![KIT logo]

Karlsruhe Institute of Technology

Determination of Seismic Waveforms Using Passive Seismic Interferometry around Landau, SW Germany

Bestimmung seismischer Wellenformen mittels passiver seismischer Interferometrie im Bereich Landau, SW Deutschland

Diploma Thesis of

Felicitas Stein

At the Faculty of Physics
Geophysical Institute (GPI)

Reviewer: PD Dr. Joachim Ritter
Second reviewer: Prof. Dr. Friedemann Wenzel
Advisor: Dr. Jörn Groos

Duration: 10th July, 2012   –   10th July, 2013

www.kit.edu

I declare that I have developed and written the enclosed thesis completely by myself, and have not used sources or means without declaration in the text.

PLACE, DATE

. . . .
(**Felicitas Stein**)

Deutsche Zusammenfassung

In der Pfalz in Südwestdeutschland treten seit 2006 vermehrt mikroseismische Ereignisse ($M_L > 3$) auf (Plenkers et al., 2013; Groos et al., 2013). Diese seismischen Ereignisse werden durch zwei Tiefengeothermie-Anlagen induziert, von denen sich die eine in Landau und die andere in Insheim befindet. Einige der stärkeren Beben (M_L 2,4 - 2,7) sollen laut Berichten der Einwohner zu geringfügigen Gebäudeschäden geführt haben. Aufgrund dieser Beben errichtete das Geophysikalische Institut des Karlsruher Instituts für Technologie zur Überwachung der induzierten Seismizität 2009 ein Netzwerk seismischer Messstationen, das TIMO2-Netzwerk (TIMO: Tiefenstruktur des mittleren Oberrheingrabens).

Vor allem im Bereich der Reservoir-Ingenieurwissenschaften ist es von großem Interesse, in welchem Zusammenhang die induzierte Seismizität mit Änderungen der physikalischen Eigenschaften des Untergrunds steht. Es stellt sich die Frage, wie diese Änderungen abgebildet und überwacht werden können. Im Rahmen dieser Diplomarbeit wurde untersucht, ob mittels passiver seismischer Interferometrie im Bereich Landau eine Abbildung des Untergrunds und dessen zeitliche Änderungen mit dem TIMO2-Datensatz bestimmt werden können.

Die grundlegende Datenbearbeitungstechnik der passiven seismischen Interferometrie besteht in der Kreuzkorrelation zweier an verschiedenen Orten aufgezeichneter Zeitreihen des seismischen Rauschens, wobei letzteres den ununterbrochenen und überall vorhandenen Vibrationen der Erdoberfläche in einem sehr breiten Frequenzspektrum entspricht (siehe Kapitel 2.2). Bei der Kreuzkorrelation werden die beiden Zeitreihen zueinander verschoben und man erhält je nach Verschiebungsrichtung negative oder positive Zeitverschiebungswerte, die akausaler und kausaler Teil der Kreuzkorrelation genannt werden. Über die Berechnung von Kreuzkorrelationen können seismische Wellen (siehe Kapitel 2.1), die zwischen zwei Messstationen propagieren, aus dem seismischen Rauschen extrahiert werden. Das heißt, ohne auf aktive kontrollierte Quellen (z.B. Sprengungen) oder auf Erdbebenaufzeichnungen angewiesen zu sein, ist es möglich, Wellenformsektionen zu erstellen, aus denen wiederum Informationen über die Eigenschaften des Mediums gewonnen werden können.

Voraussetzung für die Erstellung von Tomographien des Untergrunds (Imaging) mittels passiver seismischer Interferometrie ist, dass sich die Kreuzkorrelationsfunktion der Antwortfunktion (der Greenschen Funktion) des Untergrunds zwischen den beiden Empfängern annähert. Als Kriterien werden eine Symmetrie der Kreuzkorrelation sowie ein hohes Signal-zu-Rausch-Verhältnis (SNR) herangezogen. Die Voraussetzung für die Bestimmung zeitlicher Änderungen (Monitoring) des Mediums mittels passiver seismischer Interferometrie ist mit dem Vorhandensein zeitlich stabiler Signale in den Kreuzkorrelationen gegeben. Erfolgreich angewendet wird die passive seismische Interferometrie seit zirka zehn Jahren und hat sich seitdem zu einem sehr großen Forschungsbereich in der Seismologie entwickelt. Für detailliertere Informationen zur Theorie der passiven seismischen Interferometrie siehe Kapitel 2.3.

In dieser Arbeit wurde überprüft, ob die oben genannten Voraussetzungen vom TIMO2-Datensatz erfüllt werden. Verwendet wurden dafür zwölf TIMO2-Stationen, welche in und in der nahen Umgebung von Landau aufgestellt sind. Der minimale Abstand zwischen den

verwendeten Stationen beträgt zirka 1 Kilometer, der maximale Abstand beträgt ungefähr 22 Kilometer. Die Messinstrumente, die Teil des KArlsruher Breitband Arrays (KABBA) sind, zeichnen die Bodenschwinggeschwindigkeit in drei Richtungen auf: in vertikaler (Z) Richtung und in den horizontalen Richtungen Nord-Süd (N) und Ost-West (E). Weitere Informationen zum Datensatz können Kapitel 3 entnommen werden.

Um stabile Interferogramme zu erhalten, wurden die folgenden Datenbearbeitungsschritte auf die kontinuierlichen Zeitreihen der ausgewählten Stationen angewendet: Zuerst wurden die tageweise vorliegenden Zeitreihen der Stationen vorprozessiert. Der Mittelwert, der lineare Trend und die Instrumentenantwort wurden entfernt. Ein Hochpassfilter von 0,1 Hertz wurde angewendet und die Abtastrate auf 100 Hertz gesetzt. Danach wurden die Kreuzkorrelationsfunktionen aller möglichen paarweisen Stationskombinationen mit 20 Minuten langen (vertikale Komponenten) bzw. (aus Speicherplatzgründen) mit einer Stunde langen (horizontale Komponenten) Abschnitten der vorprozessierten Zeitreihen berechnet. Die Kreuzkorrelationen der horizontalen Komponenten wurden so rotiert, dass die eine Achse in Richtung der Verbindungslinie des Stationspaars (radial-radial, RR) zeigt und die andere Achse senkrecht dazu steht (transversal-transversal, TT). Die Rotation erlaubt, auf den einzelnen Komponentenkombinationen seismische Wellentypen aufgrund deren verschiedener Polarisationen zu identifizieren. Alle Kreuzkorrelationen wurden in einem letzten Schritt im Frequenzbereich normiert. Daraufhin konnten die Kreuzkorrelationen je nach Anwendung gestapelt (z.B. zu Tages- oder Jahres-Kreuzkorrelationen) und/oder gefiltert werden (siehe Kapitel 4). Die Datenbearbeitung orientiert sich an der Arbeit von Groos et al. (2012).

Um einen Überblick über die große Datenmenge zu erhalten, wurden die gestapelten Kreuzkorrelationen des Jahres 2012 der ZZ, der RR und der TT Komponentenkombinationen über dem jeweiligen Stationsabstand aufgetragen und in neun verschiedenen Frequenzbändern (alle Bänder liegen zwischen 0,1 Hertz und 45 Hertz) auf ihre Eigenschaften hin untersucht (siehe Unterkapitel 5.1.1). Im Allgemeinen nimmt das Verhältnis von kohärenten Signalen – Rauschen, das von gleichen Quellen stammt und sowohl an der einen als auch an der anderen Station gemessen wird – zu inkohärenten Signalen – Rauschen, das von unabhängigen, separaten Quellen stammt – mit zunehmendem Abstand und zunehmender Frequenz ab. Daraus wird geschlossen, dass sich tieffrequentere Signale kohärenter über weitere Entfernungen ausbreiten als hochfrequentere Signale. Im Frequenzbereich zwischen 0,1 Hertz und 0,8 Hertz sind überwiegend im akausalen Teil der Kreuzkorrelationen Signale vorhanden. Dies lässt sich auf eine ungleichmäßige Verteilung der Rauschquellen zurückführen sowie auf eine Ausbreitung des seismischen Rauschwellenfeldes um Landau in ungefähr West-Ost-Richtung. Im Frequenzband zwischen 0,8 Hertz und 1,6 Hertz ist das Verhältnis von Signalen im akausalen und kausalen Bereich relativ ausgeglichen. Auffallend ist, dass Signaleinsätze in den ZZ und RR Kreuzkorrelogrammen schwerer bestimmbar sind als in den TT Kreuzkorrelogrammen. Dies wurde bisher auch in anderen Arbeiten (z.B. Jay et al., 2012; Behm and Snieder, 2013) beobachtet. In höheren Frequenzbändern nimmt einerseits das Signal-zu-Rausch-Verhältnis (SNR) stark ab und andererseits tauchen auch teilweise schwebungsartige „Signale" in den Kreuzkorrelationen auf. Wie diese schwebungsartigen „Signale" entstehen, konnte im Rahmen der Diplomarbeit nicht abschließend geklärt werden. Es wird vermutet, dass die „Schwebungen" beim Kreuzkorrelationsprozess entstehen. Angenommen an zwei Stationen werden unabhängig voneinander sinusoidale Signale aufgezeichnet, die sich nur geringfügig in ihren Frequenzen unterscheiden, so ergeben sich in der Kreuzkorrelation dieser Aufzeichnungen „Schwebungen". Dies wurde anhand synthetischer Sinussignale überprüft (Anhang C).

Das Frequenzband zwischen 0,8 Hertz und 1,6 Hertz erschien für eine genauere Untersuchung der ZZ, RR und TT Kreuzkorrelationen hinsichtlich ihrer Anwendbarkeit der Imaging Technik am geeignetsten. Zum einen liegt in diesem Frequenzband ein relativ hohes Signal-zu-Rausch-Verhältnis vor und der Anteil an Signalen im kausalen
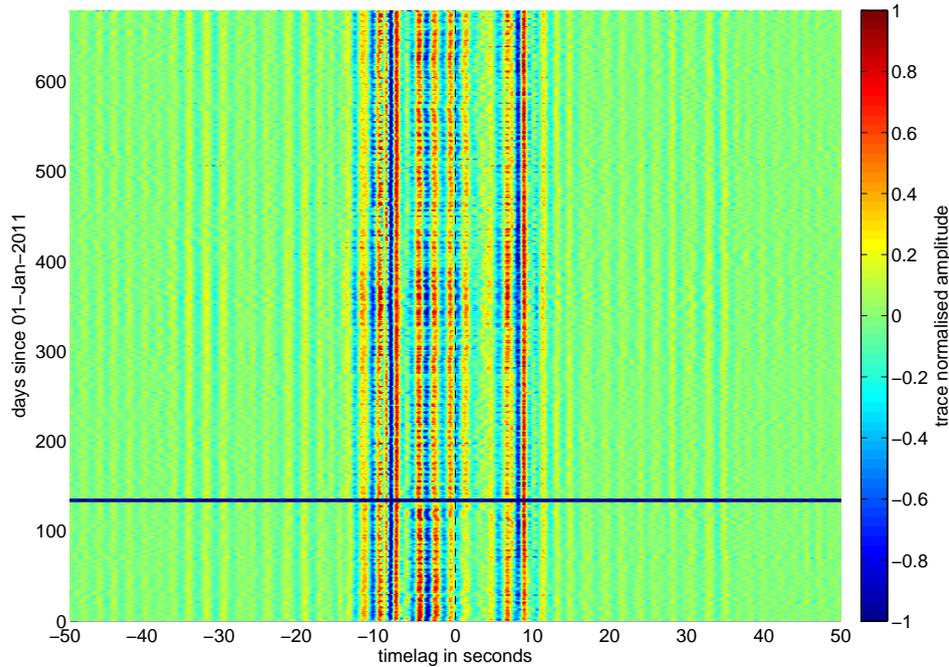
Figure 1.: Matrixplot der vorhandenen *TT* Ein-Tages-Kreuzkorrelogramme aus den Jahren 2011 und 2012 des Stationspaares TMO53-TMO54. Die Kreuzkorrelationen sind zwischen 0,4 Hertz und 0,8 Hertz gefiltert. Deutliche sowie stabile Signale tauchen um Zeitverschiebungswerte von -10 Sekunden und +10 Sekunden auf. Auch die Coda der Kreuzkorrelogramme enthält zeitlich stabile Signale. Die Tage, an denen keine Daten vorhanden sind, sind dunkelblau dargestellt.

und akausalen Teil ist ausgewogen. Zum anderen sind die Kreuzkorrelogramme nicht durch „Schwebungen" beeinflusst. Somit wurden die Kreuzkorrelationen, gefiltert zwischen 0,8 Hertz und 1,6 Hertz, dreier Stationspaare mit unterschiedlichen Stationsabständen und Azimuten näher analysiert (siehe Unterkapitel 5.1.2).

Bei dieser Untersuchung ergab sich, dass einerseits die *ZZ*, *RR* und *TT* Kreuzkorrelogramme aller drei Stationspaare innerhalb eines gewissen Zeitverschiebungsfensters (mit einer Ausnahme) symmetrische Signale aufweisen und andererseits, dass das SNR auf den *TT* Komponenten wesentlich höher als auf den *ZZ* oder *RR* Komponenten ist. Von daher erscheinen symmetrische Signale auf *TT* viel deutlicher als auf *ZZ* oder *RR*. Schlussfolgernd kann man annehmen, dass sich die *TT* Kreuzkorrelationsfunktionen am besten der Antwortfunktion zwischen jeweils zwei Stationen annähern. Das wiederum bedeutet, dass diese Komponentenkombination für die Erstellung eines (Lovewellen-) Tomographie-Modells um Landau verwendet werden kann. Die bestimmten Geschwindigkeiten der Signale zwischen zirca 300 Metern pro Sekunde und 410 Metern pro Sekunde sprechen für Oberflächenwellen. Mit Hilfe von Polarisationsdiagrammen wurde auch die Partikelbewegung der Wellen untersucht. Die für Rayleighwellen erwartete Elliptizität in der *RR-ZZ*-Ebene ergab sich nur bei einem Stationspaar eindeutig.

