismological stations is full of clues about the structure of the Earth. Seismologists have the task to unravel this information coded in the many wiggles. Different techniques have been developed to access at least part of this information from seismograms (measurement). In a next step these data have to be translated into the elastic properties of the Earth (inference). As a final target of seismology, these properties are related to thermodynamic quantities of our interest, such as temperature, composition and dynamic processes of the Earth (interpretation).

The objective of this thesis is to make some contributions to all three parts: measurements (Chapters 2-4), inference (Chapter 4) and interpretation (Chapter 5). The following is a brief outline of the topics treated in each chapter.

Chapter 1 discusses the possibilities and limitations of delay times for the study of the Earth's three-dimensional structure. Delay times are the difference between observed arrival times of body waves relative to the expected times calculated for a reference model. Since the reference model is normally a radially symmetric Earth model, delay times contain the integral information of the three-dimensional velocity perturbations along the paths of propagation. More than 28 million of delay times provided by the International Seismological Centre (ISC) in their bulletin are available to interrogate the Earth's interior. Picked onset times are only sensitive to a small volume around the fastest ray path which is limited in size by causality and is considerably smaller than the volume that influences e.g. cross-correlation measurements of arrival times.

However, there are also practical limitations connected to delay times. Picking the correct arrival time of a body wave from a seismogram with an ambient noise level poses a difficult task for analysts or automated picking algorithms. In a strongly heterogeneous medium diffracted waves cause weak first onsets. Furthermore, earthquake locations are generally calculated from the same data and errors in the location propagate into delay times. The need to calculate an earthquake location makes the relation between arrival and delay times nonlinear. These issues will help the reader to recognize and understand some of the questions that arise in the later Chapters 2-4 which are based on delay times.

The quality of seismological data limits the accuracy of many seismological results. The large quantity of P wave delay times in the ISC Bulletin is used in **Chapter 2** to check individual seismological stations for quality of station timing. Contrary to most other studies based on delay times, the structural component caused by mantle heterogeneity has to be removed for this purpose. This can be achieved by taking the median of a large number of delay times which have sampled very different parts of the Earth. This median filter is moved over the period of station operation to detect changes of the station time. Causes that do not originate at the station can be ruled out by verifying the time variations found in subsets of delay times for different source regions. This gives the surprising result that many stations exhibit systematic time-dependent variations of station timing of up to a few seconds. These variations often coincide with gaps in the data, suggesting hard- or software changes as the most likely cause. But causes can be manifold and include a strong dependence on annual noise level changes for station YKA or a move of the station MAG by about 100 km which was not reported to ISC. This analysis also casts doubts on the assumption that errors in the data are uncorrelated between different stations as for instance all stations operated by the Swiss Seismological Survey show the same time variations.

The severity of systematic errors, such as those examined in Chapter 2 arise from the fact that the large number of data available in the ISC Bulletin does not necessarily mean that the errors cancel out by averaging. Therefore, **Chapter 3** examines the detrimental effect of timing errors on global travel time tomographic models. Tomography is one of the most prominent kind of studies for which ISC Bulletin data are frequently used. Simulated timing errors are inverted and the resulting model shows only small amplitudes relative to inversions of true delay time data. The small effect, however, is due to the much lower variance of the simulated data. Relative to the input variance time variations are efficient in propagating into the model as about half of the signal is explained by the model.

This effect is contrasted by random errors caused by a finite reading precision of arrival times which is not correctly reported to the ISC. This causes a random error in the resulting delay times, also small in size relative to the structural signal. The difference to systematic errors is that random errors hardly propagate into the inverted model, since their contribution is suppressed by averaging character of the inversion.

In **Chapter 4** the heterogeneity of the mantle is described statistically as a random medium. Different realizations of a random medium differ in the spatial distribution of anomalies, but have some common statistical properties, such as r.m.s. amplitude and distribution of scale lengths. The objective of this chapter is to estimate these properties and their variation with depth from the statistical properties of delay times. The variance and the correlation of pairs of delay times from a common source as functions of source depth, epicentral distance and distance between the two stations are used to infer the covariance function between the slowness field of two points in the mantle.

This analysis is complicated by several aspects. First, the strong tails of the

distribution of delay times cause a strong dependence of second order moments, such as variance and covariance, on the chosen cut-off value. The bivariate distribution of delay times for similar rays shows that most of the residuals in the tails are not correlated implying that they are probable mispicks or other phase arrivals by mistake identified as P . The latter can explain many of the late delay times of shallow earthquakes for which the positive tails are strongest. Second, earthquake mislocation cause a systematic bias in the statistical properties of delay times. Part of this can be eliminated by using the variance of delay time differences, but the effect of horizontal mislocations is extremely difficult to estimate. Third, delay times observed at different regions show dissimilar properties even at very small scales. This casts serious doubts on the assumption that the heterogeneity can be viewed as a laterally uniform random medium. Despite these complications the data strongly suggest that the Earth has a significant amount of heterogeneity at scales of the order of 10-100 km.

Chapter 5 deals with the interpretation of seismological velocity models in terms of thermodynamical properties. Since this topic is by far too broad to be elaborated in this thesis, Chapter 5 focuses on one aspect, that is the large scale thermal structure of the continental upper mantle down to a depth of 400 km. In this depth range thermal effects dominate over compositional imprints on the seismic velocities. The S wave velocity structure in the upper 200 km correlates well with tectonic provinces and surface heat flow. Furthermore, the temperatures derived from velocity for an average continental mantle composition are in very good agreement with geotherms extrapolated from the surface truncated by a mantle adiabat. The comparison to different geotherm models shows that the upper mantle seismic velocities are determined by the amount of crustal radiogenic heat production and the thickness of the thermal boundary layer. It is suggested that the thermal blanketing effect of the crustal heat production could also be the cause for variations of the boundary layer thickness. If this is true, erosion of crustal material is the dominant cause behind the formation of fast velocity roots beneath old cratons, a feature prominent in tomographic models.

Chapter 1

Delay times

1.1 Introduction

Body waves are observed as transient signals in seismograms. Their travel times can largely be explained by one-dimensional velocity models of the Earth. Small deviations of the order of a few percent occur due to three-dimensional structure, which is directly linked to the dynamics of our planet. These deviations of the travel times, called delay times or travel time residuals, form the basic data of seismological studies of three-dimensional structure.

This chapter introduces the largest collection of delay times, namely the International Seismological Centre (ISC) Bulletin, which forms the input data of Chapters 1 - 4. Furthermore, some theoretical and practical aspects concerning the determination of delay times are discussed, which are important for the understanding of the possibilities and limits of travel time studies.

1.2 The ISC Bulletin

The ISC provides a global compilation of hypocenters and arrival times starting in January 1964. The 33 years, for which the Bulletin is currently available, contain 857,296 events – mostly earthquakes but a few explosions. The increasing number of observations is displayed in Figure 1.1. In total, there are 19,436,534 observations of earthquakes and 28,245,992 phases listed in the bulletin.

Why is the ISC Bulletin so important in seismology? (1) It is the largest collection of hypocenters and phases. This is in particular important for studies which need a widespread coverage with ray paths. No other data set is able to compete in this respect with the ISC Bulletin. (2) The large number of observations makes it a source of accurate hypocenter locations with global coverage.

The ISC calculates earthquake locations and delay times after the same method and for the same one-dimensional Earth model, namely the tables of Jeffreys and

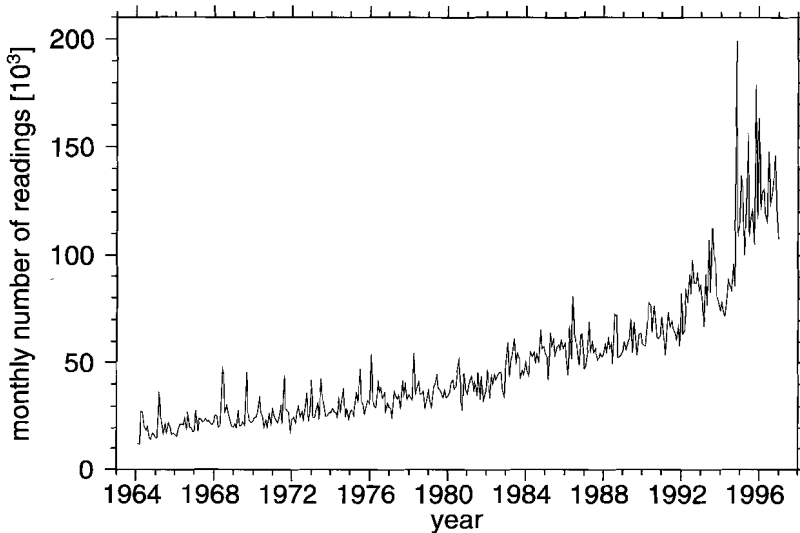


Figure 1.1: *Monthly number of readings in the ISC Bulletin. A reading comprises a set of phases from one station for one event. The number of reported phases is about 45% larger. The increasing numbers reflects the fact that the number of seismological stations is increasing and that more and more smaller earthquakes are processed.*

Bullen [1940], since the start of the Bulletin in 1964. This conservative approach is chosen to produce a data base as uniform as possible. Several efforts have been made in this decade to improve the reference model. This resulted in the models *IASP91* [Kennett and Engdahl, 1991], *sp6* [Morelli and Dziewonski, 1993] and *ak135* [Kennett *et al.*, 1995]. Since it is no problem anymore to reprocess the complete Bulletin, the ISC is presently considering to change their algorithms [Willemann and Storchak, 1999], e.g. use of better reference model, more phases for computing locations, station corrections, etc.

The delay times used for the studies described in this thesis (Chapters 1 - 4) are reprocessed by Engdahl *et al.* [1998]. It comprises nearly 100 000 events that are well constrained teleseismically by arrival time data from the ISC Bulletin. For recent years, where the ISC Bulletin is not yet published, data from the U.S. Geological Survey's National Earthquake Information Center (NEIC) has been added. The data set is complete for earthquakes of magnitude $M_w = 5.2$ and larger for the period 1964-1996. The locations and delay times are calculated for model *ak135* [Kennett *et al.*, 1995].

The quality of the ISC Bulletin is difficult to assess, because generally no independent information on earthquake location and arrival times exist. An exception are explosions for which the origin time and location is known. However, it is

questionable whether studies of explosions with their impulsive wave onsets are also representative for earthquake data [Davies, 1992]. Other tests are based on the fact that the difference between travel times should vanish in the limit of zero distance between two rays [Gudmundsson *et al.*, 1990; Davies, 1992]. Remaining differences between nearby stations then are an estimate of the noise contained in the delay time data set. Gudmundsson *et al.* [1990] have estimated in this way the signal-to-noise ratio for the original ISC *P* wave data to be about 2. However, this approach assumes that the data of different stations are statistically independent. This is doubtful because nearby stations very often belong to the same seismological array or network (see Section 2.3.2) and location errors have a similar effect on both stations (see Section 4.5).

Despite all errors in the ISC Bulletin its usefulness cannot be denied. This is best illustrated by the results of many studies, for instance travel time tomography, that reveal structural information which correlates well with independent information.

1.3 How to measure arrival times?

There are several possibilities to define arrival times. Examples are the position of the signal onset, the first extremum, the absolute extremum or the maximum of the envelope or correlations with a reference signal. Determination of the onset is often difficult, since the onset can be weak (emergent character of the wave) and masked by the natural and man-made noise level. Defining arrivals by some extremum gives a criterion that can easily be measured. However, the extrema of the waveforms are more influenced by instrument bandwidth, noise, dispersion and frequency dependent damping. Therefore, cross correlation techniques are used instead to determine the differential times between wave arrivals at two different stations [e.g. Bungum and Husebye, 1971; VanDecar and Crosson, 1990; Mao and Gubbins, 1995; Shearer, 1997] or of two phases separated in the time domain [e.g. Woodward and Masters, 1991]. Individual seismograms, as well as stacks of several recordings or synthetic seismograms, can be used for the correlation. Absolute arrival times are only needed for the determination of the origin time. The hypocenter location and delay times are independent of absolute time and can be obtained from relative arrival times between seismograms from different stations and phases as well.

Cross correlation procedures can be automated and give generally accurate results. Algorithms for numerical determination of onset times have also been developed [e.g. Kväerna, 1995], but they do not (yet?) exceed the precision of pickings of human analysts. Many of the arrival times listed in the ISC Bulletin came from analog stations with paper recording. Numerical methods would thus need the digitization of the recordings, which is practically impossible, because many of the paper recordings do not exist any more. But there are also theoretical differences between the two approaches which makes them not equally suited for

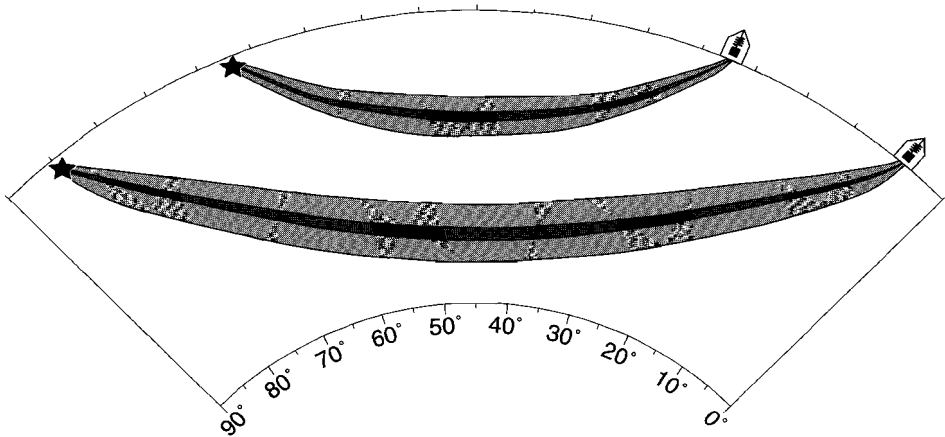


Figure 1.2: *Causality zones for two epicentral distances (45° and 80°). For each distance the outline of causality zones for a travel time difference of 0.1 s and 2 s are shown. The Fresnel zone for the dominant frequency of short period body waves (approx. 1 s) would be in between the two causality zones shown.*

all studies.

1.3.1 Picking onsets versus cross-correlation techniques

Wave onsets are mainly determined by the high-frequency content of the signal whereas cross correlations by the frequencies with the largest amplitudes. Thus, unless different filters are applied, the two techniques will give results representative for somewhat different frequency bands. The differences between the two methods are even larger due to dispersion and frequency dependent damping. Dispersion directly influences the phase difference, which is measured by the correlation. The wave onset is less influenced by it.

Many studies of body waves are based on ray theoretical approaches, e.g. travel time tomography. Ray theory is a high frequency approximation. In the limit of infinite frequency wave propagation between a source and receiver can be described by a wave traveling along the ray path with an extremum travel time (Fermat's principle). Energy that traveled along the fastest ray determines the onset of the first transient wave signal. Thus, travel times determined from onset times come closest to the ray approximation.

1.3.2 On Fresnel zones and causality zones

A wave with a finite frequency band is not only sensitive to structure on the ray path but to a certain volume around the ray. The width of this volume, which determines the minimum spatial resolution that can be achieved with a single

epicentral distance	bottoming depth	width of	
		Fresnel zone $f = 1$ Hz	causality zone $T = 0.1$ s
35°	1020 km	82 km	26 km
60°	1490 km	227 km	100 km
85°	2170 km	263 km	116 km

Table 1.1: *Width of Fresnel and causality zones for different epicentral distances and a source depth of 30 km. Calculated for model PEM-C.*

wave observation, is often linked to Fresnel zones [e.g. Kravtov, 1988; Nolet, 1992]. The first Fresnel zone is defined as the volume that is sampled by ray paths whose travel time difference with the fastest ray is small enough that the energy interferes constructively. The shapes of Fresnel zones and causality zones is shown in Figure 1.2 Thus, it is the volume sampled by all rays, not necessarily stationary ones, that have a travel time difference with the fastest ray of less than one half period.

This connection of the resolution width to the frequency can be misleading since the resolution depends also on the measurement technique. If the measurement needs only a certain time window T , starting at the arrival time of the fastest ray, causality limits the volume that can be sampled by energy contributing to the wave form within T . We will call the cross-section of this volume causality zone to distinguish it from Fresnel zone. Under good circumstances onset times can be picked with an accuracy of less than 0.1 s. This is much smaller than the dominant frequency of P wave observations, which is about 1 s. The causality zone, related to the picked onset time is thus significantly smaller than the Fresnel zone, calculated for the dominant frequency. If, on the other side, correlation techniques are used, window lengths of one period or more are needed in order to obtain stable results. In this case the causality zone is twice the size of the first Fresnel zone or even larger. Table 1.1 gives some estimates of the width of causality and Fresnel zones for teleseismic epicentral distances. In reality the achievable resolution with ISC data is certainly worse than the size of the causality zone for $T = 0.1$ s due to observational uncertainties.

From a theoretical point of view, we find again that picked onset times are in better agreement with ray theoretical approximation due to the smaller causality zones. However, we have not yet considered the difficulties related to picking onset times.

1.4 Difficulty of picking onset times

For more than 30 years seismologists have been encouraged to read onsets to 0.1 s [Bullen, 1963]. However, Douglas *et al.* [1997] have shown that this precision cannot be reached for many observations depending on the signal-to-noise ratio

and the form of the pulse radiated by the source. Filtering can be used to increase the signal-to-noise ratio [e.g. Claerbout, 1964; Douglas, 1997], but noise with a similar frequency content often makes the true onset unrecognizable on individual observations.

Figure 1.3 shows seismograms from North American stations for a South American event together with the arrival times reported to the ISC (solid vertical bars). The waveforms are aligned by cross correlating the traces [Gardien, 1997]. Three stations (HRV, TUC and SBC) have picked the onsets more than a second different from the another stations. A large positive noise wiggle just before the P wave on the trace of station SBC makes it for this station impossible to pick the correct onset without any information from other stations.

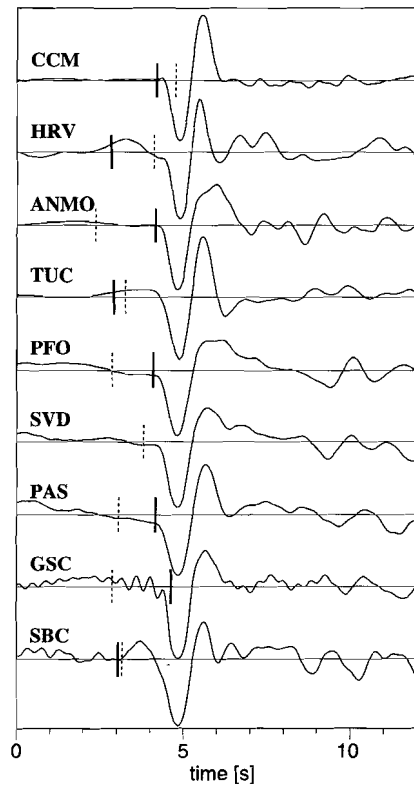


Figure 1.3: P wave observations of North American stations of an earthquake that occurred beneath Chilean-Argentinian border (24.9° S, 68.4° W, 119 km depth, $M=5.8$, 24 Feb. 1993). Dashed vertical bars show theoretical arrival time after the IASP91 model. Solid vertical bar marks the arrival times reported to the ISC. (Figure from Gardien [1997].)

Many of the ISC data are picked at individual stations and the analysts, picking the onset times, have often no other seismograms to compare with. This makes it in many cases extremely difficult to decide where the correct onsets starts.

The difficulty to determine the correct onset is connected to some wave propagational effects in random media. Wielandt [1987] studies diffraction effects for a homogeneous medium with a spherical inclusion of a different wave velocity. An exact solution of the wave equation is known for this simplified medium, which makes it possible to study the asymmetry of velocity anomalies. Fast anomalies will give a negative delay while slow anomalies give only positive delays, if their size is large relative to the path length or the heterogeneity is located close to the source or receiver. Therefore a velocity shift occurs in random media which lets the medium appear faster than the volume average velocity. This velocity shift of the expected value of the fastest ray relative to the ray in the average medium is quantified by several studies [Petersen, 1990; Müller *et al.*, 1992; Nolet and Moser, 1993; Roth *et al.*, 1993; Korn, 1993]

We want to emphasize here that it is not clear, whether this effect is in its full size present in travel time observations, since the fastest ray in random media contains little energy. Rays bend towards slow velocities; as a consequence most energy is concentrated in slow velocity regions. The amplitudes of first arriving diffracted waves decrease with distance after the anomaly [see Figure 3 and 4 and accordant text in Wielandt, 1987]. This has also important consequences for seismic imaging where the first-arrival travel time does not always yield the best model [Geoltrain and Brac, 1993; Audebert *et al.*, 1997]. The effect can be seen in Figure 1.5 which shows a finite difference simulation [Wielandt, personal communication, 1994, 1999] of an initially plane wave in the two-dimensional random medium shown in Figure 1.4. The arrow marks a part of the wave front that traveled through a fast velocity anomaly and is slightly advanced. The wavefront spreads out (convex shape) and the amplitude dies out over the next few time steps and later diffracted waves take over the role of the dominant wavefront, which is not advanced. Therefore it is questionable whether the velocity shift is measurable in real data or whether the fastest waves have amplitudes that are so small that they are masked by noise.

1.5 Relation between arrival and delay times

Delay times cannot be directly observed, but are derived from arrival times and a reference Earth model. Delay times δt are defined as

$$\delta t = \underbrace{T_{arrival} - T_{origin}}_{t_{obs}} - t_{ref} \quad (1.1)$$

where $T_{arrival}$ denotes the observed arrival time, T_{origin} the origin time of the earthquake and t_{ref} the reference travel time for the same epicentral distance and source depth in the reference model. Generally the origin time T_{origin} and

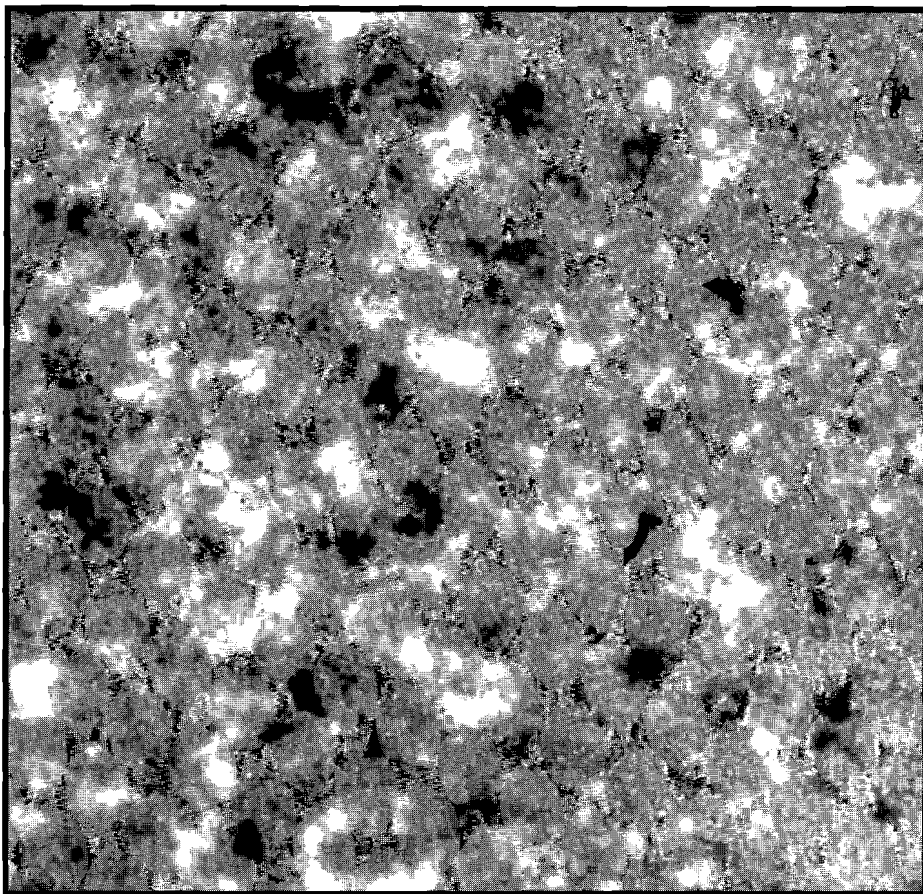


Figure 1.4: *Realization of a random medium with an exponential autocorrelation function. Dark and bright colors denote fast and slow velocities respectively. [Figure courtesy of E. Wielandt, University of Stuttgart]*

the hypocenter location, needed for the calculation of t_{ref} , must be determined from the arrival times too. In only very few cases can the hypocenter location be determined independently, e.g. by observation of the fault. Therefore, the arrival times influence all three terms of eq. (1.1), the last two nonlinearly through the hypocenter location. This relation between arrival and delay times is sketched in Figure 1.6. The closed loop illustrates the already discussed evolution of one-dimensional Earth models (Jeffreys-Bullen tables \rightarrow *IASP91* \rightarrow *sp6*, *ak135*).

The relation between delay times and the three-dimensional Earth structure is frequently made through the integration of the slowness perturbations δu along

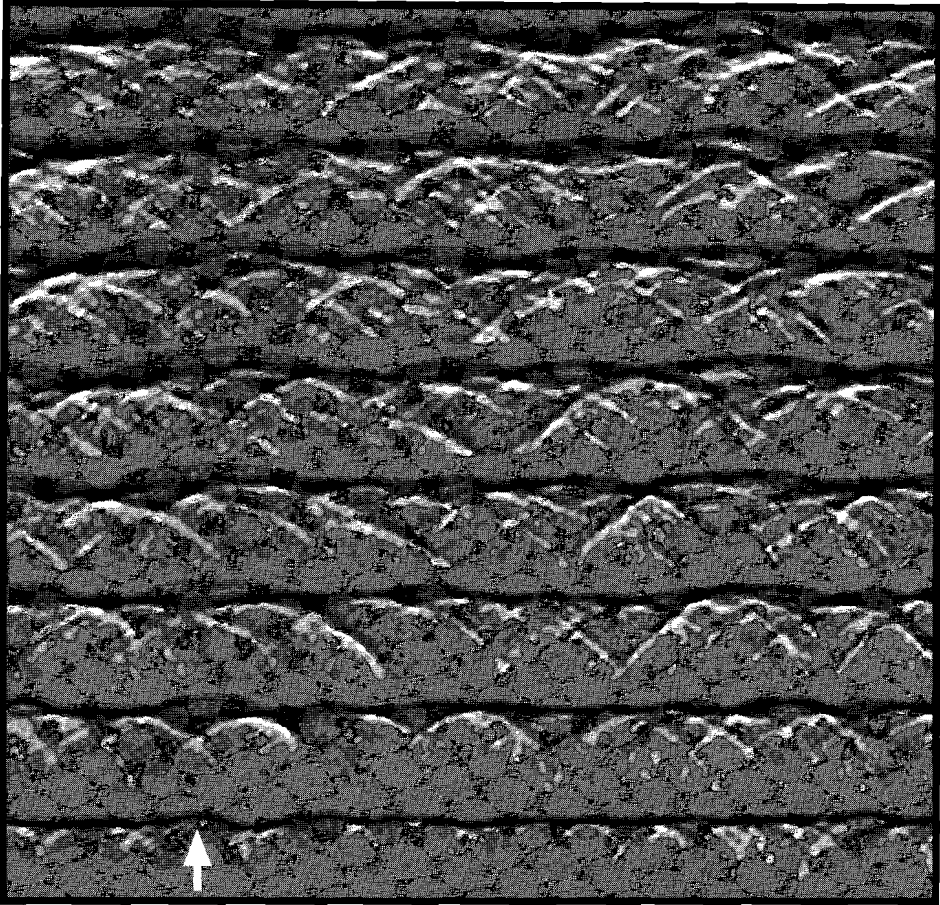


Figure 1.5: *Propagation of an initially plane wave through the medium displayed in Figure 1.4. Shown are 8 snapshots at different times. The wave propagates in the upward direction. Dark and bright colors denote positive and negative amplitudes respectively. [Figure courtesy of E. Wielandt, University of Stuttgart]*

the ray path:

$$\delta t = \int_{\text{path}} \delta u(r) dr \quad (1.2)$$

This integral can be obtained from Fermat's principle or by linearizing the eikonal equation [Aldridge, 1994]. However, this equation agrees only with eq. (1.1) if there is no error in the origin time. A non-negligible error gives an extra, nonlinear contribution.

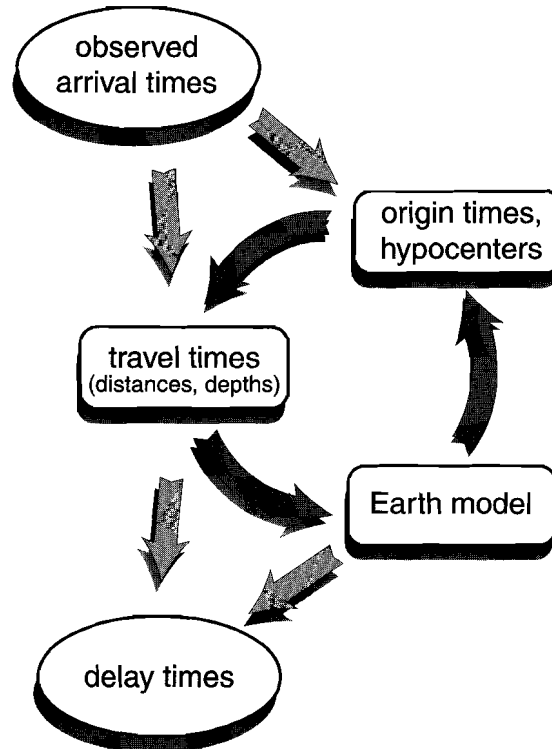


Figure 1.6: *Sketch displaying the steps to derive delay times from a large collection of arrival times, such as the ISC Bulletin. The picture is not complete. Some Earth models used in travel time seismology are constructed using some information from normal mode frequencies or known locations of explosions.*

Deficiencies of the reference model, e.g. neglecting three-dimensional structure, cause some mislocation of earthquakes. The location of an event is calculated by minimizing the variance (or some other norm) of the delay times. Hence, the variance of delay times, calculated for the true location in the same reference model, is larger than the variance for the calculated hypocenter. The variance calculated from a travel time data set underestimates the true variance caused by heterogeneity.

1.6 Conclusions

Delay times have proven to be powerful data for many seismological studies of three-dimensional Earth structure. This introducing chapter has briefly discussed some issues of their possibilities and limitations: (1) Picked onset times come

closest to the ray theoretical approximation and should therefore be preferred if ray theory is used. Differential delay times measured by cross correlation are sensitive to a larger volume (causality zone) than measured by picking onset times. (2) However, correct picking is difficult due to the presence of noise and error prone if no seismograms from other stations of the same event are available. Especially diffraction effects lead to weak onsets that can easily be covered by noise. (3) The need to locate the source with the same data causes a nonlinear relation between the observed arrival times and delay times. Errors in the location due to observational errors or three-dimensional velocity structure causes a bias to smaller delay time variance. To minimize this effect, only well constrained earthquakes should be used.

Chapter 2

Bias in reported seismic arrival times deduced from the ISC Bulletin

Abstract. The accurate timing of seismological data is crucial for most quantitative examinations in seismology. We present evidence that travel time data of many stations contain systematic variations in timing which can be identified by checking the median of station delay times as a function of time. This function is expected to be constant but many deviations are found. Several hundred stations that report arrival times to the ISC have been examined. The median station delay times of almost 8% of these stations show changes of more than 1 s and thus exceed the structural signal in the data. Temporal variations of 0.5-1 s are common. Changes in the distribution of observed earthquakes and other possible explanations of such variations have been tested and fail to explain most of the observations. Therefore, the bulk of the observed changes must be caused by flaws in the timing of the data or by biased picking of arrival times. For instance, at one station with a strong annual variation of noise level, the arrival times are on average picked several tenths of a second later during months with a high noise level.

Because of their systematic nature, these errors will not necessarily cancel out by using the large number of travel times in the ISC Bulletin and may therefore introduce a bias in many investigations. If the observed timing variations are due to the recording equipment of the stations, the errors will be present in the digital waveform data as well. Tomographic studies could potentially be affected, but in particular studies of temporal variations of Earth structure based on travel time data, e.g. inner core rotation, need to be looked at with caution as results might

This chapter has been published as Röhms, Trampert, Paulssen and Snieder, *Geophys. J. Int.*, **137**, 163-174, 1999.

be influenced by station effects. The exact nature of the bias is study-dependent and needs thorough investigation in each individual case.

2.1 Introduction

The accurate timing of recordings is crucial in seismological research. Nowadays, seismometers and acquisition systems have reached unprecedented precision over a large band of frequencies and should help to pick travel times with high accuracy limited only by ambient noise. Unfortunately, the highest possible accuracy is not always reached in practice. Random and systematic errors reduce data quality and thereby pose limits to the results of seismological research.

The International Seismological Centre (ISC) Bulletin is a primary data set for many seismological studies; these have improved our understanding of the dynamics of the Earth, including plate tectonics, mantle convection and the geodynamo. Furthermore, its hypocenter information is used in a large number of seismicity studies [e.g. Adams, 1985] and seismic risk analyses [e.g. Burton *et al.* 1984]. The many different reported phases have led to numerous studies of Earth structure, such as regional and global tomography [for the most recent studies, e.g. van der Hilst *et al.*, 1997; Vasco and Johnson, 1998; Bijwaard *et al.*, 1998], investigations of upper mantle structure and discontinuities [e.g. Krishna and Kaila, 1987; Kato and Hirahara, 1991], studies of the core mantle boundary and D" layer [e.g. Rodgers and Wahr, 1993; Obayashi and Fukao, 1997; Sylvander *et al.*, 1997], measurements of inner core anisotropy [Morelli *et al.*, 1986; Shearer *et al.*, 1988; Su and Dziewonski, 1995] and inner core rotation [Su *et al.*, 1996]. Davies *et al.* [1992], Robertson and Woodhouse [1996] and Su and Dziewonski [1997] used the Bulletin to determine the ratio of relative *S* to *P* heterogeneity, which can be used to constrain mineral physics of the mantle. Further references may be found in the articles listed.

Recently, Engdahl *et al.* [1998] have taken ISC arrival times, with data from the USGS National Earthquake Information Center (NEIC) added for recent years where the ISC Bulletins have not yet been published, and relocated all events that are teleseismically well constrained. The use of a more accurate reference model and more phases, especially depth phases, yielded a subset of improved quality which will certainly be used as a new reference data set for many studies. Recent tomographic models have already used these data and it will certainly be a source for numerous studies in the future. We used this data set for our investigation for practical reasons. However, we would like to stress that most of the problems encountered are inherent in the original data and not due to the reprocessing of Engdahl *et al.* [1998].

The quality of the data set is hard to assess since it is mainly determined by the accuracy of the millions of original arrival time picks which were independently performed at the stations or network central sites using a variety of seismometers and acquisition systems. Hypocenters for some areas can be checked against high-

quality localizations of regional networks. Errors for individual arrival times can only be estimated by statistical tests as done for random errors by Gudmundsson *et al.* [1990] for the original ISC Bulletin. In studies based on travel times quality is normally improved by forming summary rays in which random errors will partly be cancelled out. Unfortunately, this is not the case for systematic errors which therefore have a high potential to introduce a bias into investigations. Gudmundsson *et al.* [1990] have concluded that '*systematic errors are perhaps the most serious limitation of the ISC data*'. A station bias related to the gain of stations was suggested and discussed by Grand [1990].

In this chapter systematic changes of delay times at several stations are presented and various causes discussed. Variations of station time are detected by examining the temporal evolution of static station residuals. Lateral variations of upper mantle and crustal structure cause significant deviations of travel times from predictions of one-dimensional Earth models. The near-receiver effect is given by the mean station residual. A static [Cleary and Hales, 1966] plus one [Bolt and Nuttli, 1966; Herrin and Taggart, 1968; Lilwall and Douglas, 1968] or two [Dziewonski and Anderson, 1970] azimuthal terms were fitted to all residuals of a particular station and subtracted from the travel times as a correction for aspherical Earth structure. The station residual is a measure of the near-station structure sampled by the distribution of ray paths. Therefore, it is expected to be the same for different time windows if their lengths are long enough so that the spatial distribution of earthquakes is similar within each window. For many seismological stations there are a large amount of reported arrival times which can be used to search for temporal variations of the station residual.

There are several possible reasons for a change of the station residual:

- (1) an error in the station timing system, e.g. a clock error;
- (2) a systematic change in the picking of phases, including effects caused by a change of the frequency band (e.g. a short period seismometer that was replaced by a broad band instrument) or by the use of different filters;
- (3) a movement of the seismometers if the station coordinates are not updated;
- (4) changes in the earthquake distribution observed at a station.

Only the last reason implies that a change is not caused by a systematic error. It is possible to identify such a case since it will also change some other properties, as outlined in the next section. Note that a clock error will also change the time information in digital waveform data, not just in picked arrival times. However, for most stations it is impossible to distinguish between the first and second reasons from arrival times alone.

The effects on studies that are based on the ISC Bulletin depend very much on the particular methods which are used. The next chapter addresses this question for global delay time tomographic models. However, the mean delay time of all stations does not show any significant variation in time and it can therefore safely be assumed that for the construction of one dimensional Earth models a bias for one station is compensated by the large number of other stations. The hypocenter determinations of larger earthquakes, which are recorded at many stations, will

probably not suffer from any detrimental effects either. In contrast, investigations of regional differences that use the arrival times directly are much more sensitive to systematic changes at individual stations or networks and the bias will depend very much on the number of other stations in the area. In most cases the precise influence can only be estimated by extensive modeling.

Furthermore, many interesting aspects of current studies are close to the limit of what can be resolved in the data. For instance, tomographic models reveal small perturbations of only a few per cent from spherical symmetric Earth models which give important insights in the geodynamical processes. To improve these results even further it is certainly desirable to increase the accuracy of the data wherever possible.

2.2 Identification of the variations

In this study, delay times from the data set of Engdahl *et al.* [1998] are used. Delay times are the travel times (observed arrival times – origin times) minus the theoretical travel times in the 1D reference model *ak135* [Kennett *et al.*, 1995]. Additionally, corrections for ellipticity and station elevation taken from the same data set are subtracted. We verified that the bulk of the problems that are reported here are also present in the original ISC Bulletin.

About 350 stations have reported more than 5000 *P* wave arrival times during the 32 years covered by the data set. For these stations several independent estimates of the static station residual can be calculated and used to check the time information of individual stations. In order to do so we use a moving window over the period of operation for each station and calculate the median as a function of time of the window center.

We illustrate the method first for delay times from station TUC (Tucson, Arizona). The top panel of Figure 2.1 shows the mean deviation, which is defined as

$$\langle |\xi - \mu| \rangle$$

where ξ are the individual measurements, μ is their median and $\langle \rangle$ denotes the expected value. In the middle panel the median μ can be seen surrounded by a 99% confidence interval of the median [e.g. Rice, 1995] displayed as a grey shade. This is based on the assumption that the samples are drawn from the same probability density function, which is doubtful for delay times since the spatial distribution of hypocenters is not equal in each time window. Therefore, the confidence level will be somewhat lower than 99%. Nonetheless, such bounds are useful because they immediately show the relative accuracy of the median estimates for different times or stations.