Um die zeitliche Stabilität der Signale in den Kreuzkorrelationen zu untersuchen, wurden beispielhaft die Daten nur eines Stationspaars (TMO53-TMO54) näher untersucht (siehe Unterkapitel 5.1.3). Da durch das Stapeln der Kreuzkorrelationen eines Tages schon stabile Interferogramme erzeugt werden können, wurden diese Ein-Tages-Kreuzkorrelogramme in sogenannten Matrixplots abgebildet, um über einen langen Zeitraum stabile Signale identifizieren zu können (siehe Abbildung 1). Zur Erstellung der Matrixplots werden in einem ersten Schritt unter Verwendung einer Farbskala die Amplitudenwerte der Ein-Tages-

Kreuzkorrelogramme dargestellt. Das heißt, man erhält eine Aufsicht auf das Kreuzkorrelogram mit in blau dargestellten negativen Amplituden und in rot dargestellten positiven Amplituden. Diese Aufsichten der chronologisch geordneten Ein-Tages-Kreuzkorrelogramme werden Zeile für Zeile in einer Matrix dargestellt. In diesem Fall wurden alle vorhandenen Ein-Tages-Kreuzkorrelogramme aus den Jahren 2011 und 2012 verwendet. Bei den untersuchten Frequenzen zwischen 0,1 Hertz und 1,6 Hertz weisen die Kreuzkorrelogramme aller drei Komponentenkombinationen (ZZ, RR und TT) zeitlich stabile Signale über den gesamten untersuchten Zeitraum von fast zwei Jahren auf. Nicht nur stabile ballistische Wellen treten in den Kreuzkorrelogrammen auf, sondern auch in der Coda der Kreuzkorrelogramme sind zeitlich stabile Signale zu erkennen, die wahrscheinlich gestreuten und reflektierten Wellen entsprechen. Die Reproduzierbarkeit, die Wiederholbarkeit und die Stabilität der Kreuzkorrelogramme bedeuten, dass der TIMO2-Datensatz die Voraussetzungen für die Bestimmung zeitlicher Variationen der seismischen Geschwindigkeiten erfüllt.

Desweiteren wurde bei der Untersuchung der zeitlichen Stabilität der Signale in den zwischen 0,8 Hertz und 1,6 Hertz gefilterten Kreuzkorrelogrammen festgestellt, dass Werktags-/Wochenendschwankungen in der Amplitudenstärke vorhanden sind. Daraus wurde geschlossen, dass die Quelle(n), die dieses Signal in den Kreuzkorrelationen erzeugen, auf menschliche Aktivitäten zurückgeführt werden können. Nicht nur die zeitlichen, sondern auch die spektralen Eigenschaften dieser Rauschsignale weisen auf künstlich erzeugte Quellen hin. Im Allgemeinen kann man das seismische Rauschen unterhalb von ungefähr 1 Hertz natürlichen (ozeanischen, meteorologischen, ...) Aktivitäten und oberhalb von zirka 1 Hertz künstlichen Quellen zuschreiben. Allerdings stellt diese spektrale Grenze von 1 Hertz nur einen groben Anhaltspunkt dar und die tatsächliche Unterteilung des Ursprungs des Rauschens richtet sich nach den örtlichen Begebenheiten (e.g. Bonnefoy-Claudet et al., 2006; Groos and Ritter, 2009).

Wie weiter oben bereits erwähnt, weisen die Kreuzkorrelogramme bei Frequenzen zwischen 0,1 Hertz und 0,8 Hertz im akausalen Teil mehr amplitudenstarke Signale auf als im kausalen Teil. Zwischen 0,1 Hertz und 0,4 Hertz kann diese Asymmetrie durch einen ebenen Welleneinfall aus Nord-West Richtung auf das Netzwerk erklärt werden (siehe Unterkapitel 5.2.1). Aufgrund der Richtung und des relativ niedrigen Frequenzbandes kann davon ausgegangen werden, dass es sich um die Meeresmikroseismik handelt, die durch die Interaktion von Wellen mit dem Meeresboden im Atlantischen Ozean und in der Nordsee um die Britischen Inseln und Frankreich erzeugt wird. Im Frequenzband zwischen 0,4 Hertz und 0,8 Hertz ist die Meeresmikroseismik nicht mehr alleinige Ursache für die Signale in den Kreuzkorrelogrammen. Unter der Annahme, dass eine dominante Punktquelle innerhalb des Netzwerks die Signale in den Kreuzkorrelogrammen erzeugt, wurde eine sogenannte Migrationsanalyse nach Horstmann (2010) durchgeführt (Unterkapitel 5.2.2). Bei dieser Analyse wird ein Raster hypothetischer Punktquellen über das zu untersuchende Gebiet gelegt. Die Geschwindigkeit der Wellen, die von jeder Punktquelle emittiert werden, wird einheitlich auf einen festen Wert festgelegt. Anschließend wird ermittelt, wie gut die Signale in den Kreuzkorrelogrammen durch die einzelnen Punktquellen erklärt werden können. Im Frequenzband zwischen 0,8 Hertz und 1,6 Hertz kann eine weitere dominante Rauschquelle im Westen des Stationsnetzwerkes nicht ausgeschlossen werden (siehe Abbildung 2).

Um diese offene Frage zu beantworten, wurden im Januar 2013 zwei weitere seismische Stationen im Süd- und Nord-Westen des TIMO2-Netzwerkes aufgestellt. Allerdings waren bis zum Ende dieser Arbeit (Juli 2013) noch nicht genug Daten von diesen beiden Stationen vorhanden, als dass damit erneut eine Migrationsanalyse hätte berechnet werden können. Das liegt daran, dass eine der Stationen kurz nach Aufbau aufgrund eines Defekts in der Stromversorgung über einen Zeitraum von zirka zwei Monaten keine Daten aufzeichnete. Im Frequenzband zwischen 0,8 Hertz und 1,6 Hertz wurde ebenso eine Migrationsanalyse
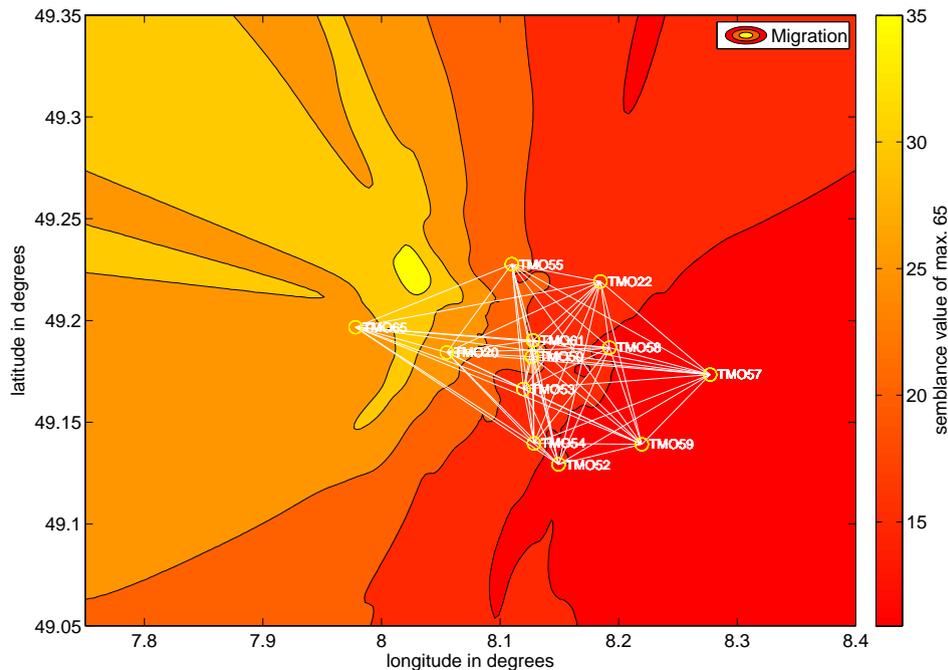
Figure 2.: Das beste Ergebnis der Migrationsanalyse mit den ZZ Jahres-Kreuzkorrelogrammen aus 2012 von 65 Stationspaaren gefiltert zwischen 0,4 Hertz und 0,8 Hertz. Die vordefinierte Geschwindigkeit beträgt 600 Meter pro Sekunde. Die Stationen, die für die Analyse verwendet wurden, sind durch Kreise dargestellt und mit Linien verbunden. Eine Quelle in dem Gebiet (gelb), das sich um $49,2277°$ N und $8,0237°$ E befindet, erklärt die Daten in den Kreuzkorrelationen am besten (zu 57%). Außerdem erstreckt sich eine Fläche, wo eine relativ große Übereinstimmung mit den realen Daten erhalten wird, nach Nordwesten. Dies kann ein Hinweis auf die Meeresmikroseismik sein.

durchgeführt (Unterkapitel 5.2.2). Diese ergab eine dominante Rauschquelle im Südosten des Netzwerks. In dem relativ kleinen Gebiet, das dabei lokalisiert wurde, liegt der Windpark von Offenbach an der Queich mit drei Windrädern. Grundätzlich ist es möglich, dass dieser Windpark eine dominante Rauschquelle für die Signale in den Kreuzkorrelogrammen darstellt. Um diese Vermutung zu bestätigen, müsste allerdings noch untersucht werden, ob auch ein zeitlicher Zusammenhang zwischen dem Auftreten der Signale in den Kreuzkorrelogrammen und den zeitlichen Variationen der Windgeschwindigkeiten besteht.

Abschließend ist zu bemerken, dass die geothermischen Reservoire in einer Tiefe von ungefähr drei Kilometern liegen und somit die in den Kreuzkorrelationen beobachteten ballistischen Oberflächenwellen für eine Abbildung oder zeitliche Überwachung dieser Reservoire nicht geeignet sind. Aufgrund der relativ kleinen Stationsabstände (maximal ~ 22 Kilometer) können bei Einhaltung des oft genannten Kriteriums, dass der Abstand zwischen zwei Stationen größer als das Dreifache der Wellenlänge λ sein sollte (Lin et al., 2008), maximal Wellen mit einer Wellenlänge von 7,3 Kilometern untersucht werden. Desweiteren nehmen die Amplituden A der Oberflächenwellen exponentiell mit der Tiefe T ab: $A \sim \exp\left(-\frac{2\pi}{\lambda} \cdot T\right)$ (Müller, 1973). Das heißt, in einer Tiefe von drei Kilometern wäre die Amplitude der Oberflächenwellen nur noch ungefähr 0,08 mal so groß wie an der Erdoberfläche, was vernachlässigbar klein ist. Allerdings stellt sich die Frage, ob ein Teil der Wellen, die in der Coda der Kreuzkorrelogramme auftreten, sich in größere Tiefen als die ballistischen Wellen erstrecken und somit anhand dieser Codawellen mittels Monitor-

ing Techniken zeitliche Variationen bei den geothermischen Reservoiren bestimmt werden
können. Außerdem sollte untersucht werden, ob mittels passiver seismischer Interferome-
trie mit dem TIMO2-Datensatz auch ballistische Raumwellen sichtbar gemacht werden
können. Deren Ausbreitung beschränkt sich nicht nur auf die nahe Oberfläche und somit
sind diese Wellen auch für die Untersuchung des Untergrunds in großen Tiefen geeignet.

Contents

1. Introduction

Since 2006, microearthquakes have occurred in the surroundings of two geothermal power plants located at the cities of Landau and Insheim in Rhineland Palatinate in Southwest Germany (Plenkers et al., 2013; Groos et al., 2013). Some of the larger earthquakes (M_L 2.4 to 2.7) are suspected to have caused minor damage to buildings as reported by residents. Thus, the Geophysical Institute of the Karlsruhe Institute of Technology has operated the seismic station network TIMO2 (TIMO: Deep structure of the Central Upper Rhine Graben) since 2009 to monitor the induced seismicity related to these geothermal systems. For reservoir engineering the relation between the induced seismicity and the changes of the physical parameters of the subsurface is of great interest. But how can these changes be imaged or monitored?

By cross-correlating the time series of two stations and applying normalisation methods to the cross-correlations, it is possible to extract waves from seismic noise, which propagate between the two stations (coherent waves). This methodology of turning passive seismic measurements into deterministic responses is called passive seismic interferometry (Wapenaar et al., 2010).

Since the past ten years, especially since the work of Campillo and Paul (2003), (passive) seismic interferometry has evolved to a large field of research in seismology. However, not only in seismology but also in ultrasonics (Weaver and Lobkis, 2001, 2002) and underwater acoustics (Roux and Fink, 2003) waves, which propagate between two sensors, are retrieved by cross-correlating noise recordings, and details about the intervening medium are derived (Wapenaar et al., 2006).

Even though the theory behind passive seismic interferometry makes many assumptions regarding the properties of the seismic wavefield, and of the medium, as well as on the distribution of the noise sources, interferometric methods have successfully been implemented in seismic data analysis. The main applications of passive seismic interferometry, which have emerged over the past decade, can be divided into the fields of noise-based imaging and noise-based monitoring. On the one hand, seismic tomography models of the subsurface are created. Very often surface wave tomography models are generated due to the dominance of surface waves in the cross-correlations. But also body wave tomography models were determined (for example Brenguier et al., 2007). The scales on which these tomography models are determined vary from local to continental scales (e.g. on smaller scales Picozzi et al., 2009; Behm and Snieder, 2013; on larger scales Lin et al., 2008; Poli et al., 2013). On the other hand, temporal petrophysical variations of the medium are inferred from relative changes of seismic velocities over time (e.g. Sens-Schönfelder and Wegler, 2006; Mainsant et al., 2012; Hobiger et al., 2012). Hadziioannou et al. (2009)

showed that the requirements for monitoring are weaker than for imaging when passive seismic interferometry is used. While for imaging the cross-correlation should converge to the response between two receivers, for monitoring temporally stable signals in the cross-correlations are sufficient.

Furthermore, ambient noise cross-correlations are not only used for imaging or monitoring purposes but also, for example, for the localisation of noise sources (e.g. Horstmann, 2010; Ma et al., 2013) or for the time synchronisation of seismic networks (Sens-Schönfelder, 2008).

One of the most important advantages of passive seismic interferometry is the fact that no active, controlled sources are necessary, and that one is independent of recordings of earthquakes. Hence, if the data set is suited for passive seismic interferometric applications, tomography models and temporal changes of the medium can be gained cost-effectively and without great field work effort. Therefore, using the continuous seismic time series, recorded by the three-component TIMO2 stations, the basic question which will be studied and answered within this thesis reads as follows: Does the TIMO2 data set meet the requirements for passive imaging and monitoring? In this context three other questions need to be answered:

1. Is it possible to obtain stable interferograms with the TIMO2 data set?

2. What are the properties – with respect to coherent signals – of the cross-correlations?

3. Do specific noise sources dominate the cross-correlations?

The following thesis starts with a short introduction to seismic waves (Section 2.1) and seismic noise (Section 2.2). Afterwards, the basics of (passive) seismic interferometry and its main applications are briefly explained (Section 2.3). In Chapter 3 the used TIMO2 data set is introduced. The fieldwork, which was done in the scope of this thesis, is described in Section 3.2. Chapter 4 deals with the data processing, which follows Groos et al. (2012). A focus is set on the rotation of the horizontal component cross-correlations (Section 4.4). The results are introduced in Chapter 5. Subsection 5.1.1 gives an overview of the properties of the cross-correlations of 65 stationpairs. In Subsection 5.1.2 more detailed information on the observed signals, which occur in the cross-correlations, are given using only three stationpairs. The temporal stability of signals, which is required for monitoring, is investigated in Subsection 5.1.3. The second part of Chapter 5 handles with the localisation of noise sources using two different approaches (Section 5.2).

2. Basic Principles of Seismic Noise Interferometry

In this chapter the basic principles of seismic interferometry are discussed with a focus on passive seismic interferometry and its applications. In comparison to active controlled sources (e.g. explosions) 'passive' means ambient ground motions. For the analysis of the ambient seismic noise field it is important to know the properties of seismic waves. Hence, the first section (2.1) of this chapter deals with seismic waves. Section 2.2 illuminates facts about seismic noise.

2.1. Seismic Waves

Basically, there are two different types of seismic waves: body waves and surface waves. Furthermore, one can distinguish between two kinds of body waves: the P- and the S-waves. The P-wave propagates in the same direction as the particles move. Therefore, the P-wave is also known as longitudinal, compressional or dilatational wave. It is comparable to sound waves. The propagation direction of the S-wave is perpendicular to the ground motion. Hence, the S-wave is also called transverse or shear wave. While the propagation of body waves is three-dimensional, surface waves – as their name already implies – propagate along the surface of a body. The amplitudes of the surface waves are significantly unequal to zero only in a depth range of the order of a few wavelengths. The most prominent surface waves are Rayleigh waves and Love waves. Rayleigh waves are typically retrograde elliptically polarised, Love waves are horizontally polarised. In contrast to body waves, the velocity of surface waves is strongly frequency dependent. This characteristic is called dispersion. For more theoretical background information please refer to Müller (1973) or Aki and Richards (2009). Figure 2.1 schematically shows the particle motions of the four different wave types.

If an earthquake occurs, the first type of wave, which will arrive at a recording station, will be the P-wave, which is therefore also called primary wave. It is followed by the S-wave (secondary wave). This order of appearance also explains the naming of the two body waves. After the S-wave, the surface waves arrive. The latter typically dominate the seismogram because the amplitude of the surface waves decreases slower with increasing distance to the source than the amplitude of the body waves. In general, the Love waves travel slightly faster than the Rayleigh waves (Müller, 1973). The absolute value of the propagation velocity of seismic waves depends on the elastic properties of the medium.

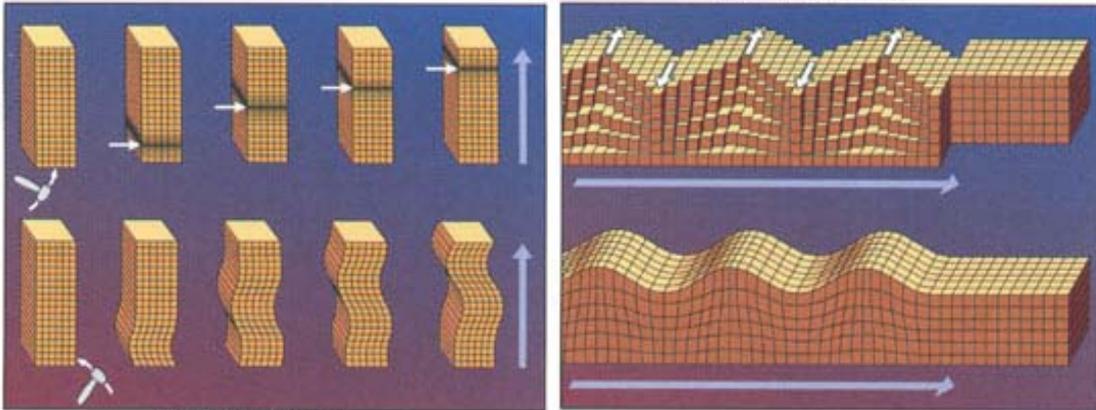Of course, not only earthquakes emit seismic waves. Seismic noise sources (see Section

Figure 2.1.: The left picture shows the particle motion of the seismic body waves. On top, the compression of the particles by an upward propagating P-wave is illustrated and denoted by the white arrows. At the bottom, an S-wave, which is also moving upwards, is depicted. The right figure shows the displacement by surface waves. On top, a Love wave is modelled. The ground motion is perpendicular to the propagation direction of the Love wave. Below you can see the Rayleigh wave with a compressional and a vertical portion of particle movement (http://earthquake.usgs.gov/learn/glossary/images/PSWAVES.JPG, 21 May 2013).

2.2), for example, also excite different types of waves. If one is able to extract signals out of seismic noise records (for example by using interferometric methods), it might be interesting which types of waves the signals consist of. One method for determining the types of waves makes use of the different polarisations of the waves. An important requirement for this method is a three-dimensional recording. Commonly, these three-dimensional recordings initially consist of one seismogram of the vertical ground motion and two other seismograms of the horizontal ground motions in north-south and in east-west direction. If these seismograms are rotated into an adequate coordinate system, the polarisation of wave groups can be determined, and in a next step, the type of wave might be identifiable (Plesinger et al., 1986). For example, if a source emits a Rayleigh wave, the vertical displacement will be recorded on the vertical component only. The compressional portion of the particle motion, however, will be recorded on both horizontal seismograms as long as one of these two horizontal axes lies not exactly in the same direction in which the Rayleigh wave propagates. By rotating the horizontal seismograms, it is possible to obtain the compressional portion of the Rayleigh wave on only one horizontal record. In Section 4.4 the rotation especially of cross-correlations is explained in more detail.

2.2. Seismic Noise

If data is acquired, noise is always superimposed on the records (Curtis et al., 2006). In seismology, where seismometers record the ground motion over time, the persistent presence of vibrations of the Earth (with its atmosphere) can be defined as seismic noise. The origin of seismic noise is often divided into two classes: in natural and in man-made - also called cultural or artificial - sources. Natural sources are, for example, tides, oceanic waves striking along the coasts, turbulent wind, effects of wind on uneven grounds, on trees or buildings, barometric pressure fluctuations, changes in temperature, and volcanic tremor (Bonnefoy-Claudet et al., 2006). Cultural noise sources are amongst others traffic on roads

or railways, pipelines, industrial machinery, producing oil or gas fields, and also transient signals like explosions e.g. in mines or human footsteps (Wilson et al., 2002). All seismic noise sources together comprise a very broad frequency band. Depending on the origin, spectral as well as temporal characteristics of the noise are different. Generally, noise induced by human activities exhibits daily and/or weekly variations. The frequencies of the cultural noise are mainly above 1 Hertz. Frequencies below 1 Hertz mostly correspond to signals emitted by natural noise sources (Bonnefoy-Claudet et al., 2006). The persistent ocean-generated microseismic noise, for example, has relatively low frequencies between 0.05 Hertz and 0.2 Hertz.

In general, the seismic noise field consists of all types of waves (see Chapter 2.1). However, the proportion between the different wave types, e.g. the proportion between Rayleigh and Love waves, depends on the source properties and the site conditions (Bonnefoy-Claudet et al., 2006). For more information on the origin and the nature of the ambient seismic noise field with many links to further reading please refer to Bonnefoy-Claudet et al. (2006) or Groos (2010).

Due to the above given information the seismic noise field might appear chaotic and random. However, it is possible to extract valuable information out of seismic noise recordings by applying simple processing methods to the data. Seismic noise recordings are often also called passive seismic recordings, because they do not contain signals emitted by active controlled sources, e.g. explosions. The methodology to turn passive seismic measurements into deterministic seismic responses is called passive seismic interferometry (Wapenaar et al., 2010). The basic principles of seismic interferometry are discussed in the following section (Section 2.3).

2.3. Seismic Interferometry

In the most general sense, interference is a superposition of waves (Lauterborn et al., 1993). Hence, interferometry can be defined as the study of interference phenomena between pairs of signals in order to obtain information from the differences between them. The underlying mathematical operation for this investigation is the cross-correlation of these two signals (Curtis et al., 2006). Assuming two time series $A(t)$ and $B(t)$, the analytical cross-correlation function (CCF) is defined as:

$$\widehat{R}_{AB}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int\limits_{-T/2}^{+T/2} A(t)B(t+\tau)dt \tag{2.1}$$

with the time window length T. Thus, the cross-correlation $\widehat{R}_{AB}(\tau)$ is a measure of the similarity between $A(t)$ and $B(t)$ as a function of the timelag τ applied to $B(t)$. The part of the cross-correlation with positive timelag ($\tau > 0$) is called the causal part of the cross-correlation. The other part ($\tau < 0$, negative timelag) is denoted as the acausal part of the cross-correlation. For more information on the calculation of the digital, linear CCF see Appendix A, and for the implementation of the computation of the CCF in the software MATLAB see Section 4.3.

Now, substitute the signals mentioned above by passive seismic observations recorded at two different receiver locations. Cross-correlating these seismic observations yields (under certain circumstances as described below) an estimate of the response that can be interpreted as the response that would be measured at one of the receiver locations if there were a delta-like pulse source at the other. This principle is called seismic interferometry (Wapenaar et al., 2010).

There are different approaches which theoretically explain the concept of extracting the response of the medium between two receivers by cross-correlating the responses of these
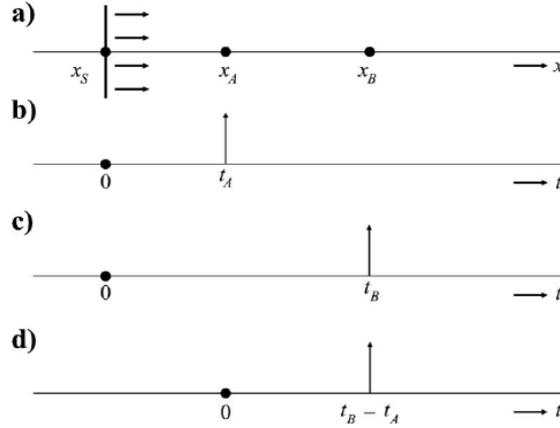
Figure 2.2.: a) An impulsive source at $x = x_s$ emits a plane wave propagating rightward, passing two receivers, one at x_A and another one at x_B. b) The response observed by the receiver at x_A at the time $t_A = \frac{x_A - x_S}{v}$. c) The response observed by the receiver at x_B. d) The cross-correlation of the two responses in b) and c). Here, the axis can also be denoted as τ instead of t (Wapenaar et al., 2010).

two receivers. All of the theories make assumptions on the wavefields, and on the medium, the waves propagate through. Furthermore, the sources have to meet several requirements. For example, in the field of ultrasonics Weaver and Lobkis (2002) showed that the cross-correlation between diffuse signals in two transducers is very nearly equal to the direct response of one transducer to an impulse applied to the other. It is assumed that all normal modes in the material are excited by uncorrelated noise sources of equal strength. Moreover, the specimen used in the study of Weaver and Lobkis (2002) can be defined as a closed system. The Earth on regional or local scale, however, should be considered as an open system (Snieder and Wapenaar, 2010). Snieder (2004) derived the theory of seismic interferometry using a stationary phase approach. His derivation is not based on normal modes (therefore, his theory is also valid for open systems) and the equipartitioning between them, but it also is only valid for scattered waves which propagate on average isotropically near the receivers. Another theory is based on time-reversal symmetry and was published by Derode et al. (2003). Wapenaar (2004) uses Rayleigh's reciprocity theorem for explaining the theory of seismic interferometry. In the paper of Wapenaar et al. (2005), the approaches of Derode et al. (2003) and Wapenaar (2004) are compared to each other and a relationship between these theories is formulated.