The median was chosen as a primary indicator since the ℓ_1 -norm gives a more robust estimation of the center for the distribution of delay times which have long tails [Pulliam *et al.*, 1993]. In principle, the estimators for the ℓ_2 -norm can be used too. To do so the median and mean variation have to be replaced by the mean

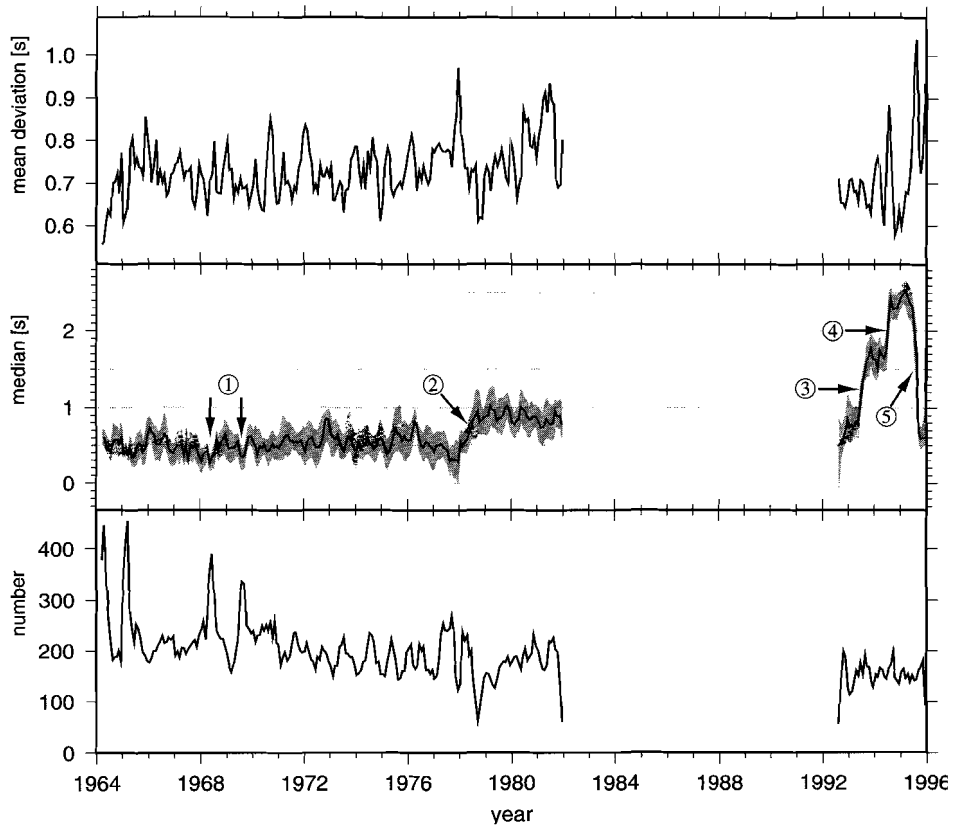


Figure 2.1: *Moving window 3 months long applied to P delay times of station TUC. Upper panel shows the mean deviation, middle panel the median surrounded by the 99% confidence interval of the median and lower panel the number of delay times falling in each window.*

and standard variation respectively. A confidence interval can be calculated from the standard error. The disadvantage of the ℓ_2 -norm estimators is their enhanced sensitivity to outliers in the data; as a result they show more scatter. We have checked these indicators as well but do not show them in the figures since they do not contain any additional information.

The lower panel of Figure 2.1 shows the number of delay times that fall in the window. Since this station reports many arrival times, a relatively short window length of 3 months is chosen for Figure 2.1. The window is shifted in intervals of 1 month. In the following, several features labeled by circled numbers will be discussed.

Label ① marks two dips of the median which extend over the same length as

the moving window. They can only be caused by anomalous residuals situated in the center of the dips. The number of monthly events in the lower panel shows clear peaks at these locations and points to the explanation, which is the occurrence of two swarms of aftershocks from regions with early arrivals.

For the time period 1994-1977 the median is almost constant with only some small fluctuations. This is expected output of the median filter for a station that operates correctly. A step of the median from values around 0.5 s to 0.8 s is labeled ②; this step is not large but is nonetheless clearly visible. From January 1982 until September 1992 the station did not report any arrival times to the ISC. After this gap the median undergoes dramatic changes in June 1993 (③), June/July 1994 (④) and August 1995 (⑤). The mean deviation (top panel) undergoes only relatively small fluctuations around a constant value and therefore does not indicate changes in the distribution of observed earthquakes. Larger jumps of the median are often accompanied by higher values of the mean deviation, as can be seen for the last two jumps (④ and ⑤).

However, the absence of large variations of the mean deviation and of the number of arrival times in Figure 2.1 alone does not prove that the jumps of the median (②, ③, ④ and ⑤) are not caused by changes in the spatial distribution of observed earthquakes as it is the case for ①. An important verification to rule out this possibility is to check whether a particular behavior of the median can also be observed for delay times from sources in several different regions of the Earth. This is demonstrated for station TUC in Figure 2.2. Because there are far fewer residuals for the individual regions, the moving window procedure gives larger fluctuations for the median. In order to compensate for this, the window length is increased to 12 months and the medians of the individual regions are smoothed over three points. As expected, each of the negative excursions ① occurs only in one of the regional curves, and this validates the idea that they originate from aftershock swarms. Further examination reveals that one happened at the Cocos-Nazca plate boundary and the other east of Hokkaido. On the other hand, jump ② is visible for all regions, as is the minimum just before it. Owing to the larger window length, the plateaus separating the jumps ③, ④ and ⑤ cannot be seen any more, but the positive excursion is characteristic of all curves.

A different person started reading the phases for TUC in the Spring of 1976. In 1978 the console and recorder were moved to the University of Arizona and the sensors were connected to them by a phoneline. The latter period corresponds well to the jump ② in the median delay time, which could be explained by delays in the telemetric system. Excursions ③ and ⑤ were due to the malfunctioning of an Omega clock which was not resolved until 1996 when a GPS system was installed. Thus, the data reported for the period 1992 through 1995 should not be used. Knowing these changes at the station allows one to shift parts of the data to probably the correct level for the years before 1978. However, from the delay times alone there is no possibility of deciding whether the correct median for station TUC is around 0.5 s (as for the period 1964 – 1977), 0.8 s (as for the period 1979 – 1981) or whether it has yet another value.

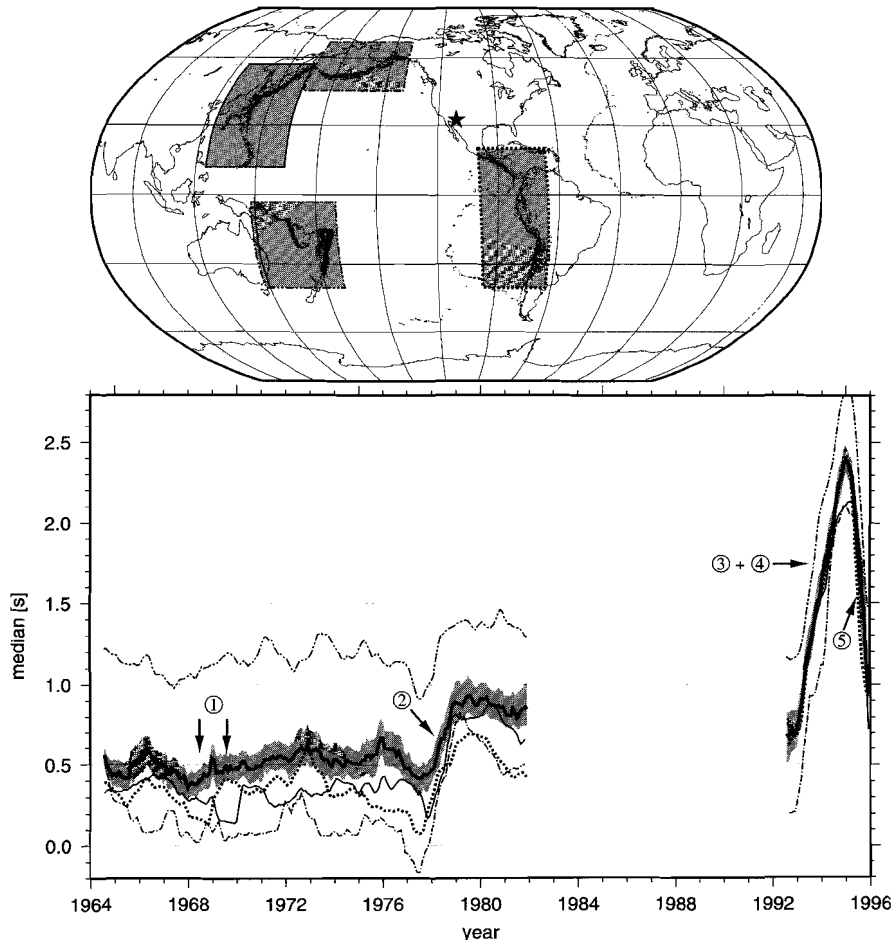


Figure 2.2: Solid line with gray shade shows median of moving window for station TUC as in Figure 2.1 but this time for a window length of 12 months. The four additional lines are the medians of the same procedure for different source regions displayed in the world map as gray shades. The line textures surrounding the boxes correspond to those of the different median curves shown in the lower plot. The station location is marked by a star in the world map. The window length was increased in order to have more residuals in each time window. In addition, medians for different regions are smoothed by averaging over three points.

The confidence band for many stations can be narrowed by shifting the delay times for different regions to a common level, in other words, by removing the azimuthal dependence of the station median. Additionally, this will slightly reduce the small random fluctuations. Figure 2.2 shows, for instance, that at station TUC all waves arriving from the southwest are delayed by about 0.8 s relative to other source regions. Subtracting this value from the delayed arrivals will lower the mean deviation and the width of the confidence band.

It should be noted that such a small step as that visible in 1978 (②) could only be detected because the station reports many phases. The period before 1978 for that station contains only very small fluctuations of the median compared to most other stations.

2.3 Results

In this section more examples are presented. Figure 2.3 shows five stations where temporal changes are found. Only the median of all *P* wave delay times is shown, but for all stations that are discussed in this chapter delay times from different regions and the other parameters described in the last section have been checked to ensure that the variations are not caused by changes in the earthquake distribution.

Station FCC (Fort Churchill, Manitoba) in Figure 2.3(a) is one of the most extreme cases. Between 1967 and 1984 the median shows very little variability around -0.4 s. After several years without any reported delay times the median is 1.5 s higher in 1992 for a few months (March until July, ①) before it further increases (②) to 2.6 s where it stays for about 1 year (July 1992 until September 1993). This is a total change of 3 s. After another break the median returns to its original value in 1995. A shorter time window reveals that the median has decreased already in October 1993, which cannot be seen with the 6 month window used in Figure 2.3. During the years 1990-1994 the station had a short period z-component seismometer and the signal was transmitted via a V-SAT satellite to Ottawa, where the time tagging was performed [B. Shannon, personal communication, 1997]. This period coincides with the anomalous median, which can therefore be explained by an incorrect estimation of the transmission delay.

Figure 2.3(b) reveals that for a 3 year period starting in 1985 (between ① and ②) the median of station LSA (Lhasa, China) is about 1 s lower than in other years. This is also evident in the residuals of several different regions, in contrast to the smaller variations in 1989 and later (③). Thus for the later variations it is not possible to decide whether or not they were caused by changes in the spatial distribution of the earthquakes that were picked.

Station WIT (Witteveen, the Netherlands) in Figure 2.3(c) reported fewer arrival times to the ISC, which has lead to a much greater uncertainty. Despite some variations in the beginning (①), two time intervals can be seen separated by a jump of the median of 0.5 s at the end of 1978/beginning of 1979 (③). At this time the seismometer and recording system were renewed (Grenet instrument

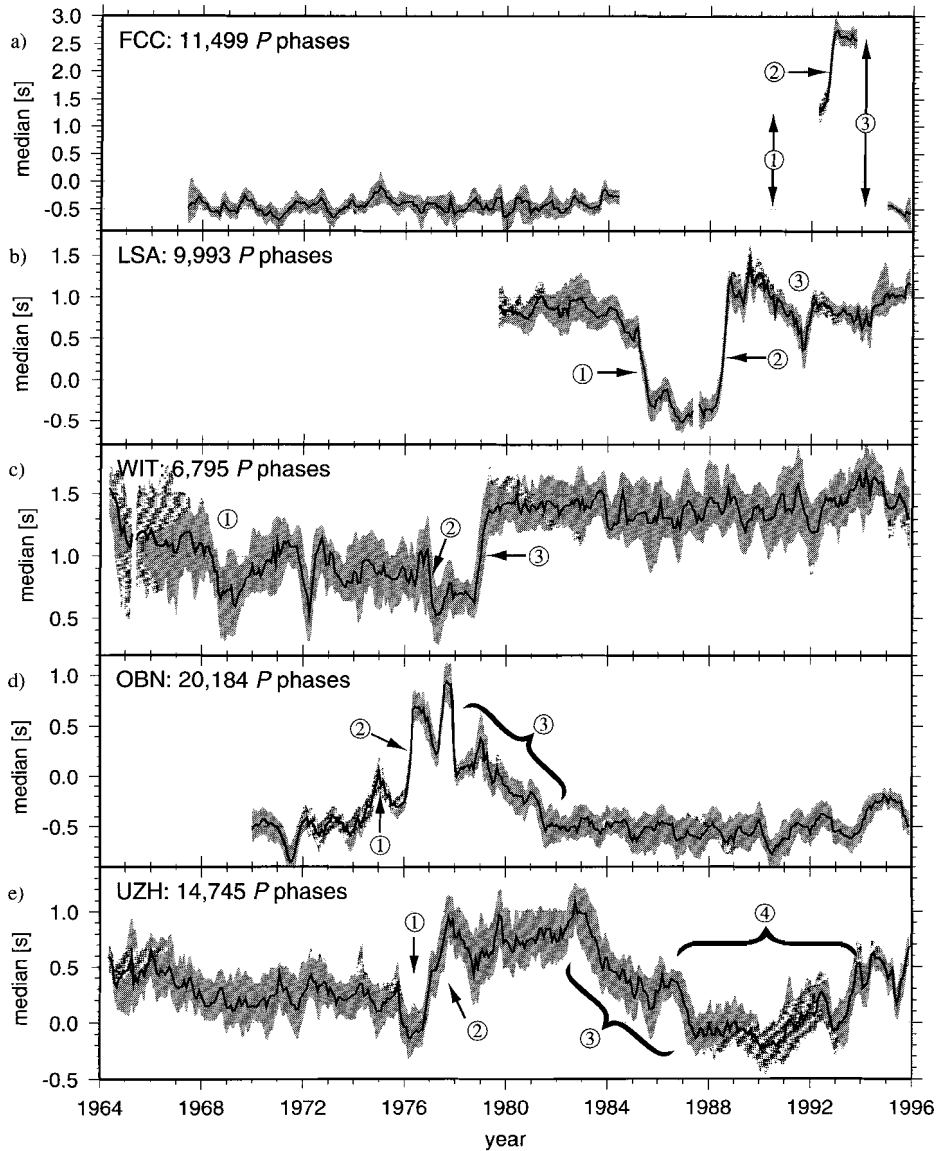


Figure 2.3: Examples of varying medians for stations FCC (Fort Churchill, Manitoba), LSA (Lhasa, China), WIT (Witteveen, Netherlands), OBN (Obninsk, Russia), and UZH (Uzhgorod, Ukraine). The station code and the number of reported P wave delay times appears on the upper left corner of each panel. Window length is 6 months.

with photographic paper to a Willmore MK-II with paper recording; [R. Sleeman, personal communication, 1997]). 2 years before that there is a very small drop (②). Again ② and ③ are seen for residuals from different regions.

Variations in the examples shown can be identified easily as sudden changes at certain dates. Unfortunately, for most stations this is not the case. Variations are very often too small to be seen with the number of residuals available, or several changes are so close to one other that they cannot be separated. Two more typical examples in this respect are the stations OBN (Obninsk, Russia) and UZH (Uzhgorod, Ukraine) displayed in Figure 2.3(d) and (e) respectively. Nonetheless, a variation of the median can clearly be seen for all residuals and different regions. However, the localization of individual changes becomes more difficult. Changes in 1974 (①) and 1976 (②) for OBN are visible for individual regions as well; however, for the decrease after 1976 (③) it is not possible to decide whether it is a drift over a period of several years or it is caused by several smaller steps. Looking at individual regions only confirms a general decrease, but there are too few residuals to draw further conclusions.

Similar difficulties arise in the interpretation of station UZH. A relatively high median for the period 1977 (②) until 1982 (③) and a small low for 1988-1991 (④) are evident. However, the trough in 1976 (①) is seen by delay times from a few regions only and it is not known whether or not the decrease between these anomalous periods (③) is caused by only two jumps in 1983 and 1989.

2.3.1 Station location

A large change of the median can also be seen for station MAG (Figure 2.4, Magadan, Eastern Siberia) in 1972. A shorter window than the 32 month window used for this Figure shows that the jump happened in April 1972. However, Figure 2.4 reveals a very different behavior of medians for delay times from different regions, thus there is not a variation of the time at the station. Using the travel times and earthquake locations after 1973 the coordinates of the station were calculated. This revealed that the station has been moved to another location approximately 50 km to the north, which cannot be found in the coordinates provided by the ISC. This replacement is in agreement with the fact that in April 1972 the station did not report any phases.

The new location is very close to station MGD (Magadan 1), which started reporting phases in March 1973. For 515 events both stations report *P* wave arrivals. For 403 events the arrival times are exactly the same and for only 2 earthquakes they differ by 1 s or more. This shows that the seismometers are located very close to each other, probably even at the same site, or that the seismograms are picked twice. The fact that most arrival times match exactly is not very surprising considering the fact that most arrival times are only picked to a reading precision of 1 s.

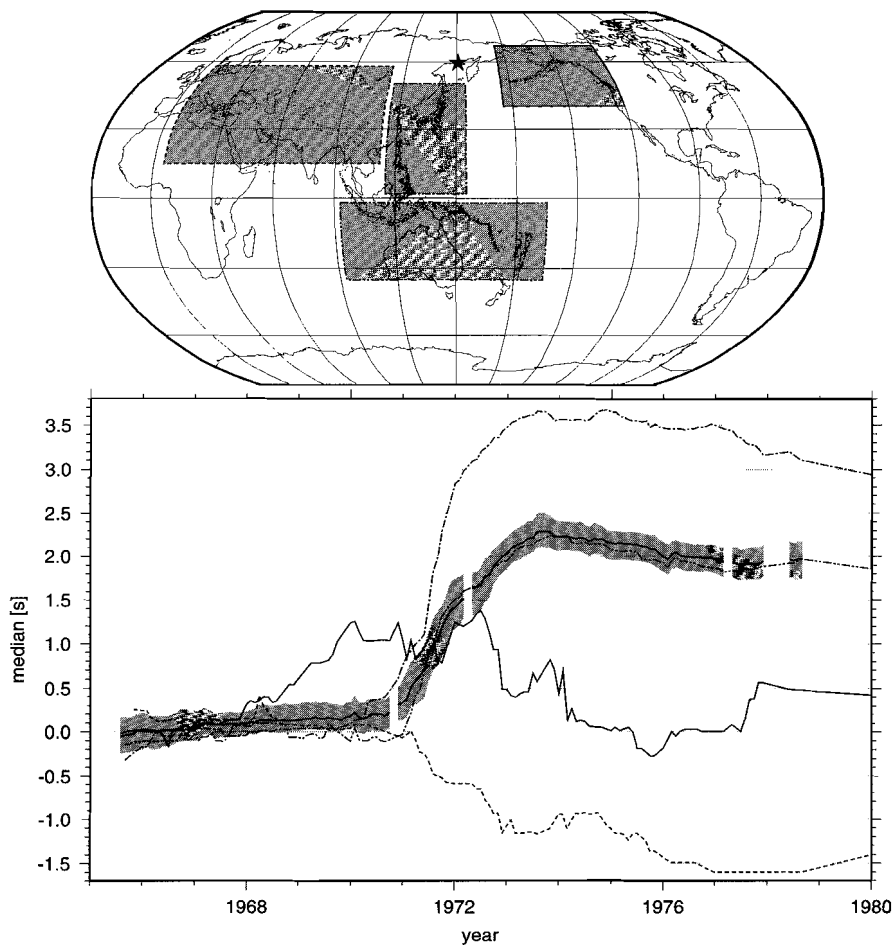
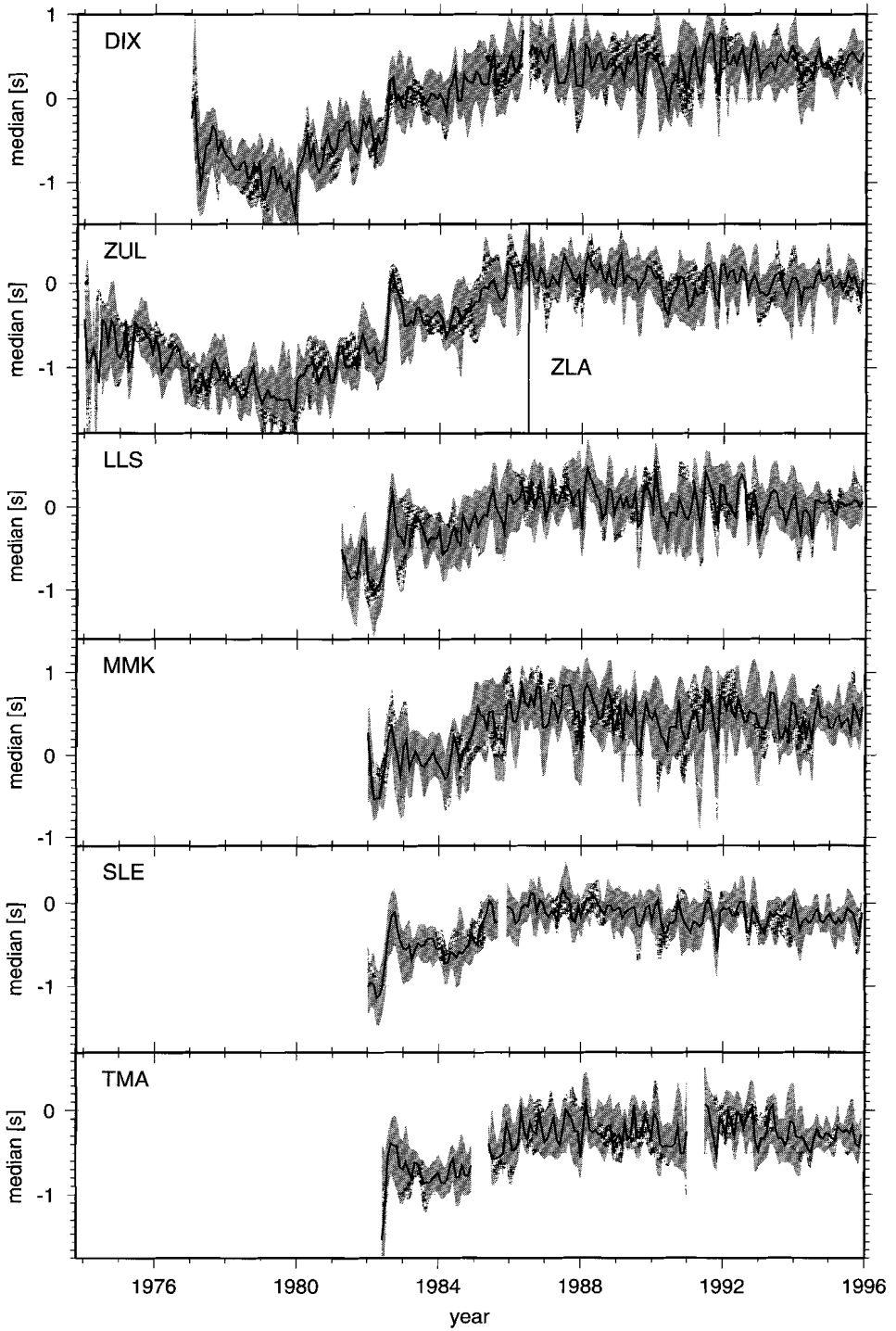


Figure 2.4: Same as Figure 2.2, but for station MAG (Magadan, Russia) and a window length of 36 months. The median of all delay times shows a change of more than 1.5 s in 1972, but medians for individual regions show different jumps. This is probably caused by a movement of the seismograph of about 50 km to the north, which cannot be found in the coordinates of this station provided by the ISC.

Figure 2.5 (next page): Medians for five stations of the network of the Swiss Seismological Service. Window length is 2 months. All stations show a jump in 1982 and an increase between 1984 and 1986. The stations are Grande Dixence (DIX), Zürich-Lageren (ZUL and ZLA), Linth-Limmern (LLS), Mattmark (MMK), Schleithem (SLE) and Tamaro (TMA). For explanation see text.



2.3.2 Station networks

Not all stations are operated separately – some are combined in arrays or networks. For instance the Swiss Seismological Survey operates a network of 27 stations. For most of these sites arrival times have been reported to the ISC. Figure 2.5 shows the 7 stations for which the largest number of P phases have been reported. Data of stations ZUL and ZLA (both Zürich) have been merged because the seismometer has been moved only by about 150 m in 1986, which does not change the median, as can be seen in Figure 2.5. All stations have a flat median after 1986, but between 1980 and 1986 a clear evolution towards later picks is visible. The most remarkable feature within this trend is a jump in 1982. Stations DIX (Grande Dixence) and ZUL show a negative trend until 1979. The similar behavior is not surprising since the analog seismometer signal of all 27 stations is telemetered to Zürich, where it is digitized, stored and analyzed. Therefore, changes at the central site will influence the phase picks of all stations of the network in the same way. The jump of approximately 0.5 s in 1982 falls in the same year as a change of the analyst who picked the phases [H. R. Maurer, personal communication, 1997].

Another network where the data are analyzed at one place is the LDG-CEA network in France. All stations belonging to this network show a slightly negative trend. In this case the variation is relatively small, and for an individual station it is close to the minimum value which can be recognized as a systematic change.

2.3.3 Changes in noise level

The Yellowknife Seismic Array (YKA, Northwest Territories) shows another interesting aspect (see Figure 2.6). In addition to a different average median after a gap in 1989-1990 when the array was completely rebuilt [*Station Book of the Federation of Digital Seismograph Networks (FDSN)*], a distinct seasonal variation can be observed. The variation of reported delay times (lower panel), especially of small earthquakes, points to a strong variation of the noise level, which agrees with the seasonal power spectral estimates in the FDSN Station Book. During months with a high noise level only about half as many phases are picked as during the period December to May. This variation is also found for different regions of hypocenters and different ranges of magnitudes. Therefore, it cannot be caused by the absence of reported small earthquakes during the second half of the year. The variation of the median is even stronger for large earthquakes, as can be seen in Figure 2.7. This Figure reveals two periods. Depending on the magnitude, the median is 0.25-0.7 s higher during July to October compared to December to May. This suggests that during periods with a higher noise level the first onset is difficult to recognize and therefore phases are picked several tenths of a second late on average. The magnitude dependence might be surprising but is probably a result of the relatively simple waveforms of small events and their generally more impulsive character. Thus, if small events are detected the resulting picks are more accurate compared to larger events with more complex waveforms and therefore

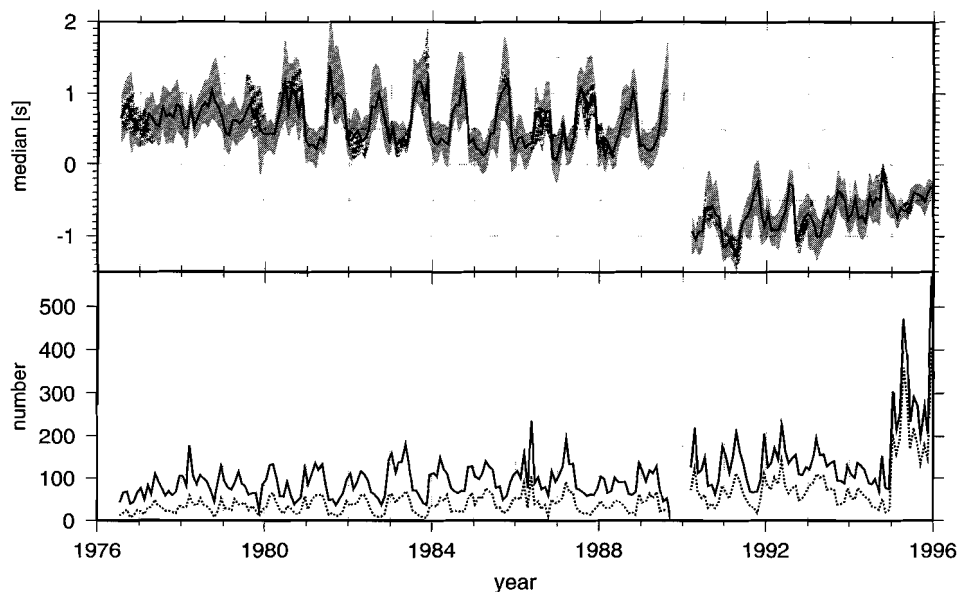


Figure 2.6: *Result of moving window technique for YKA (Yellowknife Seismic Array, Northwest Territories). The median and its 99% confidence interval is displayed in the top panel. The number of all delay times (solid line) and that from earthquakes with magnitude lower than or equal to 4.8 (dotted line) is shown in the lower panel. Window length is 2 months.*

the probability is higher that the onset is not seen correctly in the presence of noise and the arrival time is picked too late.

2.3.4 'Fake' arrival times

In the 1990s the United States National Seismograph Network (USNSN) started to generate 'fake' P wave arrivals to include surface wave magnitudes into the PDE system. This procedure was routinely applied to all USNSN stations with long-period channels whenever there was no P wave trigger but surface wave magnitude information [B. Presgrave, personal communication, 1998]. In order to filter out these unreal data points they were always listed as P waves with arrival times given integer multiples of 10 s, resulting in positive residuals between 5 and 15 s. No onset quality or first motion direction was assigned to them. The large positive residuals were chosen to prevent their use for other purposes. However, these arrivals were passed on to the ISC and can be found in their printed bulletins. Depending on the origin time of the ISC relocation a few might even have somewhat smaller residuals. Although the number of 'fake' arrival times (several thousand) is small compared to the total number of P arrivals, for some stations

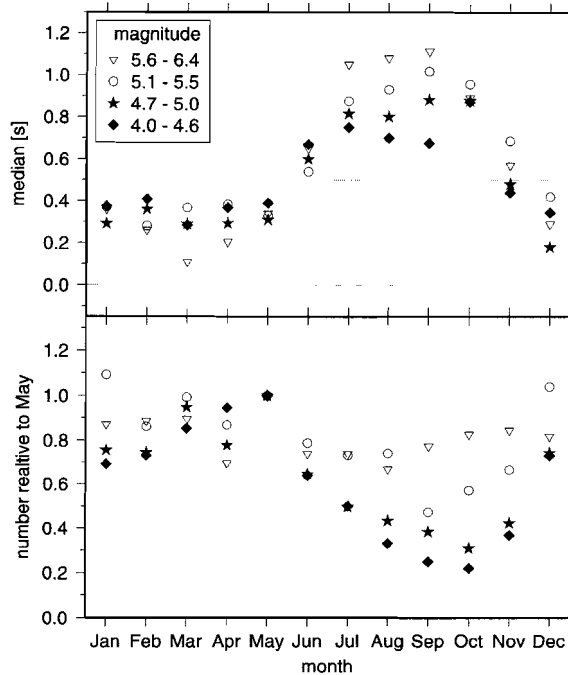


Figure 2.7: Median (upper panel) and number of delay times (lower panel) as a function of month for YKA (Yellowknife Seismic Array, Northwest Territories). Symbols denote different magnitude ranges as defined in the box. For each month, data of the period 1979 - 1988 is used.

they can be the majority of arrival times, e.g. station HON (Honolulu, Hawaii) for 1991 and later. Some of the 'fake' arrival times will match closely the predicted arrival times of later phases, such as depth phases (pP and sP) or PcP . Consequently, they might be converted to these phases whenever a phase reidentification is carried out. Several hundreds of these faked phases can indeed be found as later phases (primarily depth phases) in the data set of Engdahl *et al.* [1998] with small (positive or negative) delay times.

2.4 Discussion

Since the examples have been chosen to be illustrative, most of them show above average variations of the station residual. It is desirable to estimate the extent of these errors in the complete data set. This is, however, not an easy task since

In the meantime the ISC has changed their procedures and no 'fake P ' phases are published any more [R. Willemann, personal communication, 1999].

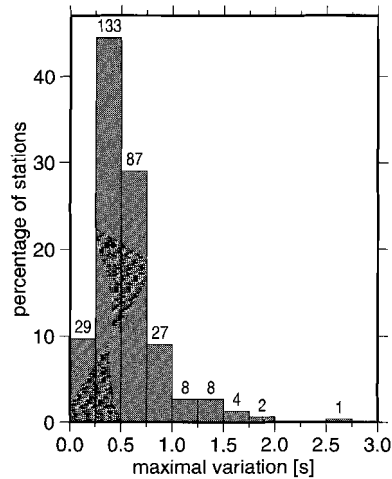


Figure 2.8: *Histogram displays the percentage of stations that have a maximum variation of the median denoted by the abscissa. The number of stations appears at the top of each column.*

each station shows a different pattern. Especially if stations which report not more than a few thousand residuals are examined and variations are not larger than 0.5 s, it becomes difficult to decide whether variations are caused by errors or unequal earthquake distributions. However, we try to give such an estimate. For the 299 stations that report at least 5000 teleseismic arrival times the maximum variation of the median for a moving window is shown in Figure 2.8. We restricted this analysis to teleseismic events, that is, epicentral distances larger than 30° , to exclude the much larger delay times of regional events, although the median is not very sensitive to them. The window length was not chosen to be a fixed period of time as in the examples shown but as the period in which 1000 residuals fall. This means that for a station that reports few P phases the window length can be several years whereas for others the window length is only several months, e.g. station TUC. The length of 1000 residuals was chosen because it recovers the maximum variations of the median in the examples shown quite well, e.g. 1.74 s for TUC, 2.64 s for FCC and 1.57 s for LSA, and still smoothes random fluctuations of the median. Note that the chosen length is larger and thus the smoothing stronger than those used in any of the examples presented. The minimum number of residuals ensures that for each station at least five completely independent values for the median are obtained.

The 299 stations examined for Figure 2.8 have contributed 64.5% of the teleseismic P phases in the data base although they represent only 7.0% of the stations. Not all variations are caused by systematic errors. Different distributions of earthquakes and random errors also cause some small fluctuations. A good

station without systematic errors has variations of only a few tenths of a second. For example, station TUC (Figs 2.1 and 2.2) has a maximum variation of the median of 0.28 s for a window length of 1000 delay times for the period 1964 - 1976, where it can be considered as a good station. Thus, variations of less than approximately 0.5 s are not necessarily caused by errors. On the other hand, systematic errors of only a few tenths of a second or which are short in time are not detected. Variations larger than 0.5 s are very unlikely to be explained by the earthquake distribution alone and at least part of the deviations stems from errors. 137 stations (45.8%) show such large variations. For 23 stations (7.7%) there are changes of 1 s or more. These numbers should be compared to the mean deviation of teleseismic P delay times, which is 0.9 s if the distribution is cut at ± 3.5 s as is usually done in tomography for teleseismic waves.

What are the possible causes for such systematic variations? Earthquake distribution cannot be responsible for at least the larger variations since they can be identified from travel time data by looking at different areas, magnitudes, depths, etc. Since the collection and processing of the arrival times by the ISC, as well as by the National Earthquake Information Center (NEIC) and the relocations of Engdahl *et al.* [1998] are carried out for all events and stations in the same way this cannot explain changes for individual stations either. The obvious possibility is that the observed changes are caused by instrumental or phase picking errors. In case of instrumental problems, all work based on the original seismograms would also be influenced except when using relative times, e.g. differential travel times between phases.

The simplest error would be an incorrect time of the station clock. Delays in the acquisition system that are not corrected would have the same effect. Such errors are most likely to occur when equipment is replaced. Indeed, many of the jumps in the examples shown coincide with gaps, which suggests that a change of equipment has been taken place.

On many seismograms the precise onset is not obvious and different people (or automatic pickers) will not always pick it at the same time. For the Swiss stations, for instance, a change of analyst probably explains as much as 0.5 s difference on average of all picks. Digital stations can be very helpful with their possibilities of filtering the data and enlarging the displayed seismogram. On the other hand, they might also be a source of errors if instrument corrections or filters are applied incorrectly, e.g. the use of zero-phase instead of causal filters. The transfer function of the instrument can also have a large influence and distort the onset.

Station YKA demonstrates that the noise level can be crucial for correct picking. A dependence of arrival times on the gain, which is normally adjusted to the noise level, for the ISC Bulletin has already been suggested by Grand [1990]. Douglas *et al.* [1997] demonstrated for some examples how the correct onset cannot be seen until the magnification is large enough so that the noise before the phase is clearly visible. If the onset is weak it will probably be missed at stations with a high noise level.

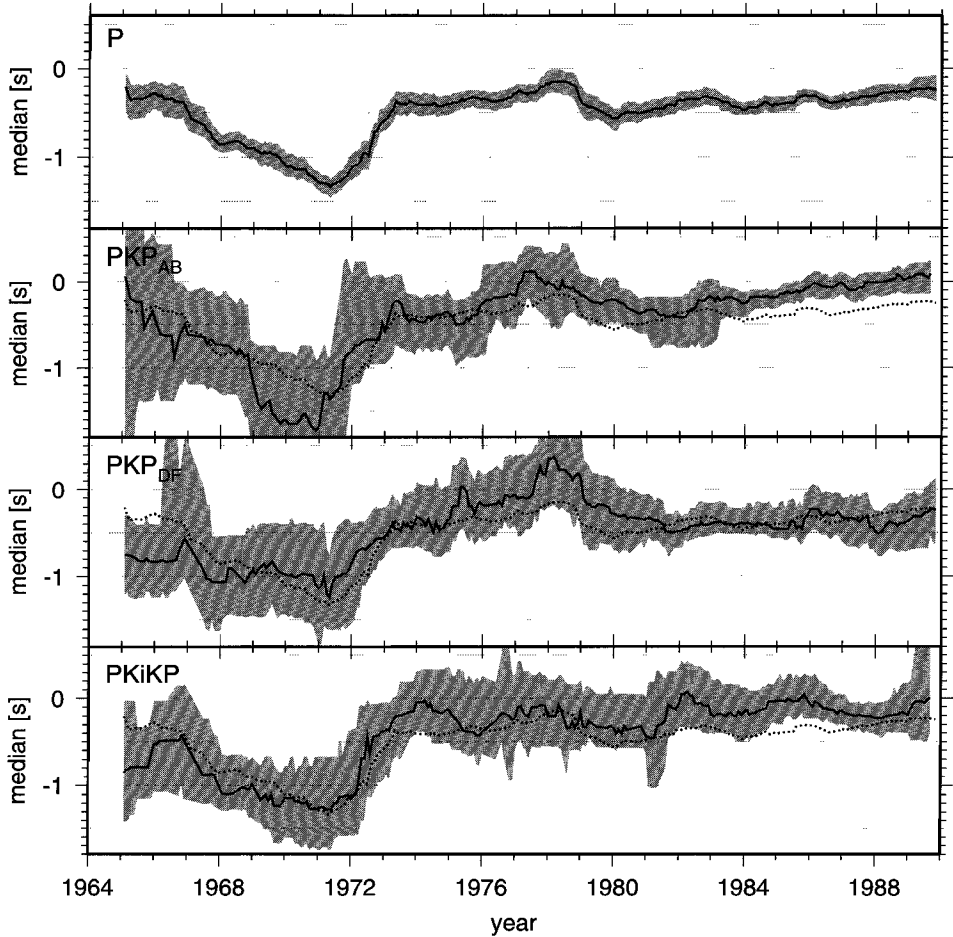


Figure 2.9: *Similar behavior of the median for delay times of 4 different phases observed at station BNG (Bangui, Central African Republic). Phase names appear in the upper left corner of each panel. Window length is 24 months. The median of the P phase (uppermost panel) repeated in the other panels as dotted lines for ease of comparison.*

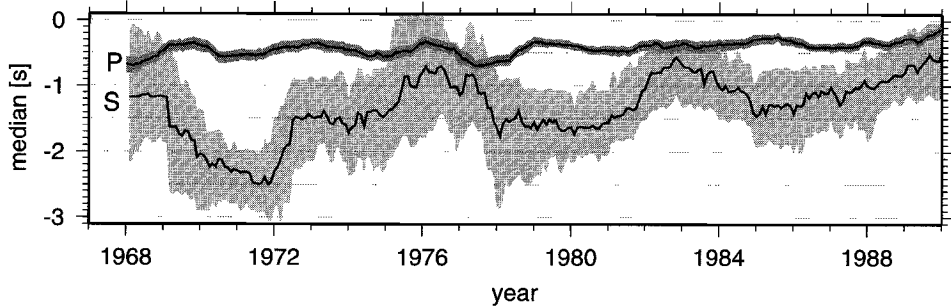


Figure 2.10: *Different behavior of P and S wave delay times recorded at HYB (Hyderabad, India). The much smaller number and generally poorer quality yield the broader confidence interval for S waves. Window length is 18 months.*

Systematic errors are present in all phases, not only P . As an example, Figure 2.9 shows the P phase together with the core phases PKP_{AB} , PKP_{DF} and $PKiKP$ for station BNG. Although the number of core phases is relatively low compared to P and the uncertainties are therefore much larger, a jump in 1972, smeared out over the window length, is present in all phases.

Errors of station time are thought to be avoided by using differential travel times of two phases from the same event recorded at the same station, but if errors are caused by the picking procedure they may still be present. Picking of onsets may depend on the noise level or the frequency content of phases and are likely to be different for first arriving and later phases. Examples where the changes of different phases are dissimilar can also be found, especially in comparisons of P and S phases. This is demonstrated in Figure 2.10 for delay times from station HYB (Hyderabad, India). Whereas no significant change in P wave delay times occurs, the median of S phases does show large variations of more than 1 s. Generally, we find that variations of S phases are much larger than those for P . This can be understood by the fact that due to scattered waves the noise level present at the S wave arrivals is higher, which results in a poorer quality of the picks. Furthermore, there is a difference in the frequency content, and whereas P phases are picked from the short period z-component, S waves may also be picked from horizontal components or long-period channels. Different behaviors may be present in other phases as well, thus even differential travel times can be influenced by changes in picking.

From our findings it is clear that searching for time-dependent variations of velocity structure inside the Earth with ISC travel times needs to be carried out very carefully. Whenever temporal variations are interpreted as changes inside the Earth, e.g. inner core rotations determined by ISC data [Su *et al.*, 1996], many different phases should be checked in order to rule out possible station biases. Unfortunately, only very few stations report enough residuals of different phases

that could be used to separate hardware errors, picking errors and changes inside the Earth. 'Fake' arrival times may have an undesired influence if a cut-off value is not chosen smaller than 5 s. If residuals of 5 s and more are included, these arrivals will bias all USNSN stations to late arrival times for the 1990s. This will be a systematic bias concerning time and region. Tomographic modeling is also potentially affected, but to what extent needs to be investigated by extensive numerical simulations.

The fact that there are stations with periods of a constant median separated by identifiable steps, e.g. TUC, FCC, LSA or WIT, raises the question whether independent station corrections for each period would correct the data. Since the times of changes can only be estimated with some uncertainty, some residuals around the steps would have to be deleted in order to be certain that no incorrect correction is applied to them. Another problem is that if the changes are caused by the picking procedure, the variations might be dependent on magnitude or other parameters, and applying constant time shifts would be incorrect.

For most other stations it is much more difficult and in many cases even impossible to construct time-dependent station corrections that remove offsets caused by errors while leaving the structural signal in the data. If the time variation is limited in time, removing part of the data would be another possibility for obtaining a smaller data set, hopefully with reduced errors. This could be done for station OBN for instance. However, for the majority of stations correcting the data is not an option since they report too few arrival times in order to identify individual changes. In all cases, whether the data are corrected or not, the average station residual has lost the meaning of being a measure of near-receiver structure. Rather, it is for many stations a sum of this and systematic time errors, with both parts being of a comparable size.

2.5 Conclusions

Delay times of the database of Engdahl *et al.* [1998] have been tested for systematic variations in time. For 46% of the stations examined changes of the median delay time of more than 0.5 s have been found. Several examples have been presented and analyzed to rule out causes that do not originate at the stations. Consequently, our conclusion concerning these stations is that either the station timing system or the phase picking procedure introduced systematic errors for certain periods in time.

The causes are manifold and include picking errors as well as hardware changes. Different behavior of the median for different phases points to systematic picking errors. Unfortunately, this can also bias differential travel times of different phases from the same event as shown in Figure 2.10 for P and S waves observed at station HYB. Since some of the variations are very large, e.g. at station LSA (Figure 2.3), it seems unlikely that they could be explained by picking errors alone but rather are caused by the instrumentation. In this case the time errors will not only be

inherent in reported arrival times but also in the waveform recordings. However, it is impossible to find out the exact cause from the reported arrival times without any additional information.

Although each station affects only a small number of phases compared to the total number of the data in the ISC Bulletin, the errors potentially influence quantitative seismological results because the errors discussed are not random but systematic for individual stations or even networks, e.g. the Swiss regional network. Studies of temporal variations of Earth structure in particular are prone to the reported biases due to the systematic changes in time of the errors. Tomographic studies will probably also suffer from these effects although the exact extent can only be detected with intensive modeling. We anticipate that they are most affected in regions with sparse station coverage where model parameters are resolved by few summary rays and systematic errors cannot be compensated by rays to other stations. In general, the implications for individual studies will depend strongly on the methods applied.

Observational seismologists should be aware of this and test results for possible biases induced by time errors. We recommend that seismological stations implement more checks on the accuracy of timing in order to detect malfunctions and to reduce errors in the future.

Chapter 3

Effects of arrival time errors on travel time tomography

Abstract. The ISC Bulletin is the primary data set for travel time tomography since it comprises the largest collection of arrival times. The large number of stations and events gives the best currently attainable coverage of the Earth's mantle with ray paths. But the data also contain a considerable amount of noise which one tries to suppress by averaging delay times for similar paths and/or parameterization and damping of tomographic inversions. In this chapter we focus on two different types of errors in body wave arrival times and estimate their effects on global tomographic models. Tests with synthetic input data are performed for the tomographic model of Bijwaard *et al.* [1998].

The first type of error stems from a finite reading precision of arrival times and is equivalent to a round-off error. This yields a random contribution to delay times. The influence of a reading precision of 1 s or better on tomography is almost insignificant since (1) its variance is very small compared to the total variance of ISC delay times and (2) only less than 5% of that variance maps into the tomographic model. A few stations report arrival times with an indicated reading precision of 0.1 s that are in reality only picked to the closest 10 s or 1 min which results in a S/N ratio much lower than 1.

The second type of error, described in Chapter 2 causes systematic variations of delay times as a function of time. A test reveals that the rms amplitude due to these systematic errors is between 4.2% (0–35 km depth) and 14.4% (1800–2000 km depth) of the model rms amplitude. This blurs the tomographic model to some degree but does not change the overall amplitude or shape of seismic anomalies. Luckily, the size of the error in the input data is one order of magnitude smaller than the standard deviation of the ISC delay times because a significant amount maps into the tomographic structure.

This chapter has been submitted by Röhm, Bijwaard, Trampert to *Geophys. J. Int.*

3.1 Introduction

Travel times of the ISC Bulletin are the primary data set for global and regional body wave tomography, since it comprises a large number of earthquakes and stations, needed to give the best possible ray coverage. The arrival times originate from local agencies or individual stations and the quality varies significantly. The effect of random errors is mainly reduced by averaging data for similar ray paths and by the regularization of tomographic inversions.

We demonstrate on a statistical basis that the reading precision, provided for each arrival time pick, is often not correct. A finite reading precision causes random round-off errors in the data. We evaluate the effect of such random errors on seismic tomography.

The effect of systematic errors is more difficult to quantify and depends on the type of error. Grand [1990] has pointed out the presence of systematic differences in delay times related to instrument gain. Chapter 2 discusses temporal variations of station timing, probably related to changes in instrumentation or picking procedure. They can be detected by moving a median filter over all delay times of a station. Constant time errors over the complete period of station operation can completely be absorbed in the station residual and yield no bias to the imaged structure. The effect of temporal variations is more difficult to assess. It depends on the duration and amplitude of the wrong timing as well as on the distribution of ray paths. In order to test the effect of temporal variations of delay times, a tomographic model is constructed with synthetic input data that simulate time errors. The test presented in this chapter is carried out for the same parameterization and regularization as used in the P tomographic model of Bijwaard *et al.* [1998] and thus directly shows the effect of the tested error. This study is based on delay times of Engdahl *et al.* [1998] (will be referred to as EHB), who relocated a subset of teleseismically well constrained earthquakes of the ISC data, supplemented with NEIC data for years where the ISC Bulletin is not yet available.

3.2 Random errors due to finite reading precision

The onset time of a seismic wave arrival cannot always be determined with an accuracy of a tenth of a second or better due to noise and the emergent character of some phases. Therefore, an extra parameter is included in the data set indicating whether the arrival time was read to the nearest 0.01 s, 0.1 s, 1 s, 10 s or 60 s. Of all P phases in the EHB data set, 79.3% have an indicated reading precision of 0.1 s or better, 20.7% of 1 s and less than 0.01% of 10 s or 60 s. Phases with a reading precision worse than 1 s are not included in the tomographic model of Bijwaard *et al.* [1998] and most other studies. A finite reading precision can be considered as a round-off error in the arrival times.

Whether arrival times are indeed picked with the indicated precision can easily be checked. Earthquake origin times are randomly distributed over time and so

should be the picked arrival times. If only the full second and tenths of a second information is analyzed we expect to find a uniform distribution between 0.0 and 59.9 s where the data are binned in intervals of a width of 0.1 s. Significant deviations from a uniform distribution reveal that not all onset times are determined to the nearest 0.1 s. On the other side, a uniform distribution does not prove that the accuracy of individual onsets is as high as 0.1 s. To show that the differences with a uniform distribution can be large, the distribution of arrival times for station AFI (Afimuela, Samoa Islands) is shown. Microseismicity is generally high at an island station, such as AFI, which makes it more difficult to read the onsets accurately. Still 4320 (18.8%) of all reported arrival times are listed to have a reading precision of at least 0.1 s and their distribution is shown in Figure 3.1. The distribution immediately reveals that arrival times are dominantly reported at full seconds and the highest bar at zero seconds strongly suggests that some of the arrival times are only determined to the nearest minute. This pattern is not consistent with the reported reading precision of 0.1 s.

The significance of this can statistically be shown, if we consider that the number of arrival times in each bin follows approximately a Poisson distribution, for which the standard deviation is given by the square root of the number. This yields (arrival times on full minutes are removed) for full seconds 42.0 ± 6.5 and for tenth seconds 3.1 ± 1.8 . Clearly the means are significantly different.

This is not an exception and interesting patterns can be found by examining different stations. Figure 3.2 shows more examples of distributions. Again, only arrival times that are listed with a reading precision of a tenth second or better are used. In this figure the nearest tenth of a second of the arrival time is binned. It is expected that, if the reading precision was indeed not worse than $1/10$ s, all columns have the same height. However, the pattern shown by stations GBA and LAT are by far the most frequent ones encountered for stations of the ISC Bulletin.

All stations that have reported at least 500 P wave arrival times with an indicated reading precision of at least a tenth of a second have been tested for the correctness of the reading precision parameter. For each station the distribution of arrival times has been binned according to the nearest tenth of a second in the arrival time. With 99.9% certainty the distributions of 83% of examined stations is not consistent with an underlying uniform probability density function. This is found by using a chi-square test to disprove the null hypothesis that the distribution is drawn from a uniform probability density function. For half of the stations the value of χ^2 is above 500 which means that the deviations from a uniform distribution are very large.

A rough estimate of the number of arrival times, for which the reading precision is worse than one tenth second, can be obtained from the distribution for all stations together as shown in Figure 3.3. The columns for half and full seconds are clearly too high and all other anomalous patterns have almost cancelled each other out. The number of arrival times with the correct precision of 0.1 s can be estimated from bins 1-4 and 6-9. From this and the heights of the bins 0 and

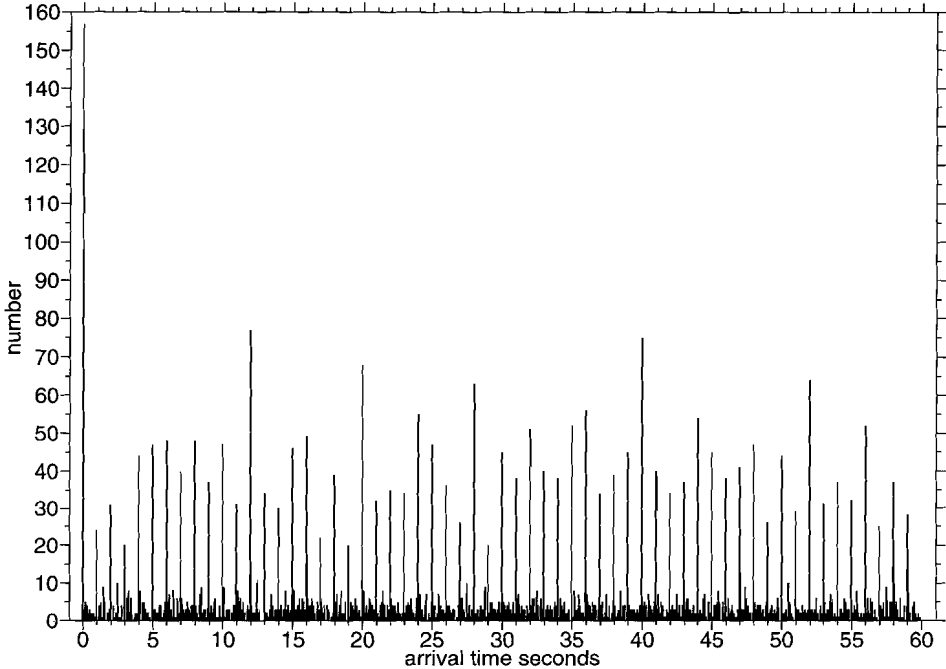


Figure 3.1: *Distribution of full seconds and tenths of a second of all arrival times from station AFI for which the reported reading precision is equal or better than a tenth of a second. Note the highest bin at a full minute.*

5 we find that 12.8% of the data should have a reading precision parameter of 0.5 s and 14.7% have a reading precision of 1.0 s (see brackets in Figure 3.3). These estimates are lower bounds since other anomalous distributions, e.g. too few arrivals on full seconds (see GEC2 in Figure 3.2), partially cancel out the effect of stations with too many arrival times on full seconds. It is also not considered that the reading precision of a small portion is even worse than 1 s.

3.2.1 Variance due to finite reading precision

Since this error in the arrival time is uncorrelated to other errors or structural information the variances of the finite reading precision (σ_{RP}^2) and of other sources (σ_{other}^2) add up to the total variance (σ_{tot}^2):

$$\sigma_{tot}^2 = \sigma_{RP}^2 + \sigma_{other}^2.$$

The variance of σ_{RP}^2 is easily calculated from the probability density function of the shifts that the true arrival times get through the finite precision. If the reading precision is T , these shifts are uniformly distributed between $-T/2$ and $T/2$. The

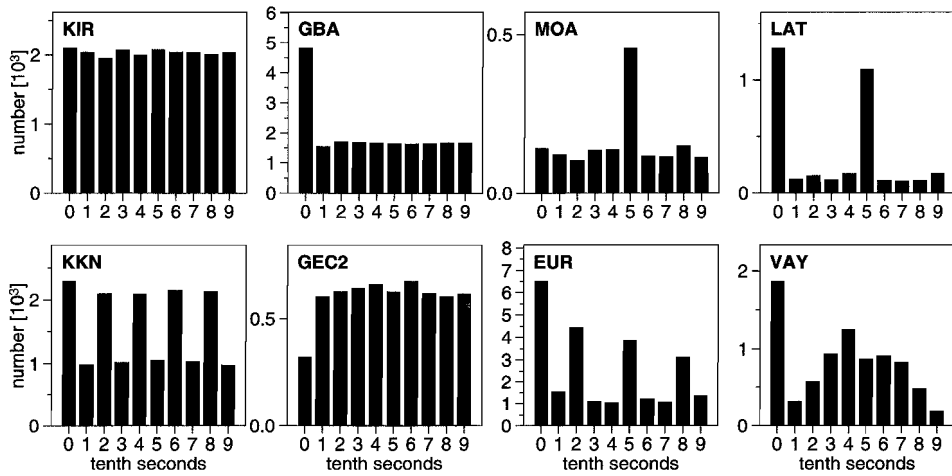


Figure 3.2: *Examples of arrival time distributions for different stations. The station code is indicated in the upper left corner. Station KIR is an example of a station where the arrival times are uniformly distributed as expected for a reading precision of at least 0.1 s. GBA has reports too many arrival times coinciding with full seconds. MOA does the same for 0.5 s and LAT for 0.1 and 0.5 s. About half of the arrivals of KKN are picked to the closest 0.2 s. GEC2, EUR and VAY show other possible distributions which do not show the expected pattern for the indicated reading precision.*

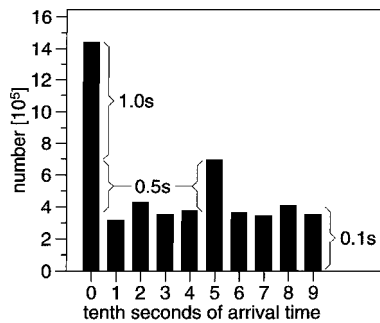


Figure 3.3: *Number of P wave arrivals of all stations according to the first decimal of the arrival time. The brackets show the parts of the histogram used to estimate the amount of phases for the reading precision indicated by the number.*

variance is obtained by integrating the probability density function:

$$\sigma_{RP}^2 = \frac{1}{T} \int_{-\infty}^{\infty} t^2 (H(t + \frac{T}{2}) - H(t - \frac{T}{2})) dt = \frac{1}{12} T^2 \quad .$$

$H(t)$ denotes the Heaviside step function. For a reading precision of 1 s this yields a variance of approx. 0.084 s^2 , much smaller than the total variance of the delay times, which is higher than 1 s^2 for all different types of phases in the EHB data. Data with a reading precision worse than 1 s is normally not used in seismological studies.

Not only is the variance caused by the finite reading precision small as compared to the variance caused by heterogeneity, the random nature of this error prevents a significant bias in the three-dimensional structure determined in the tomographic inversion. Tests with random noise reveal that less than 5% of the variance is explained by the tomographic model.

3.2.2 Reading precision less than 1 second

The fact, that in Figure 3.1 the number of arrival times for a full minute is much larger than for any other time, implies that the reading precision of P phases probably is sometimes even worse than 1 s. It is of special interest to specify the amount of the arrival times that have a precision of only 10 s or 1 min because the resulting delay time variance σ_{RP}^2 exceeds by far the variance caused by three-dimensional Earth structure.

The following gives the test results for full minutes, the values of the test for full ten seconds are appended in brackets. For the latter, arrival times falling on full minutes are not taken into account, otherwise most stations with too many full minute picks automatically fail the test for full ten seconds too. If a station reports n arrival times at full seconds (with a reading precision of at least 1 s) the number X of arrival times coinciding with full minutes (10 s) is a random variable. The probability that $X = k$ is given by the binomial distribution

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

where $p = 1/60$ (5/59) is the probability that one arrival time coincides with a full minute (10 s). The probability that a station reports l or more arrival times on full minutes (10 s) by coincidence is thus given by the cumulative binomial distribution

$$P(X \geq l) = \sum_{k=l}^n \binom{n}{k} p^k (1 - p)^{n-k}$$

where very small values of $P(X \geq l)$ indicate that the reading precision of part of the arrival times is probably less than 1 s. If we take stations that report too many arrival times at full minutes (10 s) and test whether the probability of obtaining

this by coincidence (null hypothesis) is less than 1%, 131 (110) stations are found. They contribute only 0.18% (0.45%) of the arrival times of all examined stations. Fortunately, these numbers are small, but since these data can easily be removed from the data set, our advice is not to include these data in studies based on delay times.

3.3 Systematic errors in station time

We want to offset the effects of random timing errors against a systematic error recently identified in the ISC Bulletin [Röhm *et al.*, 1999, Chapter2]. These errors give systematic variations of delay times as a function of time of up to a few seconds. They are identified by examining the average station residual determined by a moving median window.

The influence of systematic errors on global tomography is examined by constructing a tomographic model for synthetic input data. The input data of the test consist of exactly the same rays as used for the construction of the tomographic model of Bijwaard *et al.* [1998]. But instead of using the delay times of the EHB data set, synthetic delay times are generated according to teleseismic (epicentral distance greater than 30°) P wave (P_{tele}) residuals. First, all phases from stations with at least 20 observed P_{tele} data receive the median of all P_{tele} delay times. Stations with less than 20 P_{tele} data always receive zero delay times. This simulates delays from local Earth structure or a constant time offset. All stations with at least 2000 P_{tele} arrival times get in addition a time varying component depending on the month and year of the observation. This component is determined by smoothing the observed P_{tele} residuals with a moving window of length of 800 residuals in order to imitate timing errors [see Chapter 2]. The input data simulate thus a constant station residual and, for stations with many observations, the time variations that we want to study. The estimated errors are assigned to all phases used in the tomography, i.e. P , pP and pwP . However, only teleseismic P waves are used to determine the timing errors because they have a much smaller variance than regional phases and therefore yield the most robust estimate of these errors.

According to these criteria, data from 2446 and 2145 stations receive, respectively, a zero or constant delay time and data from 636 stations obtain delay times varying with time. Only 12.2% of the stations receive a residual dependent of time. This may seem a small number. However, these stations contribute 76.0% of all P , pP and pwP wave arrival times. For the majority of stations the limited number of observations does not allow to test the presence of time variations [see Chapter 2]. The simulated errors are therefore probably an underestimation. However, if these 636 stations do not give a significant bias in the imaged Earth structure, the stations with much less observations probably will not either.

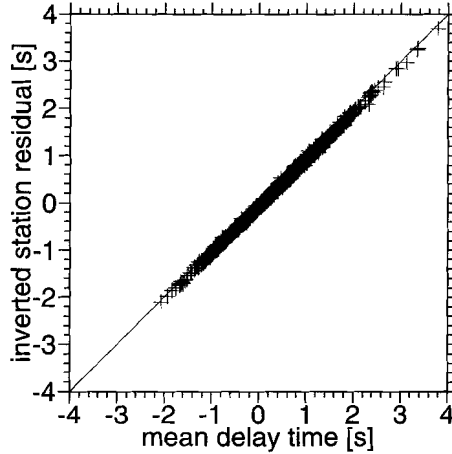


Figure 3.4: The mean of all delay times of each station versus the inverted station residual.

	tomography	synthetic test	optimum
total variance	1.79 s ²	0.35 s ²	0.35 s ^{2†}
composite rays	16.7%	2.2%	2.2% [†]
inversion	47.6%	94.2%	91.7%
unexplained	35.7%	3.6%	6.1%

[†]value taken from the synthetic test

Table 3.1: Total variance of the input data and its division into three parts: (1) the combination of individual rays into composite rays, (2) the tomographic inversion, explaining part of the variance by hypocenter corrections, station residuals and three-dimensional Earth structure, and (3) the remaining unexplained variance. The first column is the outcome of the tomographic model of Bijwaard *et al.* [1998], the second column the result of the described test for timing errors and the last column is the fictive optimum outcome of the test.

3.3.1 Test results

The tomographic model of Bijwaard *et al.* [1998] inverts for slowness perturbations in blocks, hypocenter corrections and station residuals. The mean delay time of all composite rays ending at the same station is almost completely absorbed by the station residual as can be seen from Figure 3.4. The mean absolute deviation between the two quantities is only 62 ms. Thus, a constant time error which has existed throughout the period of station operation gives no bias to the three-dimensional structure.

Table 3.1 lists the variance of all input data, i.e. the P , pP and pwP delay times, and the variance reductions achieved by bundling the individual rays into composite rays [see Bijwaard *et al.*, 1998; Spakman and Nolet, 1988] and by the tomographic inversion itself. Values for the tomographic model of Bijwaard *et al.* [1998] are listed in the left column for comparison. The right column shows the expected values if no bias to the three-dimensional tomographic structure would result from the time variations. They are calculated in the following way. The variance reduction due to the grouping of similar rays into composite rays of 2.2% is the same as for the synthetic test, because this value depends only on the geometrical distribution of ray paths and the size of the event cluster volumes ($30 \text{ km} \times 30 \text{ km} \times 30 \text{ km}$). The optimum result for the remaining 97.8% would be that the mean of all composite rays for the same station is explained by the inverted station residual and the time varying part is not projected into the model but remains completely unexplained. This yields 91.7% and 6.1% respectively. This optimum result is not achieved because part of the time varying errors are mapped in the inversion to the model parameters. Thus, the unexplained variance of the test is less than for the optimum case and some bias in the three-dimensional structure is expected.

We only find a small amplitude of the 3D structure for the synthetic test. The model rms-amplitude of the test is only 6.9% of the rms-amplitude of the tomographic model. The relative amplitude is with a value of 14.4% largest for the layer between 1800 and 2000 km depth where the amplitudes for the tomography are small. The smallest relative amplitude (4.2% of the tomographic model) is found in the shallowest layer although the absolute amplitudes are strongest there. Figure 3.5 shows an example of the resulting bias (lower panel) together with the tomographic model of Bijwaard *et al.* [1998] (upper panel). The colour scale is adapted to the different amplitudes. But not only the amplitudes are different. The pattern, obtained from the test inversion, is much less structured and appears almost random with respect to the tomography that shows anomalies extending over several cells.

There could be two reasons why the amplitudes of the synthetic test are much smaller: either the amplitudes of the time variations in the input data are very small or the errors are hardly mapped into the model parameters in the inversion. To reveal this, we compare the standard deviation of the two input data sets. Because a constant time offset does not influence the model, but is absorbed in the inverted station residuals (Figure 3.4), the mean delay time for each station was subtracted from the data. The standard deviations after forming summary rays are 0.15 s and 1.28 s for the test and the real input data respectively. Thus, the standard deviation of the test input data is about 8.5 times smaller, meaning that, relative to their strength, the systematic errors efficiently propagate into the model. (The colour scale for the synthetic test displayed in Figure 3.5 has been chosen to exactly be smaller by a factor of 8.5.) This is also evident if the input variance due to time varying errors (0.022 s^2) is compared with the unexplained variance after the inversion (0.013 s^2). In the ideal case the two should be equal but

the test outcome shows that about 40% of the error variance leaks into the model. To sum up: systematic errors have a large potential to bias tomographic studies, but – fortunately – their size in the input data is about an order of magnitude smaller than the total signal. Therefore, their detrimental effects are very limited.

3.4 Discussion

The studied random and systematic errors directly affect arrival times. Seismic tomography, however, uses delay times as input data. In order to perform the tests we have assumed that the arrival time errors are equal to the delay time errors. This is, strictly speaking, not true, since the errors will also affect the location of the hypocenter and the origin time and therefore all other delay times of the same event. The determination of the hypocenter is a nonlinear process which makes it difficult to predict error propagation. The resulting hypocentral error depends very much on the number and distribution of observations. However, for earthquakes constrained by many arrival times, as is the case for all events of the EHB data set, the hypocenter location will suffer little and the errors remain restricted to the delay times derived from erroneous arrival times.

The random and systematic errors have a very different effect on global tomography although the standard deviations of both are small relative to the total signal in the EHB delay times. Less than 5% of the (already small) variance of the random errors propagates into the model. This makes their detrimental effect insignificant. On the other side, about 40% of the variance of the systematic errors transfers into the tomographic model. Therefore, they yield a small, but visible, blurring effect.

Other error sources, e.g. mispicks of onsets or erroneous phase association, are probably also restricted to a relatively small number of data which makes the ISC Bulletin a well-suited data set for many seismological studies. In densely sampled regions, redundant information suppresses the noise in the model parameters even further. Therefore, seismic tomography is able to reveal very small structures down to length scales of about 100 km in some regions. For other regions, which are less well illuminated by ray paths, the sum of all errors lowers the achievable resolution. Errors in the arrival times reported to the ISC can normally not be corrected later since the original seismograms are not available for most stations. At best, they can be detected and the affected data can be removed, which will probably also exclude some correct data, as the example of reading precision of less than 1 s (section 3.2.2) shows. The studied errors do not stem from the processing of the ISC nor from the earthquake relocation of Engdahl *et al.* [1998]. Therefore, they can only be avoided at the seismological stations where they are introduced. There are unavoidable limits to the accuracy of picking onset times, mainly caused by the natural and man-made seismic noise level at the station site. It should be aimed to avoid additional error sources, such as a erroneous station time or a wrong indication of the reading precision.

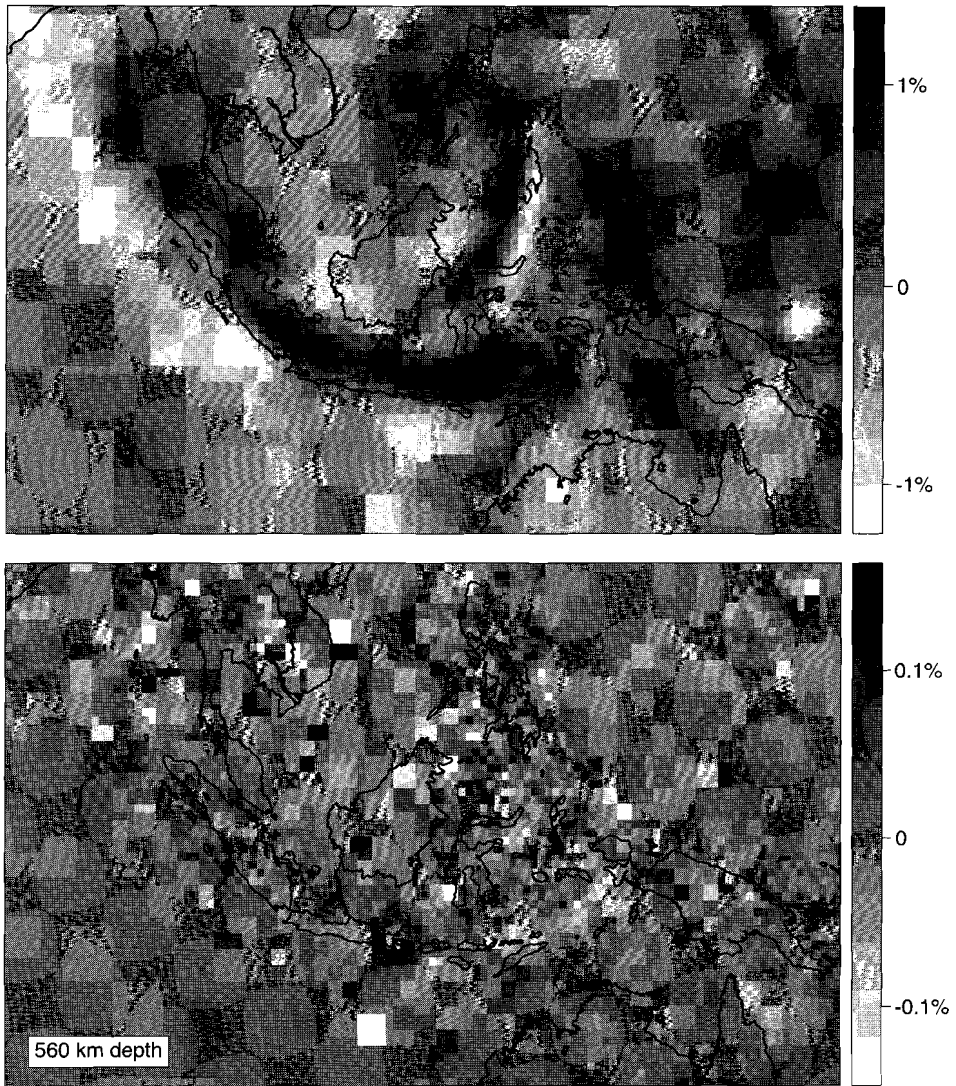


Figure 3.5: The tomographic model resulting from inversion of the EHB data set (upper panel; Bijwaard et al., 1998) and the synthetic delay times simulating the bias (lower panel) for South-East Asia at a depth of 560 km. Note the different scales of the two panels, indicating the velocity perturbation relative to model *ak135* [Kennet et al., 1995].

3.5 Conclusion

The following conclusions can be drawn from this study of two different error sources: (1) The reading precision parameter for many phases in the ISC Bulletin is not correct. Some data have a reading precision of only 10 s or worse, although they are listed to have an accuracy of at least 0.1 s. The S/N ratio of such data is much less than 1. Fortunately their number is very small compared to the total number of phases and they can easily be filtered out. (2) Random errors due to the finite reading precision are small and ineffective in propagating into tomographic model parameters. Therefore, they do not yield any significant bias. (3) Systematic timing errors, on the other hand, bias model parameters relative effectively. However, because of their small size relative to the signal in the delay times, they contribute a random pattern with, in absolute measure, low amplitudes to the tomographic model. This gives only a small, but still visible, blurring effect. The well resolved structures of tomographic models and their interpretations are therefore not significantly influenced by the errors studied here.

Chapter 4

Correlation functions of mantle structure inferred from P wave delay times

Although the number of P wave delay times of the International Seismological Centre (ISC) Bulletin with values larger than a few seconds is small compared to the total number of P data they exert considerable influence on second order moments. Variance estimates strongly depend on the chosen cut-off value if the data are truncated which is always done in practise. Therefore, knowledge of the distribution of noise is critical for a statistical examination of delay times. The majority of delay times with absolute values greater than 2-3 s are not correlated between similar ray paths, even if the distance between the paths is very small. This suggests that most data situated in the tails of the residual distribution are caused by erroneous arrival time picks or phase misidentifications. On the other hand, delay times with absolute values smaller than 2 s show a correlation which depends on the distance of the rays.

The variance and correlation of (correct) delay times depend on the scale lengths and strength of the velocity heterogeneity. A formalism is developed to invert for the covariance function of mantle structure as a function of depth. The loss of correlation between delay times from a common source recorded at two stations decreases already considerably for a distance between the stations of a few 10 km. This implies heterogeneity at the same scales. The inverted model has the best fit if this heterogeneity is distributed in the upper as well as the lower mantle.

4.1 Introduction

Many seismological studies indicate that heterogeneity in the Earth's mantle decreases in amplitude and increases in scale length with depth. However, it is not clear whether this result is not partially caused by a low-pass filter implicit in the model construction and observations, e.g. model restrictions due to the finite sampling, parameterization and regularization of inversions. Seismological data also have an averaging character, not only along the ray path, but also perpendicular to it (see the discussion of Fresnel zones and causality zones in Chapter 1), which generally increases with depth. Long-wavelength structure dominates in most tomographic models (an exception is perhaps upper mantle structure inferred by travel time tomography) [Su and Dziewonski, 1991]. However, it has been questioned whether tomographic models represent unbiased estimates of all scale lengths [e.g. Snieder *et al.*, 1991; Passier and Snieder, 1995]. Statistical analysis of body wave travel times suggested that in the upper mantle the spectrum of lateral heterogeneity is white, down to wavelengths as short as 100-200 km [Gudmundsson *et al.*, 1990; Davies *et al.*, 1992]. Furthermore, there are some reports of small-scale heterogeneity. Hedlin *et al.* [1997] analyze precursors to the core phase *PKP*, which they explain by heterogeneity distributed throughout the mantle with scales of about 8 km and 1% r.m.s amplitude. Kaneshima and Helffrich [1998] describe two localized lower mantle scatterers and estimate their possible size to about a few 10 km.

Power spectra are often used to describe the statistical properties of a medium. Alternatively, the autocovariance function (or autocorrelation function and variance) can be used. In Euclidean spaces the autocovariance function is related to the power spectrum by the Fourier transform. In a spherical geometry the autocovariance function is given by the coefficients of the Legendre expansion of the angular power spectrum. An advantage of a statistical description of mantle heterogeneity is the possibility to compare seismological models and numerical convection simulations. Puster and Jordan [1994] and Puster *et al.* [1995] discuss the properties of the spatial autocovariance function of the radial and angular velocity fields and the temperature field.

Due to improved computational capacity, travel time tomography resolves structures of more and more smaller size. Currently the smallest cell size used in global models is $0.6^\circ \times 0.6^\circ \times 35$ km in regions with the highest ray coverage [Bijwaard *et al.*, 1998]. The question appears, whether such models are capable to image most of the mantle heterogeneity or whether there remains a significant part with even smaller scale lengths unresolved. The answer has implications on questions related to the methods needed to correctly model wave propagation in the Earth's mantle and to the dynamics of our planet.

Gudmundsson *et al.* [1990] and Davies *et al.* [1992] have developed a stochastic approach to infer the scales and amplitudes of mantle heterogeneity from the variance of ISC delay times within summary rays of different sizes. The extrapolation of the variance of ray bundles for zero width gives the variance of incoherent

noise. The difference in variance between zero bundle width and the largest size (whole Earth) is due to lateral heterogeneity. The shape of increase of variance as a function of summary ray width, epicentral distance and source depth contains information about the distribution of energy among different scale lengths. Davies [1992] uses a similar approach, to use the variance of ray bundles from a common source to estimate an upper bound on the temporal and spatial mislocation of earthquakes. An estimate of incoherent noise based on the approach of Gudmundsson *et al.* [1990] for the relocated data set of Engdahl *et al.* [1998] can be found in Bijwaard *et al.* [1998].

In this chapter we take an approach similar to that of Gudmundsson *et al.* [1990] and analyze the covariance between delay times for two rays from a common source. However, we also test two basic assumptions made by statistical studies: (1) Is the noise identically, independently distributed among the measured arrival times? The correctness of this assumption affects estimates of the signal-to-noise ratio that are based on the correlation of delay times for similar ray paths. (2) Can the Earth's mantle adequately be described as a medium with random fluctuations whose statistical properties (characterized by the covariance function) are laterally uniform? We will refer to such a random medium as "laterally uniform random medium". If the statistical properties vary also laterally a parameter estimation derived from the very inhomogeneous sampling of global travel time data sets might deviate from the spherically averaged properties. This is also important for the question, whether power spectra characterize all regions adequately and do not mainly reflect the imprint of one or a few very heterogeneous regions.