A one-dimensional analysis of direct-wave interferometry (based on the example given by Wapenaar et al., 2010) shall illustrate the basic theory of seismic interferometry. For the following analysis, a lossless medium and a constant seismic velocity v are assumed. Figure 2.2 shows a planar wave emitted by an impulsive source at $x = x_S$ propagating rightward. At the time $t_A = \frac{x_A - x_S}{v}$ the wave is observed by a receiver located at x_A. The response (see Figure 2.2 b) is denoted as the Green's function $G(x_A, x_S, t)$. As the response is an impulse, one can also express the Green's function by the Dirac delta function $\delta(t - t_A)$. Similarly, $G(x_B, x_S, t) = \delta(t - t_B)$ is written for the response observed at x_B at the time $t_B = \frac{x_B - x_S}{v}$ (Figure 2.2 c). Using Equation (2.1), the cross-correlation of the two responses recorded at x_A and x_B is defined as $\int G(x_A, x_S, t)G(x_B, x_S, t + \tau)dt$. By substituting the delta functions, one obtains:

$$\int \delta(t - t_A)\delta(t - t_B + \tau)dt = \delta(\tau - (t_B - t_A)) = \delta\left(\tau - \frac{x_B - x_A}{v}\right). \qquad (2.2)$$

Following the pattern above, the delta function on the right side of Equation (2.2) can be identified with the observation at x_B of an impulsive wave emitted at x_A. Returning to the Green's function formalism, one can write:

$$G(x_B, x_A, \tau) = \int G(x_A, x_S, t) G(x_B, x_S, t + \tau) dt. \qquad (2.3)$$

This equation represents the principle of seismic interferometry: The cross-correlation of two responses is equal to the response that would be measured at one of the receiver locations (x_B) if there were a source at the other (x_A).

Seismic interferometry has many advantages: Due to the cross-correlation process neither the location x_S of the actual source nor the velocity v of the emitted signal nor the time t_S at which the source emits the signal need to be known. Given the case of an increased signal velocity the response in Figure 2.2 d) would just occur at an earlier time, because the traveltime between the two receivers would be shorter. If the source was located right of x_B and if the signal propagated leftward, the response obtained in the cross-correlation process would occur at a negative time (see Wapenaar et al. (2010)).

Wapenaar et al. (2010) also showed that the source has not necessarily to emit an impulsive signal. If a wavelet $s(t)$ is radiated, the cross-correlation of the responses observed at x_A and x_B corresponds to the Green's function between the two receivers convolved with the autocorrelation of the source function $s(t)$.

Taking one step further to reality, the two-dimensional case of direct-wave interferometry is considered now. Figure 2.3 (graphic on the left) shows two receivers located at x_A and x_B. The black dots around these two receivers represent small point sources emitting one after another transient signals with a central frequency of 30 Hertz. Each source location is defined by a distance (r_S) to the centre of the connecting line between the two receivers and by an angle (ϕ_S, which is depicted in Figure 2.3). If the sources radiate one after another and if the responses to a single source observed at x_A and x_B are cross-correlated, the so called correlation gather in the middle of Figure 2.3 will be obtained. Summing all the cross-correlations depicted in the correlation gather yields the cross-correlation on the right side of Figure 2.3. The sources located in the Fresnel zones (indicated by the thick dashed lines on the left as well as in the middle of Figure 2.3) mainly contribute to the two signals in the finally obtained cross-correlation. The signals of the sources outside the Fresnel zones interfere destructively. A mathematical explanation for this using the stationary phase approach is given by Snieder (2004). Regarding simultaneously acting, uncorrelated noise sources the cross-correlation of the responses of the two receivers would look very nearly the same as the one shown on the right of Figure 2.3. Due to the simultaneity of the signal radiation from the different noise source locations a summation of different cross-correlations is not required, and the computation of only one cross-correlation is sufficient (Wapenaar et al., 2010).

Assuming a three-dimensional noise distribution, the analysis above is still valid. The Fresnel zones only have to be translated into Fresnel volumes.

In the field, a noise distribution like the one shown on the left of Figure 2.3 is very unlikely. Moreover, all the noise sources will probably not act simultaneously. Also the frequency content or rather the amplitude spectrum of the wavefields emitted by various noise sources differs (see Section 2.2). To reduced these effects in order to obtain a good estimate of the Green's function, the real data can be normalised in the frequency domain. This normalisation yields an equipartitioning of the amplitude values for all frequencies of the analysed frequency band. Furthermore, using long passive seismic recordings (months to years) might enlarge the number of not simultaneously acting noise sources at different locations. How the data are processed in this study, is described in Chapter 4.

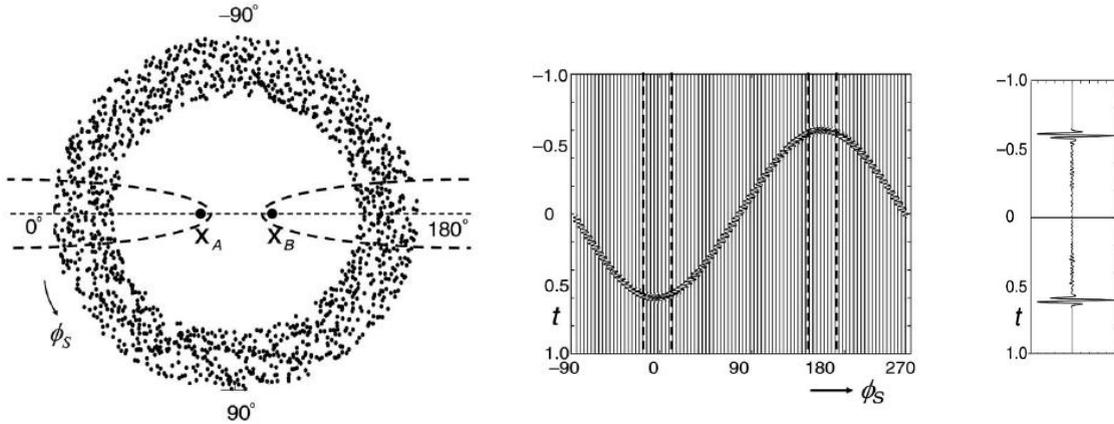Since the work of Campillo and Paul (2003), in which the interreceiver impulse responses

Figure 2.3.: Numerical example of two-dimensional direct-wave interferometry. On the left: Two receivers located at x_A and x_B isotropically surrounded by point sources (black dots). In the middle: Every sixteenth cross-correlation of the responses observed at x_A and x_B. (The sources emitted transient signals one after another.) On the right: The sum of all the computed cross-correlations. Two symmetric signals clearly emerge in this cross-correlation (Wapenaar et al., 2010).

were estimated from the correlations of seismic coda waves[1], many studies have been conducted using passive seismic interferometry (Curtis et al., 2006).

The scales, on which passive seismic interferometry is applied, vary between continental (Groos et al., 2012) and local scales (Picozzi et al., 2009).

Furthermore, two main applications of passive seismic interferometry have been established: imaging and monitoring.

Imaging stands for the calculation of tomographies of the seismic velocity distribution in the subsurface. The basis for the tomographic inversion is the retrieval of seismic waves between pairs of seismometers by cross-correlating the noise records. By dividing the interstation distance by the traveltime of the retrieved signals, the propagation velocity between the seismometers can be estimated (Wapenaar et al., 2010). Most often, surface wave tomographies are computed, because mainly surface waves are obtained in the cross-correlations of seismic noise records. The reason for the preponderance of surface waves can be related to the predominant surface origin of seismic noise (see Section 2.2, Bonnefoy-Claudet et al. (2006)). An example of a surface wave tomography shows Figure 2.4.

In the case of monitoring, temporal changes of the seismic velocities are analysed. The basic idea is that waves, which have propagated a long distance in a medium due to many reflections and scattering, are more sensitive to small variations in this medium than direct (also called ballistic) waves with shorter travel paths. Under the assumption that the late part (coda) of the cross-correlations contains such scattered and reflected waves, it is possible to monitor relative velocity variations in a medium by comparing the codas. Figure 2.5 points out an example of a temporal variation of the relative seismic velocity $\frac{dv}{v}$, where the seasonal velocity variations are connected to precipitation (Sens-Schönfelder and Wegler, 2006).

Actually, the noise-based monitoring requires only temporally stable sources, which contribute to signals in the cross-correlations. For the noise-based imaging, however, the

---

[1]In an earthquake recording the coda is that part of the seismogram which follows the distinct arrivals of, for example, the P-wave or the S-wave. The coda contains multiply scattered waves (Aki and Chouet, 1975).
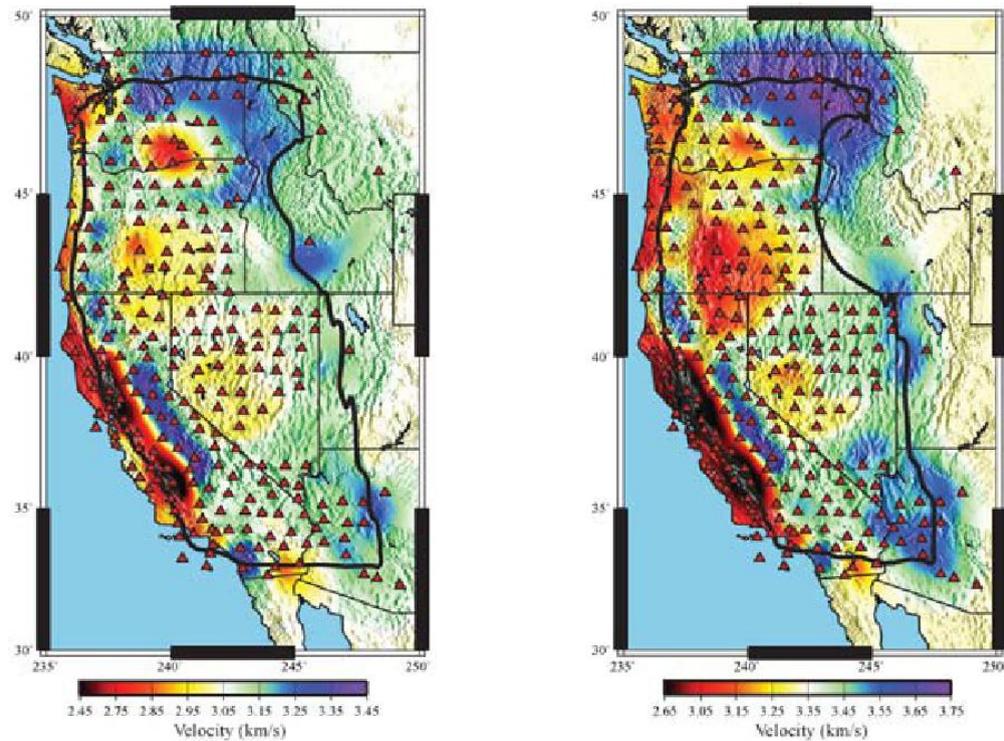
Figure 2.4.: Phase velocity tomographies of the western United States from ambient seismic noise. On the left: The estimated Rayleigh wave velocity map at a period of 8 s. On the right: The estimated Love wave velocity map at a period of 8 s (Lin et al., 2008).

cross-correlations should converge to the Green's functions to obtain reliable results (Hadzi-ioannou et al., 2009).

Since a few years, the so-called Monitoring and Imaging based on Interferometric Concepts (MIIC) processing suite based on Python code has been established for interferometric processing. For more information on the MIIC project refer to the website http://theo1.geo.uni-leipzig.de/wordpress/.

Next to the two main applications of imaging and monitoring, the localisation of seismic noise sources is possible by using seismic noise cross-correlations (Ma et al., 2013; Horstmann, 2010).
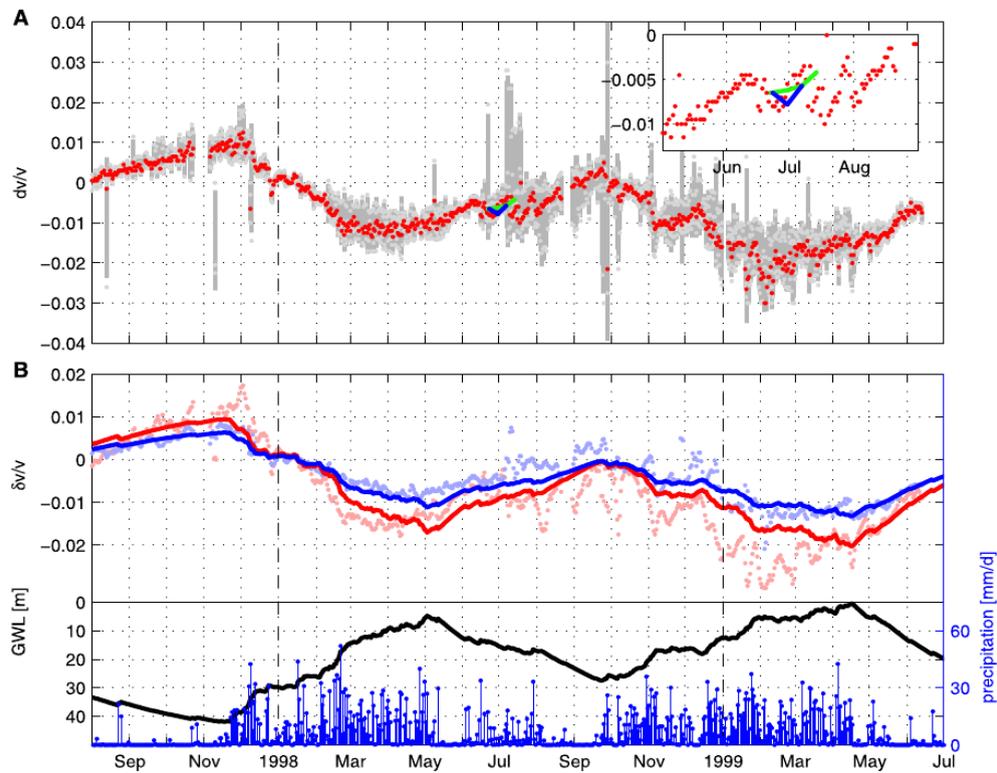
Figure 2.5.: On top (A), and in the middle (top of B): The measurements of the relative seismic velocity variations with records of stations located at Merapi volcano. At the bottom: The daily precipitation rate (blue) and the modeled ground water level (black) at Mount Merapi (Sens-Schönfelder and Wegler, 2006). A seismic velocity increase (from April to October 1998, and from April to June 1999) is connected to decreasing ground water levels in the same time spans.

3. Data Set and TIMO2-Project

The first part of this chapter gives an overview of the data set, which is analysed by passive seismic interferometry. As the used data are obtained within the TIMO2-project, the project itself is also introduced. Moreover, existing noise sources in the study area are shortly listed. In the second part of this chapter, the fieldwork, which was done in the scope of this thesis, is described.

3.1. General Information

Since 2006, many microearthquakes have occurred in the surroundings of two geothermal power plants in Rhineland Palatinate in Southwest Germany (Plenkers et al., 2013; Groos et al., 2013). One of the power plants is located in the city of Landau and has been operated since 2007, the other one was build up in Insheim and brought into service five years later in 2012. To analyse the microseismicity of these two deep geothermal systems, a network of seismic stations was deployed by the Geophysical Institute of the Karlsruhe Institute of Technology in the summer of 2009 (see Figure 3.1). The seismic stations have been operated within the TIMO2-project (Tiefenstruktur des mittleren Oberrheingrabens - Deep Structure of the Central Upper Rhine Graben), which is the second project phase of TIMO. Furthermore, TIMO2 contributes to the joint research project MAGS - concepts of Mitigating microseismic Activity of deep Geothermal Systems. The TIMO-project was conducted from December 2004 until May 2006. The interstation distances of the first network were larger in comparison to those of the TIMO2-network. However, some of the station locations remained the same (for example, the locations of the stations TMO20 and TMO22).

This thesis focuses on the analysis of the data of the TIMO2-project. As this project started in 2009 and is expected to last until August 2013, about four years of continuous, seismic data are available. Hence, the requirement of long passive seismic recordings for a better estimate of the Green's function are met (see Chapter 2.3). The seismic data are ground velocity recordings recorded by seismic stations that are part of the KArlsruhe BroadBand Array (KABBA). All of the stations record the ground velocity with three components, on one vertical component – also called the Z component – and on two horizontal components, the north-south (N) and the east-west (E) components. Broadband (e.g. Streckeisen STS-2, Guralp CMG-40T) as well as short-periodic sensors (e.g. Lennartz LE-3Dlite 1s) have been installed (differences between the sensors need to be accounted for in the (pre-)processing of the data (see Section 4.1)). During the time
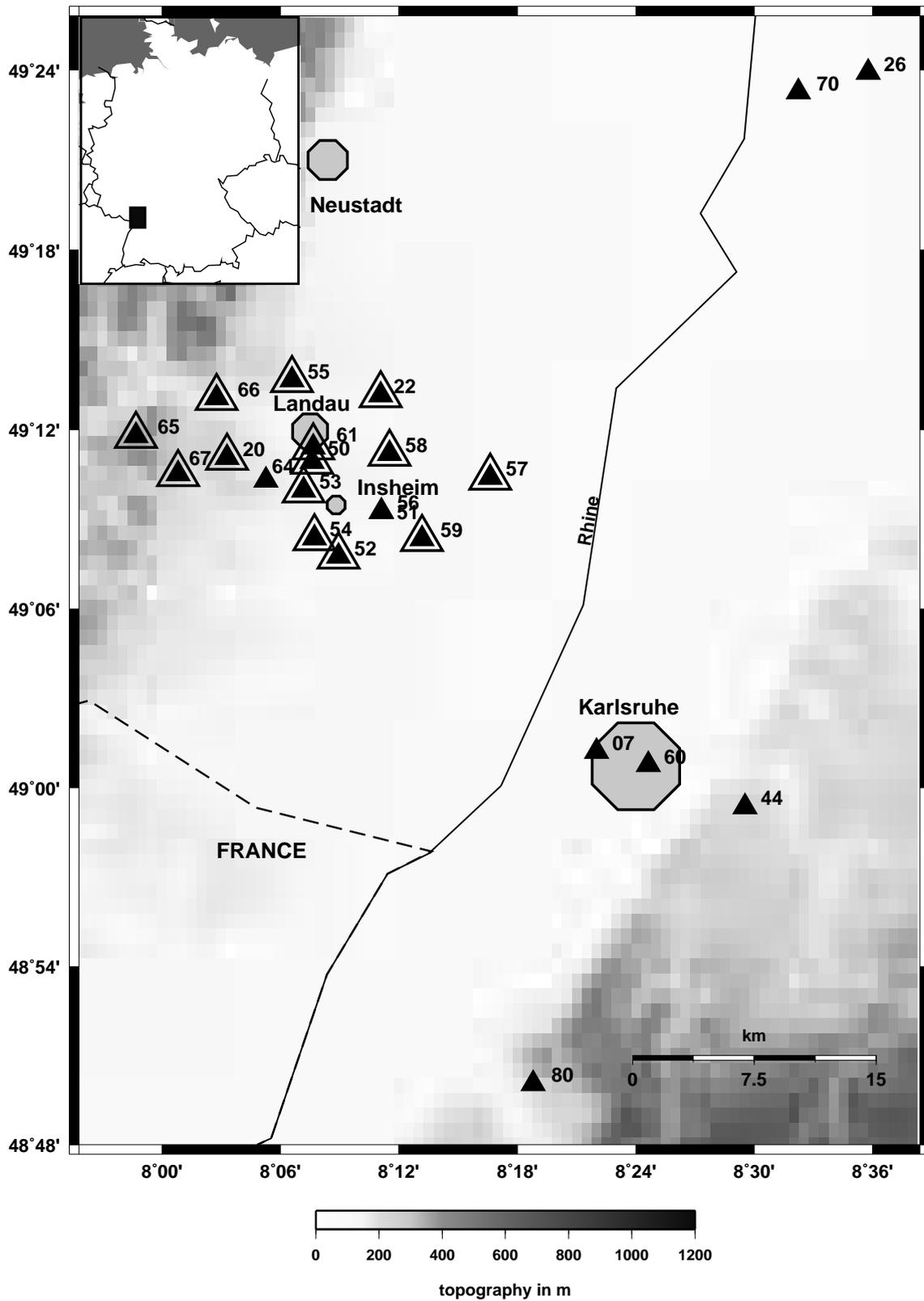
Figure 3.1.: The TIMO2-network. The framed triangles mark the stations used in this thesis. Note that not all of the stations shown in this map recorded data at the same time!

Table 3.1.: Geographic coordinates of all stations of the TIMO2-project listed by the station abbreviations in ascending order. The bold printed stations are the ones used in this thesis. Borehole stations are marked in blue.

Station	Latitude in Degrees	Longitude in Degrees	Altitude in Metres	Installed on dd-mm-yyyy	Removed on dd-mm-yyyy
TMO07	49.02059	8.36724	129	12-05-2009	-
TMO20	**49.18432**	**8.05512**	**214**	**25-03-2010**	**-**
TMO22	**49.21921**	**8.18467**	**135**	**29-06-2009**	**-**
TMO26	49.39860	8.59629	109	27-10-2009	-
TMO44	48.98935	8.49214	185	05-03-2009	-
TMO50	**49.18219**	**8.12681**	**155**	**29-06-2009**	**25-05-2012**
TMO51	49.15432	8.18515	134	03-07-2009	26-11-2009
TMO52	**49.12925**	**8.14927**	**129**	**03-07-2009**	**-**
TMO53	**49.16631**	**8.11971**	**171**	**03-07-2009**	**-**
TMO54	**49.13964**	**8.12865**	**144**	**03-07-2009**	**-**
TMO55	**49.22764**	**8.10988**	**203**	**15-10-2009**	**-**
TMO56	49.15452	8.18539	135	26-11-2009	10-03-2010
TMO57	**49.17344**	**8.27718**	**135**	**21-12-2009**	**-**
TMO58	**49.18662**	**8.19223**	**134**	**11-02-2011**	**-**
TMO59	**49.13915**	**8.21947**	**120**	**04-03-2011**	**-**
TMO60	49.01294	8.41063	120	26-05-2011	24-11-2011
TMO61	**49.19023**	**8.12781**	**152**	**20-07-2011**	**-**
TMO64	49.17170	8.08780	114-70	28-02-2012	-
TMO65	**49.19668**	**7.97808**	**482**	**25-06-2012**	**-**
TMO66	**49.21769**	**8.04607**	**287**	**22-01-2013**	**-**
TMO67	**49.17541**	**8.01357**	**262**	**24-01-2013**	**-**
TMO70	49.38788	8.53728	101	27-10-2009	-
TMO80	48.83433	8.31325	181	10-08-2010	-

period from 2009 to 2013 overall 23 stations were operated. Table 3.1 lists the seismic stations with their geographic coordinates (provided by the Global Positioning System (GPS)), the date of installation, and - where applicable - the date of removal. For this study only the recordings of the stations close to the city of Landau are used (printed in bold in Table 3.1). Altogether, these are fourteen stations. Although the stations TMO51 and TMO56 are located close to Landau (see the map in Figure 3.1) they were not used because of their relatively short operating times.

With the following equation (3.1) the number of all pairwise station combinations $P(n)$ can be computed:

$$P(n) = \frac{n \cdot (n-1)}{2} \tag{3.1}$$

with n being the number of stations. $P(n)$ is of interest, because - as described later in Section 4.2 - the cross-correlation functions of all possible pairs of stations are calculated. Note that not all of the fourteen stations mentioned above recorded data simultaneously. In 2012, for example, only eleven stations recorded data at the same time (see also Table 3.1). This example yields 55 different stationpairs.

 The maximum interstation distance of the analysed stations, operated at the same time, is ~ 22 kilometres (TMO65 and TMO57), the minimum interstation distance is ~ 0.9 kilometres (TMO50 and TMO61). Hence, the study area has a local character. A histogram of the interstation distances is shown in Figure 3.2.

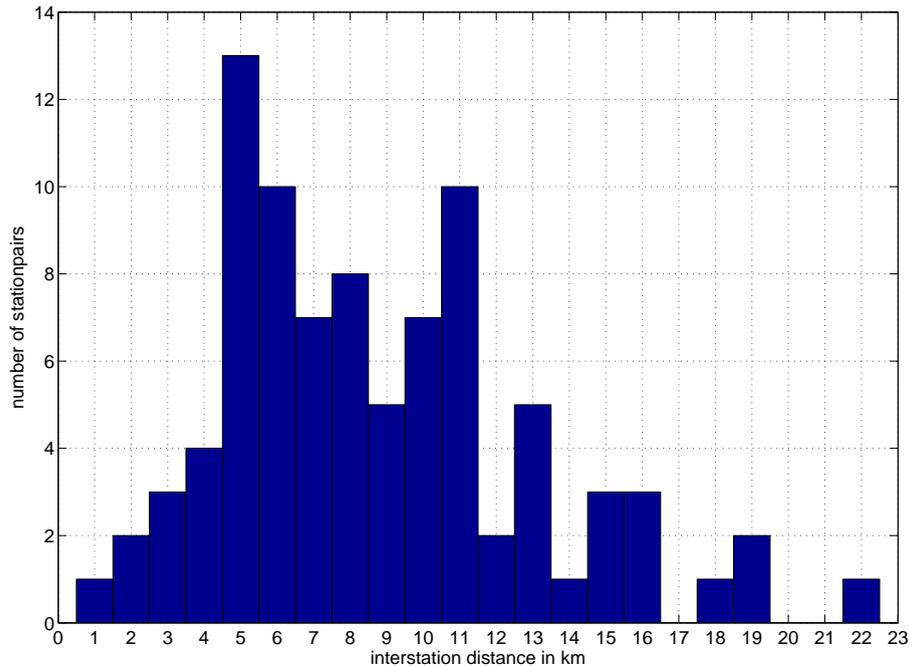Another characteristic of the area in and around the city of Landau is a high man-made

Figure 3.2.: The number of the used stationpairs plotted against their interstation distance, which is divided into one kilometre intervals. As not all of the fourteen stations were operated at the same time, overall 88 pairs of stations (instead of 91) are displayed.

noise level. The region is densely populated and intensively agriculturally used. Furthermore, the traffic in the Upper Rhine Graben (URG) contributes to the seismic noise field around Landau. A few kilometres westward of Landau there are also two quarries. In the north of Landau there are producing oil fields. However, not only the large number of seismic noise sources accounts to the high noise level in the URG, but also the wave amplification due to unconsolidated URG sediments of the Quaternary layers.

3.2. Installation of Two Additional Seismic Stations in 2013

In order to improve the network for the localisation of seismic noise sources, two more seismic stations were deployed in the western part of the station network in January 2013. Former locations of the TIMO-project turned out to be more or less inappropriate for detecting the noise source.

In Chapter 5.2 the results for the localisations of the point source without the two additional stations are shown and discussed. Unfortunately, the data of the two new seismic stations could not be integrated into the localisation analyses yet, because there were only three days of data available in the data archive at the end of this study in July 2013.

On 22nd January 2013, the first of the two seismic stations (TMO66) was installed in the cellar of the Julius Kühn-Institute for vine cultivation at the Geilweilerhof (see left picture in Figure 3.3). Two days later, the TMO67 station was set up in the southwestern part of the station network in the cellar of the Sportheim Eschbach (see right picture in Figure 3.3)). It was also planned to install a third station even further westward of the network at the Fachklinik im Eusserthal. But unfortunately, we did not get the permission for deploying a seismic station there. Figure 3.4 shows a map where the stations TMO66 and TMO67 are located, and where it was planned to set up the third station.