4.2 Method

Delay times contain the integral information along the ray paths. They can be used to invert for the anomalies that are sampled by the waves. This can be done for deterministic models by a tomographic inversion; but our interest here is to infer the statistical properties of mantle heterogeneity as a function of depth. Unfortunately, the delay time variance as a function of epicentral distance alone gives not enough information to discriminate between the variance of the medium and the size of its anomalies. This can easily be demonstrated for a two-dimensional case with a modified Gaussian random medium, for which in Appendix A the delay time variance as a function of medium r.m.s. amplitude σ , correlation width a and path length L is derived as

$$\text{Var}(\delta t) = \frac{\sigma^2 a L \sqrt{\pi}}{2} \text{erf} \left(\frac{L}{a} \right) . \quad (4.1)$$

The correlation between delay times as a function of the distance between the ray paths is needed to resolve this trade-off. We use the correlation of delay times originating from a common source. Under the assumption of a laterally uniform random medium, this geometry is uniquely determined by four parameters (see

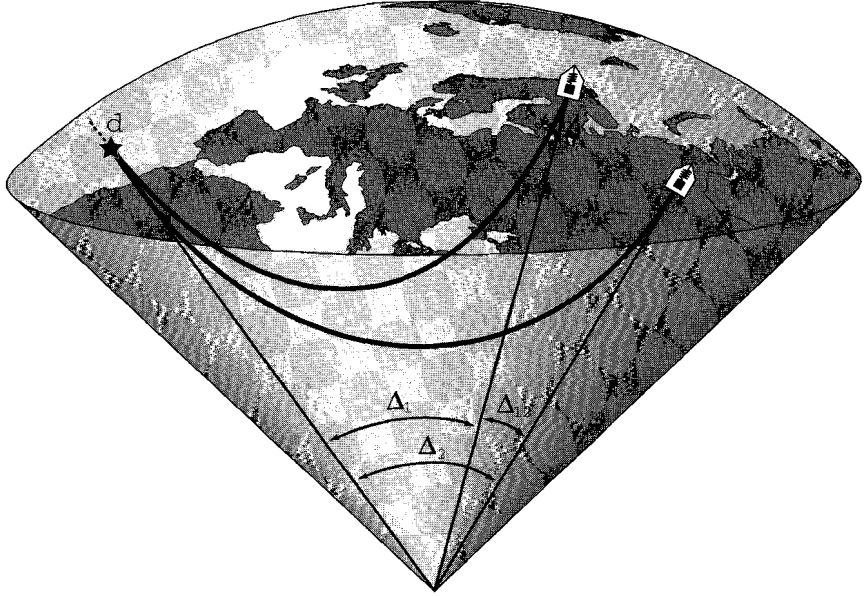


Figure 4.1: *Geometry of pairs of rays that originate from the same earthquake. The geometry is described by the source depth d and the three angles Δ_1 , Δ_2 and Δ_{12} .*

Figure 4.1): two epicentral distances, Δ_1 and Δ_2 , the distance between the two recording stations Δ_{12} and the hypocentral depth d . For small distances Δ_{12} both waves travel along a similar path and sample the same anomalies. Hence their delay times are correlated. For larger Δ_{12} the waves have sampled more different structures and the correlation of delay times decreases. The epicentral distances Δ_1 , Δ_2 and, to a smaller amount, the source depth d determine the maximum depth down to which the mantle is probed.

It is not clear if the Earth can be adequately described as a laterally uniform random medium, i.e. where the correlation function does not change laterally. If this is not the case the very inhomogeneous sampling of the ISC data may cause differences between the properties determined by our statistical approach and the spherically averaged properties of the Earth. In order to test the assumption of a laterally uniform medium, we compare the inferred correlation functions for the whole Earth and 3 subsets for different regions with each other. Two regions are defined by the location of receivers. One subset consists of all delay times

recorded in North America ($160^{\circ}\text{W} < \text{longitude} < 45^{\circ}\text{W}$, $20^{\circ} < \text{latitude}$) and the other of all phases observed in Europe and the Mediterranean ($15^{\circ}\text{W} < \text{longitude} < 40^{\circ}\text{W}$, $25^{\circ}\text{N} < \text{latitude} < 70^{\circ}\text{N}$). The two regions are selected because each contains many stations, which is needed for the analysis, and their difference in tectonic setting. North America, except for the West Coast, is a stable continental shield area, whereas Europe is tectonically much more active. The third subset is defined by the source location and consists of all data from earthquakes in Eastern Pacific region including the plate boundaries around the Philippine plate and South-East Asia ($90^{\circ}\text{E} < \text{longitude} < 170^{\circ}\text{W}$, $50^{\circ}\text{S} < \text{latitude} < 60^{\circ}\text{N}$). These three subsets therefore contain data that sample very different tectonic settings and are well suited to test the hypothesis that the Earth's mantle can be considered as a laterally uniform random medium whose statistical properties vary only with depth.

4.3 Data

The data used in this study consist of the P phases of the data set of Engdahl *et al.* [1998] (hereafter referred to as EHB). This is a subset of the ISC Bulletin in which all teleseismically well-constrained events are relocated using the spherical symmetric Earth model *ak135* [Kennett *et al.*, 1995]. Several phases were removed from the data because their quality is doubtful for the following reasons.

- (1) 'Fake' P arrival times introduced by the NEIC in order to include surface wave magnitudes are removed [Chapter 2].
- (2) For 1849 events, two or more reported arrival times fall in the expected time window for P phases and have been assigned as P by Engdahl *et al.* [1998; Gardien, personal communication, 1997]. Since it is impossible to decide which arrivals are the correct ones all multiple entries are removed.
- (3) Only data with a reading precision of 1 s or better are used. Since the reading precision parameter does not always indicate the correct value [Chapter 3] a probability test is carried out for each station. If the amount of arrival times on full minutes or full ten seconds is so large that the probability of obtaining this result by coincidence is less than 1%, these arrival times are removed.
- (4) Several stations where gross systematic time variations as described in Chapter 2 have been removed. This concerned 2.6% of the data.

4.4 Variance

The delay times of the ISC Bulletin or the EHB subset are not normally distributed but form a distribution with relatively strong tails. It is often described as the sum

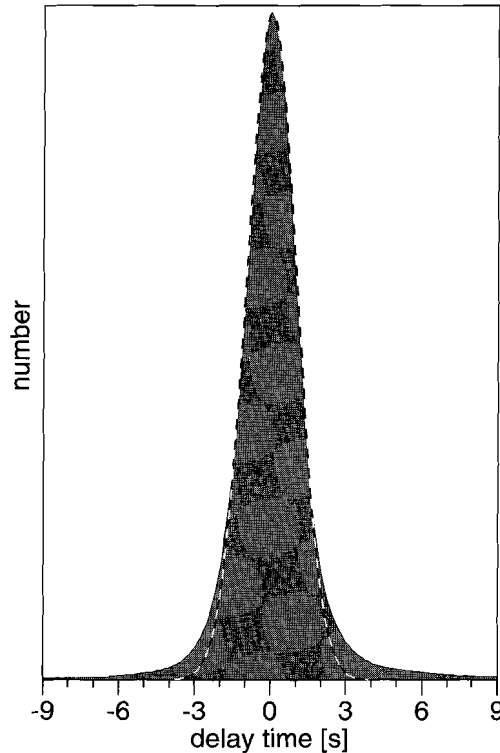


Figure 4.2: *Distribution function of P wave delay times, which has strong tails when it is compared with a normal distribution that fits the center of the distribution (dashed line).*

of two distributions [Jeffreys, 1932; Buland, 1984; Pulliam *et al.* 1993]. Figure 4.2 shows the distribution of all P wave delay times of the EHB data set. A normal distribution (dashed line) is shown for comparison. Although the number of delay times in the tails is very small compared to the number in the central region, they exert a strong influence on the variance of the distribution. Since a large part of the delay times in the tails is generally attributed to errors in the data set (e.g. due to mispicks of the onsets of arrivals, wrong phase associations, faulty instruments, source mislocations and filter effects [e.g. Scherbaum and Bouin, 1997]) the distribution is truncated prior to any analysis. Examples of employed cut-off values are 5 s [Dziewonski, Hager & O'Connell, 1977] or 3.5 s for teleseismic P waves and 7.5 s for epicentral distances shorter than 25° [Engdahl *et al.*, 1998; Bijwaard *et al.*, 1998]. Pulliam *et al.* [1993] discuss the application of the ℓ_1 -norm instead of the more often used ℓ_2 -norm to decrease the sensitivity on outliers.

Figure 4.3 shows the variance of the distribution as a function of chosen cut-off

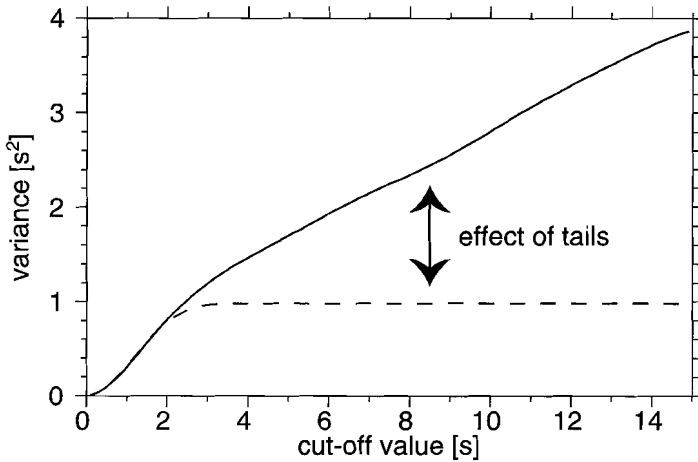


Figure 4.3: Measured variance of P wave delay times as a function of cut-off value. The tails exert a strong influence on the variance. The dashed line shows the variance of a normal distribution truncated at the same cut-off value. The finite second moment of the normal distribution ($\sigma^2 = 1\text{s}^2$) is approached rapidly within a few standard deviations and cut-off values larger than a few seconds do not significantly influence the result.

value. The dashed line shows the variance if the normal distribution of Figure 4.2 (dashed line) is truncated. In contrast to a normal distribution the variance of the ISC data does not flatten out with increasing the cut-off value. Values for the variance of delay times therefore strongly depend on the chosen cut-off value. Fortunately, we are interested in the relative variances for different epicentral distances and hypocentral depths. Their relative values depend much less on the cut-off value. We adopt a cut-off value of 3.5 s, commonly used in tomographic studies.

For many combinations of source depth and epicentral distance, the sampling is very inhomogeneous and dominated by a few source regions and certain stations. In order to suppress artifacts from the inhomogeneous sampling, weights are introduced according to the source density. For each station all phases from earthquakes receive a weight that is proportional to the inverse number of earthquakes in the same event-cluster volume. The event-cluster volumes have a size of $100\text{ km} \times 100\text{ km} \times 100\text{ km}$. The importance of the weighting increases with depth due to the fact that the sources are restricted to a few subduction zones and the sampling becomes even more inhomogeneous relative to shallow source depths. The effect of weighting is demonstrated for events with source depths larger than 500 km in Figure 4.4. The distribution of earthquakes at these depths are limited to a few subduction zones. Whereas a trend of the variance with epicentral distance is still visible for the weighted estimates (filled triangles), the unweighted

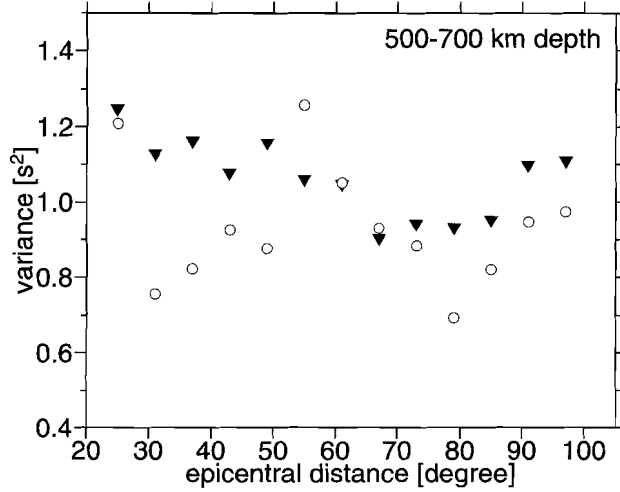


Figure 4.4: Variance of P wave delay times for epicentral distances between 22° and 100° , averaged over intervals of 6° . Hypocentral depths are larger than 500 km. The open circles represent the result, if the data are not weighted, whereas the filled triangles show the outcome, if the data are weighted according to the number of earthquakes in each event-cluster volume.

estimates do not show any trend.

Figure 4.5 shows the variance of the whole data set and the three regions, specified in Section 4.2, for 4 depth intervals. Symbols are missing if the weighted number of delay times was less than 200. General features are a decrease of the variance with source depth and with epicentral distance until 80° and an increase for larger distances. Since the variance for a uniform random medium increases with path length, Gudmundsson et al. [1990] interpret the decrease with epicentral distance as a concentration of the strength of heterogeneity at shallow depths. The variance for distances less than about 30° might partially be due to picks of later phases of upper mantle triplications and wrong phase associations between P and depth phases, e.g. pP . At very large distances the relatively strong variance is probably caused by an increase of heterogeneity close to the CMB (D'' layer) and wrong phase association between P and PcP .

Before we continue with the measurement of the covariance for pairs of rays we want to make a general comment on delay times and their variance, which may seem trivial but is sometimes overlooked.

Delay times are defined as the difference between the actual travel time (i.e., the difference between the observed arrival time and the origin time) and the reference travel time in a one-dimensional reference model for the same distance between the hypocenter and the station. The optimal situation for determining the

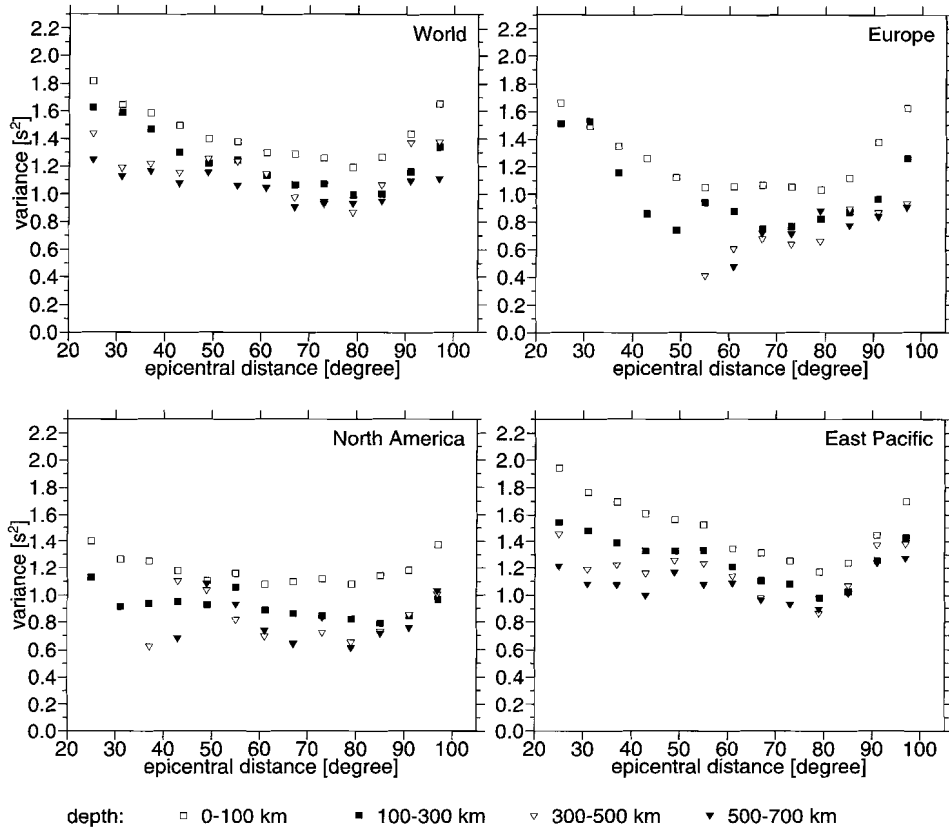


Figure 4.5: Variance estimates for four different source depth and the regions indicated in the upper left corner. Regions are specified in Section 4.2. The measurements are only used, if the number of weighted delay times is at least 200.

three-dimensional Earth structure would be if we had the true hypocenter and the true origin time. In reality, both must be estimated, normally using the same data that are used later for inferring the three-dimensional Earth structure. The true source parameters are not obtained due to the effects of unknown heterogeneity and errors in the observed arrival times. The best estimate of the hypocenter and origin time that can be determined by earthquake location procedures is the point in space and time for which the variance (or some other norm) of the delay times is minimal. Thus, the error in the hypocenter determination decreases the delay time variance and the variance for the true hypocenter location would be larger. This is very different from picking errors of arrival times which generally increase the delay time variance.

source depth [km]	total number	early arrivals		late arrivals	
		number	percentage	number	percentage
0 – 700 (Fig. 4.6)	10 897 882	89 563	(0.82%)	137 679	(1.26%)
0 – 33	5 259 133	48 545	(0.92%)	98 762	(1.88%)
33 – 100	3 572 823	28 205	(0.79%)	34 372	(0.96%)
100 – 700 (Fig. 4.8)	2 065 926	12 813	(0.62%)	4 545	(0.22%)

Table 4.1: *The number of pairs contributing to the bivariate distributions split in the total number and the numbers contributing to the negative (early arrival times) and positive (late arrival times) uncorrelated tails. The regions that define the uncorrelated tails are outlined in Figure 4.8.*

4.5 Correlation

The correlation coefficient is measured as a function of the four parameters describing the geometry shown in Figure 4.1 (d , Δ_1 , Δ_2 and Δ_{12}). If Δ_{12} is small, both rays are similar and sample the same heterogeneity. The bivariate distribution of all pairs of delay times with $\Delta_{12} < 3^\circ$ is displayed in Figure 4.6. Since the two stations are indistinguishable in this plot, each pair is taken twice to make the plot symmetric about the diagonal. If all pairs had the same correlation the isolines would be ellipses with the principal axes falling on the diagonals. The center of the plot is in agreement with this but the tails split in diagonal and horizontal (equals vertical) tails. The horizontal tail consists of pairs whose delay times are not correlated at all. Since the strength of the horizontal (and vertical) tails decays slower than the correlated residuals close to the diagonal the use of cut-off values for ISC delay times improves the signal-to-noise ratio.

Although the number of pairs in the tails is only about 1% of the total number they considerably influence the correlation coefficient. The dependence of the correlation on the cut-off value for the data of Figure 4.6 is shown in Figure 4.7. The largest correlation coefficient is obtained if no residuals with absolute values larger than 2.2 s are used in the calculation.

Two principle sources contribute to the tails of uncorrelated delay times: wrong picks and wrong phase associations. The latter cause only contributes to late arrival times (positive delay times) and certain combinations of source depth and epicentral distance, at which later arriving waves follow shortly after P . Possible phases that could accidentally be identified as P are (1) depth phases for shallow earthquakes, e.g. pP or sP , (2) PcP for large epicentral distances and (3) later phases of upper mantle triplications for short distances. Because the positive tails (late arrivals) are much stronger than the negative ones, we examine the possibility of phase misidentification further and split the data set depending on the source depth. Figure 4.8 shows the same bivariate distribution as Figure 4.6 but only for earthquakes that originated deeper than 100 km. Clearly, the positive tails are relatively smaller if the shallow earthquakes are excluded. In order to make

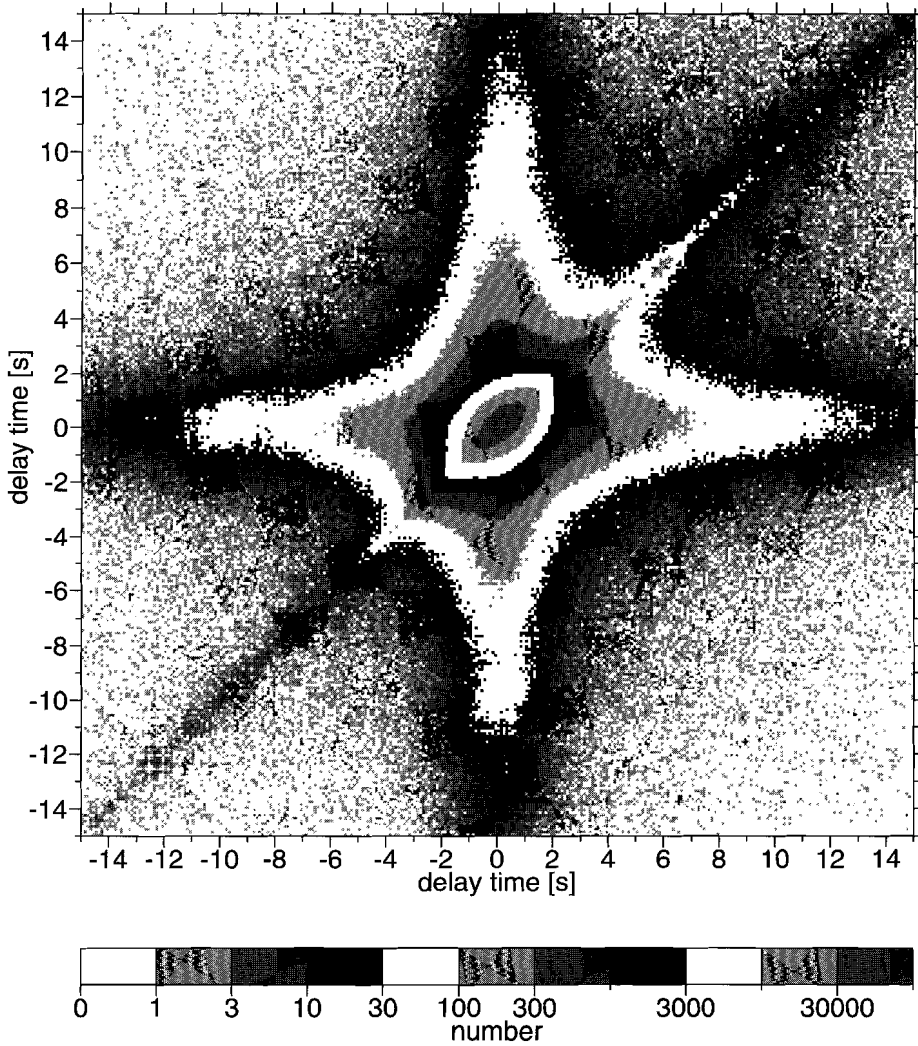


Figure 4.6: *Bivariate distribution of pairs of P wave delay times that originate from the same earthquake and are observed by stations that are closer than 3° . The gray scale is logarithmic and repeating gray tones mean an increase by a factor of 100.*

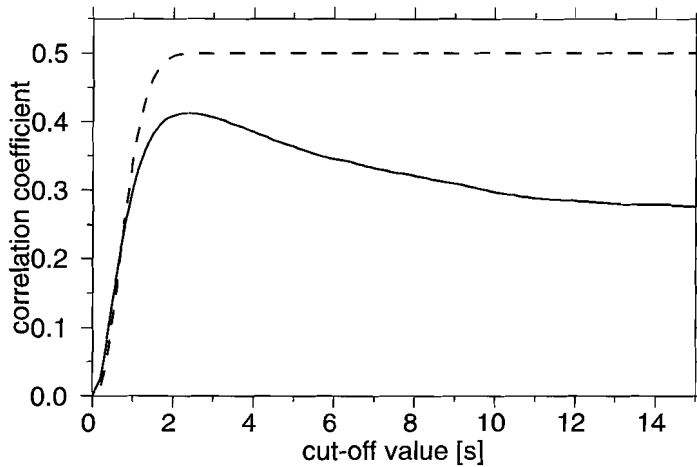


Figure 4.7: Correlation coefficient of all pairs of P wave delay times with a common source and stations closer than \mathcal{F} (data of Figure 4.6) as a function of cut-off value. The dashed line shows the result for a bivariate normal distribution with a correlation coefficient of 0.5.

this result more quantitative we counted the number of pairs in the positive and negative tails, indicated by the rectangles in Figure 4.8, for three different source depth intervals. The numbers are listed in Table 4.1. It can be seen that both tails become weaker with increasing source depth. This characteristic is much more pronounced for the positive tails (late arrival times). Thus, misidentification of depth phases as P can explain a large part of the positive tails because only for shallow depths the difference in travel times between these phases is small enough that both phases fall in the same time window

Another, surprising feature in Figures 4.6 and 4.8 is, that the correlated tails (diagonal) are very narrow for large absolute delay times. Points of a bivariate distribution with the same correlation coefficient as in the central region should be scattered in a broader region. Therefore, we are suspicious that also very large delay times in the correlated tails often are not correct. Several causes are possible: (1) Errors in the hypocenter location and origin time influence the delay times of nearby stations in the same way. (2) Positive delay times could originate, if for both stations later phases (pP , sP , pwP , PcP) are misidentified as P . (3) Both stations picked a wave of an undetected local earthquake. (4) Since the data contributing to Figures 4.6 and 4.8 come from stations that are close to each other, there is a fair chance that the stations belong to the same network and that the seismograms are not picked independently [Chapter 2]. Therefore, it is more likely

A pwP wave is a compressional wave that leaves the source in upward direction, is reflected at the water surface and travels to the station as a P wave.

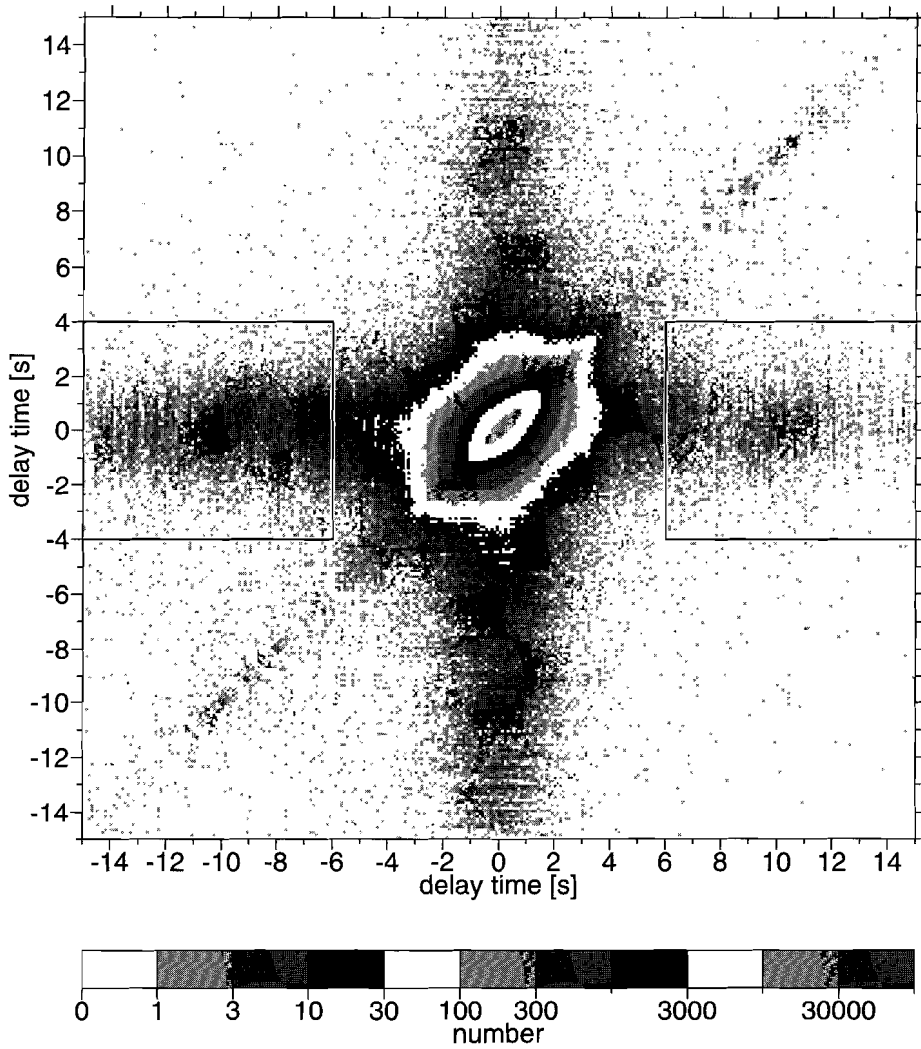


Figure 4.8: Same as Figure 4.6 but only for earthquakes with source depth larger than 100 km. Note that the positive tails are less prominent even if the lower total number of pairs is considered.

that the same mispick is done for both seismograms or that a prominent noise peak (or arrival from another earthquake) was picked twice.

The influence of the uncorrelated tails on the correlation coefficient and the subjectiveness of choosing a small cut-off value prevents us from using

$$r = \frac{\sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y)}{\sqrt{\sigma_x^2 \sigma_y^2}} \quad (4.2)$$

for the calculation of the correlation between pairs of delay times, where μ is the mean and σ the standard deviation. Instead we fit a theoretical distribution function to the observed bivariate distributions as displayed in Figure 4.6 and 4.8. We model the probability density function of observed delay times by a Jeffrey distribution, i.e. the sum of two normal distributions

$$\begin{aligned} \text{pdf}(\delta t) &= p \text{pdf}_{\text{signal}}(\delta t) + (1-p) \text{pdf}_{\text{noise}}(\delta t) \\ &= \frac{p}{\sqrt{2\pi} \sigma_s} \exp\left(-\frac{1}{2} \frac{(\delta t - \mu_s)^2}{\sigma_s^2}\right) + \frac{1-p}{\sqrt{2\pi} \sigma_n} \exp\left(-\frac{1}{2} \frac{(\delta t - \mu_n)^2}{\sigma_n^2}\right) \end{aligned} \quad (4.3)$$

where p is the probability that a delay time belongs to a correct pick. The probability density function $\text{pdf}_{\text{signal}}(\delta t)$ describes the distribution of correct picks and includes small random errors, which are present in the correlated delay times and broaden the probability density function $\text{pdf}_{\text{signal}}(\delta t)$, whereas $\text{pdf}_{\text{noise}}(\delta t)$ is the distribution of delay times that show no correlation with other delay times at all.

The joint probability density function of pairs of delay times δt_1 and δt_2 are obtained by considering the fact that only the signal pdf's are correlated. If the correlation coefficient between delay times from $\text{pdf}_{\text{signal}}(\delta t_1)$ and $\text{pdf}_{\text{signal}}(\delta t_2)$ is given by r the joint probability density function $\text{pdf}(\delta t_1, \delta t_2)$ is

$$\begin{aligned} \text{pdf}(\delta t_1, \delta t_2) &= \\ &= \frac{p^2}{2\pi\sigma_s^2\sqrt{1-r^2}} \exp\left(-\frac{1}{2} \frac{(\delta t_1 - \mu_s)^2 + (\delta t_2 - \mu_s)^2 - 2r(\delta t_1 - \mu_s)(\delta t_2 - \mu_s)}{(1-r^2)\sigma_s^2}\right) \\ &+ \frac{p(1-p)}{2\pi\sigma_s\sigma_n} \exp\left(-\frac{1}{2} \left\{ \frac{(\delta t_1 - \mu_s)^2}{\sigma_s^2} + \frac{(\delta t_2 - \mu_n)^2}{\sigma_n^2} \right\}\right) \\ &+ \frac{p(1-p)}{2\pi\sigma_s\sigma_n} \exp\left(-\frac{1}{2} \left\{ \frac{(\delta t_1 - \mu_n)^2}{\sigma_n^2} + \frac{(\delta t_2 - \mu_s)^2}{\sigma_s^2} \right\}\right) \\ &+ \frac{(1-p)^2}{2\pi\sigma_n^2} \exp\left(-\frac{1}{2} \frac{(\delta t_1 - \mu_n)^2 + (\delta t_2 - \mu_n)^2}{\sigma_n^2}\right) \end{aligned} \quad (4.4)$$

It is assumed here that the distributions of delay times δt_1 and δt_2 are indistinguishable, i.e. that they have the same means μ_s and μ_n and the same standard deviations σ_s and σ_n . This is not necessarily fulfilled if the two epicentral distances are different, since the variance varies with this parameter as is shown in

Figure 4.5. However, since the correlation coefficient approaches rapidly zero with increasing distance between the two stations Δ_{12} , we are mainly interested in very small Δ_{12} . For small Δ_{12} , however, the two epicentral distances are similar and the assumption is correct. For the few measurements with large difference between Δ_1 and Δ_2 we will underestimate the absolute value of r .

The correlation coefficient is measured for the global data set and the three different regions. Four depth intervals, 6 bins for the two epicentral distances (from 28° to 100°) and 13 bins with exponentially increasing width for Δ_{12} are used. This gives, under the constraint that $|\Delta_1 - \Delta_2| < \Delta_{12}$, 636 measurements of the correlation coefficient for each region. Each measurement is done in the following way. First all pairs are constructed that fall in the same bins for d , Δ_1 , Δ_2 and Δ_{12} . Inhomogeneous ray path distribution is reduced by weighting each pair with the inverse number of other pairs in the same ray pair cluster. A ray pair cluster corresponds to all pairs of rays, whose source is in the same event-cluster volume (size $100 \text{ km} \times 100 \text{ km} \times 100 \text{ km}$) and the same two station-cluster areas (size $100 \text{ km} \times 100 \text{ km}$). This gives pairs of delay times and corresponding weights. The bivariate distribution is calculated by taking each pair twice as $(\delta t_1, \delta t_2)$ and $(\delta t_2, \delta t_1)$ since there is no difference whether δt_1 was measured at station A and δt_2 at station B or vice versa. Taking each pair twice makes the bivariate distribution symmetric and, in the case of few measurements, gives more stable estimates for r . This procedure corresponds to the above assumption that the two delay times δt_1 and δt_2 are indistinguishable.

The joint probability density function of eq. (4.4) is fitted to the distribution. A Monte Carlo method is used to give the least-squares fit to the logarithm of the distribution for absolute delay times smaller than 5 s. Two examples of these measurements are shown in Figure 4.9. The maximum amplitudes are somewhat underestimated but the general shape, which determines the correlation coefficient, is generally fitted accurately. The fit becomes worse if less pairs contribute to the bivariate distribution and measurements are only performed if at least 500 pairs could be found.

This procedure gives generally larger estimates of r than eq. (4.2) because the influence of the uncorrelated tails is removed. Uncertainty estimates of r are obtained by measuring the width of the global minimum of the merit function. This is done by randomly sampling of the model space around the global minimum and accepting all models whose misfit is not larger than 105% of the minimum.

The result of the correlation measurements for two source depth intervals is displayed in Figure 4.10. The correlation coefficient for two subregions and the shallowest source depth interval, which contains the most data, is shown in Figure 4.11 for comparison. The decrease of the correlation coefficient with growing station separation is by far the strongest dependence. A slight dependence on the source depth and the two epicentral distances is not much larger than the uncertainties of the measurements. The larger scatter and error bars for North America (Figure 4.11, upper panel) stems from the relatively few pairs for large station separation. The correlation coefficients do not approach unity for the limit

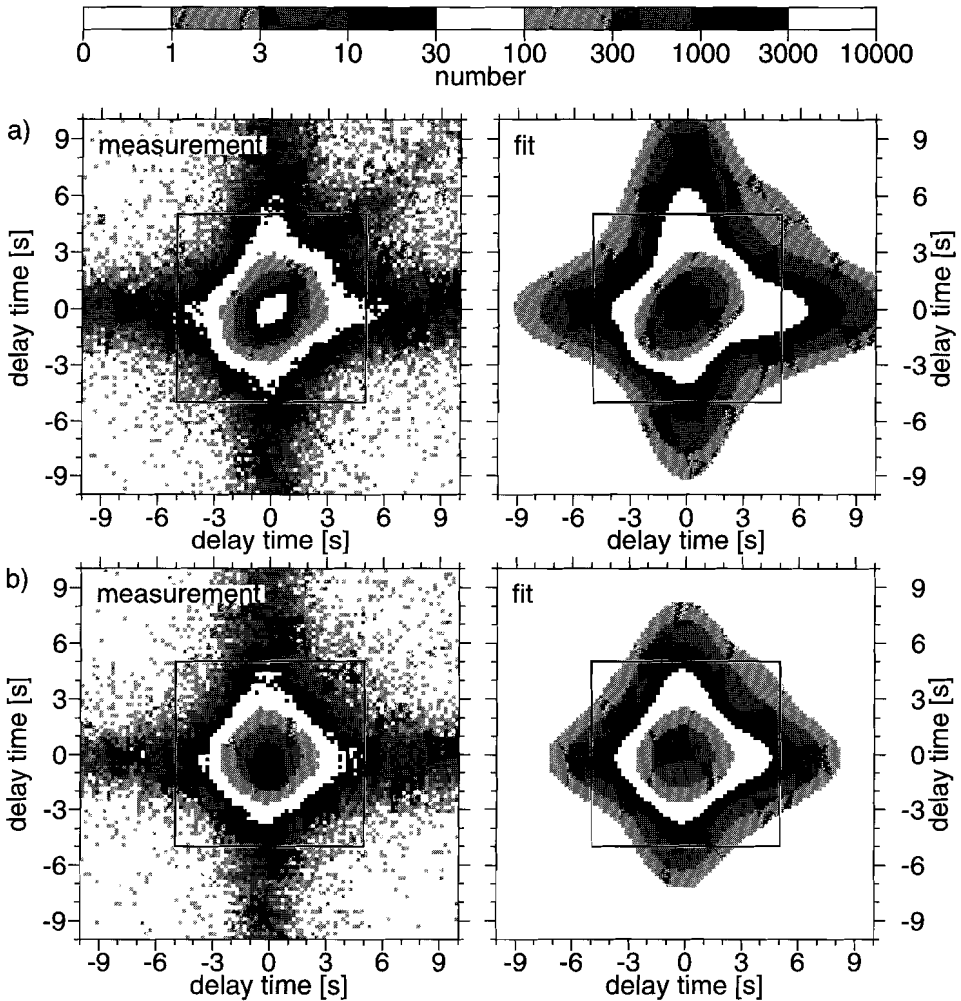


Figure 4.9: *Examples of measured bivariate distributions and the best fit used to determine the correlation coefficient. The square indicates the region used for the least-squares fit. The parameter combination is: $0 < d < 100$ km, $52^\circ < \Delta_1 < 64^\circ$, $52^\circ < \Delta_2 < 64^\circ$. (a) Distance between the stations is small ($1.94^\circ < \Delta_{12} < 4.19^\circ$) and the correlation coefficient estimate of the delay times is 0.46. (b) The correlation for larger distances between the stations ($27.74^\circ < \Delta_{12} < 34.16^\circ$) is absent and the estimate of r yields 0.04.*

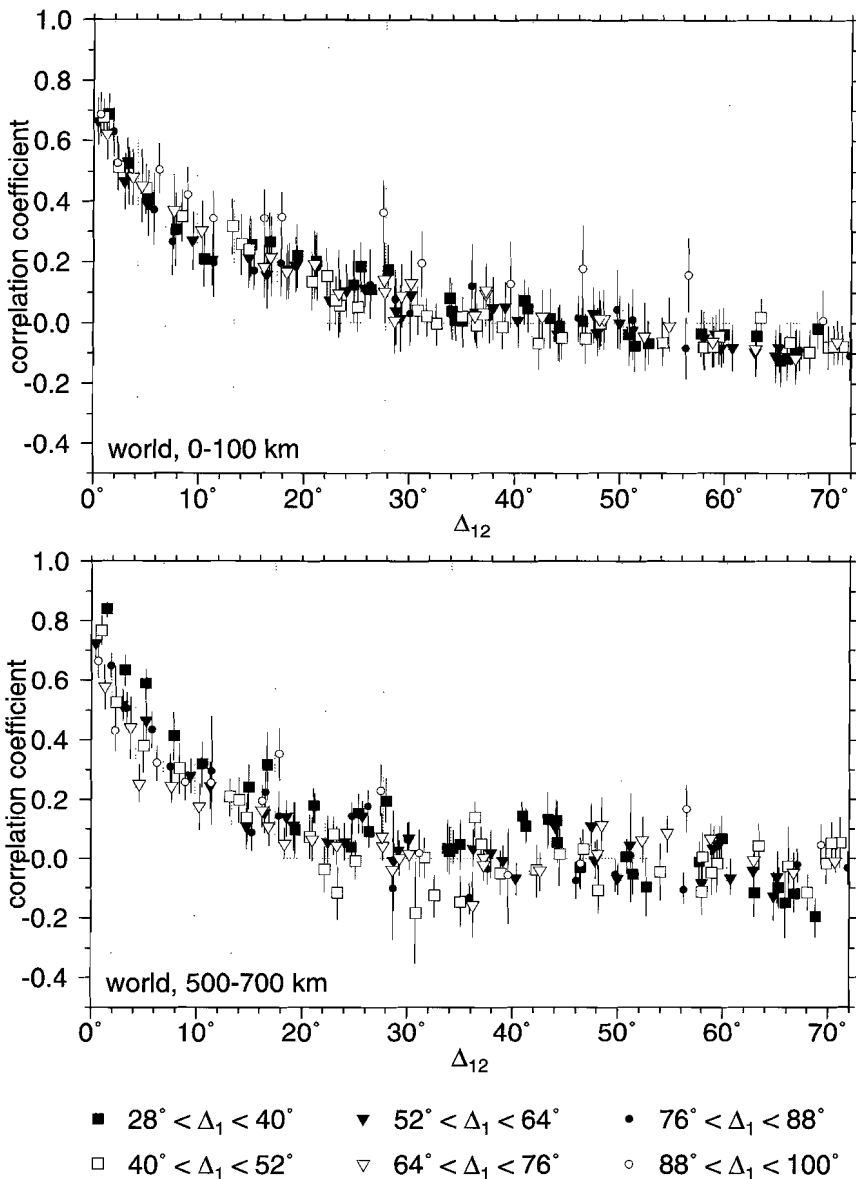


Figure 4.10: Measured correlation coefficient of the global data set as a function of the distance between the two stations Δ_{12} . The source depth intervals are written in the lower left corners. The symbols display the smaller value of the two epicentral distances ($\Delta_1 \leq \Delta_2$). The dotted vertical grid lines show the intervals used for the binning. They have exponentially increasing widths. In order to make the symbols better visible, they are horizontally randomly shifted within each bin. This is done only for visualization.