Figure 3.3.: Installation of TMO66 in the cellar of the Julius Kühn-Institute (left). Installation of the TMO67 station in the cellar of the Sportheim Eschbach (right). On both pictures you can see the big Streckeisen STS-2 and the smaller Trillium Compact broadband sensors.
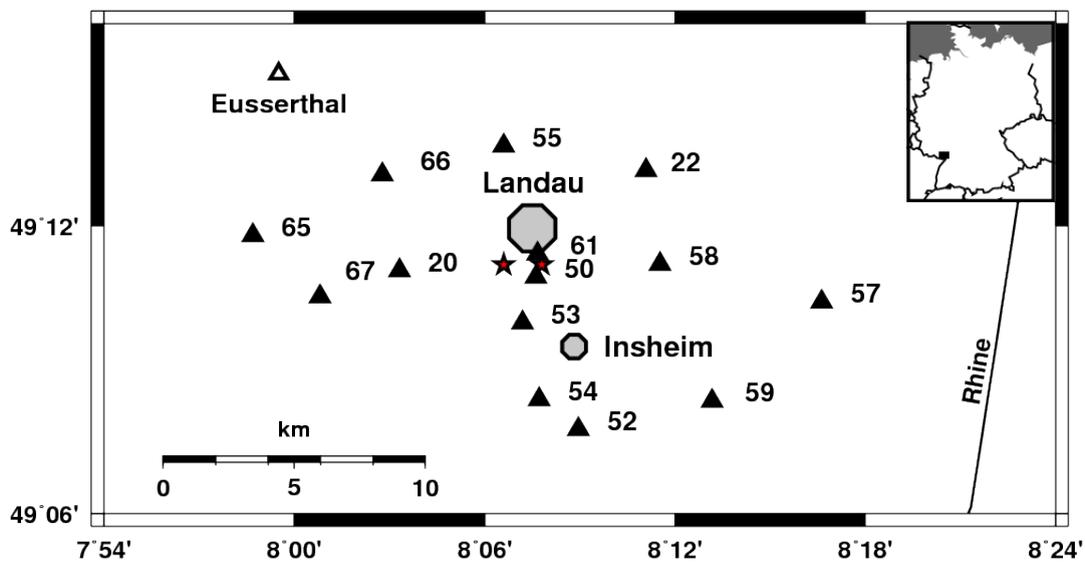

Figure 3.4.: This map shows only the used stations around the city of Landau. In the western part, there are the two added stations TMO66 and TMO67. The location, where the third station was planned to be deployed, is denoted by the not filled triangle named 'Eusserthal'. The two red stars mark the landing points of two boreholes of the geothermal power plant in Landau.

4. Data Processing

The data processing of the vertical and the horizontal component time series is divided into three main steps:

(1) the preprocessing of all available daily time series,

(2) the calculation of the cross-correlation functions (CCFs) of either 20 minutes long time windows (for the vertical component CCFs) or one hour long time windows (for the horizontal component CCFs) and

(3) the subsequent normalisation of the CCFs in the frequency domain accompanied by stacking, and filtering of the data.

This procedure follows Groos et al. (2012). A big advantage of this processing scheme is that the waveforms and the amplitudes of the time series are not non-linearly modified prior to the calculation of the cross-correlation functions.
The horizontal component CCFs are additionally rotated to achieve transverse and radial components (see Section 4.4).
Figure 4.1 gives an overview of the order of the processing steps.
The data processing is implemented by the software MATLAB (www.mathworks.com). Many used functions are part of the Karlsruhe Seismology Processing (KaSP) toolbox for MATLAB (see Appendix B).

4.1. Preprocessing of the Daily Time Series

As the waveform data of all TIMO2 stations are stored at the KABBA data centre in files on a daily basis, the preprocessing of the data is applied to one-day long time series. Actually, not only the desired daily time series is read into MATLAB but also – if available – the waveforms of the day before and after. This is done for two reasons:

(i) Artefacts due to tapering or filtering or the removal of the instrument response at the beginning and the end of the extended daily time series can be eliminated by cutting off the affected part of the time series.

(ii) To avoid a loss of data due to the stacking process of the cross-correlations, the first 20 minutes of the time series of the next day are added to the end of the time series of the desired day. (For the calculation of the CCFs, the time series are segmented into 20 minutes long windows. If the daily time series were not enlarged, the CCF computed with the last 20 minutes time series of one day would contain many zeros.
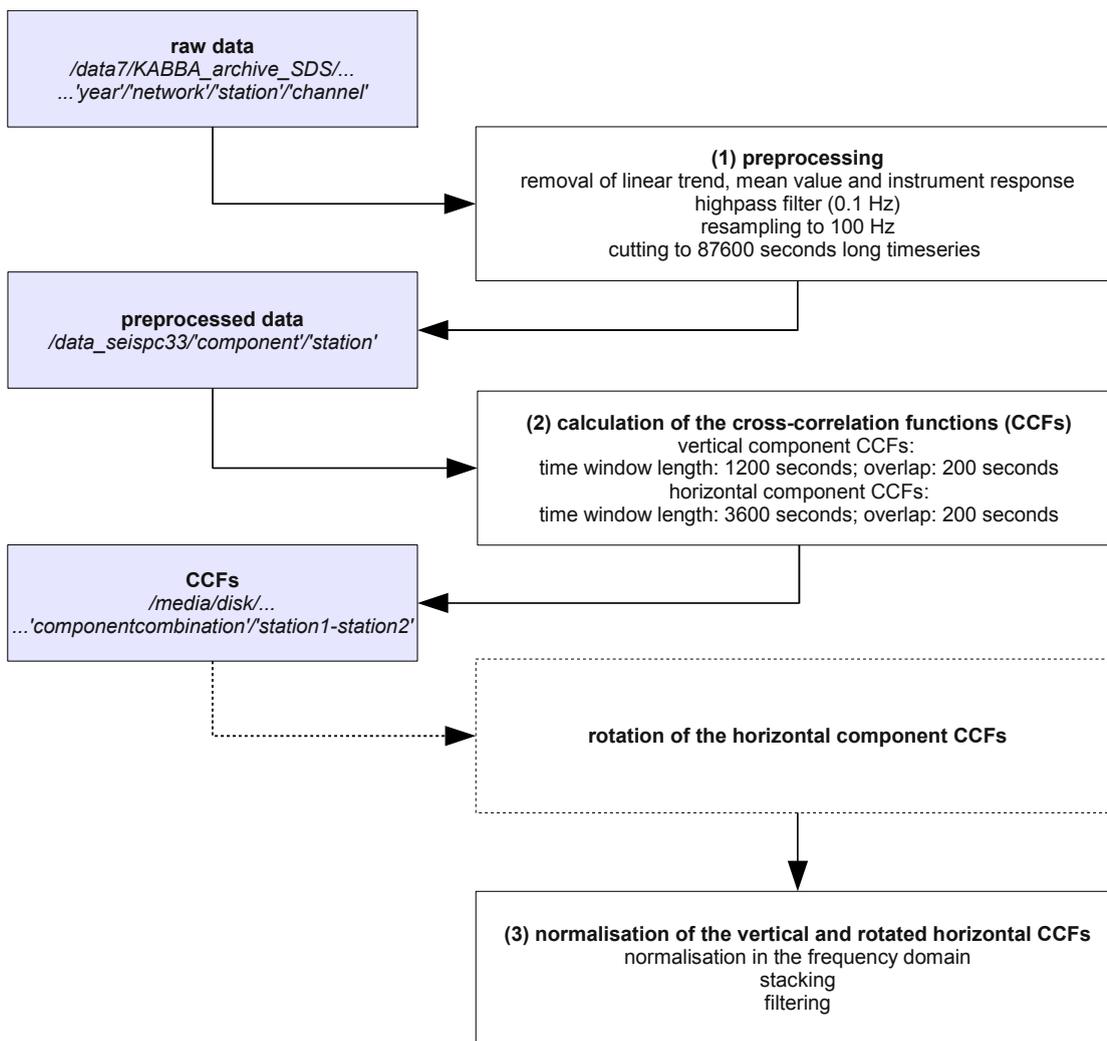
17

Figure 4.1.: Flow chart for the data processing. The dashed box applies to the horizontal
            components only. The paths of the folders, where the data are stored, are
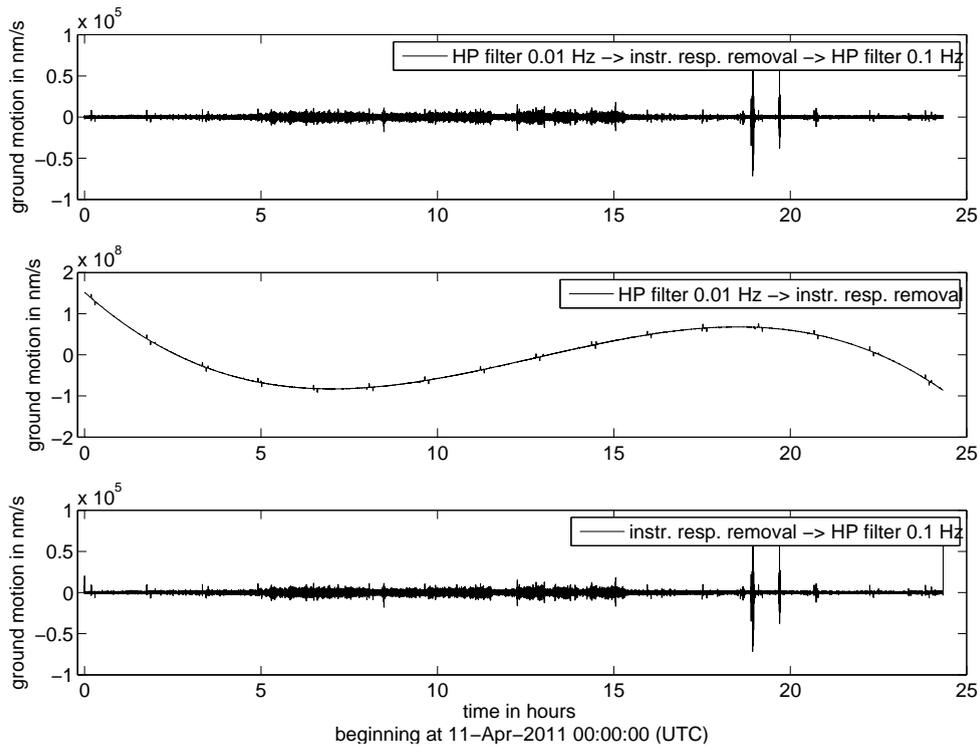   printed in italics.

Figure 4.2.: The differently processed daily time series of the Z component of the TMO50 station (Lennartz LE-3Dlite 1s) of Julian day 101.2011. In the middle, the low frequency artefacts due to the removal of the instrument response (instr. resp. removal) are clearly visible. The unwanted, equidistant (one and a half hours apart) small peaks are due to the regular boot of the hard disk of the digitiser. At the bottom, there are very high peaks at the beginning and the end of the waveform. On top, the chosen preprocessing is depicted examplarily.

The data sample combinations of the last 20 minutes of this one day and the following day would be lost, and not contribute to the stacked CCF.) Although the obtained waveform is not exactly one day long ($24 \cdot 60 \cdot 60 \text{ s} + 20 \cdot 60 \text{ s} = 87600 \text{ s}$), I will refer to this waveform as the daily time series in the following.

At first, the ground velocity unit is converted from counts to nm/s. Afterwards, the mean value and the linear trend of the daily time series are removed. After tapering, a zerophase 0.01 Hertz high-pass filter is applied. Then, the instrument response is removed. Possible low frequency artefacts caused by the removal of the instrument response are eliminated by applying a zerophase 0.1 Hertz high-pass filter. These steps are necessary, because the time series of different sensors are used in this study. Due to the Lennartz LE-3Dlite 1s instruments the recordings have to be filtered twice. Figure 4.2 shows, how a daily time series of a Lennartz LE-3Dlite 1s instrument (TMO50) due to different preprocessing schemes changes. Figure 4.3 shows the same for the time series of the same day recorded by a Guralp CMG-40T seismometer (TMO20).

In a last step of the preprocessing, the sampling rate is set to 100 Hertz. No corrupt time series (for example, time series containing large offsets) are excluded. This follows the recommendation of Groos et al. (2012), who do not see an improvement of the cross-correlation functions (CCFs), where corrupt time series are excluded prior to the calculation of these CCFs, in comparison to normalised (!) CCFs, where no corrupt time series are eliminated. The preprocessing of the data is implemented by the function KABBA_TMO_NCC_preprocessing_station_reduced.m (see Appendix B.1).
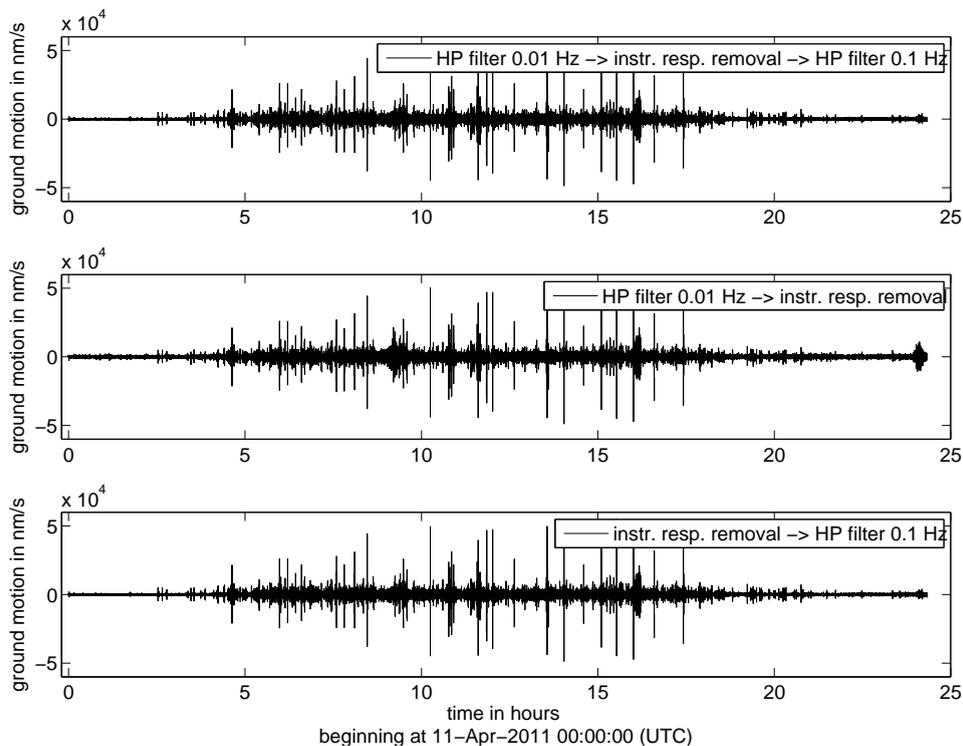
Figure 4.3.: The differently processed daily time series of the Z component of the TMO20
station (Guralp CMG-40T sensor) of Julian day 101.2011. There are almost no
differences between the waveforms obtained with the three different processing
schemes.