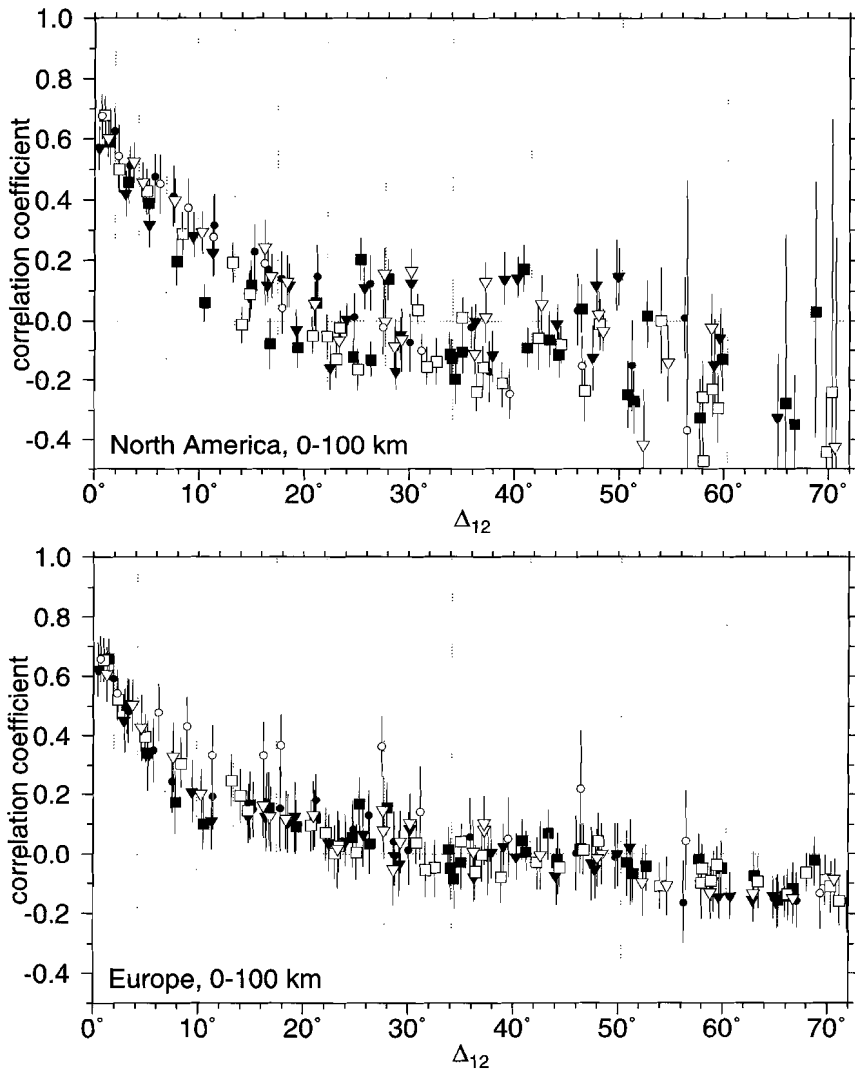


Figure 4.11: The correlation coefficient for two subregions. All stations are located in North America and Europe for the upper and lower panel respectively. Source depths are shallower than 100 km. Symbol legend same as in Figure 4.10.

of $\Delta_{12} \rightarrow 0$ because of measurement and hypocenter location errors which are present in $\text{pdf}_{\text{signal}}(\delta t)$ of eq. (4.3). This is equivalent to the findings of Gudmundsson *et al.* [1990] that the variance of delay times from the same ray bundle does not tend to zero for infinitesimal width of the summary rays.

4.6 Inversion for mantle variance and correlation functions

In the ray theoretical approach a delay time can be written as the integral of the slowness perturbations along the ray path (eq. (1.2)):

$$\delta t = \int_{\text{path}} \delta u(\mathbf{r}) ds \quad . \quad (4.5)$$

This yields for the covariance between two delay times:

$$\begin{aligned} \langle \delta t_1 \delta t_2 \rangle &= \left\langle \int_{\text{path1}} \delta u(\mathbf{r}_1) ds_1 \int_{\text{path2}} \delta u(\mathbf{r}_2) ds_2 \right\rangle \\ &= \int_{\text{path1}} \int_{\text{path2}} \langle \delta u(\mathbf{r}_1) \delta u(\mathbf{r}_2) \rangle ds_1 ds_2 \quad . \end{aligned} \quad (4.6)$$

The indices 1 and 2 are used to distinguish the two stations of each pair and $\langle \rangle$ denotes the expected value. Eq. (4.6) defines the inverse problem between the covariance of delay times $\langle \delta t_1 \delta t_2 \rangle$ and the autocovariance function of slowness perturbations $\langle \delta u(\mathbf{r}_1) \delta u(\mathbf{r}_2) \rangle$. If the amplitudes of the heterogeneity are small, such that ray bending effects can be neglected and the ray paths are well approximated by the paths in a one-dimensional reference model, this inverse problem is linear.

Eq. (4.6) could be further simplified, if we would make certain assumptions about the autocorrelation function of the mantle heterogeneity. For instance, Gudmundsson *et al.* [1990] assume that the autocorrelation function of the medium is Gaussian and its scale length is small relative to the ray curvature. This allows to restrict the integrals to small parts of the rays in which they are approximated by straight lines.

We do not think that any of these assumptions are appropriate for the Earth. Global tomography has revealed the existence of large scale variations and our knowledge about the mantle is insufficient to assume *a priori* a certain shape of the autocorrelation function. Therefore, we do not simplify eq. (4.6) any further. For a laterally uniform random medium, the autocovariance function of the medium is a function of three variables, for which we choose the depth z of the upper of the two points \mathbf{r}_1 and \mathbf{r}_2 , the horizontal and vertical distance between them Δ_h and Δ_v respectively:

$$\langle \delta u(\mathbf{r}_1) \delta u(\mathbf{r}_2) \rangle =: C(z, \Delta_h, \Delta_v) \quad . \quad (4.7)$$

The autocovariance function can be split in the variance and correlation function:

$$C(z, \Delta_v, \Delta_h) = \sqrt{\text{Var}(z) \text{Var}(z + \Delta_v)} \text{Cor}(z, \Delta_h, \Delta_v) \quad . \quad (4.8)$$

Eq. (4.6) would be enough to invert the measured variances $\langle \delta t^2 \rangle(d, \Delta)$ and covariances $\langle \delta t_1 \delta t_2 \rangle(d, \Delta_1, \Delta_2, \Delta_{12})$ if no noise was present in the data. That this is not the case can directly be seen from Figures 4.10 and 4.11 where the correlation does not approach unity for $\Delta_{12} \rightarrow 0$. The fit of two separate probability density functions (eq. 4.4) removes the influence of outliers on the correlation coefficient but not small measurement and location errors. Therefore, we model the delay times δt from station i by the sum of three contributions:

$$\delta t_i = \delta t_{is} + \delta t_{in} + \delta t_{im} \quad . \quad (4.9)$$

The delay caused by three-dimensional structure δt_{is} forms the signal of our main interest. The second term δt_{in} represents an incoherent noise component which is uncorrelated with any other term:

$$\begin{aligned} \langle \delta t_{in} \delta t_{js} \rangle &= \langle \delta t_{in} \delta t_{jm} \rangle = 0 & \forall i, j \\ \langle \delta t_{in} \delta t_{jn} \rangle &= 0 & \forall i, j, i \neq j \quad . \end{aligned} \quad (4.10)$$

The last term in eq. (4.9) represents noise that is the same for all observations of the same event. It is completely correlated between different stations (we drop the station subscript therefore in the following equations) and also not independent from the structural signal:

$$\begin{aligned} \delta t_{im} &= \delta t_{jm} = \delta t_m & \forall i, j \\ \langle \delta t_m \delta t_{is} \rangle &\neq 0 \end{aligned} \quad (4.11)$$

This term is introduced to model origin time errors (and possibly some effects by vertical mislocations which has a strong trade-off with the origin time). As an example of the importance of this term, consider a mid-oceanic ridge event which occurs generally in a slow velocity anomaly. This results in late arrival times at all stations. Since the location is determined by minimization of the delay time variance, the mean delay is absorbed by a late origin time and is not reflected in the calculated delay times. The same effect occurs at the receiver side, if station residuals are used as a correction to delay times. As a result, delay times contain little information about long wavelength upper mantle heterogeneity. With the partition of delay times in eq. (4.9) we obtain for the variance

$$\begin{aligned} \langle \delta t^2 \rangle &= \langle (\delta t_s + \delta t_n + \delta t_m)^2 \rangle \\ &= \langle \delta t_s^2 \rangle + \langle \delta t_n^2 \rangle + \langle \delta t_m^2 \rangle + 2\langle \delta t_s \delta t_m \rangle \quad . \end{aligned} \quad (4.12)$$

Here, the station index i has been omitted because in a uniform random medium the expected value of the variance is for every station location the same and only depends on the epicentral distance and hypocenter depth.

The term δt_m is negatively correlated with the structural information δt_s . This follows from the fact that the variance is minimized during the earthquake location procedure. For the correct source location the delay time variance would be $\langle \delta t_s^2 \rangle + \langle \delta t_n^2 \rangle$ and $\delta t_m = 0$. The estimated source location will be different only if another location exists for which $\langle \delta t^2 \rangle$ is smaller. This is the case only if

$$\langle \delta t_m^2 \rangle + 2\langle \delta t_s \delta t_m \rangle \leq 0 \quad (4.13)$$

This reasoning is, strictly speaking, only true if the variance is calculated for the same set of delay times as is used for the location procedure. Since P is by far the most frequently observed phase and P and S wave velocities are correlated, this result probably also holds in our case where we consider only P wave delay times.

The covariance accordingly gives

$$\begin{aligned} \langle \delta t_1 \delta t_2 \rangle &= \langle (\delta t_{1s} + \delta t_{1n} + \delta t_m) (\delta t_{2s} + \delta t_{2n} + \delta t_m) \rangle \\ &= \langle \delta t_{1s} \delta t_{2s} \rangle + \langle \delta t_{1s} \delta t_m \rangle + \langle \delta t_{2s} \delta t_m \rangle + \langle \delta t_m^2 \rangle \end{aligned} \quad (4.14)$$

The incoherent noise δt_n does not influence the covariance whereas the coherent noise δt_m affects both the variance and covariance. The negative correlation between δt_m and δt_s (eq. (4.13)) gives an alternative interpretation for the slope of the correlation coefficient of delay times, which is negative even at very large station separation Δ_{12} (Figures 4.10 and 4.11). Without the mislocation term δt_m this would be interpreted as structure with a wavelength of the order of the station separation at which it is observed. However, according to eq. (4.13) the last three terms of eq. (4.14) are negative. Therefore, eq. (4.14) shows that the structural correlation $\langle \delta t_{1s} \delta t_{2s} \rangle$ could be larger than the measured correlation $\langle \delta t_1 \delta t_2 \rangle$, which contains all terms.

The term δt_m could be included as a parameter in the inversion but we choose to minimize a quantity that is independent of δt_m , which is the variance of differential travel times

$$\begin{aligned} \langle (\delta t_1 - \delta t_2)^2 \rangle &= \langle \delta t_1^2 \rangle + \langle \delta t_2^2 \rangle - 2\langle \delta t_1 \delta t_2 \rangle \\ &= \langle \delta t_{1s}^2 \rangle + \langle \delta t_{2s}^2 \rangle + \langle \delta t_{1n}^2 \rangle + \langle \delta t_{2n}^2 \rangle - 2\langle \delta t_{1s} \delta t_{2s} \rangle \end{aligned} \quad (4.15)$$

We refer to the above expected value as structure function, an expression used in meteorology where analogous functions are used to describe turbulent flows [Stull, 1991].

Gudmundsson *et al.* [1990] observe a small, but consistent, drop of the summary ray variance for scale lengths between 60° and 90° (their Figure 2; text on p. 27) which they cannot explain. This could be the same effect as the drop to negative correlation values in our Figures 4.10 and 4.11 at about 60° and can partly be explained by the mislocation term δt_m . The effect of the terms containing δt_m is increasing with the size of the summary rays.

4.6.1 Theoretical constraints on the autocorrelation function

The autocorrelation and autocovariance functions possess two symmetries in the case of a layered, laterally uniform random medium:

$$C(z, \Delta_h, \Delta_v) = C(z, -\Delta_h, \Delta_v) \quad (4.16)$$

$$C(z, \Delta_h, \Delta_v) = C(z + \Delta_v, \Delta_h, -\Delta_v) \quad (4.17)$$

It is assumed here that z and Δ_v are positive in the same (downward) direction.

The mean velocity for all depths is described by the one-dimensional reference model. The covariance function characterizes the random fluctuations δu around the velocity of the reference model. The mean of all fluctuations on a spherical layer is zero

$$\langle \delta u \rangle = 0 \quad (4.18)$$

which translates directly into the following condition for the covariance (and correlation) function:

$$\int_0^\pi C(z, \Delta_h, \Delta_v) \sin(\Delta_h) d\Delta_h = 0 \quad . \quad (4.19)$$

The derivation is given in Appendix B.1. Because we are using only direct P phases the mantle is only sampled to a maximum distance of about 100° at the surface and even less in deeper portions of the mantle. The part of the mantle that is unconstrained by the data can always take values such that eq. (4.19) is satisfied. Therefore, this condition cannot be incorporated in the inversion.

Another constraint has been pointed out by Chernov [1967]. If the random fluctuations are a continuous function, the gradient at zero separation vanishes. For a laterally uniform medium on a sphere, with no preferred orientation we obtain:

$$\left[\frac{dC(z, \Delta_h, \Delta_v)}{d\Delta_h} \right]_{\Delta_h=0} = 0 \quad . \quad (4.20)$$

The proof can be found in appendix B.2. We do also not incorporate this constraint in the inversion for two reasons: (1) If heterogeneity in the mantle is (partially) caused by compositional differences it is well possible that $\delta u(\vartheta, \phi)$ is discontinuous. (If it was purely thermally induced heat diffusion would prevent discontinuous properties.) (2) If the smallest scales of heterogeneity is smaller than what we can measure in the data or resolve in the inversion parameterization, the true correlation function might be better approximated by a correlation function that does not satisfy eq. (4.20).

The measured correlation functions as shown in Figures 4.10 and 4.11 strongly suggest that at least one of the arguments is true. It is derived in Appendix B.3 that if the derivative of C at zero distance would equal zero, then the covariance (and correlation) function of delay times as a function of distance Δ_{12} had a zero slope for $\Delta_{12} \rightarrow 0$. This is not found in Figures 4.10 and 4.11. We can thus conclude from these figures that the heterogeneity is either discontinuous or has at least heterogeneity at scales smaller than about 80 km, which is the smallest bin size used for the correlation estimates. A third option would be a bias in the correlation measurements due to some sort of correlated noise (other than δt_m).

4.6.2 Model parameterization and inversion algorithm

The function $C(z, \Delta_v, \Delta_h)$ is parameterized by natural cubic splines (second derivative equals zero at boundaries). In order to minimize the number of nodes and still be able to describe narrow correlation functions the distance between the nodes is exponentially increasing in Δ_v and Δ_h and linear in z . In total, 150 coefficients are used.

Additionally, 9 parameters describing the random noise component δt_n in the variance measurements are used. The dependence on epicentral distance is assumed to be the same for each source depth and is described by 6 parameters and three additional parameters describe the overall dependence with depth.

We solve for the autocovariance function of mantle structure by minimizing the misfit for the structure function which can be split in the sum of three double integrals and two noise parameters (see eq. (4.15)). We neglect ray bending effects due to three dimensional structure. In this case the inverse problem is linear. Linear problems are well suited to be solved on computers. However, it is hardly possible to constrain the solution of linear least-squares algorithms. It can be seen from eq. (4.8) that the autocovariance function of the medium must fulfill certain criteria because the variance is non-negative and the correlation function limited between -1 and 1. Since

$$\text{Var}(z) = C(z, 0, 0) \quad (4.21)$$

the covariance must fulfill

$$C(z, 0, 0) > 0 \quad (4.22)$$

$$|C(z, \Delta_v, \Delta_h)| \leq \sqrt{C(z, 0, 0) C(z + \Delta_v, 0, 0)} \quad (4.23)$$

Furthermore, we constrain the solution by searching for smooth models. This is achieved by adding an extra term to the misfit function depending on the second derivatives of C and the variance. The constraints on variance and correlation coefficient are forced by returning a large misfit value on violation. The downhill simplex method [Nelder and Mead, 1965; Press *et al.*, 1992] is used to find the minimum. The symmetries (4.16) and (4.17) are implicitly taken into account by only parameterizing the autocovariance function for $\Delta_h > 0$ and $\Delta_v > 0$.

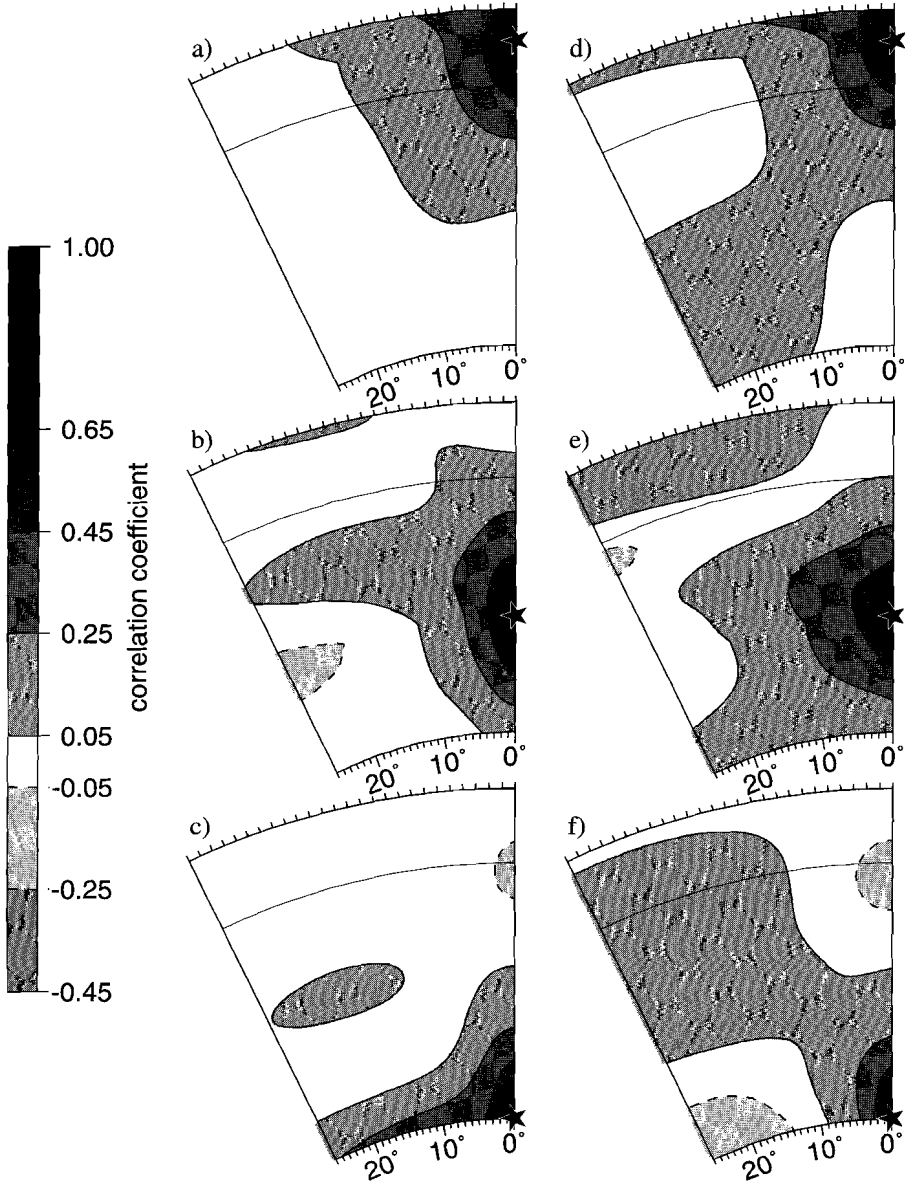


Figure 4.12: The autocorrelation function determined for the global data set [a), b) and c)] and for receivers in North America [d), e) and f)]. Each slice shows the correlation coefficient of a point at the depth indicated by the star and a second point at the vertical and lateral distance inside the frame. The upper and lower boundary of each panel is given by the Earth's surface and the core-mantle boundary respectively. The boundary between upper and lower mantle at 660 km is also drawn. Negative contour lines are dashed.

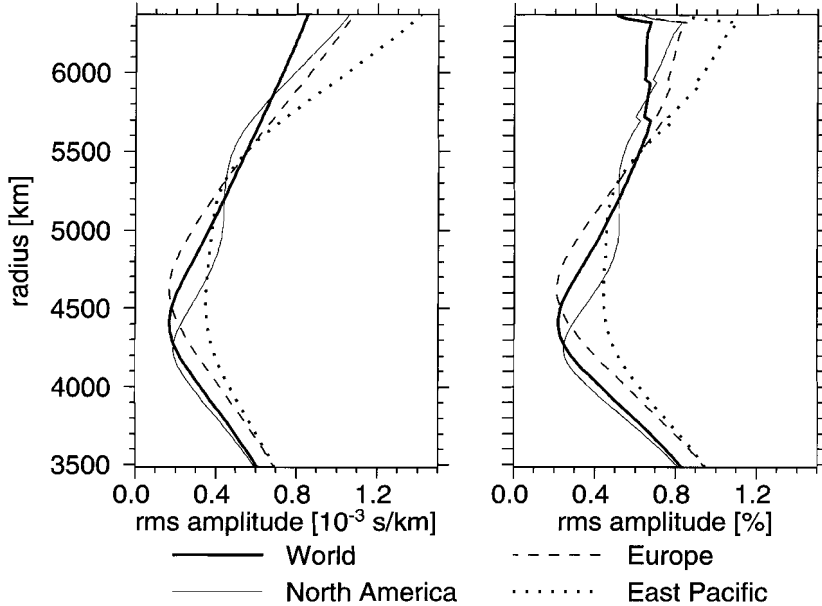


Figure 4.13: *Inverted r.m.s. amplitude of mantle heterogeneity as a function of depth in absolute values (left panel) and relative to the reference model ak135 (right panel).*

4.7 Inversion results

The inverted correlation functions for the global data set and one subregion are shown in Figure 4.12. The two other regions give results similar to the ones shown. The corresponding r.m.s. amplitude of mantle structure for all 4 regions is shown in Figure 4.17. The most prominent features in these figures are the rapid decay to small correlation values with increasing distances Δ_v and Δ_h throughout the mantle. This is a surprising result, since tomographic models and the stochastic models of Gudmundsson *et al.* [1990] and Davies *et al.* [1992] find a large increase in correlation length with depth. The variance curves show a pronounced minimum in the lower mantle, which agrees well with the studies mentioned.

Before we discuss more details we want to make some remarks concerning the stability of the results obtained through the inversion. If we do not use any regularization in the inversion, the width of the correlation functions tends towards the smallest possible values allowed by the spline parameterization throughout the mantle. Furthermore the solution oscillates between significant positive and negative values throughout the mantle. The variance is much more stable and is not much influenced by regularization. Since the oscillations are not very realistic a smoothness constraint is applied and the correlation function is damped to zero

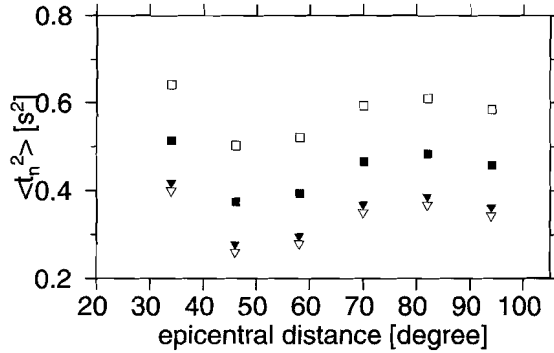


Figure 4.14: Estimates of the incoherent noise component $\langle \delta t_n \rangle$. Symbol legend for the source depth is the same as in Figure 4.5. The dependence on epicentral distance is the same for each depth.

at $\Delta_h = 90^\circ$. No regularization is imposed on the variance and the 9 parameters estimating the incoherent noise. The inversion results in models with correlation half-width of the order of 100 km unless we increase the regularization to values that the data misfit becomes significantly worse. However, the lower mantle correlation values are more poorly constrained by the data. If we force the correlation lengths to larger values at only one depth the misfit increases less rapidly if this depth is in the lower mantle than if it is in the upper mantle. Somewhat larger correlation values in the lower mantle are possible without causing a much worse data misfit. In this case variance in the lower mantle drops even more.

4.7.1 Regional differences

In the inversion of delay times for the covariance function of mantle structure we have assumed that the Earth can be described as a lateral uniform random medium, that is the statistical properties of heterogeneity, such as amplitudes and scale lengths, vary only with depth. The division of the data set according to three different regions (Section 4.2) can be used as a first test for this hypothesis. Such a test cannot be carried out for all scale lengths, because from a statistical point of view the Earth is just one realization and furthermore of finite size. Therefore, it is impossible to test the hypothesis for scale lengths that are not significantly smaller than the examined region. Fortunately, the rapid decay of the correlation value with increasing distance between two stations (Figure 4.10) shows that much of the heterogeneity is of the scale of a few degrees. For these scales the regions are large enough (remember that the area sampled by the rays is larger than the defined regions for source or receivers) to obtain a sampling of many anomalies. Furthermore, the structure function is not only insensitive about temporal mislocation, but at small values of Δ_{12} also against spatial mislocation since both delay

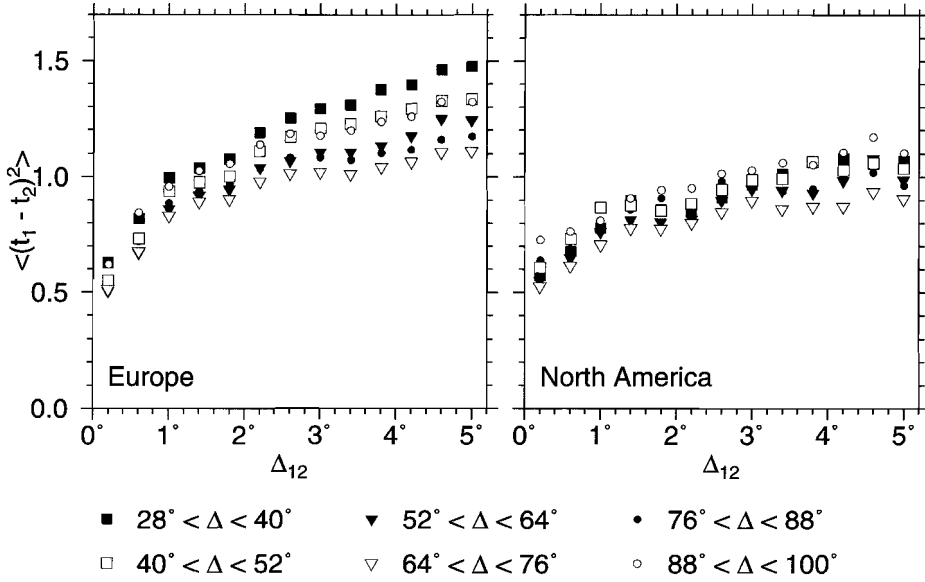


Figure 4.15: *Structure function for small station separation Δ_{12} .*

times are altered in the same way.

Figure 4.15 shows measured values of the structure function for small distances between the two rays. The two panels show the estimates for stations in Europe (and the Mediterranean) and for North American stations. Clearly, the structure function in the regions appears to be very different. Europe shows a much larger increase, especially within the first 1.5° implying more heterogeneity at that scale. Also the spread between short and large epicentral distances Δ is at $\Delta_{12} = 5^\circ$ about twice as large for Europe.

If heterogeneity is not laterally uniform distributed, this casts doubts on the usefulness of comparing power spectra or correlation functions for models that are not sampled in the same way. Power spectra can be determined by a few, strongly heterogeneous regions. A Dirac δ -function and white noise exactly possess the same power spectra, although their other properties are very different. Power spectra do not contain any information in the spatial distribution of anomalies, which is equally important in the case of nonuniform media. Furthermore, if the sampling of nonuniform models is different, there is no reason why the spectra should be the same.

4.7.2 Comparison with tomographic models

For comparison, we calculated the autocorrelation function of mantle heterogeneity for two tomographic models. Figure 4.16 (left column) shows the result for

the S wave tomographic model of Woodhouse and Trampert [1995]. This model does not use any ISC travel times. The use of surface waves gives a good vertical resolution, whereas the lateral resolution is limited to the first 16 degrees of a spherical harmonic expansion. This is directly visible in the autocorrelation function which has a much shorter width in the vertical than in the lateral direction. The three slices in the right column show the autocorrelation function of the P velocity model of Bijwaard *et al.* [1998] based on the same relocated ISC travel times as this study. Both models have a broader correlation function than our inversion results. The r.m.s. amplitudes as a function of depth, shown in Figures 4.17, show a large contrast in amplitude between upper and lower mantle, which is larger than our result. A small increase close to the core-mantle-boundary (D'' layer) is visible in all models.

The difference between the models reflects the fact that inverted models $\delta\tilde{u}$ are not a one-to-one representation of the true Earth δu , but a version filtered by a limited resolution:

$$\delta\tilde{u}(\mathbf{r}) = \int R(\mathbf{r}, \mathbf{r}') \delta u(\mathbf{r}') dV' \quad . \quad (4.24)$$

In the ideal case the resolution kernel $R(\mathbf{r}, \mathbf{r}')$ would be identical to a Dirac δ -function and each point in the model space $\delta\tilde{u}(\mathbf{r})$ would reflect only information of one point of the studied object $\delta u(\mathbf{r})$. But the resolution is limited due to the finite number of observations, measurement errors, nonuniqueness, model parameterization, approximations of the theory used, etc. Therefore, a calculated covariance function C is also filtered by these kernels:

$$\tilde{C}(\mathbf{r}_1, \mathbf{r}_2) = \iint R(\mathbf{r}_1, \mathbf{r}_1') R(\mathbf{r}_2, \mathbf{r}_2') C(\mathbf{r}_1', \mathbf{r}_2') dV_1' dV_2' \quad (4.25)$$

These kernels are different for waveform tomography, travel time tomography and our statistical approach and have to be considered when models are compared with each other or with numerical simulations of mantle convection.

4.8 Discussion

The structure function (eq. (4.15)) is in principle similar to the variance of summary rays used by Gudmundsson *et al.* [1990]. Both measures estimate the variations of structure between the rays and are insensitive to structure sampled by all rays. There are also some differences: (1) Because the pairs of delay times used for the calculation of the structure function stem from the same event the structure function is insensitive towards temporal mislocation. The same holds approximately for the vertical mislocation since we are only considering teleseismic distances; thus all rays leave the source almost in the downwards direction. Horizontal mislocation has an imprint on the structure function only, if the stations are in very different directions (large Δ_{12}). The same arguments hold for the

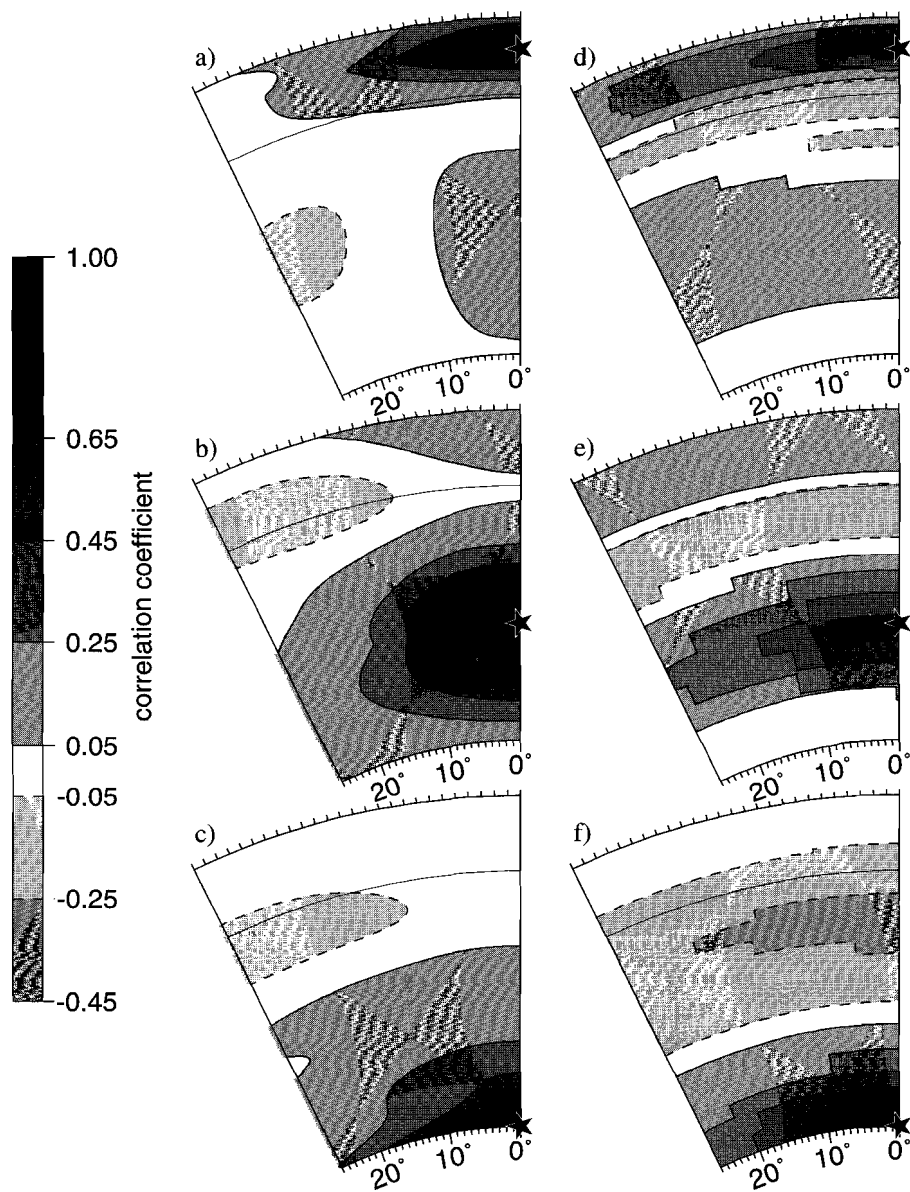


Figure 4.16: *The autocorrelation function for the waveform tomographic model of Woodhouse and Trampert [1995] (left column) and for the travel time tomographic model of Bijwaard et al. [1998] (right column), which is based on the same delay time data set as this study.*

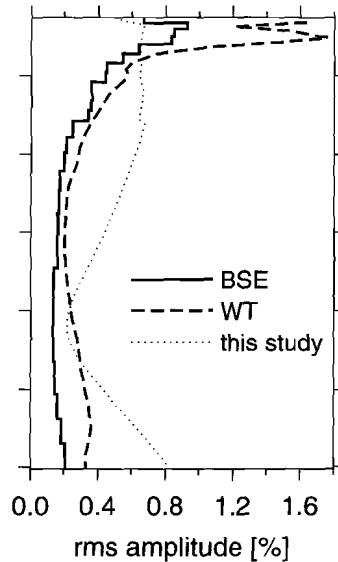


Figure 4.17: *Relative r.m.s. amplitude of the tomographic models of Bijwaard et al. [1998] (BSE) and Woodhouse and Trampert [1995] (WT). The reference models are ak135 and PREM accordingly. In addition, the S wave velocity perturbation of the Woodhouse and Trampert model have been scaled by 0.5 to allow a better comparison to P velocity heterogeneity [Robertson and Woodhouse, 1996]. Correlation functions for the tomographic models are shown in Figure 4.16. The dotted line shows the result of this study for the global data set (taken from Figure).*

variance of summary rays only, if the all hypocenters for each summary ray are affected in a similar way. This is partially true since mislocation is mainly caused by local structure, but it is opposed by the fact that the number of seismological stations has strongly increased and may give systematic differences between early and late events in the ISC data. (2) Gudmundsson *et al.* [1990] had to make the assumption that the rays are approximately parallel and straight lines in order to be able to solve the double integral of eq. (4.6) analytically. With the correlation lengths they obtained one cannot sustain that the rays are straight over a correlation length. (3) To estimate the variance of summary rays a minimum number of rays is needed (Gudmundsson *et al.* [1990] used summary rays consisting of at least 4 individual rays). For the calculation of the structure function any two rays with a common source can be used. This helps to sample a mantle volume as large as possible since also sparse intraplate events can be used which often miss in summary rays of small size. However, we have to use an explicit downweighting for densely sampled paths. Summary rays implicitly avoid bias from inhomogeneous sampling.

If we calculate the variance $\langle \delta t_s^2 \rangle$ for the inverted model and add the random noise $\langle \delta t_n^2 \rangle$ and compare this prediction with the observations of Figure 4.5 we find that our models favor a variance which is significantly larger

$$\underbrace{\langle \delta t_s^2 \rangle + \langle \delta t_n^2 \rangle}_{\text{model prediction}} > \underbrace{\langle \delta t^2 \rangle}_{\text{observation}} = \langle \delta t_s^2 \rangle + \langle \delta t_n^2 \rangle + \langle \delta t_m^2 \rangle + 2\langle \delta t_s \delta t_m \rangle \quad (4.26)$$

This can be reconciled if the last two terms on the right hand side describing the origin time errors give a significant negative contribution (see eq. (4.13)). This implies that origin time errors should not be neglected. Unfortunately, the contribution of spatial mislocations to delay times is likely to be important too. This contribution is not taken into account in our modeling. The spatial mislocation gives a negative contribution to the variance. The argument is the same as for the origin time errors discussed on page 71. To examine the horizontal mislocation in more detail, we split the delay times again in separate contributions. The term δt_h describes the contribution due to a horizontal mislocation and all other parts are represented in $\delta t_x = \delta t_s + \delta t_n + \delta t_m + \delta t_v$ where the last term gives the contribution of a vertical mislocation. With this division the covariance becomes

$$\langle \delta t_1 \delta t_2 \rangle = \langle \delta t_{1x} \delta t_{2x} \rangle + \langle \delta t_{1h} \delta t_{2h} \rangle + \langle \delta t_{1x} \delta t_{2h} \rangle + \langle \delta t_{1h} \delta t_{2x} \rangle \quad (4.27)$$

If the horizontal mislocation was random and not correlated with heterogeneity the last two terms on the right hand side would be zero and $\langle \delta t_{1h} \delta t_{2h} \rangle$ could be calculated from the geometry of the considered ray pair. However, this cannot be true, since the contribution of random, uncorrelated mislocations would be positive to the variance of delay times, which is not allowed according to the above arguments. Nonetheless, we can draw some qualitative conclusions. The term $\langle \delta t_{1h} \delta t_{2h} \rangle$ is positive if the rays leave the source in similar directions and negative for geometries for which the stations are located at opposite sides from the source. This could be an additional explanation (see also page 71), why the slope of the correlation function in Figure 4.10 stays negative even at large Δ_{12} . The alternative explanation in terms of structure would be a very long-wavelength heterogeneity.

Many tomographic models allow a correction term for the hypocentral coordinates that can absorb the effects of mislocations due to an imperfect reference model. If such corrections are not included in travel time tomographic inversions, mislocation effects can be mapped in long-wavelength structure through the trade-off discussed.

The large change of the structure functions at distances between the stations Δ_{12} of less than 1° in Figure 4.15 implies that there is also significant heterogeneity at scales of less than 100 km (except if all of that increase in the structure function is caused by noise that becomes correlated at such scales, which is unlikely). At scales smaller than 100 km scattering effects and other wave phenomena become increasingly important and it is questionable whether ray theory is a good approximation of wave propagation in such a medium.

4.9 Conclusions

Although we have not been able to obtain reliable estimates of the scale lengths of mantle P wave velocity heterogeneity, several important conclusions can be drawn, either directly from the data or from the more robust features of the inversion.

(1) Noise is not uniformly distributed among P wave arrival times in the ISC Bulletin. We verified the idea that the distribution of delay times can best be explained by the sum of two distributions (Jeffrey distribution): a normal distribution with a standard deviation of 1 s^2 containing correct picks with small measurement errors and a much broader one. Most of the delay times of the second distribution show no correlation with other residuals. The positive tails contain probably many picks of depth phases and PcP phases that have been mistakenly identified as P . (2) Although the number of delay times in the tails of the distribution is small relative to the total number they exert a strong influence in moments of second and higher order. Therefore, some results critically depend on the chosen cut-off value. We find, for instance, the highest correlation between teleseismic P wave delay times of similar rays, if all data with an absolute delay larger than 2.2 s are excluded. (3) Furthermore, delay times are significantly influenced by mislocation effects. Mislocation probably poses the most serious limitation to the estimation of statistical properties of mantle structure from the properties of delay times, since the mislocation is correlated with the structural signal in the delay times and hence the estimation of statistical properties of mantle structure from the properties of delay times is affected by unknown mislocation errors. (4) A significant loss of correlation between delay times of two stations separated by only a few tens of km implies the presence of heterogeneity with scales of similar size. The exact amount of structure at such scales is hard to assess since stations are more and more likely to be of the same network or array, in which case it becomes more probable that also part of the noise becomes correlated. This still leaves the question open of what is the strength and distribution of length-scales in the Earth. (5) The delay times of rays that sample different regions of the Earth possess different statistical properties. The Earth can therefore not adequately be described as a laterally uniform random medium and it is questionable whether correlation lengths and power spectra are useful tools to compare models sampled in different ways.

Chapter 5

Thermal structure of continental upper mantle inferred from S wave velocity and surface heat flow

Abstract. Results from seismic tomography provide information on the thermal structure of the continental upper mantle. This is borne out by the good agreement between tectonic age, surface heat flow and a tomographic S wave velocity model for depths less than 180 km. The velocity anomalies of tomographic layers deeper than 230 km have relatively small amplitudes and show little correlation with surface heat flow or shallow velocities. We associate the drop in correlation and amplitude of the velocity perturbations between 180 and 230 km depth with the maximum thickness of the thermal boundary layer (TBL), in which larger variations in temperature and possibly composition than in the underlying convecting mantle can be sustained.

Velocity profiles for different tectonic provinces are converted to temperature using mineralogical data. Both anharmonic and anelastic effects on the wave speeds are taken into account. The resulting geotherms differ most at depths of 60 - 120 km with variations of up to 900°C. Below 230 km differences do not exceed 300°C. These geotherms agree well with one-dimensional conductive geotherms for the observed range of continental heat flow values using the empirical relationship that 40% of the surface heat flux stems from upper crustal radiogenic heat

This chapter has been submitted by Röhms, Snieder, Goes and Trampert to *Earth Planet. Sci. Lett.*

production.

The S wave velocity in the continental upper mantle appears to be adequately explained (within the uncertainties of the tomography and the conversion to temperature) by a thermal signature. A compositional component can, however, not be ruled out as it may have only a minor effect on the velocity and the heat flow.

The surface heat flow is controlled by the shallow heat production and the thickness of the TBL. Seismology helps to determine the relative importance of the two factors and our results confirm the similar importance of both factors previously inferred from observations of surface heat flow and upper crustal heat production. Variations of TBL thickness could be controlled by compositional differences or by varying blanketing effects of the crust due to crustal radiogenic heat production.

5.1 Introduction

Oceanic lithosphere is continuously recycled by mantle convection. It is formed at mid oceanic ridges, thickens when it cools and finally subducts at plate boundaries. In contrast, continental crust does not participate in recycling. Old continental regions are many times older than the oldest existing ocean floor. They contain Archean nuclei around which younger material is accreted. Several basic questions about the continents are not yet fully answered, e.g. what is the growth rate as a function of geological time (see Windley [1995] for a compilation of different models). The mantle part of the continental lithosphere raises some additional questions. Is it always of the same age as the continental crust and, if so, what causes this long-term stability? What is its present-day thickness? Answers to these questions depend on the adopted definition of lithosphere, which can be based on very different quantities (see Anderson [1990], Schmeling and Marquart [1991], Anderson [1995] and references herein). Definitions which can be found in the literature refer to the mechanical, elastic or flexural lithosphere, which can support stresses over long time scales, the chemical lithosphere, which is an isolated reservoir separated from mantle convection, the electrical lithosphere, which has a low electrical conductivity relative to the partially molten asthenosphere [e.g. Jones, 1982], the thermal lithosphere, which is equivalent to the thermal boundary layer (TBL) and the seismic lithosphere, which is either defined as the lid above the seismic low-velocity zone [e.g. Anderson, 1990] or as the region with higher than average velocities [e.g. Ricard *et al.*, 1996]. The thicknesses associated with the various definitions can be significantly different, e.g. the lower part of the thermal lithosphere (TBL) is probably very weak and is not part of the mechanical lithosphere [Anderson, 1995].

This study is concerned with the seismic upper mantle structure (down to the transition zone beginning at about 400 km depth) and the corresponding thermal structure. Since the low-velocity zone is often absent in stable continental regions [Nataf *et al.*, 1986, Lerner-Lam and Jordan 1987] and since we consider smooth

perturbations from a seismic reference model, we define the seismic lithosphere in this chapter as the region where the seismic velocity perturbations correlate with tectonic provinces and the surface heat flow. This definition includes a larger part of the mantle than, for example, the mechanical lithosphere as will become clear later when the seismic lithosphere is related to the thermal boundary layer (TBL). The TBL is viewed as the region with predominantly conductive heat transfer.

Many different values for the thickness of the seismic continental keels are reported in the literature, ranging from several tens of kilometers beneath tectonically active regions to a few hundred kilometers below ancient cratons. There is no consensus on the depth extent of continental high velocity roots and estimates range from less than 200 km to more than 400 km [e.g. Jordan, 1975; Anderson, 1990; Polet and Anderson, 1995].

Variations in continental lithospheric thickness (all definitions) can be caused by several mechanisms which have distinct spatio-temporal signatures that can be used to distinguish between them.

- (1) Plate tectonic events, such as continental collision or rifting, thicken or thin the lithosphere accompanied by thermal disturbances and possibly also by accretion or delamination of a chemically distinct layer. This causes variations mainly at active plate boundaries. The thermal anomalies decrease with time after the last tectonic event. Schmeling and Marquart [1991] find the thermal decay times of lateral TBL thickness undulations to be between 15 and 235 Ma if their mean depths are between 50 and 200 km. Vitorello and Pollack [1980] estimate the contribution of the last tectonothermal event to the surface heat flow. This contribution halves approximately every 200 Ma [see their Figure 7]. Hence, tectonic events have a significant influence for several hundred million years, but old stable regions should be close to thermal equilibrium.
- (2) Mantle convection can cause undulations of the lithosphere [Schmeling and Marquart, 1991]. Downwellings go along with a thick lithosphere whereas upwellings make the lithosphere thinner. Convection cell patterns typically change within transition periods of 100 to 200 Ma but can remain stable for much longer periods of time. It is, however, possible that continental roots and underlying mantle convection are coupled [Gurnis, 1988]. In this case the relative movement (plate vs. convection pattern) might be significantly slower than the absolute movement of either the plate or the convection pattern relative to a fixed reference frame (e.g. hotspots).
- (3) Erosion has a twofold effect on the heat flow. First, it is accompanied by uplift which brings deeper isotherms closer to the surface. For high erosion rates of tectonically active regions with pronounced topography this might be responsible for up to 50 % of the surface heat flow [see Vitorello and Pollack, 1980, and references herein]. Second, it removes part of the radiogenic crust, thus permanently decreasing the surface heat flux and the shallow

temperature gradient. This change in the crustal blanketing effect may also influence the thickness of the TBL. This issue is addressed below in more detail.

- (4) A compositional root of depleted mantle material can originate through extraction of partial melt [Jordan, 1988; de Smet *et al.*, 1998]. The reduced weight of depleted material presents an explanation for the proposed stability of the continental lithosphere. Partial melting needed for this mechanism consumes latent heat and thus accelerates the formation of a cold TBL. This was especially important in the early stages of continental evolution.

The temporal behavior is one difference between these mechanisms. A second one is the original depth of the disturbances which is below the mechanical lithosphere for (1), below the TBL for (2), close to the surface for (3), or mainly in the lower part of the TBL for (4). Of course, these explanations do not exclude each other and a combination of the different sources is possible. From this, it should be clear that young continental regions can have a wide diversity of thermal structures, while old continental regions are expected to show less thermal variation.

Several studies have used seismic models to infer the thermal structure. Yan *et al.* [1989] have used the S wave velocity perturbations at 150 km depth to predict the global surface heat flow and found a reasonable agreement with observations. An alternative explanation through strong compositional differences (iron content) is considered by them to be less probable. Sobolev *et al.* [1996] converted P wave velocity anomalies of a teleseismic tomographic model of the French Massif Central to a temperature field using data from xenoliths to constrain the composition. Goes *et al.* [1999] used both P and S wave velocities to calculate the temperatures under Europe. The latter two studies both find that the expected influence of composition is much smaller than the temperature effect. Nataf and Ricard [1996] took the opposite approach and calculated synthetic seismic velocities for a thermal model based on tectonic constraints and have found a good agreement with seismic data [Ricard *et al.*, 1996].

In this chapter, we first show the good accordance between the tomographic model of Woodhouse and Trampert [1995], the compilation of surface heat flow measurements of Pollack *et al.* [1993] and the continental regionalization of Nataf and Ricard [1996]. Encouraged by the high correlation of S wave velocities and heat flow, seismic velocity profiles for different tectonic provinces are converted to temperature. The resulting geotherms span a wide range of temperatures for the upper 200 km which agrees well with conductive geotherms for typical values of continental surface heat flow. Our results do not support the hypothesis that the convection pattern dictates the thermal structure of the TBL. On the contrary, they point to a top-down cause-effect relation in which variations in the shallow radiogenic heat production produce the large-scale velocity anomalies in the shallow mantle. Beneath old cratons, where erosion has removed most of the upper

crustal radiogenic elements, a thick TBL, associated with fast velocity anomalies, has formed.

5.2 Seismic tomography and tectonic regionalization

Surface wave dispersion measurements and seismic tomography have revealed the correlation between tectonic setting and lithospheric seismic velocities, with old cratons showing the highest, and tectonically active regions having relatively low velocities [e.g. Nakanishi and Anderson, 1982; Woodhouse and Dziewonski, 1984; Nataf *et al.* 1986]. In this study the tomographic model of Woodhouse and Trampert [1995] is used to analyze the continental upper mantle structure above 400 km depth. This model is constructed from fundamental mode surface wave phase velocity measurements and additional waveform fitting of whole seismograms to improve the vertical resolution at greater depths. The model is restricted to the first sixteen degrees of the spherical harmonic representation which limits the lateral resolution.

We did not use a model obtained from travel time tomography because waveform tomographic models possess (1) a superior vertical resolution in the upper mantle and (2) a more homogeneous lateral resolution. Travel time tomographic models have a much higher lateral resolution, mostly in tectonic regions with a dense coverage of stations or earthquake hypocenters. In other areas, however, specifically in many stable continental regions, which are the regions of our main interest, they provide little information about the shallow mantle. Additional constraints on *P* wave velocities would be helpful for our investigation [Goes *et al.*, 1999]. However the only *P* wave models available are based on travel times.

To investigate the correlation between tectonic region and deeper mantle structure the tomographic model is compared with the regionalization of model 3SMAC [Nataf and Ricard, 1996], which divides continental areas into three different tectonic types: Archean cratons, stable platforms and tectonic continents (Figure 5.1). The resolution of the tomographic model does not justify a division into more types which would yield smaller provinces. The $2^\circ \times 2^\circ$ lateral discretization of the model is fine enough for the comparison with the tomographic model.

We calculated an average velocity profile for each region from the tomographic model of Woodhouse and Trampert [1995]. Since the tomographic model has a finite lateral resolution, only the area of each province, which is located at a distance of at least 5° from the boundary to another province is used. These areas are shown in Figure 5.1 in gray shades. Figure 5.2 shows the corresponding velocity profiles. The strongest feature is the decrease in variation of the seismic velocity from more than 10% at shallow depths to less than 4% below a depth of 230 km. All 7 Archean profiles (no. 1–7 in Figure 5.1) from 5 different continents show a striking similarity. In the upper 200 km they are clearly separated from the tectonic regions which show a much larger diversity. The profiles for stable platforms span

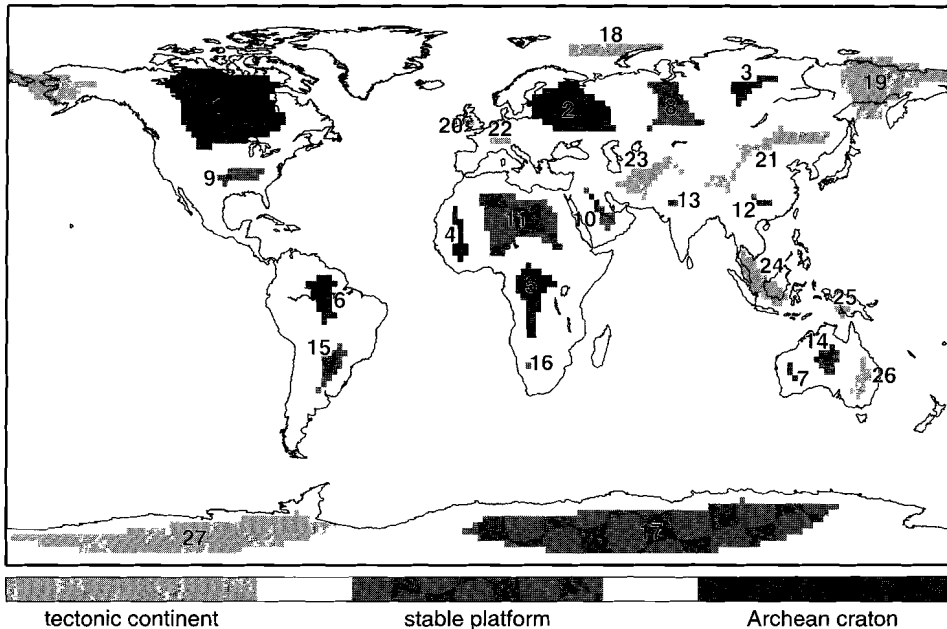


Figure 5.1: *Continental tectonic provinces after the 3SMAC regionalization [Nataf and Ricard, 1996] marked by the gray shades. Shown is only the inner part of each province that is at least 5° away from any other province.*

almost the complete range from the fast profiles of Archean cratons to the slow profiles of tectonic continents. Closer inspection of the stable platform provinces reveals possible causes for this broad range. The three fastest profiles, which show the same velocities as the Archean provinces, are no. 9 (Eastern USA), which also shows a lower median surface heat flow of 47mW/m^2 compared to other provinces of the same type (Table 5.1), no. 14 (Central Australia), which forms in the tomographic model one fast anomaly together with the Archean region no. 7, and no. 17 (Eastern Antarctica), which is a Precambrian shield. On the other side are the three slowest profiles of stable platforms: no. 10 (Arabian peninsula), which is surrounded at all sides by active plate boundaries, no.11 (Northern Africa), which contains the Hoggar and Tibesti hotspots [Richards *et al.*, 1988; Yamaji, 1992], and no. 12 (China), which is in active extension [Gilder *et al.*, 1991]. We find from the tomographic model (Figure 5.2) that the thickness of the seismic lithosphere, defined as the upper mantle region which shows strong variations of seismic velocity related to surface tectonics, is slightly more than 200 km.

The average profile of all continental regions with a distance of at least 5° from oceanic lithosphere, shown as the thick gray line in Figure 5.2, is positive at all depths down to 400 km. Since the global average is subtracted this implies that

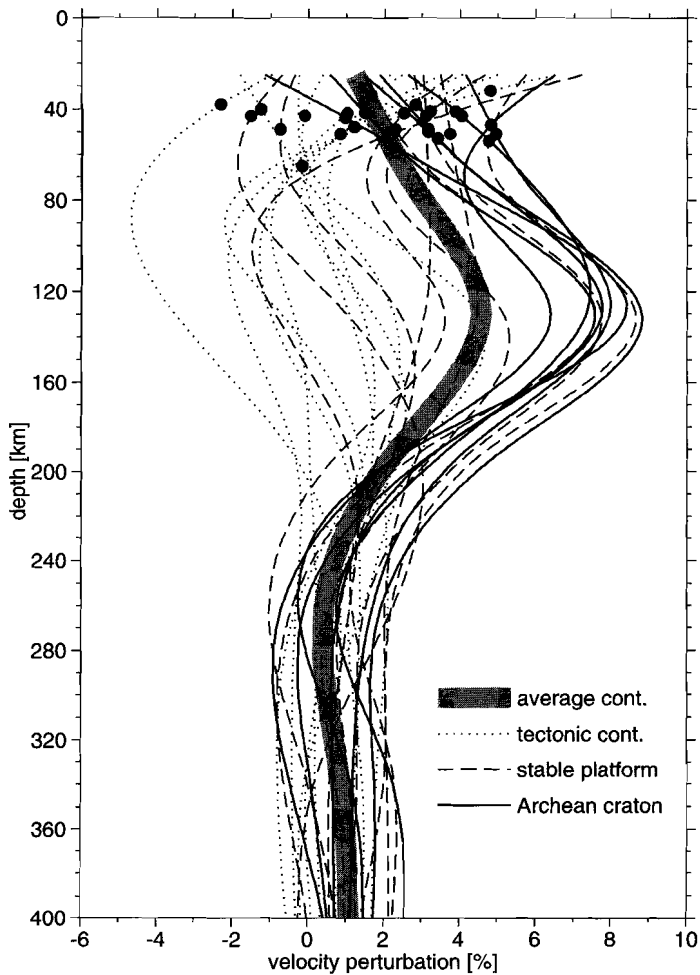


Figure 5.2: Velocity profiles are obtained by averaging the velocity perturbations of the global tomographic model of Woodhouse and Trampert [1995] over the interior regions of the provinces from the 3SMAC continental regionalization shown in Figure 5.1. The dots denote the maximum crustal thickness for each province which are used for the crustal correction of the tomography. The global average velocity perturbation (degree 0 term of the spherical harmonic expansion) for each depth is subtracted. Different line textures denote the three types of continental provinces. The very thick gray line displays the mean profile of all continental regions with a distance of at least 5° from oceanic lithosphere.

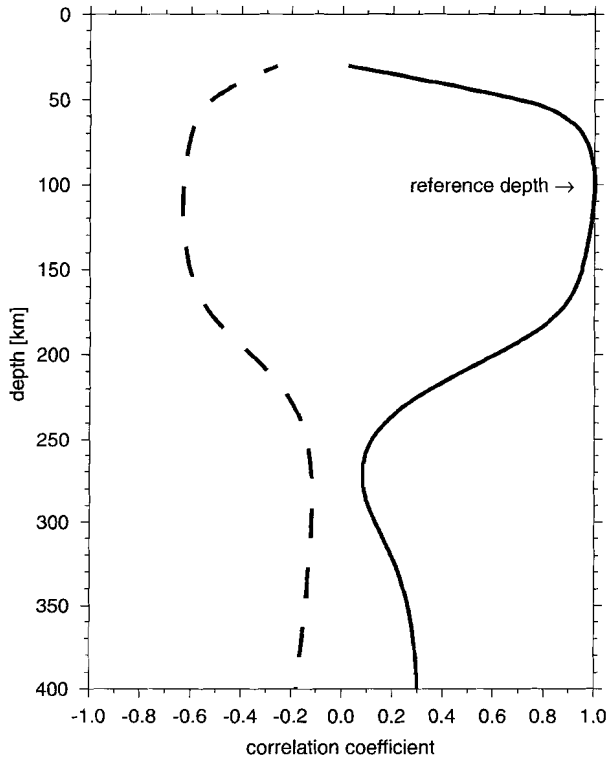


Figure 5.3: Correlation coefficient between S wave velocity at the depth given on the ordinate and the S wave velocity at 100 km depth (solid line) and averaged surface heat flow (dashed line). Only continental areas with a distance of at least 5° from oceanic crust are included in the calculation. The averaging of heat flow values is explained in the text.

the S wave velocity below oceans is on average slower than the subcontinental velocity. Relative to the global average, i.e. zero in Figure 5.2 for all depths, some provinces have fast velocity perturbations of up to about 2% throughout the upper mantle. This can partly explain why other studies [e.g. Jordan, 1975; Polet and Anderson, 1995] found much deeper roots than the 200 km found in this study. However, this type of anomaly is not correlated with age (Figure 5.2). Partly, the difference probably also originates from the improved vertical resolution of the tomographic model of Woodhouse and Trampert [1995] as compared to older models.

The difference between profiles above and below a depth of about 200 km are even more striking in the correlation of the velocity perturbations at two different depths. Figure 5.3 shows the correlation of the velocity field of continental area

(with a distance of at least 5° from the oceans) between a depth of 100 km and a second depth indicated by the vertical axis. The correlation profile can be divided into two regions. Between 60 and 180 km depth the correlation coefficient is higher than 0.8. Below 230 km depth the correlation drops to values between 0.1 and 0.3. The transition from high to low correlations marks the depth of the seismic lithosphere, which we have defined as the region where the velocity correlates with surface tectonics. The low values of the correlation coefficient at depths less than 60 km is caused by crustal effects since the crustal correction is not made in the calculation for Figure 5.3. The very small correlation between the seismic lithosphere and the underlying structure suggests that the strong shallow anomalies are not controlled from below by the convection but have a shallow origin, i.e. within the upper 200 km.

5.3 Surface heat flow

Surface heat flow measurements form another, completely independent, data source giving valuable constraints on the thermal structure of the crust and shallow mantle. Many studies [e.g. Chapman and Pollack, 1975; Sclater *et al.*, 1980; Sclater *et al.*, 1981; Pollack *et al.*, 1993] have found a systematic decrease of surface heat flow with the age of tectonic provinces. For this study the global heat flow compilation of Pollack *et al.* [1993] is used.

Three points concerning the heat flow data have to be considered in the comparison: (1) there is a considerable amount of scatter in the measurements due to environmental problems, e.g. water circulation and past climatic changes, and difficulties with early conductivity determinations [Sclater *et al.*, 1980], (2) the surface heat flow can have large changes over short distances due to the local geology and topography, and (3) the measurements are highly unevenly distributed over the Earth's surface. A simple pointwise comparison for all measurement sites of surface heat flow and volume averaged quantities, such as seismic velocities revealed by waveform tomography, is complicated by each of the three points listed above. Appropriate averaging methods have to be used to construct average heat flow values for larger areas that can be used for the comparison. In order to suppress any overweighting from clustered measurement points we first averaged all heat flow measurements within $20 \text{ km} \times 20 \text{ km}$ cells.

Figure 5.4 shows histograms of heat flow values for the three different types of continental provinces and for all continental values. Just as for the velocity profiles of Figure 5.1 only measurements with a distance of at least 5° from the tectonic boundaries are used. Median values and absolute deviation about the median are listed in Table 5.1.

For comparison with the tomographic model the heat flow values were averaged a second time with the aim to estimate the average heat flow of an area similar in size to the resolution of the velocity model. For this a heat flow value was assigned to each node of a $2^\circ \times 2^\circ$ coordinate grid if at least one $20 \text{ km} \times 20 \text{ km}$

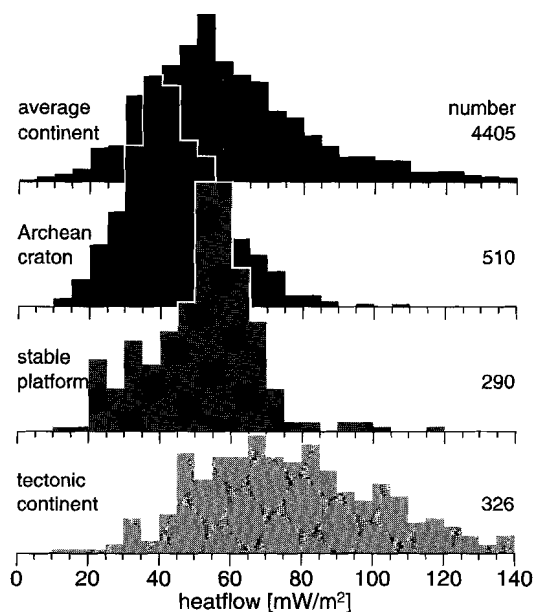


Figure 5.4: Histograms for continental heat flow measurements and HF measurements for the interiors of the three different type of tectonic provinces. Histograms are normalized; the number of measurements is given on the right side of each histogram.

average value was within 110 km and 3 values were within 550 km distance of the node. All nodes on continents that were more than 5° away from the oceans and that received an averaged heat flow estimate were used in the subsequent calculation of the correlation coefficient between surface heat flow and S wave velocity perturbation as a function of depth. The result is shown in Figure 5.3 as the dashed line.

	median [mW/m ²]	99% conf. interval [mW/m ²]	mean abs. deviation [mW/m ²]
continental area	55	54 - 56	26.1
Archean craton	42	40 - 44	10.8
stable platform	54	51 - 56	11.3
tectonic cont.	76	70 - 81	24.3

Table 5.1: Median values for surface heat flow of different provinces, 99% confidence intervals around the median and the mean absolute deviations about the median.

This correlation function has almost the same shape as the correlation function of two different depth layers of the tomographic model but has the opposite sign and a lower amplitude. The negative sign is due to the anticorrelation between temperature and seismic velocities. Considering the scatter in the heat flow measurements and a nonlinear relation between temperature and S wave velocity, the anticorrelation is remarkably large. The strong anticorrelation between heat flow and velocity between 60 km and 180 km depth suggests that the velocity perturbations are predominantly caused by temperature effects. The small correlation coefficients above 60 km depth are again caused by crustal effects. Yan *et al.* [1989] have used S wave velocities at 150 km depth to predict global heat flow and found a qualitative agreement to the observed global heat flow pattern for the lowest 6 degrees of a spherical harmonic expansion. The strong anticorrelation in Figure 5.3 shows that this is also valid for continental regions only and at smaller scales.

5.4 Constraints on temperatures from S wave velocity

The strong anticorrelation between heat flow and seismic velocity favors a thermal interpretation of the tomographic model. Other studies [e.g. Jordan, 1979; Jackson and Rigden, 1998; Goes *et al.*, 1999] also confirm that the expected effect of reasonable variations in upper mantle composition as found in xenoliths is of minor importance compared to the expected temperature effect and is probably of similar magnitude as the uncertainties in the tomography. In our subsequent analysis we concentrate on the mantle below 60 km depth to avoid complications due to the presence of crustal material and give an interpretation in terms of temperatures alone.

Following the procedure developed by Goes *et al.* [1999] the conversion of S wave velocity to temperature is carried out by using laboratory measurements of density and the elastic moduli for various mantle minerals. These values are extrapolated to mantle temperature and pressures under the infinitesimal strain approximation. (Finite strain theory would be more correct, but a comparison has shown that the error introduced is relatively small for our depth range of interest [Vacher, personal communication]). The elastic moduli for an average mantle composition are determined by using the Voigt-Reuss-Hill averaging method. Anelasticity, which is an important effect for finite frequencies [Karato, 1993] but unfortunately less well constrained by experimental data, is also taken into account. The importance of anelasticity increases with temperature and does not allow the application of a linear relation between velocity and temperature anomalies. For further details, including the mineralogical data used, we refer to Goes *et al.* [1999].

Due to the nonlinear relation between velocity and temperature the conversion has to be done using absolute velocities rather than velocity perturbations.

This causes an extra difficulty, since the reference model used for the tomography (PREM; Dziewonski and Anderson 1981) incorporates an anisotropic layer down to the Lehmann discontinuity at 220 km depth. The equivalent isotropic PREM model also contains this discontinuity in form of a 5 percent step of S wave velocity. Many different explanations have been proposed for the Lehmann discontinuity, such as a chemical boundary or the base of a partially molten layer [see Karato, 1992 for a brief summary and references]. An anisotropic layer above the discontinuity is confirmed by several seismological studies [e.g. Nataf et al., 1986; Montagner and Tanimoto, 1991]. Karato [1992] proposed that a change from dislocation creep to diffusion creep causes the discontinuity. Dislocation creep aligns the olivine crystals, which have highly anisotropic elastic properties, whereas diffusion creep does not. The global occurrence of the discontinuity is questioned and little seismological evidence exists that the corresponding isotropic velocities are discontinuous at this depth, e.g. the global waveform stacks of Shearer [1990] show the 410, 520 and 660 km discontinuities but not the Lehmann discontinuity.

It would be best to include this discontinuity in our modeling, e.g. through a discontinuous composition or rheology. However, the lack of a clear thermodynamic understanding of it makes this approach impossible. We have decided, as a second best approach, to slightly modify the seismic reference model. If the changes of the reference model are small the phase velocities and the sensitivity kernels of the surface waves remain similar and a tomographic inversion will yield only slightly different velocity perturbations. We verified that these changes are much smaller than the velocity perturbations displayed in Figure 5.2. Using PREM as reference model, the absolute S wave velocity V for a profile i is given by

$$V_i = V_{PREM} + \delta V_{global} + \delta V_i$$

where V_{PREM} is the PREM S wave velocity, δV_{global} is the global mean velocity perturbation from PREM found in the inversion (i.e. the degree zero term of the spherical harmonic expansion) and δV_i is the velocity perturbation as displayed in Figure 5.2. Using the absolute velocities V_i would yield discontinuous temperatures due to the discontinuity in V_{PREM} . Therefore, we have chosen to introduce a new, thermally based, reference model. Since all Archean profiles are very similar (Figure 5.2) and also Archean heat flow values show the least scatter (Table 5.1) an obvious choice is to specify a geotherm representative for Archean cratons. This geotherm is calculated using the method of Chapman [1986] for the median Archean heat flow of 42 mW/m². Below the intersection of the conductive geotherm with the adiabat (potential temperature of 1200°C) at about around 200 km depth, temperatures along the adiabat are taken for the geotherm. This gives new absolute velocity profiles V'_i :

$$V'_i = V(T_{Archean}) - \delta V_{Archean} + \delta V_i.$$

$V(T_{Archean})$ is the velocity profile calculated from the Archean reference geotherm and $\delta V_{Archean}$ the mean Archean velocity perturbation (weighted by the area of

each province). The relatively cold adiabat was chosen for two reasons. First, as can be seen from Figure 5.2 continental regions at depths below 200 km are approximately 1% faster than the global average and Archean cratons are on average slightly faster than average continental regions (small positive correlation coefficient in Figure 5.3). Therefore, if no lateral variation in composition is assumed temperatures below Archean provinces are somewhat lower than the global average. The second and more important reason is, that for a colder adiabat the new reference model ($V(T_{Archean}) - \delta V_{Archean} - \delta V_{global}$) is closer to the reference model of the tomography (V_{PREM}). For each profile δV_i in Figure 5.2 the absolute velocity profile V'_i was calculated and converted to temperature. The result is shown in Figure 5.5.

To demonstrate the influence of different compositions and the anelastic effect we selected three velocity profiles for provinces 1, 8 and 19 (see Figure 5.1), one of each tectonic type. The corresponding temperatures were calculated with respect to two different compositions and two Q models. One composition represents an average continental lherzolite based on xenolith data which is depleted relative to a primitive mantle, the other a primitive garnet lherzolite [Jordan, 1979]. The compositions are described in Table 5.2. The Q models represent an average model (Q_1 of Goes *et al.* [1999]; after Sobolev *et al.* [1996, their model 2]) calibrated to seismic data and a fully experimental Q model which is a relatively strong estimate of temperature dependence (Q_2 of Goes *et al.* [1999]; after Berckhemer *et al.* [1982]). Figure 5.6 shows the resulting geotherms. The solid lines show the three profiles for the preferred composition (average continental garnet lherzolite, Q_1), also used for Figure 5.5. A change of composition to a primitive mantle (dotted line, same Q model) shows only negligible effects. A change to the anelastic model Q_2 (long dashed line) has also only small effects, except for the very high temperatures of the tectonic profile around 100 km depth. The very slow velocities of tectonic provinces in this depth region give temperatures which are about 100°C lower for the stronger anelastic effect of model Q_2 than the temperatures computed for model Q_1 . It should be noted that the Archean profiles show almost no difference because the profiles are so close to the fixed new reference model. The short dashed line illustrates a shift of the reference model; the temperatures of the reference Archean geotherm, which has been converted to the reference velocity profile, were increased by 100°C everywhere. This results in a shift of the Archean profile by the same amount. The other profiles, especially of tectonic provinces at shallow depths, show a somewhat lower increase due to the anelastic effect. This shows that a wrong estimate of the average continental composition has only little effect on the relative position of the computed geotherms, whereas a wrong thermal reference model can shift all geotherms systematically to higher or lower temperatures. Of course, we have not yet discussed the effect of a laterally varying composition. This could shift some of the geotherms by amounts comparable to the variation of the reference model (by 100°C), which is much less than the total observed variation.