4.2. Calculation of the Cross-Correlation Functions

The digital linear cross-correlation functions (CCFs) (Appendix A) of the time series of
the vertical components (ZZ) and all possible pairs of the horizontal North (N) and
East (E) components, i.e. EE, EN, NN and NE, are computed for each station-
pair[1]. The functions XCORRprocessingSLWDW_NOfilter.m, cc_SLWDW_preparation.m,
and CCF_new.m, which are used for the calculation of the CCFs, are listed and briefly
explained in Appendix B.1.

At first, only the CCFs of the vertical component time series are computed. For this
processing step, the length of the time series used for the calculation of the CCFs, and the
maximum analysed timelag of the cross-correlation functions have to be defined:

For the definition of the maximum analysed timelag of the CCFs, I computed the travel
time of a very slow propagating seismic wave (~ 100 metres per second) between these two
stations which are farthest apart. At the time, the processing for this thesis started (in
the summer of 2012), the maximum interstation distance was ~ 16 kilometres (distance
between TMO20 and TMO57)[2]. By dividing the interstation distance by the minimum
velocity a travel time of ~ 160 seconds is obtained. So, I decided to analyse a maximum
lag time of ± 180 seconds. Given this maxlag value, the length of the time series was set
to 20 minutes. This length is about six times the maximum analysed lag time, thus it
follows the advice of Groos (2010). I even use a time window overlap of 200 seconds to

---

[1]For lack of time, all available time series (depending on the run time of the used stations) of the Z
  component between the year 2009 and Julian day 300 of the year 2012 are cross-correlated. For the
  calculation of the cross-correlations of the horizontal components only the available time series of the
  years 2011 and 2012 are used.

[2]Later on, other stations were installed so that the maximum interstation distance increased to ~ 22
  kilometres (distance between TMO65 and TMO57).

ensure that no sample combinations are missed in the cross-correlation process (Groos, 2010). (Of course, an overlap of 180 seconds would also have been adequate.)
For every pair of stations, the 20-minutes long segments of the time series are cross-correlated in the frequency domain. Hence, one obtains on average 87 CCFs per one-day long time series. As the data of all ZZ-CCFs occupy a very large disk storage of 957.6 Gigabytes, I decided to change the processing for the subsequent calculation of the horizontal component CCFs. Therefore, the segmentation length of the preprocessed daily time series was modified from 20-minutes to one-hour (keeping the overlap of 200 seconds). Per one-day long time series one now obtains only 25 CCFs for the horizontal components instead of 87 CCFs for the vertical component. In this way, for all CCFs of one pair of the horizontal components, e.g. NN, only about a third of the disk storage of the ZZ-CCFs is needed. The difference between the CCFs obtained with 20 minutes and one-hour long time windows is discussed in Section 4.5.
So far, no corrupt data (now, i.e. corrupt CCFs) have been excluded.

4.3. The Working Principle of the xcorr-Function in MAT-LAB

For a better understanding of how the cross-correlation functions are calculated in MATLAB (see Appendix A) the following procedure must be faced:
Two time series $A(t)$ and $B(t)$ are recorded. $A(t)$ is recorded at the location x_A, $B(t)$ at x_B. The length of each digitised time series is 10 s; the sampling rate is 1 Hertz. If a pulsed signal of an amplitude of 1 first arrives at x_A and afterwards unmodified at x_B, the signal will appear in the cross-correlation function, computed by xcorr($A(t), B(t)$), at a negative timelag. If the signal arrives first at x_B and then at x_A, xcorr($A(t), B(t)$) yields a signal in the causal part of the CCF (positive timelag). Both examples are shown in Figure 4.4 and Figure 4.5.

A descriptive reconstruction[3] of the computation of each data point of the CCF with the xcorr-function can be taken from Figure 4.6. This figure shows the same time series as the ones in Fig. 4.4. It illustrates how the CCF xcorr($A(t), B(t)$) evolves by shifting the time series $A(t)$ across $B(t)$. In the case shown, for example, the first data point of the CCF at timelag $\tau = -9$ s is calculated by multiplying the first value of $A(t)$, which is $A(10 \text{ s} - 9 \text{ s}) = A(1 \text{ s})$, with the last value of $B(t)$, which is $B(10 \text{ s})$. The second CCF value ($\tau = -8$ s) is the sum of the multiplications $A(10 \text{ s} - 8 \text{ s}) = A(2 \text{ s})$ times $B(10 \text{ s})$ and $A(9 \text{ s} - 8 \text{ s}) = A(1 \text{ s})$ times $B(9 \text{ s})$, and so on. At $\tau = -5$ s the sum of the products is equal to 1 (see Fig. 4.4), because at this timelag the signal is 'in line'.
Hence, xcorr($A(t), B(t)$) is a realisation of the following analytical definition of the cross-correlation function:

$$\widehat{R}_{AB}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int\limits_{-T/2}^{+T/2} A(t + \tau)B(t)dt. \tag{4.1}$$

As I always choose the western station as the first and the eastern station as the second input variable of the xcorr-function, a signal in the causal part of the cross-correlation means that the signal primarily arrives at the eastern station.

---

[3] Remember, that the following explanation of the operating mode of the xcorr-function is only descriptive, also in view of the fact that MATLAB calculates the cross-correlation in the frequency domain and everything shown here happens in the time domain.
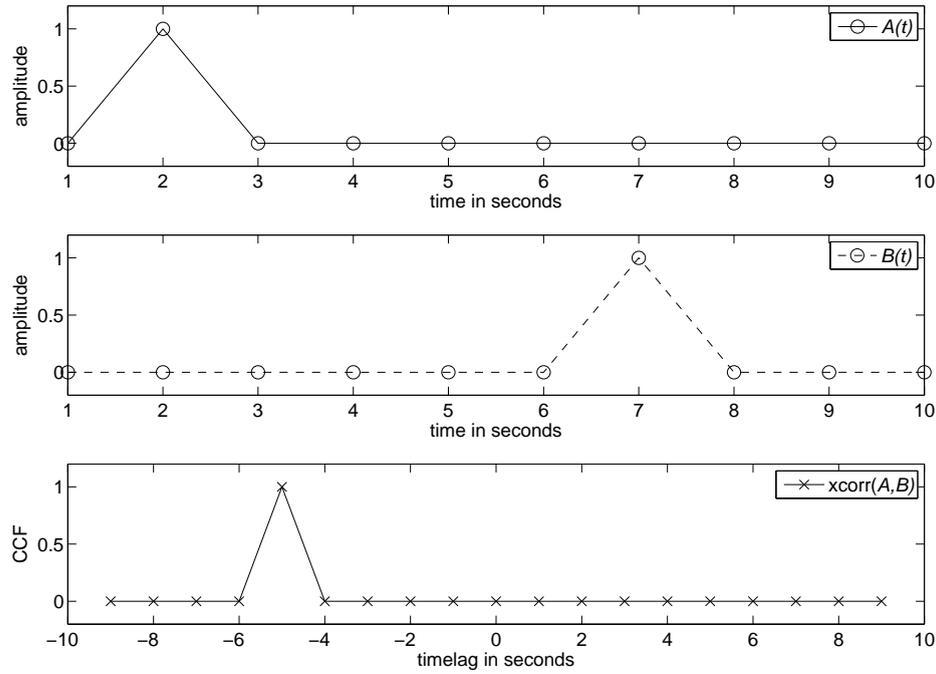
Figure 4.4.: Two time series $A(t)$ (top) and $B(t)$ (middle) are recorded at different locations x_A and x_B. A signal with an amplitude of 1 is recorded first at x_A and 5 seconds later at x_B. The calculation of the cross-correlation function in MATLAB ($\text{xcorr}(A(t), B(t))$) yields a signal in the acausal part of the CCF at timelag $\tau = -5$ s (bottom).
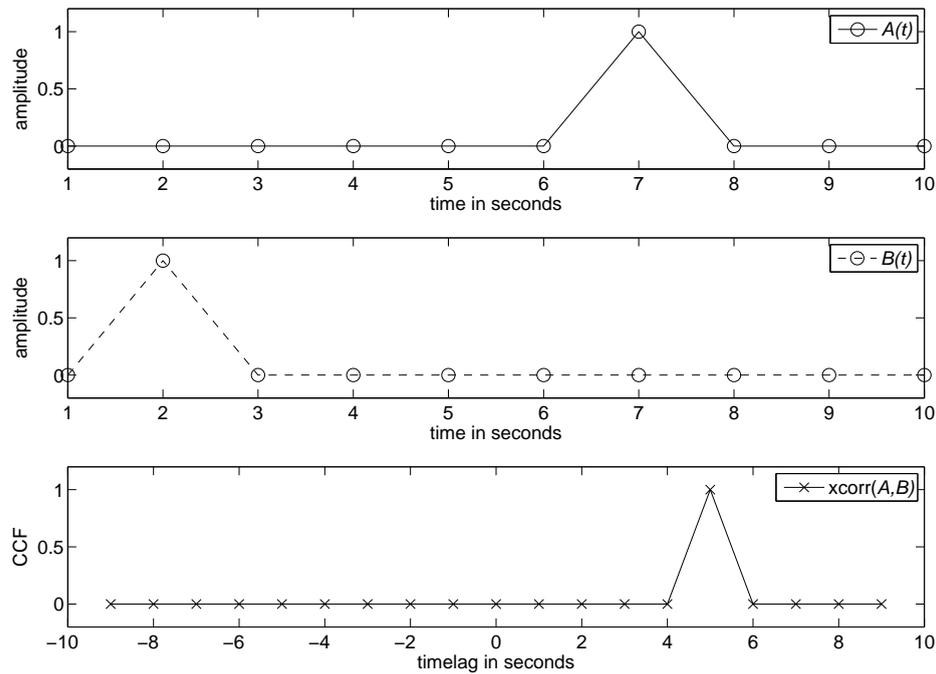

Figure 4.5.: In comparison to Figure 4.4, the signal now arrives first at x_B (middle) and then at x_A (top). Therefore, the signal appears in the causal part of the CCF (bottom).
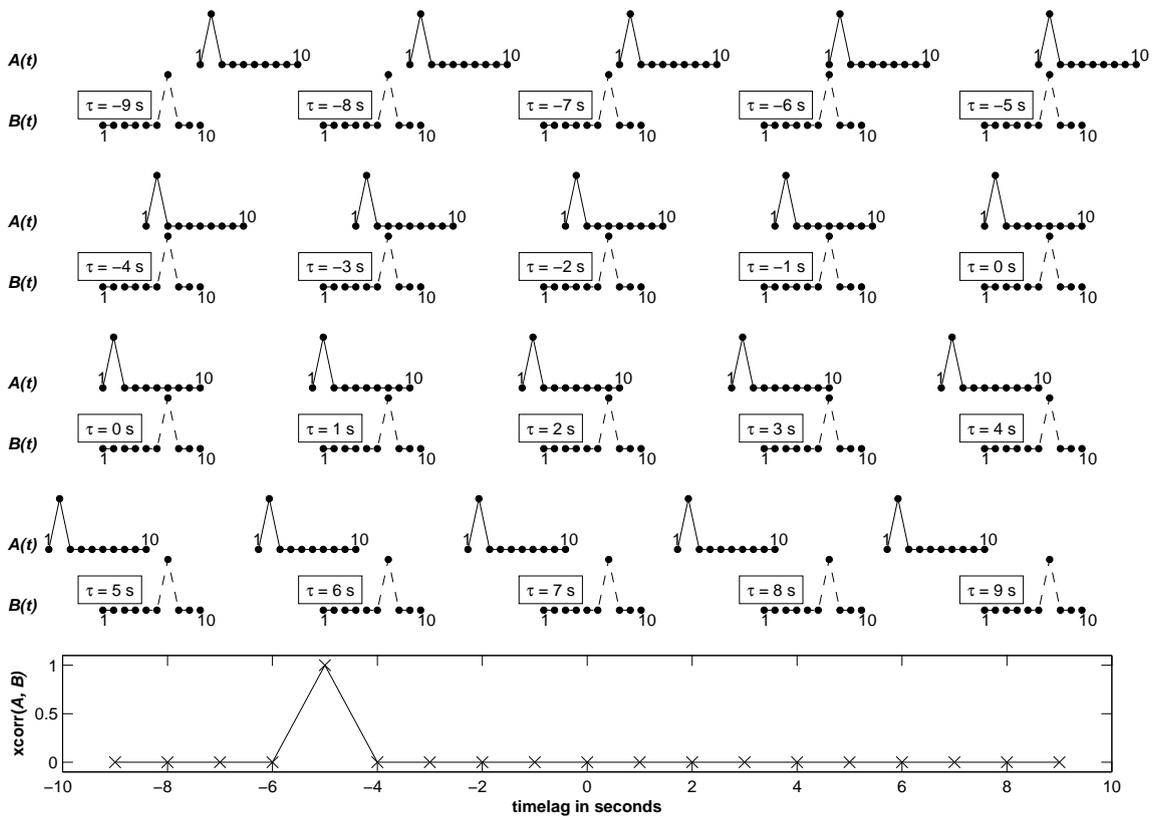
Figure 4.6.: This figure shows, starting in the upper left corner and moving line-by-line to the lower right corner, the shift of the time series $A(t)$ by τ across the time series $B(t)$. For every timelag the data points that lie upon each other are first multiplied. Afterwards these products are added and yield the value of the CCF at the given timelag. The CCF calculated by MATLAB is shown at the bottom.
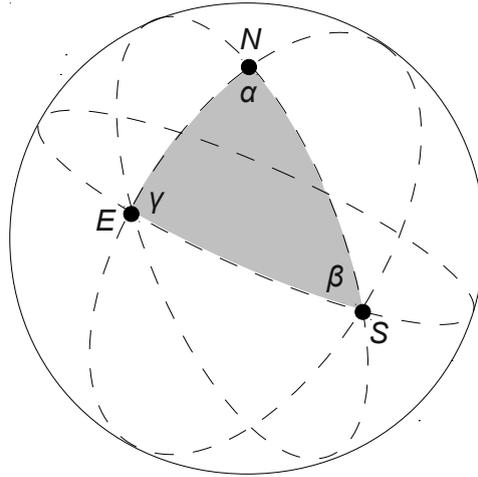
Figure 4.7.: The earth (slightly tilted forward) with the north pole N and with an earthquake epicentre at the location denoted by E plus a seismic station positioned at S recording the earthquake. The three great circle paths enclose a spherical triangle on the earth's surface with the three internal angles α, β and γ.