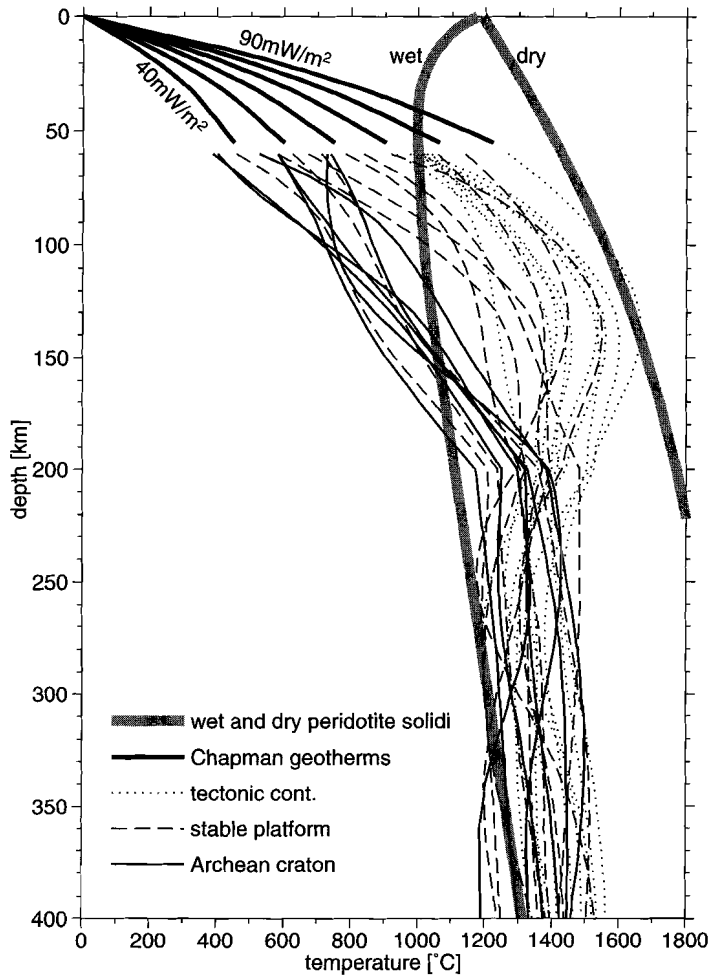


Figure 5.5: Temperature profiles calculated for the velocity profiles from Figure 5.2. In the upper 55 km geotherms after Chapman [1986] for surface heat flows of 40 to 90 mW/m² are shown. The thick gray lines show a wet and dry peridotite solidus taken from Thompson [1992]

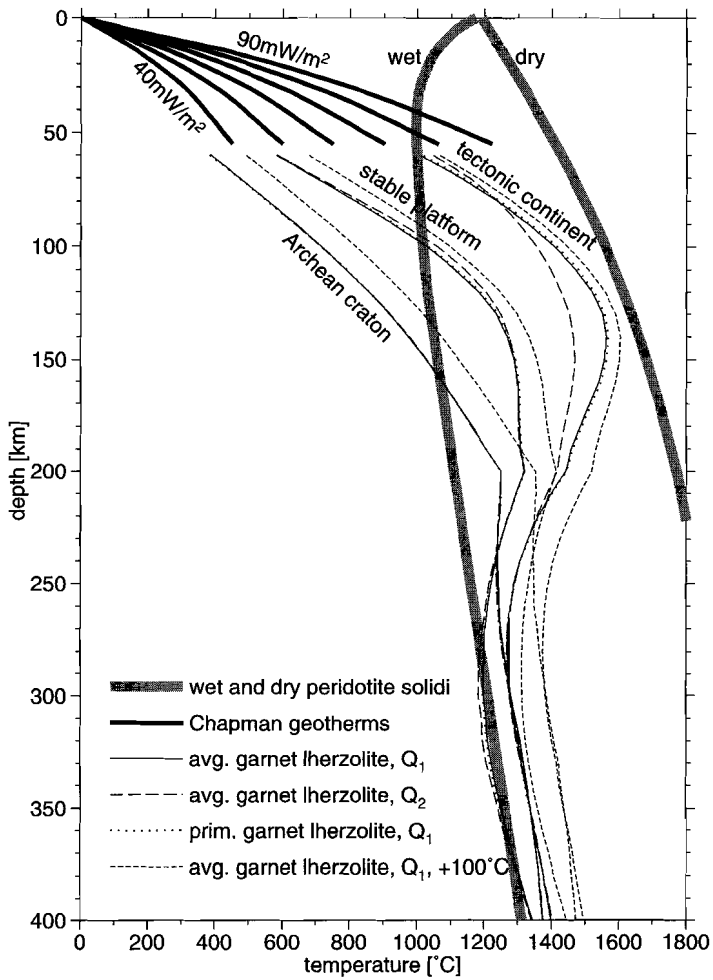


Figure 5.6: The three profiles for regions no. 1, 8 and 19 of Figure 5.1 converted to temperatures for different compositions, anelastic effects and a change of reference geotherm. Geotherms in upper 55 km and solidi are the same as in Figure 5.5.

garnet herzolite	olivine Mg_2SiO_4	orthopyroxene MgSiO_3	clinopyroxene $\text{CaMgSi}_2\text{O}_6$	garnet $\text{Mg}_3\text{Al}_2\text{Si}_3\text{O}_{12}$
average	0.65	0.28	0.03	0.04
	volume fraction of Fe for all minerals: 0.09			
primitive	0.58	0.18	0.10	0.14
	volume fraction of Fe for all minerals: 0.11			

Table 5.2: *Volume fraction of minerals in the two mantle rocks tested in Figure 5.6.*

5.5 Comparison with heat flow derived geotherms

The conversion of velocity to temperature is made under the assumption that no melt is present, which is questionable for tectonic continents. If the mantle is partially molten the calculated temperatures are overestimated. For an upward heat flow temperatures must increase monotonically with depth, but some of the relatively hot geotherms in Figure 5.5 for stable platforms and tectonic continents show a maximum at a depth of about 120 km. This discrepancy can be explained by the presence of partial melt. Since the maxima are close to the dry solidus this is a likely explanation. About 4% of melt would be needed to obtain geotherms which are about 200°C colder around the maxima. However, Figure 5.6 demonstrates that also a stronger anelastic effect, such as in model Q_2 , is able to decrease the temperatures significantly in this temperature-depth region, decreasing the need for partial melt.

Kinks can be seen in the temperature profiles of Figure 5.5 and 5.6 at 200 km depth which are artifacts that originate from the abrupt transition from a conductive geotherm to an adiabat in the reference profile. In reality, the geotherms are expected to be smooth across this transition. Some of the details of the calculated geotherms could also point to compositional variations, but the uncertainties of individual profiles are too large to draw such conclusions. Therefore, we concentrate on the gross pattern in Figure 5.5 and point out two robust features. First, there is a large range of temperatures at shallow depths of about 900°. Second, tectonic regions reach the adiabat at depths shallower than 80 km as opposed to Archean regions which reach the adiabat at depths of around 200 km. The latter is, of course, prescribed by the thermal reference model but this depth is required to explain the seismic anomalies of Figure 5.2.

Figure 5.7 shows the geotherms for surface heat flow values every 10 mW/m² between 40 and 90 mW/m² for three different one-dimensional thermal models which differ only in the distribution of radiogenic elements. The frame in the center shows geotherms as proposed by Chapman [1986] and one of these was used as a reference for the calculations of the profiles shown in Figure 5.5. These geotherms for steady state conductive heat transfer use the surface heat flow as the only

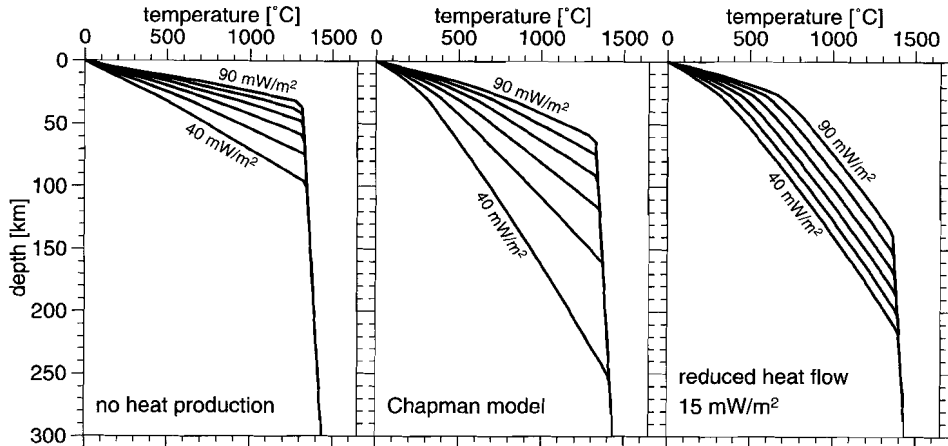


Figure 5.7: Three different geotherm families for surface heat flows of 90, 80, 70, 60, 50 and 40 mW/m^2 (top down). All geotherms are truncated by the 1300°C adiabat. The left panel displays geotherms for a crust without any heat production. The middle panel shows geotherms proposed by Chapman [1986]. Geotherms in the right panel have all the same reduced mantle heat flow of 15 mW/m^2 .

parameter to which the heat production is linked by the relationship that 40% is attributed to upper crustal radiogenic heat sources and 60% to deeper sources (the total crustal contribution is 57% and 44% for a surface heat flow of 40 mW/m^2 and 90 mW/m^2 respectively). Such a linear relation between surface heat flow and shallow heat production is approximately observed for several regions [e.g. Roy *et al.*, 1968; Pollack and Chapman, 1977]. It led to the view that heat production decreases with depth, which is incorporated in the Chapman model by an exponential decrease in the upper crust. Although the exact distribution of radiogenic elements is unknown, a decrease with depth is required. Otherwise the crustal heat production would account for the complete surface heat flow or even exceed it [Rudnick and Fountain, 1995]. The empirical 40-60% partition [Pollack and Chapman, 1977; Vitorello and Pollack, 1980] has been questioned by Pujol *et al.* [1985] for statistical reasons. Especially for stable regions the shallow radiogenic contribution could be larger. A compilation of Rudnick and Fountain [1995] also indicates that for Archean regions the shallow heat production contributes probably more than 50% to the total heat flow. However, the Chapman [1986] model with the 40-60% partition appears to give a good first order estimate of the possible range of continental geotherms. Their variation in shallow temperatures and their range in thickness of the conductive layer agree well with the seismically determined geotherms. The Chapman [1986] geotherms diverge in the crust and shallow mantle and predict temperature differences of more than 500°C below the crust. This large variation is also found in geothermobarometric data obtained of

xenoliths from different regions [cf. O'Reilly and Griffin, 1985; Bertrand *et al.*, 1986, Rudnick *et al.*, 1998]. At high temperatures in greater depths heat transfer is dominated by advection which is incorporated by truncating the conductive geotherms with an adiabat with a potential temperature of 1300°C. This value was also adopted from Chapman [1986] but is not very well constrained. Other proposed values of the average potential temperature are e.g. 1400°C by Anderson and Bass [1984] or 1280°C by McKenzie and Bickle [1988]. These uncertainties in the thermal properties could cause a systematic shift of all geotherms of Figure 5.5 by as much as 100-200°C.

Next to the middle panel of Figure 5.7 two extreme end members of possible geotherm families are shown. They are calculated for the same parameters with the only difference in the amount of crustal heat production. The left plot shows geotherms for the case without any heat production. These geotherms are almost straight lines (some deviation from straight lines are caused by changes in the thermal conductivity) until they cross the adiabat and their slope is proportional to the surface heat flow. In other words, if the temperature of the convecting mantle follows an adiabat, the surface heat flow is solely determined by the thickness of the conductive boundary layer. If it is assumed that the seismic anomalies shown in Figure 5.2 are mainly caused by temperature effects (since compositional variations are very unlikely to explain such large variations) these geotherms can be ruled out as an explanation because they are not apt to explain the differences in seismic velocity below 100 km depth.

Another extreme family is the case in which the mantle heat flow for all geotherms is the same. Sclater *et al.* [1980] estimated the non-radiogenic component of the surface heat flow to be between 21 and 25 mW/m² for lithospheric ages greater than 400 Ma. Such geotherms would reach the adiabat at depths shallower than 150 km, if the mantle heat production was not significantly higher than in the Chapman model. Other reported values of reduced mantle heat flow in the literature go down to 6 mW/m² for Archean regions [see Table 3 of Rudnick and Fountain, 1995]. Since we are interested in differences that exist down to about 200 km a value of 15 mW/m² was chosen for this example. The different surface heat flow is only caused by different crustal heat production. Comparison of these geotherms with the ones from Figure 5.5 shows that the differences in temperature are far too small. Additionally, these geotherms would not allow partial melting for depths less than 100 km which is expected for many tectonically active regions.

Therefore, continental geotherms, if no lateral variation in the thermal conductivity (caused by variable composition) is assumed, are determined by a combination of (1) the crustal heat production and (2) the thickness of the conductive boundary layer. A variable thickness of the TBL can be seen in Figure 5.5 but this thickness alone is not sufficient to explain the large variation of observed heat flow values. The Chapman geotherm family, which includes a correlated variation in both, gives temperatures very similar from what is found in Figure 5.5 and is thus consistent with the tomographic *S* velocities. The temperature variations are largest at shallow depths around 60 km where they reach a maximum of about

900°C. Below the seismic lithosphere temperature variations do not exceed 300°C.

5.6 Discussion

Different heat fluxes lead to diverging geotherms but at depths where heat transfer is dominated by material transport and not by conduction the temperatures come closer to a mean mantle adiabat. The Chapman [1986] geotherm family for surface heat flow values between 42 and 90 mW/m² covers a broad temperature-depth space (down to 225 km depth and up to 900°C temperature variation). This depth range is in good accordance with the seismic lithosphere, which was defined as the region of strong velocity variation related to tectonic provinces, obtained from Figures 5.2 and 5.3. Additionally, the temperature range explains the velocity difference of up to 12% found in the lithosphere.

Variations of the surface heat flow are sometimes mainly ascribed to the thickness of the continental roots [Nyblade and Pollack, 1993] or to differences in the crustal heat production [Morgan, 1985; Lenardic, 1997]. Measurements of surface heat flow and concentrations of heat producing elements generally give not enough information to quantify the relative importance of the two factors very well [Ballard and Pollack, 1987]. Seismic studies yield important extra information to resolve this. The good correspondence of our results with the Chapman geotherms suggests that both causes are of similar importance and correlated. The extreme geotherm families depicted in the left and right panels of Figure 5.7, which involve only one cause, cannot explain the seismic velocity variations. Therefore, we want to make some comments on both mechanisms. Tectonically active regions can have transient temperature disturbances, but in the following we consider only regions that were stable for longer times and thus are close to thermal equilibrium. The crust is enriched in radiogenic elements, mainly Cs, Rb, Th and U, relative to the depleted mantle because most of these elements are incompatible, which means that they do not fit readily into crystal lattices, and thus are easily removed from the mantle by the extraction of partial melt, in which they are concentrated. Erosion, in turn, removes the enriched crustal rocks exposed to the surface from a particular region. Thus, differentiation of mantle material and erosion are the factors determining the crustal heat production.

The thickness of the TBL influences the temperature gradient and thus the conductive heat flow. One question which remains is what causes the variation in thickness. Since the seismic velocities in the lithosphere have only a small correlation with the velocities below the TBL (Figure 5.3) a deep cause, e.g. mantle convection, is unlikely. Two other possibilities exist. First, this boundary layer could be compositionally different. For instance, differentiation processes could generate a mantle residue, which has a lower density, a higher solidus temperature and a higher viscosity [e.g. Jordan, 1988; de Smet *et al.*, 1998] and thus forms a layer that does not participate in convection. Second, the thickness could thermally be determined by the amount of crustal thermal blanketing. Little shallow

heat production results in a smaller temperature increase through the crust and thus more efficient cooling of the lower lithosphere resulting in a thick TBL (enhanced by a temperature dependent viscosity). The opposite would be true for high concentrations of heat producing elements in the crust. If this explanation is correct, then the concentration of radiogenic elements would be the dominant controlling factor for the TBL in thermal equilibrium by influencing the temperature directly through heat production and indirectly by determining the thickness of the TBL. However, a counteracting effect may exist too. A smaller crustal blanketing yields a larger temperature difference between the top and the bottom of the upper mantle which enhances convection that will oppose the formation of a thick boundary layer. This needs not be true for a temperature dependent viscosity. Quantitative modeling of boundary layers for a temperature dependent viscosity is needed to verify whether the observed variations in crustal heat production can generate variations in the thickness of the TBL consistent with the seismic velocities.

5.7 Conclusions

From this investigation we can draw the following conclusions: (1) For continental regions larger than 10° in diameter examined in this study the strong velocity variations of the upper 180 km show a high correlation with tectonic provinces and with surface heat flow. This correlation almost completely vanishes at depths larger than 230 km. We propose that this transition marks the maximum depth of the TBL. (2) The temperature profiles derived from the tomographic *S* wave velocities agree well in amplitude and depth range with the surface heat flow controlled conductive geotherms proposed by Chapman [1986] as a preferred continental thermal model. The Chapman geotherms include a connection between crustal heat production and mantle heat flow described by a 40% – 60% contribution of upper crustal and deeper sources, respectively, to the surface heat flow. The presence of compositional variations is neither required nor can it be ruled out on the basis of observed velocity perturbations. (3) Lateral temperature variations below the seismically defined lithosphere do not appear to control the thermal and seismic structure of the lithosphere since there is only little correlation as indicated by the *S* wave velocity on Figures 5.2 and 5.3. On the other hand, radiogenic heat production in the crust has a important influence on the underlying mantle. This implies a top down cause-effect relation. Most of the crustal radiogenic elements of old continental regions have been removed by erosion. This shrinking of the crustal thermal blanketing is accompanied with cold roots which are visible as fast velocity anomalies in tomographic models. (4) The thickness of the TBL (sometimes called the thermal lithosphere), defined as the upper boundary layer with predominant conductive heat transfer, ranges between between 60 km (for a surface heat flow of 90 mW/m^2) and 225 km (42 mW/m^2). The variations in thickness could be caused compositionally, e.g. by depletion processes that remove

the minerals with the lowest melting points, or thermally through different blanketing effects of the crust according to the amount of heat producing elements. On the basis of seismic velocities alone these two options cannot be distinguished. But a correlation between the thickness of the TBL and the amount of crustal heat production seems required to explain the continental geotherms determined from seismic velocities.

Appendix A

Variance of delay times in a two-dimensional random medium

In a two-dimensional Euclidean space with a uniform, isotropic random medium is characterized by its mean velocity and a covariance function $C(x, y) = C(r)$, describing the statistical properties of the fluctuations. A Gaussian function, slightly modified so that it represents fluctuations with zero mean (see Müller et al., [1992] for an accordant modified exponential function), is given by

$$C(r) = \sigma^2 \left(1 - \frac{r^2}{a^2}\right) \exp\left(-\frac{r^2}{\sigma^2}\right) \quad (\text{A.1})$$

where σ is the standard deviation of the fluctuations and a is a scale length describing the size of anomalies. According to eq. (4.6) the variance of delay times for a path length L is given by

$$\langle \delta t^2 \rangle = \int_0^L \int_0^L \langle \delta u(\mathbf{s}_1) \delta u(\mathbf{s}_2) \rangle ds_1 ds_2 \quad (\text{A.2})$$

$$= \int_0^L \int_0^L C(s_1 - s_2) ds_1 ds_2 \quad . \quad (\text{A.3})$$

Inserting the covariance function and solving the double integral yields

$$\langle \delta t^2 \rangle = \frac{\sigma^2 a L \sqrt{\pi}}{2} \operatorname{erf}\left(\frac{L}{a}\right) \quad . \quad (\text{A.4})$$

The result depends on the variance σ^2 as well as on the length scales a of the medium and on the path length. For curved rays in the Earth an additional factor describing the geometry in the reference medium is needed. If L is much larger than a the error function approaches unity and the resulting delay time variance is approximately linear in σ , a and L .

Appendix B

Constraints on the correlation function

B.1 Fluctuations with zero mean

A uniform, isotropic random medium with a zero mean poses a constraint on its covariance function $C(\Delta)$. For spherical layers (in the notation of Chapter 4: $\Delta_v = 0; \Delta = \Delta_h$) this translates in eq. (4.19):

$$\int_0^{\pi} C(\Delta) \sin(\Delta) d\Delta = 0 \quad .$$

Proof: The covariance function is given by

$$C(\Delta) = \langle \delta u(\theta, \varphi) \delta u(\theta', \varphi') \rangle \quad (\text{B.1})$$

$$= \frac{1}{8\pi^2} \int_0^{\pi} \int_0^{2\pi} \int_0^{2\pi} \delta u(\theta, \varphi) \delta u(\theta', \varphi') d\alpha d\varphi \sin(\theta) d\theta \quad (\text{B.2})$$

where θ and φ is the colatitude and longitude respectively. The point (θ', φ') is separated from point (θ, φ) by the angle Δ in direction α . Inserting this expression

in the integral of eq. (4.19) and rearranging the order of the integrals gives

$$\begin{aligned}
 & \int_0^\pi C(\Delta) \sin(\Delta) d\Delta \\
 &= \frac{1}{2\pi} \int_0^\pi \int_0^{2\pi} \delta u(\theta, \varphi) \underbrace{\frac{1}{4\pi} \int_0^\pi \int_0^{2\pi} \delta u(\theta', \varphi') d\alpha \sin(\Delta) d\Delta}_{= \langle \delta u \rangle = 0} d\varphi \sin(\theta) d\theta \\
 &= 0 \qquad \qquad \qquad \square \qquad \text{(B.3)}
 \end{aligned}$$

Of course, this constraint is also valid for the correlation function, which differs from the covariance only in the normalization with the variance. It follows from eq. (4.19) that an exponential function $\exp(\frac{\Delta}{a})$ or a Gaussian function $\exp(\frac{\Delta^2}{a^2})$ does not describe a medium with zero-mean.

The generalization of the condition to other spaces is straightforward by replacing the sine in eq. (4.19) by the appropriate weighting function. Müller et al. [1992] derive the equivalent formula for Euclidean spaces of dimension D from the Fourier transform of the medium and the Wiener-Khinchin theorem:

$$\int_0^\infty R(r) r^{(D-1)} dr = 0 \quad . \qquad \text{(B.4)}$$

B.2 Continuous fluctuations

A second constraint can be obtained if the covariance function represents a continuous medium. Chernov [1967] gives the proof that in this case the autocorrelation function must have a zero gradient at zero distance between two points. Since the proof is somewhat different for a spherical geometry, we give it here for this case.

We consider an isotropic, continuous medium on a spherical shell. In this case the covariance function is a function of angular distance only. Since the covariance function of the sphere equals the expected value of all covariance functions of great circles we can restrict our proof for one great circle. Let the position on the circle be described by an angle $\xi \in [0, 2\pi]$. The covariance function is given by

$$C(\xi') = \langle \delta u(\xi) \delta u(\xi + \xi') \rangle \qquad \text{(B.5)}$$

Since

$$\frac{d\delta u(\xi + \xi')}{d\xi'} = \frac{d\delta u(\xi + \xi')}{d\xi} \qquad \text{(B.6)}$$

we can write the derivative of the covariance function as

$$\frac{dC(\xi')}{d\xi'} = \frac{d}{d\xi'} \langle \delta u(\xi) \delta u(\xi + \xi') \rangle \quad (\text{B.7})$$

$$= \left\langle \delta u(\xi) \frac{d\delta u(\xi + \xi')}{d\xi'} \right\rangle \quad (\text{B.8})$$

$$= \left\langle \delta u(\xi) \frac{d\delta u(\xi + \xi')}{d\xi} \right\rangle \quad (\text{B.9})$$

The derivative at $\xi' = 0$ can further be simplified as

$$\left. \frac{dC(\xi')}{d\xi'} \right|_{\xi'=0} = \left\langle \delta u(\xi) \frac{d\delta u(\xi)}{d\xi} \right\rangle \quad (\text{B.10})$$

$$= \left\langle \frac{1}{2} \frac{d\delta u^2(\xi)}{d\xi} \right\rangle \quad (\text{B.11})$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{2} \frac{d\delta u^2(\xi)}{d\xi} d\xi \quad (\text{B.12})$$

$$= \frac{1}{4\pi} [\delta u^2(2\pi) - \delta u^2(0)] \quad (\text{B.13})$$

$$= 0 \quad (\text{B.14})$$

The last step follows from the 2π periodicity of $\delta u(\xi)$. This proves eq. (4.20).

The constraint is however not valid in the radial direction since the size of the medium is not infinite (needed for the proof given by Chernov [1967]) and we do not assume the medium is uniform in this direction. Eq. (B.11), which is the expectation of the derivative of the variance, does generally not equal zero in this case so that the vertical derivative of the covariance function does not necessarily vanish for zero separation.

B.3 Relation with delay time covariance

The covariance function of two delay times is linked to the covariance function of the medium. To evaluate the consequence of eq. (4.20) for delay times we consider a two-dimensional, uniform, isotropic random medium characterized by the covariance function $C(x, y) = C(r)$. The most simple case is that two rays of length L are parallel with the separation Δ . The covariance T between pairs of delay times is given by eq. (4.6) as

$$T(\Delta) = \int_0^L \int_0^L C(r) dx dx' \quad (\text{B.15})$$

where the distance between two points on the rays is

$$r = \sqrt{(x - x')^2 + \Delta^2} \quad . \quad (\text{B.16})$$

The derivative of eq. (B.15) is

$$\frac{dT(\Delta)}{d\Delta} = \int_0^L \int_0^L \frac{dC(r)}{d\Delta} dx dx' = \int_0^L \int_0^L \frac{dC(r)}{dr} \frac{dr}{d\Delta} dx dx' \quad (\text{B.17})$$

$$= \int_0^L \int_0^L \frac{dC(r)}{dr} \frac{\Delta}{\sqrt{(x - x')^2 + \Delta^2}} dx dx' \quad (\text{B.18})$$

We now find the value of the derivative at $\Delta = 0$:

$$\left. \frac{dT(\Delta)}{d\Delta} \right|_{\Delta=0} = \int_0^L \int_0^L \left. \frac{dC(r)}{d\Delta} \right|_{\Delta=0} \delta(x - x') dx dx' \quad (\text{B.19})$$

Since $r = 0$ if $x = x'$ and $\Delta = 0$ we obtain

$$\left. \frac{dT(\Delta)}{d\Delta} \right|_{\Delta=0} = \int_0^L \left. \frac{dC(r)}{d\Delta} \right|_{r=0} dx = L \left. \frac{dC(r)}{d\Delta} \right|_{r=0} \quad (\text{B.20})$$

The right hand side of the last equation is zero for a continuous medium (of infinite size or periodic boundary condition) and therefore the derivative of the delay time covariance function in the limit of zero distance between the two stations equals zero.

The derivation above can be generalized for the geometry of Figure 4.1 with a common source and curved ray paths. To do this the distance r between two points on the rays (eq. (B.16)) needs to be modified according to the geometry. However, this does not change the principle of the derivation, since for small distances between the stations the leading term of r is linear in Δ . The only difference in the result is a geometry dependent factor in the right hand side of eq. (B.20).

References

- Adams, A.D. Seismicity of Western Europe as determined from ISC data files. In P. Melchior, editor, *Seismic activity in Western Europe with particular consideration to the Liège earthquake of November 8, 1998*, pages 57–69. NATO ASI Series, D. Raidel, 1985.
- Aldridge, D.F. Linearization of the eikonal equation. *Geophysics*, **59**, 1631–1632, 1994.
- Anderson, D.L. Geophysics of the continental mantle: an historical perspective. In M.A. Menzies, editor, *Continental mantle*, pages 1–30. Clarendon Press, Oxford, United Kingdom, 1990.
- Anderson, D.L. Lithosphere, asthenosphere, and perisphere. *Rev. Geophys.*, **33**, 125–149, 1995.
- Anderson, D.L. and J.D. Bass. Mineralogy and composition of the upper mantle. *Geophys. Res. Lett.*, **11**, 637–640, 1984.
- Audebert, F., D. Nichols, T. Rekdal, B. Biondi, D.E. Lumley, and H. Urdaneta. Imaging complex geologic structure with single-arrival Kirchhoff prestack depth migration. *Geophysics*, **62**, 1533–1543, 1997.
- Ballard, S.B. and H.N. Pollack. Diversion of heat by Archean cratons: a model for southern Africa. *Earth Planet. Sci. Lett.*, **85**, 253–264, 1987.
- Berckhemer, H., W. Kampfman, E. Aulbach, and H. Schmeling. Shear modulus and Q of forsterite and dunite near partial melting from forced oscillation experiments. *Phys. Earth Planet. Inter.*, **29**, 30–41, 1982.
- Bertrand, P., C. Sotin, J.-C.C. Mercier, and E. Takahashi. From the simplest chemical system to the natural one: garnet peridotite barometry. *Contrib. Mineral. Petrol.*, **93**, 168–178, 1986.
- Bijwaard, H., W. Spakman, and R.R. Engdahl. Closing the gap between regional and global travel time tomography. *J. Geophys. Res.*, **103**, 30055–30078, 1998.
- Bolt, B.A. and O.W. Nuttli. P wave residuals as a function of azimuth. *J. Geophys. Res.*, **71**, 5977–5985, 1966.
- Buland, R. Residual statistics. *Terra Cognita*, **4**, 268, 1984.
- Bullen, K.E. *An introduction to the theory of seismology*. Cambridge University Press, Cambridge, 3rd edition, 1963.
- Bungum, H. and E.S. Husebye. Errors in time delay measurements. *Pageoph*, **91**, 56–70, 1971.

- Burton, P.W. Seismic risk in Turkey, the Aegean and the eastern Mediterranean; the occurrence of large magnitude earthquakes. *Geophys. J. R. Astron. Soc.*, **78**, 475–506, 1984.
- Chapman, D.S. Thermal gradients in the continental crust. In J. Hall J.B. Dawson, D.A. Carswell and K.H. Wedepohl, editors, *The nature of the lower continental crust*, volume 24, pages 63–70. Geological Society Special Publication, 1986.
- Chapman, D.S. and H.N. Pollack. Global heat flow: a new look. *Earth Planet. Sci. Lett.*, **28**, 23–32, 1975.
- Chernov, L.A. *Wave propagation in a random medium*. Dover Publications, New York, 1967.
- Claerbout, J.F. Detection of *P*-waves from weak sources at great distances. *Geophysics*, **29**, 197–211, 1964.
- Cleary, J. and A.L. Hales. An analysis of the travel times of *P* waves to North American stations, in the distance range 32° to 100°. *Bull. Seismol. Soc. Am.*, **56**, 467–489, 1966.
- Davies, J.H. Lower bound of average earthquake mislocation from variance of travel-time residuals. *Phys. Earth Planet. Inter.*, **75**, 89–101, 1992.
- Davies, J.H. and O. Gudmundsson. Spectra of mantle shear wave velocity structure. *Geophys. J. Int.*, **108**, 865–882, 1992.
- de Smet, J.H., A.P. van den Berg, and N.J. Vlaar. Stability and growth of continental shields in mantle convection models including recurrent melt production. *Tectonophysics*, **296**, 15–29, 1998.
- Douglas, A. Bandpass filtering to reduce noise on seismograms: Is there a better way? *Bull. Seismol. Soc. Am.*, **87**, 770–777, 1997.
- Douglas, A., D. Bowers, and J.B. Young. On the onset of *P* seismograms. *Geophys. J. Int.*, **129**, 681–690, 1997.
- Dziewonski, A.M. and D.L. Anderson. Preliminary reference Earth model. *Phys. Earth Planet. Inter.*, **25**, 297–356, 1981.
- Dziewonski, A.M. and D.L. Anderson. Travel times and station corrections for *P* waves at teleseismic distances. *J. Geophys. Res.*, **88**, 3295–3314, 1983.
- 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.
- Engdahl, E.R., R.D. van der Hilst, and R.P. Buland. Global teleseismic earthquake relocation with improved travel times and procedures for depth determination. *Bull. Seismol. Soc. Am.*, **88**, 722–743, 1998.
- Gardien, W. Accurate measurement of relative *P*-wave traveltimes and comparison with the ISC. Master's thesis, Utrecht University, 1997.
- Geoltrain, S. and J. Brac. Can we image complex structures with first-arrival travel-time? *Geophysics*, **58**, 564–575, 1993.
- Gilder, S.A., G.R. Keller, M. Luo, and P.C. Goodell. Timing and spatial distribution of rifting in China. *Tectonophysics*, **197**, 225–243, 1991.
- Goes, S., R. Govers, and P. Vacher. Shallow mantle temperatures under Europe from *P* and *S* wave tomography. *J. Geophys. Res.*, **in press**, 1999.
- Grand, S.P. A possible station bias in travel time measurements reported to ISC. *Geophys. Res. Lett.*, **17**, 17–20, 1990.

- Gudmundsson, O., J.H. Davies, and R.W. Clayton. Stochastic analysis of global traveltimes data: mantle heterogeneity and random errors in the ISC data. *Geophys. J. Int.*, **102**, 25–43, 1990.
- Gurnis, M. Large-scale mantle convection and the aggregation and disposal of supercontinents. *Nature*, **332**, 695–699, 1988.
- Hedlin, M.A.H., P.M. Shearer, and P.S. Earle. Seismic evidence for small-scale heterogeneity throughout the Earth's mantle. *Nature*, **387**, 145–150, 1997.
- Herrin, E. and J. Taggart. Regional variation in *P* travel times. *Bull. Seismol. Soc. Am.*, **58**, 1325–1337, 1968.
- Jackson, I. and S.M. Rigden. Composition and temperature of the Earth's mantle: seismological models interpreted through experimental studies of Earth minerals. In I. Jackson, editor, *The Earth's mantle: composition, structure and evolution*, pages 405–460. Cambridge Univ. Press, U.K., 1998.
- Jeffreys, H. An alternative to the rejection of observations. *Proc. R. Soc.*, **137A**, 78–87, 1932.
- Jeffreys, H. and K.E. Bullen. *Seismological tables*. British Association for the Advancement of Science, London, 1940.
- Jones, A.G. Observations of the electrical asthenosphere beneath Scandinavia. *Tectonophysics*, **90**, 37–55, 1982.
- Jordan, T.H. The continental tectosphere. *Rev. Geophys. Space Phys.*, **13**, 1–12, 1975.
- Jordan, T.H. Mineralogies, densities and seismic velocities of garnet lherzolites and their geophysical implications. In F.R. Boyd and H.O.A. Meyer, editors, *The mantle sample: inclusions in kimberlites and other volcanics*, pages 1–14. AGU, Washington D.C., 1979.
- Jordan, T.H. Structure and formation of the continental tectosphere. *J. Petrol.*, **Special Lithosphere Issue**, 11–37, 1988.
- Kaneshima, S. and G. Helffrich. Detection of lower mantle scatterers northeast of the Mariana subduction zone using short-period array data. *J. Geophys. Res.*, **103**, 4825–4838, 1998.
- Karato, S. On the Lehmann discontinuity. *Geophys. Res. Lett.*, **19**, 2255–2258, 1992.
- Karato, S. Importance of anelasticity in the interpretation of seismic tomography. *Geophys. Res. Lett.*, **20**, 1623–1626, 1993.
- Kato, M. and K. Hirahara. Precursory arrivals to *PP*. *Geophys. J. Int.*, **106**, 551–557, 1991.
- Kennett, B.L.N. and E.R. Engdahl. Traveltimes for global earthquake location and phase identification. *Geophys. J. Int.*, **105**, 429–465, 1991.
- Kennett, B.L.N., E.R. Engdahl, and R. Buland. Constraints on seismic velocities in the earth from traveltimes. *Geophys. J. Int.*, **122**, 108–124, 1995.
- Korn, M. Seismic waves in random media. *J. Appl. Geophys.*, **29**, 247–269, 1993.
- Kravtsov, Y.A. Rays and caustics as physical objects. In E. Wolf, editor, *Prog. in Optics, XXVI*, pages 227–348. Elsevier, Amsterdam, 1988.
- Krishna, V.G. and K.L. Kaila. Seismological study of the elevation of phase boundaries within the subducting lithospheric slabs. *Tectonophysics*, **134**, 201–213, 1987.

- Kværna, T. Automatic onset time estimation based on autoregressive processing. *Semi-annual Technical Summary, 1 April - 30 September, NOR SAR Sci. Rep. No. 1 95/96*, pages 113–133, 1995.
- Lenardic, A. On the heat flow variation from Archean cratons to Proterozoic mobile belts. *J. Geophys. Res.*, **102**, 709–721, 1997.
- Lerner-Lam, A.L. and T.H. Jordan. How thick are continents? *J. Geophys. Res.*, **92**, 14007–14026, 1987.
- Lilwall, R.C. and A. Douglas. Estimation of *P*-wave travel times using the joint epicentre method. *Geophys. J. R. Astron. Soc.*, **19**, 165–181, 1970.
- Mao, W. and D. Gubbins. Simultaneous determination of time delays and stacking weights in seismic array beam forming. *Geophysics*, **60**, 491–502, 1995.
- McKenzie, D. and M.J. Bickle. The volume and composition of melt generated by extension of the lithosphere. *J. Petrol.*, **29**, 625–679, 1988.
- Montagner, J.-P. and T. Tanimoto. Global upper mantle tomography of seismic velocities and anisotropies. *J. Geophys. Res.*, **96**, 20337–20351, 1991.
- Morelli, A. and A.M. Dziewonski. Body-wave traveltimes and a spherically symmetric *P*- and *S*-wave velocity model. *Geophys. J. Int.*, **112**, 178–194, 1993.
- Morelli, A., A.M. Dziewonski, and J.H. Woodhouse. Anisotropy of the inner core inferred from *PKIKP* travel times. *Geophys. Res. Lett.*, **13**, 1545–1548, 1986.
- Morgan, P. Crustal radiogenic heat production and the selective survival of ancient continental crust. *Proceedings of the fifteenth lunar and planetary science conference, JGR*, **90**, C561–C570, 1995.
- Müller, G., M. Roth, and M. Korn. Seismic-wave traveltimes in random media. *Geophys. J. Int.*, **110**, 29–41, 1992.
- Nakanishi, I. and D.L. Anderson. World-wide distribution of group velocity of mantle Rayleigh waves as determined by spherical harmonic inversion. *Bull. Seismol. Soc. Am.*, **72**, 1185–1194, 1982.
- Nataf, H.-C., I. Nakanishi, and D.L. Anderson. Measurements of mantle wave velocities and inversion for lateral heterogeneities and anisotropy, 3. Inversion. *J. Geophys. Res.*, **91**, 7261–7307, 1986.
- Nataf, H.-C. and Y. Ricard. 3SMAC: an a priori tomographic model of upper mantle based on geophysical modeling. *Phys. Earth Planet. Inter.*, **95**, 101–122, 1996.
- Nelder, J.A. and R. Mead. A simplex method for function minimization. *Computer J.*, **7**, 308–313, 1965.
- Nolet, G. Imaging the deep earth: technical possibilities and theoretical limitations. In A. Rocca and D. Mayer-Rosa, editors, *Proceedings of the 25th ESC Symposium Barcelona 1990*, pages 107–115, Barcelona, Spain, 1992. Servei Geològic de Catalunya.
- Nolet, G. and T.-J. Moser. Teleseismic delay times in a 3-D Earth and a new look at the *S* discrepancy. *Geophys. J. Int.*, **114**, 185–195, 1993.
- Nyblade, A.A. and H.N. Pollack. A comparative study of parameterized and full thermal-convection models in the interpretation of heat flow from cratons and mobile belts. *Geophys. J. Int.*, **113**, 747–751, 1993.
- Obayashi, M. and Y. Fukao. *P* and *PcP* travel time tomography for the core-mantle boundary. *J. Geophys. Res.*, **102**, 17825–17841, 1997.

- O'Reilly, S.Y. and W.L. Griffin. A xenolith-derived geotherm for southeastern Australia and its geophysical implications. *Tectonophysics*, **111**, 41–63, 1985.
- Passier, M.L. and R.K. Snieder. On the presence of intermediate-scale heterogeneity in the upper mantle. *Geophys. J. Int.*, **123**, 817–837, 1995.
- Peterson, N.V. Inverse kinematic problem for a random medium in geometric optics approximation. *Pageoph*, **132**, 417–437, 1990.
- Polet, J. and D.L. Anderson. Depth extent of cratons as inferred from tomographic studies. *Geology*, **23**, 205–208, 1995.
- Pollack, H.N. and D.S. Chapman. On the regional variation of heat flow, geotherms, and lithospheric thickness. *Tectonophysics*, **38**, 279–296, 1977.
- Pollack, H.N., S.J. Hurter, and J.R. Johnson. Heat flow from the earth's interior: analysis of the global data set. *Rev. Geophys.*, **31**, 267–280, 1993.
- Press, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical recipes in FORTRAN*. Cambridge University Press, Cambridge, U.K., 2nd edition, 1992.
- Pujol, J., D.M. Fountain, and D.A. Anderson. Statistical analysis of the mean heat flow / reduced heat flow relationship for continents and its tectonothermal implications. *J. Geophys. Res.*, **90**, 11335–11344, 1985.
- Pulliam, R.J., D.W. Vasco, and L.R. Johnson. Tomographic inversions for mantle P wave velocity structure based on the minimization of l^2 and l^1 norms of international seismological centre travel time residuals. *J. Geophys. Res.*, **98**, 699–734, 1993.
- Puster, P. and T.H. Jordan. Stochastic analysis of mantle convection experiments using two-point correlation functions. *Geophys. Res. Lett.*, **21**, 305–308, 1994.
- Puster, P., T.H. Jordan, and B.H. Hager. Characterization of mantle convection experiments using two-point correlation functions. *J. Geophys. Res.*, **100**, 6351–6365, 1995.
- Ricard, Y., H.-C. Nataf, and J.P. Montagner. The three-dimensional seismological model a priori constrained: Confrontation with seismic data. *J. Geophys. Res.*, **101**, 8457–8472, 1996.
- Rice, J.A. *Mathematical statistics and data analysis*. Duxbury Press, Belmont, 2nd edition, 1995.
- Richards, M.A., B.H. Hager, and N.H. Sleep. Dynamically supported geoid highs over hotspots: Observation and theory. *J. Geophys. Res.*, **93**, 7690–7708, 1988.
- Robertson, G.S. and J.H. Woodhouse. Ratio of relative S to P velocity heterogeneity in the lower mantle. *J. Geophys. Res.*, **101**, 20041–20052, 1996.
- Rodgers, A. and J. Wahr. Inference of core-mantle boundary topography from ISC PcP and PKP traveltimes. *Geophys. J. Int.*, **115**, 991–1011, 1993.
- Roth, M., G. Müller, and R. Snieder. Velocity shift in random media. *Geophys. J. Int.*, **115**, 552–563, 1993.
- Roy, R.F., D.D. Blackwell, and F. Birch. Heat generation of plutonic rocks and continental heat flow provinces. *Earth Planet. Sci. Lett.*, **5**, 1–12, 1968.
- Rudnick, R.L. and D.M. Fountain. Nature and composition of the continental crust: A lower crustal perspective. *Rev. Geophys.*, **33**, 267–309, 1995.
- Rudnick, R.L., W.F. MacDonough, and R.J. O'Connell. Thermal structure, thickness and composition of continental lithosphere. *Chem. Geol.*, **145**, 395–411, 1998.

- Scherbaum, F. and M.-P. Bouin. FIR filter effects and nucleation phases. *Geophys. J. Int.*, **130**, 661–668, 1997.
- Schmeling, H. and G. Marquart. The influence of second-scale convection on the thickness of continental lithosphere and crust. *Tectonophysics*, **189**, 281–306, 1991.
- Slater, J.G., C. Jaupart, and D. Galson. The heat flow through oceanic and continental crust and the heat loss of the Earth. *Rev. Geophys.*, **18**, 269–311, 1980.
- Slater, J.G., B. Parsons, and C. Jaupart. Oceans and continents: similarities and differences in the mechanisms of heat loss. *J. Geophys. Res.*, **86**, 11535–11552, 1981.
- Shearer, P.M. Seismic imaging of upper-mantle structure with new evidence for a 520-km discontinuity. *Nature*, **344**, 121–126, 1990.
- Shearer, P.M. Improving local earthquake locations using the L1 norm and waveform cross correlation: Application to the Whittier Narrows, California, aftershock sequence. *J. Geophys. Res.*, **102**, 8269–8283, 1997.
- Shearer, P.M., K.M. Toy, and J.A. Orcutt. Axi-symmetric earth models and inner-core anisotropy. *Nature*, **333**, 228–232, 1988.
- Snyder, 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.
- Sobolev, S.V., H. Zeyen, G. Stoll, F. Werling, R. Altherr, and K. Fuchs. Upper mantle temperatures from teleseismic tomography of French Massif Central including effects of composition, mineral reactions, anharmonicity, anelasticity and partial melt. *Earth Planet. Sci. Lett.*, **139**, 147–163, 1996.
- Spakman, W. and G. Nolet. Imaging algorithms, accuracy and resolution in delay time tomography. In N.J. Vlaar et al., editor, *Mathematical Geophysics*, pages 155–187. D. Reidel, Norwell, Mass., 1988.
- Stull, R.B. *An introduction to boundary layer meteorology*. Kluwer, Dordrecht, 1991.
- Su, W.-j. and A.M. Dziewonski. Predominance of long-wavelength heterogeneity in the mantle. *Nature*, **352**, 121–126, 1991.
- Su, W.-j. and A.M. Dziewonski. Inner core anisotropy in three dimensions. *J. Geophys. Res.*, **100**, 9831–9852, 1995.
- Su, W.-j. and A.M. Dziewonski. Simultaneous inversion for 3-D variations in shear and bulk velocity in the mantle. *Phys. Earth Planet. Inter.*, **100**, 135–156, 1997.
- Su, W.-j., A.M. Dziewonski, and R. Jeanloz. Planet within a planet: rotation of the inner core of earth. *Science*, **274**, 1883–1887, 1996.
- Sylvander, M., B. Ponce, and A. Souriau. Seismic velocities at the core-mantle boundary inferred from *P* waves diffracted around the core. *Phys. Earth Planet. Inter.*, **101**, 189–202, 1997.
- Thompson, A.L. Water in the Earth's upper mantle. *Nature*, **358**, 295–302, 1992.
- van der Hilst, R.D., S. Widiyantoro, and E.R. Engdahl. Evidence for deep mantle circulation from global tomography. *Nature*, **386**, 578–584, 1997.
- VanDecar, J.C. and R.S. Crosson. Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares. *Bull. Seismol. Soc. Am.*, **80**, 150–169, 1990.
- Vasco, D.W. and L.R. Johnson. Whole earth structure estimated from seismic arrival times. *J. Geophys. Res.*, **103**, 2633–2671, 1998.

- Vitorello, I. and H.P. Pollack. On the variation of continental heat flow with age and thermal evolution of continents. *J. Geophys. Res.*, **85**, 983–995, 1980.
- Wessel, P. and W.H.F. Smith. New, improved version of generic mapping tools released. *Eos Trans. AGU*, **79**, 579, 1998.
- Wielandt, E. On the validity of ray approximation for interpreting delay times. In G. Nolet, editor, *Seismic tomography*, pages 85–98. D. Reidel Publishing Company, 1987.
- Willemann, R.J. and D.A. Storchak. Updating the ISC Bulletin. In *abstract*, Birmingham, 1999. XXII General Assembly of the IUGG.
- Windley, B.F. *The evolving continents*. John Wiley & Sons, 1995.
- 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**, 5953–5986, 1984.
- Woodhouse, J.H. and T. Trampert. Global upper mantle structure inferred from surface wave and body wave data. *Eos Trans. AGU*, page F422, 1995.
- Woodward, R.L. and G. Masters. Global upper mantle structure from long-period differential travel times. *J. Geophys. Res.*, **96**, 6351–6377, 1991.
- Yamaji, A. Periodic hotspot distribution and small-scale convection in the upper mantle. *Earth Planet. Sci. Lett.*, **109**, 107–116, 1992.
- Yan, B., E.K. Graham, and K.P. Furlong. Lateral variations in the upper mantle thermal structure inferred from three-dimensional seismic inversion models. *Geophys. Res. Lett.*, **16**, 449–452, 1989.

Samenvatting

Dagelijks doorlopen elastische golven, die door aardbevingen worden veroorzaakt, de aarde. Daarbij verzamelen de verschillende golftypes (oppervlakte- en ruimtegolven, transversale en longitudinale golven) langs verschillende voortplantingswegen informatie over de inwendige aarde. Ieder seismogram, dat door één van de honderden seismologische stations wordt geregistreerd, bevat daarmee een overvloed aan aanwijzingen over de opbouw van de inwendige aarde. Seismologen hebben de taak de hierin aanwezige informatie te ontcijferen. De wijze waarop dit gebeurt kan in drie stappen worden opgedeeld:

- (1) **Meting:** Omdat het tot nu toe niet mogelijk is om alle informatie te gebruiken, worden eerst relevante gegevens uit de seismogrammen geëxtraheerd. Aankomsttijden van aardbevingsgolven, dispersiecurven van oppervlaktegolven en frequenties van eigentrillingen van de aarde zijn voorbeelden van zulke gegevens.
- (2) **Gevolgtrekking :** In de volgende stap moeten de gegevens worden omgezet in informatie over het inwendige van de aarde. Dan is men er echter meestal nog niet, omdat de verkregen resultaten vooral iets zeggen over de verdeling van de seismische snelheden en gedeeltelijk - met een grotere onbetrouwbaarheid - over de dichtheid.
- (3) **Interpretatie:** De gegevens die vooral van interesse zijn, zijn echter naast de dichtheid de chemische samenstelling van het gesteente en de temperatuurverdeling. Hierover kunnen alleen met aanvullende informatie (bijv. metingen aan gesteente) en/of bepaalde veronderstellingen uitspraken worden gedaan.

Dit onderzoek houdt zich bezig met alle drie de stappen: meting (hoofdstuk 1-4), gevolgtrekking (hoofdstuk 3 en 4) en interpretatie (hoofdstuk 5). In de eerste vier hoofdstukken worden looptijdresiduen van het Bulletin van het *International Seismological Centre* (ISC) gebruikt. Looptijdresiduen van aardbevingsgolven zijn de verschillen tussen de "feitelijke" en de theoretische looptijden in een (ééndimensionaal) referentiemodel van de aarde. De "feitelijke" looptijden worden gevormd uit de geobserveerde aankomsttijden en uit de hieruit bepaalde haardtijd. Onnauwkeurigheden in de bepaling van de haardtijd hebben daarom fouten in de "feitelijke" looptijden tot gevolg. Terwijl *P*-golven (longitudinale ruimtegolven) ca. 20 min. nodig hebben om de aarde eenmaal te doorlopen, zijn de looptijdresiduen in de orde van enkele secondes. **Hoofdstuk 1** behandelt verschillende algemene aspecten van looptijdresiduen, in het bijzonder de mogelijkheden en grenzen om daarmee de driedimensionale structuur van de aarde te onderzoeken. Meer dan 28 miljoen aankomsttijden en bijbehorende looptijdresiduen worden door het

ISC beschikbaar gesteld om het inwendige van de aarde te onderzoeken. Voor een gedeelte van deze gegevens bestaan verbeterde hypocentrabepalingen en vervolgens nauwkeurigere looptijdresiduen [Engdahl *et al.*, 1998]. Deze data vormen de basis voor hoofdstuk 2-4.

Looptijdresiduen bevatten informatie over snelheidsafwijkingen van het gebruikte (ééndimensionale) model van de aarde geïntegreerd langs de voortplantingsweg. Deze worden beïnvloed door elke snelheidsanomalie binnen een volume rond de straal (waarbij de straal wordt gedefinieerd door het principe van Fermat als de lijn die de bron en de ontvanger met de extreme looptijd verbindt). De grootte van het volume wordt daarbij begrensd door het causaliteitsprincipe. Dat betekent dat het verschil tussen de looptijd van de bron naar een punt binnen het volume en verder naar de ontvanger en de looptijd langs de Fermatstraal kleiner is dan het meetkader. Het aflezen van de eerste inzet van een golf gebeurt in een veel kleiner tijdsbestek dan bijvoorbeeld looptijdmetingen met behulp van correlatiemethodes. Onder goede signaal-ruis-verhoudingen kan het begin van de *P*-golf met een betrouwbaarheid van 0.1 s worden gemeten. Het bijbehorende causaliteitsvolume is aanzienlijk kleiner dan dat voor correlatiemethoden, waarvoor tijdsbestekken van meerdere secondes worden gebruikt. Ook het Fresnelvolume berekend voor de dominante periode van ca. 1 s is aanzienlijk groter. ISC-data zijn gebaseerd op begintijden. Daardoor hebben ze een relatief klein causaliteitsvolume en komen in de buurt van de straaltheoretische benadering, die zegt dat de golf uitsluitend door structuur op de straal wordt beïnvloed. Aan de andere kant zijn begintijden in seismogrammen door een natuurlijk en kunstmatig ruisniveau vaak moeilijk te herkennen. In sterk heterogene gebieden komen gediffracteerde golven, die over het algemeen slechts kleine amplitudes hebben, vaak als eerste aan. Bovendien moet rekening worden gehouden met het feit dat de hypocentra en de haardtijden door aankomsttijden worden vastgesteld die ook voor de berekening van de looptijdresiduen worden gebruikt. Hierdoor is de fout in één enkel looptijdresidu niet alleen afhankelijk van de fout in de meting van de betreffende aankomsttijd, maar ook van de andere data van deze aardbeving en van de kwaliteit van het referentiemodel.

De maximaal bereikbare nauwkeurigheid van seismologisch onderzoek wordt door de kwaliteit van de seismologische gegevens begrensd. Worden in het algemeen *P*-looptijdresiduen gebruikt om de structuur van de aarde te bepalen, in **hoofdstuk 2** wordt door een gemiddelde van de gegevens de invloed van de driedimensionale snelheidsstructuur zoveel mogelijk geëlimineerd en in plaats daarvan wordt de kwaliteit van de stationstijd en het aflezen van de looptijden gecontroleerd. Hiertoe wordt een mediaanfilter over de *P*-data van een station geschoven. De mediaan zou voor een correct functionerend station nagenoeg constant moeten zijn. Er moet echter ook worden gecontroleerd of de ruimtelijke verdeling van de geobserveerde aardbevingen niet is veranderd. Dit wordt bereikt door het mediaanfilter tevens op de gegevens van afzonderlijke regio's toe te passen en de resultaten met elkaar te vergelijken. Komen de variaties van alle gegevens met die van de regio's overeen, dan moet de oorzaak bij het station worden gezocht.

Met deze methode kunnen alleen systematische veranderingen en geen statistische fouten in de afzonderlijke metingen worden opgespoord. Het resultaat van deze testen is verrassend; bij ongeveer de helft van alle onderzochte stations worden variaties gevonden van meer dan 0.5 s. In de meeste gevallen gaat een sterke verandering van de mediaan samen met een tijd zonder gegevens. Deze observatie doet vermoeden dat veranderingen aan seismometers, opname- en analysemethodes een veel voorkomende oorzaak zijn. Inderdaad kon dit in verschillende gevallen worden bevestigd door de betreffende stations. Voor enkele stations komen echter ook andere oorzaken voor. Het Yellowknife-Array (YKA, Canada) laat bijvoorbeeld veranderingen per jaargetijde zien die ook in het daar gemeten ruisniveau en in het aantal geobserveerde aardbevingen terug te vinden is. Vooral begintijden van kleine aardbevingen worden tijdens de zomermaanden duidelijk minder en later afgelezen dan in de winter. De waarschijnlijkste interpretatie hiervoor is dat in de zomer de eerste inzetten in de ruis ten ondergaan en daardoor te laat of helemaal niet worden bepaald. Bij het station MAG (Magadan 1, Siberië) laten de looptijdresiduen van verschillende regio's een duidelijk ander gedrag zien vanaf april 1972. Dit kan alleen worden verklaard door een verplaatsing van het station over een afstand van ca. 100 km die niet aan het ISC is doorgegeven. Daardoor zijn alle looptijdresiduen vanaf deze datum voor onjuiste stationscoördinaten berekend.

Systematische fouten kunnen, meer dan statistische fouten, onderzoeken negatief beïnvloeden, omdat ze, afhankelijk van het betreffende onderzoek, zich maar voor een deel onderling opheffen. Daarom wordt in **hoofdstuk 3** de invloed op globale looptijdtomografie onderzocht. Tomografie is waarschijnlijk één van de meest interessante en prominente onderzoeken waarvoor ISC-gegevens worden gebruikt. De tomografische inversie van synthetische data die de in het vorige hoofdstuk besproken fouten simuleren, toont aan dat de snelheidsamplitudes relatief klein zijn ten opzichte van de inversieresultaten voor ISC-data [Bijwaard *et al.*, 1998] en dat het model weinig systematische patronen laat zien. De variantie van de gesimuleerde looptijdresiduen is echter ook aanzienlijk kleiner. De fouten propageren in relatie tot deze variantie net zo sterk in het driedimensionale model als de structuurinformatie die in de echte data aanwezig is.

Dit resultaat is anders dan wat men door de inversie van statistische fouten verkrijgt. Een statistische fout ontstaat bijvoorbeeld door een geringe afleesnauwkeurigheid. Er kan worden aangetoond dat bij veel stations een deel van de eerste inzetten slechts op de volgende seconde afgerond worden aangegeven, hoewel de betreffende parameter in de ISC-data een afleesnauwkeurigheid van 0.1 s aangeeft. Daardoor ontstaat een fout die toevallig is verdeeld tussen -0.5 s en 0.5 s. Ook deze variantie van de fout is in vergelijking met de totale variantie van alle P -looptijdresiduen zeer klein. Bovendien wordt zij door de middeling voor en tijdens de inversie zeer sterk onderdrukt en is te verwaarlozen voor tomografische doeleinden. In **hoofdstuk 4** wordt de aarde statistisch beschreven als een toevalsmedium. Er wordt getracht de amplitudes en correlatielengtes van de snelheidsanomalie als functie van de diepte te inverteren uit de statistische eigenschappen van de looptijdresiduen. Hiertoe wordt eerst de variantie en de correlatie tussen naburige stralen

als functie van de epicentraalafstand en de haarddiepte gemeten. Deze waarden kunnen middels een lineaire inversie worden omgezet in de covariantiefunctie van de mantelstructuur.

Het onderzoek in dit hoofdstuk wordt echter door verschillende zaken bemoeilijkt. (1) Het relatief vaak voorkomen van zeer grote (absolute) looptijdresiduen heeft ten opzichte van een normaalverdeling een sterke afhankelijkheid van de momenten van de tweede orde, namelijk de variantie en de covariantie van de gekozen cut-off waarde tot gevolg. De bivariante verdeling van looptijdresiduen van vergelijkbare stralen laat zien dat de meeste residuen in de uitlopers (grote absolute residuen) niet zijn gecorreleerd. Dit impliceert dat deze van onjuist afgelezen begintijden stammen of het begin aan de *P*-fase toegeschreven werd, wat niet juist is. Het laatste is vooral bij ondiepe bevingen en positieve looptijdresiduen te verwachten, omdat de aan de oppervlakte gereflecteerde golven (bijv. *pP*, *sP*) korte tijd na de *P*-golven binnenkomen. (2) Fouten in de hypocentrabepalingen hebben systematische afwijkingen van de looptijdstatistieken tot gevolg. Dit kan gedeeltelijk worden onderdrukt door gebruik te maken van verschillen van de looptijdresiduen. De invloed van horizontale mislocaties is echter moeilijk in te schatten en hangt af van de ruimtelijke verdeling van de seismologische stations. (3) De statistische eigenschappen van looptijdresiduen variëren van gebied tot gebied. Deze constatering trekt de aanname in twijfel dat de amplitudes en waarden van de laterale snelheidsanomalieën alleen van de diepte afhankelijk zijn. Ondanks deze complicaties wijzen de gegevens op de aanwezigheid van heterogeniteiten ter grootte van minder dan 100 km in de bovenste aardmantel.

Hoofdstuk 5 is duidelijk te onderscheiden van de voorafgaande hoofdstukken, zowel wat de uitgangsgegevens als de gebruikte methode betreft. Bovendien beperkt zich het onderzoek op de continentale bovenmantel (tot een diepte van ca. 400 km). In dit hoofdstuk wordt gezocht naar oorzaken voor de grootschalige positieve snelheidsanomalieën onder oude continentale schilden. In golfvormtomografische modellen zijn dit veruit de opvallendste structuren. In de eerste plaats wordt de seismische structuur vergeleken met een verdeling van de continenten in drie verschillende types van tectonische provincies en met de in boorgaten gemeten warmtestroom. Er is een duidelijke correlatie te herkennen tussen tectonische provincie en snelheid en een anticorrelatie tussen snelheid en warmtestroom voor de bovenste 200 km. Omdat kan worden verwacht dat de verschillen in de chemische samenstelling een duidelijk kleinere invloed op de seismische snelheden hebben dan de temperatuurvariaties, worden de snelheidsprofielen, gemiddeld over de afzonderlijke provincies, omgerekend in temperatuurprofielen onder de aanname van een constante chemische samenstelling. Hieruit resulteren geothermen die op een diepte van 100 km temperatuurverschillen tot 900°C aantonen en die op nog grotere dieptes weer zeer veel dichter naar een gemeenschappelijke manteladia-baat toelopen. Bij diepten groter dan 230 km wordt een temperatuurverschil van 300°C niet overschreden. Een vergelijking toont een zeer goede overeenkomst met thermische modellen die de oppervlaktewarmtestroom voor 40% toeschrijven aan radioactieve elementen in de bovenste korst en de overige 60% aan een zeer diepe

bron. Voor deze indeling is een hechte samenhang tussen radioactiviteit nabij de oppervlakte en dikte van de thermische grenslaag nodig. Deze is gedefinieerd als de afstand tussen aardoppervlak en snijpunt tussen een conductive geotherm en de manteladiabaat. Om de koppeling tussen deze twee grootheden te verklaren wordt voorgesteld dat de dikte van de grenslaag wordt veroorzaakt door verschil in afscherming van de radioactieve korst. Het verwijderen van radioactieve elementen door erosie heeft een dikker wordende grenslaag tot gevolg. Ondiepe processen bepalen in dit scenario de thermische structuur van de continentale mantel tot op een diepte van meerdere 100 km.

Zusammenfassung

Die Erde wird täglich von elastischen Wellen, die von Erdbeben ausgehen, durchleuchtet. Die verschiedenen Wellentypen (Oberflächen- und transversale/longitudinale Raumwellen) sammeln dabei Information über das Erdinnere entlang verschiedener Fortpflanzungswege. Jedes Seismogramm, das von einer der mehreren hundert seismologischen Stationen aufgezeichnet wird, enthält damit eine Fülle an Hinweisen über den inneren Aufbau der Erde. Seismologen haben die Aufgabe, die in den vielen Ausschlägen enthaltene Information zu entschlüsseln. Die Vorgehensweise kann dabei in drei Schritte unterteilt werden:

- (1) **Messung:** Da es bisher nicht möglich ist, alle Informationen zu verwenden, werden zuerst relevante Größen aus den Seismogrammen extrahiert, z.B. die Ankunftszeiten von Raumwellen, die Dispersion von Oberflächenwellen oder die Frequenzen von Eigenschwingungen der Erde.
- (2) **Folgerung:** In einem weiteren Schritt wird diese Information umgesetzt in Aussagen über das Erdinnere. Das ist jedoch nicht genug, da die gewonnenen Resultate vor allem Aussagen über die Verteilung der seismischen Geschwindigkeiten und teilweise, allerdings mit einer größeren Unsicherheit, über die Dichte machen.
- (3) **Interpretation:** Die wirklich interessanten Größen sind aber, neben der Dichte, die chemische Zusammensetzung der Gesteine und die Temperaturverteilung. Aussagen darüber können nur mit Hilfe zusätzlicher Information (z.B. Labormessungen an Gesteinen) und/oder unter bestimmten Annahmen gewonnen werden.

Die vorliegende Arbeit befaßt sich mit allen drei Gebieten: Messung (Kap. 1 – 4), Folgerung (Kap. 3 und 4) und Interpretation (Kap. 5).

Die ersten vier Kapitel verwenden Laufzeitresiduen des Bulletins des *International Seismological Centre* (ISC). Laufzeitresiduen von Erdbebenwellen sind die Differenz der „tatsächlichen“ Laufzeiten und der theoretischen Laufzeiten in einem (eindimensionalen) Referenzmodell der Erde. Die „tatsächlichen“ Laufzeiten werden aus den beobachteten Ankunftszeiten und der aus den Ankunftszeiten bestimmten Herdzeit berechnet. Ungenauigkeiten in der Bestimmung der Herdzeit

führen dabei zu Fehlern in den „tatsächlichen“ Laufzeiten. Während *P*-Wellen (longitudinale Raumwellen) ca. 20 min benötigen, um die Erde einmal komplett zu durchlaufen, sind die Laufzeitresiduen nur in der Größenordnung von wenigen Sekunden. **Kapitel 1** behandelt verschiedene allgemeine Aspekte von Laufzeitresiduen im Zusammenhang mit den Möglichkeiten und Grenzen, um damit die dreidimensionale Erdstruktur zu erkunden. Mehr als 28 Millionen Ankunftszeiten und zugehörige Laufzeitresiduen werden dafür durch das ISC zur Verfügung gestellt. Für einen Teil dieser Daten existieren verbesserte Hypozentrenbestimmungen und damit auch genauere Laufzeitresiduen [Engdahl *et al.*, 1998]. Diese Daten bilden die Grundlage der Kapitel 2-4.

Laufzeitresiduen enthalten Information über Geschwindigkeitsabweichungen von dem verwendeten (eindimensionalen) Erdmodell integriert entlang des Ausbreitungswegs der seismischen Wellen. Sie werden beeinflusst durch alle Geschwindigkeitsanomalien innerhalb eines Volumens, welches den Strahl (Strahl der nach dem Fermat-Prinzip die verbindende Linie zwischen Quelle und Empfänger mit extremer Laufzeit darstellt) umgibt. Die Größe dieses Volumens ist durch das Kausalitätsprinzip begrenzt, d.h. die Laufzeit von der Quelle zu einem Punkt innerhalb des Volumens und weiter zum Empfänger darf nur soviel größer sein als die Laufzeit entlang des Fermat-Strahls, daß die Ankunftszeit noch innerhalb des Meßfensters liegt. Dadurch ergibt sich, daß diese Größe von der Meßmethode abhängig ist. Das Ablesen des Ersteinsatzes einer Welle benötigt ein viel kürzeres Zeitfenster als z.B. Laufzeitmessungen mit Hilfe von Korrelationsmethoden. Unter guten Signal-Rausch-Verhältnissen kann der *P*-Wellenanfang mit einer Genauigkeit von 0.1 s gemessen werden. Das zugehörige Kausalitätsvolumen ist bedeutend kleiner als dasjenige für Korrelationsmethoden, wofür typischerweise Zeitfenster von mehreren Sekunden verwendet werden und auch signifikant kleiner als das Fresnel-Volumen, berechnet für die dominante Periode von etwa 1 s. ISC Daten basieren auf Einsatzzeiten. Dadurch haben sie ein relativ kleines Kausalitätsvolumen und kommen nahe der strahlentheoretischen Näherung, wonach die Welle ausschließlich durch Struktur auf dem Strahl beeinflusst wird. Andererseits sind Einsatzzeiten in Seismogrammen oft nur schwer zu erkennen, abhängig von dem natürlichen und künstlichen Rauschniveau. Vorallem in stark heterogenen Gebieten kommen diffraktierte Wellen, die häufig nur sehr kleine Amplituden haben, oft als erste an. Es muß außerdem beachtet werden, daß sowohl das Hypozentrum als auch die Herdzeit von den selben Ankunftszeiten ermittelt werden, welche auch zur Berechnung der Laufzeitresiduen verwendet werden. Hierdurch hängt der Fehler in einem einzelnen Laufzeitresiduum nicht nur von dem Fehler in der Meßung der entsprechenden Ankunftszeit ab, sondern auch von den anderen Daten dieses Erdbebens und der Qualität des Referenzmodells.

Die Qualität seismologischer Daten begrenzt die maximal erreichbare Genauigkeit seismologischer Untersuchungen. Während üblicherweise *P*-Laufzeitresiduen verwendet werden, um die Erdstruktur zu bestimmen, wird in **Kapitel 2** durch geeignete Mittelung der Daten der Einfluß der dreidimensionalen Geschwindigkeitsstruktur weitgehend eliminiert und an Stelle davon die Güte der

Stationszeit und des Ablesens der Einsatzzeiten überprüft. Dazu wird ein zeitlich begrenzter Medianfilter über die P -Daten einer Station geschoben. Für eine korrekt arbeitende Station sollte der Median für alle Zeitfenster nahezu konstant sein. Allerdings muß auch kontrolliert werden, ob sich die räumliche Verteilung der beobachteten Erdbeben nicht auch verändert hat. Dies wird erreicht, indem der Medianfilter zusätzlich auf die Daten aus einzelnen Teilregionen angewendet wird und die Ergebnisse miteinander verglichen werden. Stimmen die Variationen aller Daten und die der Teilregionen überein, so liegt die Ursache bei der Station. Es können mit dieser Methode nur systematische Veränderungen aufgespürt werden, nicht aber statistischen Fehler in den Einzelmessungen. Überraschender Weise zeigen ungefähr die Hälfte aller untersuchten Stationen Variationen von mehr als 0.5 s. In den meisten Fällen geht eine starke Veränderung des Median mit einer zeitlichen Lücke in den Daten einher. Diese Beobachtung läßt vermuten, daß Veränderungen an Seismometern, Aufnahme- oder Analyseverfahren eine häufig vorkommende Ursache sind. Tatsächlich konnte dies in verschiedenen Fällen nach Rücksprachen mit den jeweiligen Stationen bestätigt werden. Allerdings sind auch andere Ursachen für einzelne Stationen wichtig. Das Yellowknife-Array (YKA, Kanada) z.B. zeigt jahreszeitliche Veränderungen, welche auch in dem dortigen Rauschniveau und der Anzahl der beobachteten Erdbeben wiederzufinden sind. Vor allem Einsatzzeiten von kleinen Erdbeben werden in den Sommermonaten deutlich weniger und gemittelt später abgelesen als im Winter. Die wahrscheinlichste Interpretation ist, daß im Sommer Ersteinsätze im Rausch untergehen und dadurch zu spät oder gar nicht bestimmt werden. Für die Station MAG (Magadan 1, Sibirien) weisen die Laufzeiten von verschiedenen Regionen ein deutlich unterschiedliches Verhalten auf ab April 1972. Dies kann nur durch eine Versetzung der Station um ca. 100 km erklärt werden, welche nicht an den ISC durchgegeben wurde. Somit wurden alle Laufzeitresiduen ab diesem Datum für falsche Stationskoordinaten berechnet.

Systematische Fehler können, im Gegensatz zu statistischen, Untersuchungen eher negativ beeinflussen, da sie sich, abhängig von der jeweiligen Untersuchung, nur bedingt gegenseitig aufheben. Darum werden ihre Auswirkungen auf globale Laufzeittomographie in **Kapitel 3** untersucht. Tomographie ist wahrscheinlich eines der interessantesten Anwendungsgebiete, wofür die ISC-Daten verwendet werden. Die tomographische Inversion von synthetischen Daten, welche die im vorigen Kapitel diskutierten Fehler simulieren, zeigt, daß die Geschwindigkeitsamplituden relativ klein sind gegenüber den Inversionsresultaten für ISC-Daten [Bijwaard *et al.*, 1998]. Außerdem zeigt das Modell wenig systematische Muster. Allerdings ist die Varianz der simulierten Laufzeitresiduen auch sehr viel kleiner als die der echten Daten. Relativ zu dieser Varianz propagieren die Fehler genauso stark in das dreidimensionale Modell wie die Strukturinformation, welche in den echten Daten vorhanden ist.

Dieses Ergebnis unterscheidet sich deutlich von dem, welches man durch die Inversion von statistischen Fehlern bekommt. Ein statistischer Fehler entsteht z.B. durch die endliche Ablesegenauigkeit. Es kann gezeigt werden, daß bei vielen

Stationen Ersteinsätze nur gerundet auf die nächste Sekunde angegeben werden, obwohl der entsprechende Parameter in den ISC-Daten eine Ablesegenauigkeit von 0.1 s angibt. Hierdurch entsteht ein Fehler, welcher zufällig zwischen -0.5 s und 0.5 s verteilt ist. Auch die Varianz dieses Fehlers ist, verglichen zu der Gesamtvarianz aller *P*-Laufzeitresiduen, sehr klein. Zusätzlich werden die Fehler durch die Mittelung vor und während der Inversion sehr stark unterdrückt und sind vernachlässigbar für tomographische Zwecke.

Im **Kapitel 4** wird die Erde statistisch beschrieben als ein Zufallsmedium und es wird versucht, die Amplituden und Korrelationslängen der Geschwindigkeitsanomalien aus den statistischen Eigenschaften der Laufzeitresiduen als Funktion der Tiefe zu invertieren. Hierzu wird zunächst die Varianz der Daten abhängig von Epizentralabstand und Herdtiefe und die Korrelation zwischen benachbarten Strahlen gemessen. Diese Größen können durch eine lineare Inversion in die Kovarianzfunktion der Mantelstruktur umgerechnet werden.

Die Untersuchungen in diesem Kapitel werden jedoch durch verschiedene Punkte erschwert. (1) Das relativ häufige Vorkommen von sehr großen (absoluten) Laufzeitresiduen gegenüber einer Normalverteilung führt zu einer starken Abhängigkeit der Momente zweiter Ordnung (Varianz und Kovarianz) vom gewählten Abschneidewert. Die bivariate Verteilung der Laufzeitresiduen von ähnlichen Strahlen zeigt, daß die meisten Residuen in den Ausläufern (große absolute Residuen) nicht korreliert sind. Dies impliziert, daß diese von falsch abgelesenen Einsatzzeiten stammen oder der Einsatz fälschlicherweise der *P*-Phase zugeordnet wurde. Letzteres ist vor allem für flache Beben und positive Laufzeitresiduen zu erwarten, da an der Oberfläche reflektierte Wellen (z.B. *pP*, *sP*) kurz nach den *P*-Wellen eintreffen. (2) Fehllokalisierungen der Hypozentren führt zu systematischen Abweichungen der Laufzeitstatistiken. Dies kann teilweise unterdrückt werden durch das Verwenden von Differenzen der Laufzeitresiduen. Die Auswirkung horizontaler Fehllokalisierungen ist allerdings nur sehr schwer abzuschätzen, da sie von der räumlichen Verteilung der seismologischen Stationen abhängt. (3) Die statistischen Eigenschaften von Laufzeitresiduen variieren von Gebiet zu Gebiet. Dies zieht die Annahme in Zweifel, daß die Amplituden und die Größe der lateralen Geschwindigkeitsanomalien lediglich von der Tiefe abhängen. Trotz dieser Komplikationen weisen die Daten auf die Anwesenheit von Heterogenitäten mit einer Größe von weniger als 100 km im oberen und unteren Erdmantel hin.

Kapitel 5 unterscheidet sich deutlich von den vorangegangenen Kapiteln sowohl in den Ausgangsdaten als auch in den verwendeten Methoden. Außerdem beschränkt sich die Abhandlung auf den kontinentalen oberen Mantel (bis zu einer Tiefe von etwa 400 km). In diesem Kapitel wird nach der Ursache für die großräumigen positiven Geschwindigkeitsanomalien unter alten kontinentalen Schilden gesucht. In wellenformtomographischen Modellen sind dies die mit Abstand auffallendsten Strukturen. Zunächst wird die seismische Struktur verglichen mit einer Aufteilung der Kontinente in drei verschiedene Typen tektonischer Provinzen und mit dem in Bohrlöchern gemessenen Wärmestrom. In

den obersten 200 km besteht eine deutliche Korrelation zwischen tektonischer Provinz und Geschwindigkeit, sowie eine Antikorrelation zwischen Geschwindigkeit und Wärmestrom. Da erwartet werden kann, daß Unterschiede in der chemischen Zusammensetzung einen deutlich geringeren Einfluß auf die seismischen Geschwindigkeiten haben als Temperaturvariationen, werden die Geschwindigkeitsprofile, gemittelt über die einzelnen Provinzen und unter der Annahme einer konstanten chemischen Zusammensetzung, in Temperaturprofile umgerechnet. Dies ergibt Geothermen, welche in 100 km Tiefe Temperaturunterschiede von bis zu 900°C aufweisen. In größerer Tiefe nähern sie sich einer gemeinsamen Manteladiabaten an. Unterhalb von 230 km überschreitet der maximale Temperaturunterschied 300 °C nicht. Ein Vergleich zeigt eine sehr gute Übereinstimmung mit thermischen Modellen, welche den Oberflächenwärmestrom zu 40% radioaktiven Elementen in der oberen Kruste und die verbleibenden 60% einer sehr tiefen Quelle zuschreiben. Diese Partition benötigt eine feste Relation zwischen oberflächennaher Radioaktivität und Dicke der thermischen Grenzschicht, definiert als die Distanz zwischen Erdoberfläche und ungefährem Schnittpunkt der konduktiven Geotherme mit der Manteladiabate. Um die Kopplung dieser zwei Größen zu erklären wird vorgeschlagen, daß die Dicke der Grenzschicht durch unterschiedliche Abschirmung der radioaktiven Kruste verursacht wird. Dabei führt die Beseitigung von radioaktiven Elementen durch Erosion zu einer dicker werdenden Grenzschicht. Oberflächennahe Prozesse bestimmen in diesem Szenario die thermische Struktur des kontinentalen Mantels bis zu einer Tiefe von mehreren 100 km.

Acknowledgments

I would like to take this opportunity to sincerely thank my promoter Roel Snieder for his guidance, constructive criticism and stimulating advice. Furthermore, in times when progress was slow his encouragement was extremely valuable.

I am equally indebted to Jeannot Trampert. I especially appreciate that he invested much of his time to understand my questions when he took over the supervision during Roel's sabbatical. During this time and thereafter we had many fruitful discussions and excellent cooperation on all parts of this thesis.

On many occasions I profited from Hanneke Paulssen's expertise on questions concerning travel time measurements, a topic where I first had a naive view on but which turned out to be of much importance for my research. Her creativity and help for finding the correct information are very much appreciated.

I very much enjoyed the collaboration with Harmen Bijwaard on Chapter 3 and Saskia Goes on Chapter 5. Many thanks also to Wybo Gardien for providing Figure 1.3 and the interesting discussions about the quality of arrival time measurements.

Joop Hoofd and Theo van Zessen were always helpful with all kinds of computer related questions. Erhard Wielandt was so kind to provide Figures 1.4 and 1.5. Bob Engdahl allowed me to use the reprocessed data set and provided helpful information on it. Wim Spakman placed his computer to my disposal that I could retrieve the travel time data from it. Their help is greatly acknowledged.

I am grateful to Ulrich Hansen, Roland Klees and Kabir Roy-Chowdhury of my supervision committee and to John Scales, Chris Spiers, Cees Wapenaar and Rinus Wortel of the dissertation committee.

Dirk Kraaijpoel, Sandra Richwalski and Martin Schimmel always had time for my – trivial as well as difficult – questions and I benefited very often from their help. Many other people contributed to my research by informative discussions and to the good atmosphere I enjoyed at Utrecht University: Dolors Alsina, Caroline Beghein, Susanne Buitter, Eugenio Carminati, Hana Cizkova, Andrew Curtis, Frederic Deschamps, Rob Devilee, Harm Dorren, Paul Fleuren, Stephanie Godey, Grzegorz Krupa, Anthony Lomax, Henk Maquering, Paul Meijer, Everhard Muijzert, Filip Neele, Sander van Ouwerkerk, Tijn Passier, Jörg Schmalzl, Stan Schoofs, Jeroen de Smet, Jesper Spetzler, Volker Steinbach, Bernadine Vet, Arie van Wetum and many others.

I thank Irid Röhm-Kloster and Martin Kloster for proofreading most of this thesis. My parents in law were always helpful in all questions about living in the Netherlands and contributed greatly to the pleasant time I had from the first day I came to this country. I am grateful to my parents for the very best support I could think of throughout my life. Finally, I thank Hanny for her love, understanding, encouragement and help.

Almost all figures in this thesis have been produced with the Generic Mapping Tools (GMT) of Wessel and Smith [1998]. This research was supported by the Netherlands Organization for Scientific Research (NWO) through the Pionier project PGS 76-144.

Curriculum Vitae

December 19, 1967 Born in Stuttgart – Bad Cannstatt, Germany
1974 - 1978 Primary school in Schorndorf, Germany
1978 - 1987 Secondary school (Gymnasium) in Schorndorf, Germany
1987 - 1988 Military service in Ingolstadt, Germany
1988 - 1991 Study of Geophysics, University of Karlsruhe, Germany
1991 - 1995 Study of Physics, University of Karlsruhe, Germany
1995 - 1999 PhD student at Utrecht University, The Netherlands