4.4. Rotation of the Horizontal Component Cross-Correlation Functions

In contrast to the vertical component data (ZZ) the cross-correlation functions of the horizontal components are rotated. Therefore, the north-east coordinate systems are transformed into radial-transverse (R-T) coordinate systems after Lin et al. (2008). With the radial component (R) lying on the connecting line of a stationpair and the transverse component being always perpendicular to R, the rotation should enable us to discriminate between different wave types (e.g. between Rayleigh and Love waves) propagating between two seismic stations. Signals in the cross-correlation functions (CCFs) of the radial-radial components, for example, might be an indicator for Rayleigh and/or P-waves. In contrast, signals in the CCFs of the transverse-transverse components might result from Love and/or S-waves (Plesinger et al., 1986).

Note, that the coordinate systems of the two stations of a cross-correlation pair are rotated about different angles. Hence, it must be defined, which station's horizontal components are rotated by which angle. Furthermore, the orientations of the axes of the new coordinate systems have to be defined.

4.4.1. Definition of Azimuth and Backazimuth

As described in Section 2.3, in seismic interferometry one seismic station can be interpreted as the source and the other as the receiver. This setup is comparable to earthquake seismology, where the source is an earthquake E and the receiver is a seismic station S. Figure 4.7 schematically shows a spherical triangle on the earth's surface with the basic points at the north pole (N), at the location of the earthquake (E) and at the location of the recording seismic station (S). In general, the sum of the internal angles α, β and γ does not equal 180°. The angle measured clockwise from north to the direction the ray must leave the earthquake to arrive at the station is called azimuth, denoted by θ (in Figure 4.7: $\theta = \gamma$). The backazimuth ψ is the angle measured at the station clockwise from north to the direction from where the energy of the earthquake arrives at the recording station. In the case shown (Figure 4.7), the backazimuth can be calculated by: $\psi = 360° - \beta$. Note, that generally $\psi \neq \theta + 180°$ due to the spherical geometry (Lay and Wallace, 1995).

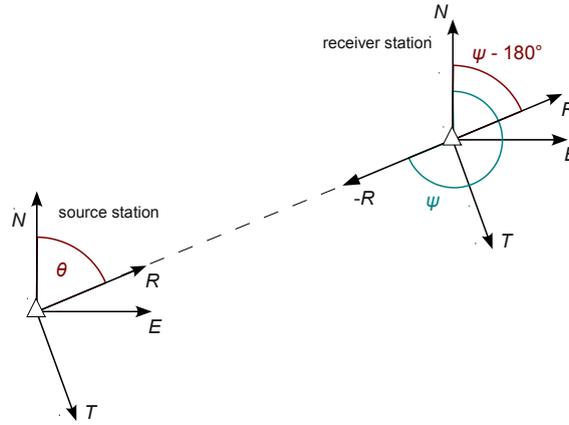Translating the above into seismic interferometry, it must be defined which one of the two

Figure 4.8.: Orientation of the radial (R) and transverse T components of a stationpair with one station at the source location (rotation about θ) and the other station at the receiver location (rotation about $\psi - 180°$)

stations of a stationpair is supposed to be the source and which one the receiver. If there is a signal in the acausal part (negative timelag) of the CCF, then I would choose this station as the source which records the signal first. If there is a signal in the causal part of the CCF, then the station which records the signal last would be the source. Transferred to the implementation in MATLAB the first input variable of the xcorr-function is the one that is recorded by the source station. Using the example for the calculation of the xcorr-function in section 4.1, the source station would record $A(t)$ and the receiver station $B(t)$.

4.4.2. Orientation of the Axes of the Rotated Coordinate Systems

For each stationpair the Cartesian coordinate system Z-N-E of the source station is rotated about the Z axis by the interstation azimuth θ into the Z-R-T system. The R or radial axis is orientated along the connecting line of the two stations and it points into the direction of the receiver location. The transverse component (T) is always perpendicular to the R and the Z component. In contrast, the Z-N-E system of the receiver station is rotated by $\psi - 180°$. By doing so, the R and T axes of both stations point to the same direction. This is made in analogy to Lin et al. (2008). Figure 4.8 illustrates, how the radial and transverse component of one stationpair are orientated.

4.4.3. Derivation of the Rotation Operator

For the derivation of an equation that represents the transformation of the horizontal component cross-correlations EE, EN, NE and NN into the transverse-transverse, radial-radial, transverse-radial and radial-transverse CCFs and that is used by Lin et al. (2008), I first analyse the rotation of the coordinate system of the source station. In the following, the source station is denoted as station 1.
Figure 4.9 shows a data point P_1 recorded in the N-E system of station 1. The initial coordinates of P_1 are: (E_1, N_1). By a linear combination of these two coordinates with coefficients related to the interstation azimuth θ the new coordinates of P_1 in the rotated system (R_1, T_1) can be calculated.

$$T_1 = E_1 \cos\theta - N_1 \sin\theta \qquad (4.2)$$
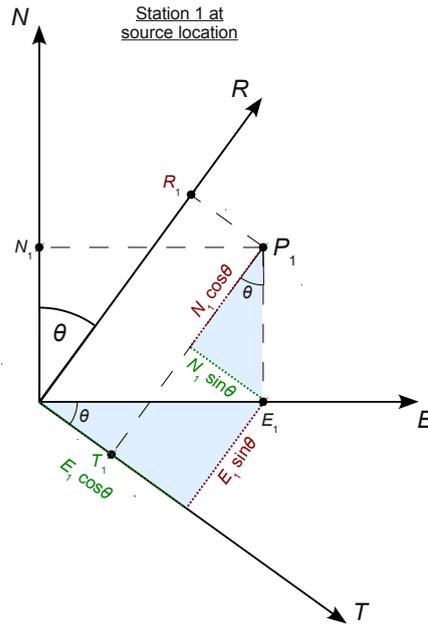$$R_1 = E_1 \sin\theta + N_1 \cos\theta \qquad (4.3)$$

Figure 4.9.: Geometrical drawing for the derivation of the two dimensional rotation matrix. The initial N-E coordinate system of the source station is rotated by the azimuth angle θ into the R-T coordinate system. For the computation of the new coordinates of the data point P_1 auxiliary right-angled triangles (blue areas) are inserted. For the computation of the T_1 component the green doted lines are subtracted. Adding the red doted lines yields the R_1 component.
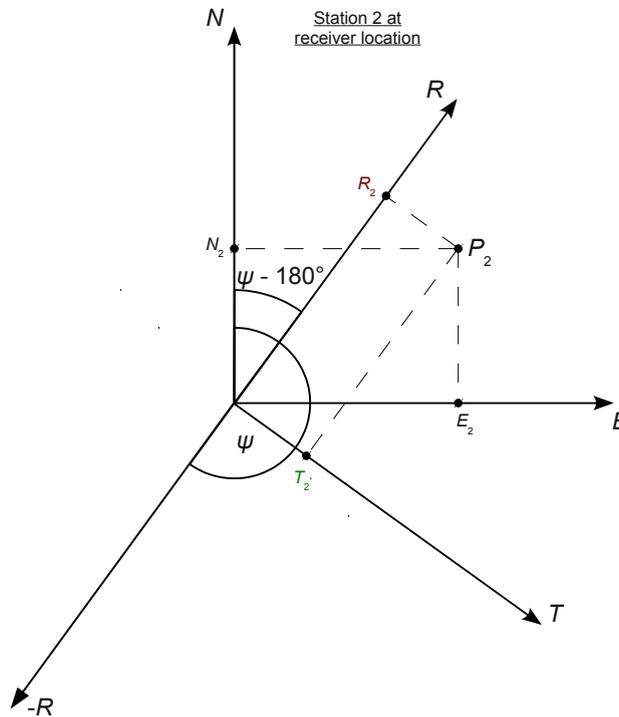

Figure 4.10.: The initial N-E coordinate system of the receiver station is rotated by $\psi -$ $180°$ into the R-T coordinate system. This setup is the same as the one in Figure 4.9 except for the different angles.

The rotation of the coordinate system of the receiver station (station 2) is similar to that of station 1. As described in section 4.4.2 the radial and transverse components point to the same direction at both stations. Hence, the *N-E* coordinate system of station 2 is not rotated by the backazimuth angle ψ, but by $\psi - 180°$. Figure 4.10 shows the coordinate systems of the receiver station. Compared to Figure 4.9 the initial situation is the same except for the different angles. Now, the rotation is not about θ anymore, but about the angle $\psi - 180°$. If one substitutes θ by $\psi - 180°$ in formulas (4.2) and (4.3), one obtains:

$$T_2 = E_2 \cos(\psi - 180°) - N_2 \sin(\psi - 180°) \tag{4.4}$$
$$R_2 = E_2 \sin(\psi - 180°) + N_2 \cos(\psi - 180°) \tag{4.5}$$

Using the addition theorems for cosine and sine

$$\cos(\alpha \pm \beta) = \cos\alpha \cos\beta \mp \sin\alpha \sin\beta$$
$$\sin(\alpha \pm \beta) = \sin\alpha \cos\beta \pm \cos\alpha \sin\beta \tag{4.6}$$

formulas (4.4) and (4.5) can be written as:

$$T_2 = -E_2 \cos\psi + N_2 \sin\psi \tag{4.7}$$
$$R_2 = -E_2 \sin\psi - N_2 \cos\psi \tag{4.8}$$

Multiplying the transverse and radial component of station 1 with the T and R components of station 2 yields four equations. The following formula compactly represents these four equations for the transformation of the initial horizontal CCFs into the rotated CCFs.

$$\begin{pmatrix} T_1 T_2 \\ R_1 R_2 \\ T_1 R_2 \\ R_1 T_2 \end{pmatrix} = \begin{pmatrix} -\cos\theta\cos\psi & \cos\theta\sin\psi & -\sin\theta\sin\psi & \sin\theta\cos\psi \\ -\sin\theta\sin\psi & -\sin\theta\cos\psi & -\cos\theta\cos\psi & -\cos\theta\sin\psi \\ -\cos\theta\sin\psi & -\cos\theta\cos\psi & \sin\theta\cos\psi & \sin\theta\sin\psi \\ -\sin\theta\cos\psi & \sin\theta\sin\psi & \cos\theta\sin\psi & -\cos\theta\cos\psi \end{pmatrix} \cdot \begin{pmatrix} E_1 E_2 \\ E_1 N_2 \\ N_1 N_2 \\ N_1 E_2 \end{pmatrix} \tag{4.9}$$

The rotation operator above is also used by Lin et al. (2008).

4.4.4. Permutability of Cross-Correlation and Rotation

Analytically, the processing steps of rotating and cross-correlating the data commute with each other, because the angles θ and ψ are neither dependent on t nor on τ:

$$R_1 R_2(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{+T/2} R_1(t + \tau) R_2(t) dt =$$

$$\overset{(4.3),(4.8)}{=} \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{+T/2} (E_1(t + \tau) \sin \theta + N_1(t + \tau) \cos \theta)(-E_2(t) \sin \psi - N_2(t) \cos \psi) dt =$$

$$= -\sin \theta \sin \psi \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{+T/2} (E_1(t + \tau) E_2(t) dt$$

$$- \sin \theta \cos \psi \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{+T/2} E_1(t + \tau) N_2(t) dt$$

$$- \cos \theta \sin \psi \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{+T/2} N_1(t + \tau) E_2(t) dt$$

$$- \cos \theta \cos \psi \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{+T/2} N_1(t + \tau) N_2(t) dt =$$

$$= -\sin \theta \sin \psi E_1 E_2(\tau) - \sin \theta \cos \psi E_1 N_2(\tau)$$
$$- \cos \theta \sin \psi N_1 E_2(\tau) - \cos \theta \cos \psi N_1 N_2(\tau) =$$

$$= \begin{pmatrix} -\sin \theta \sin \psi & -\sin \theta \cos \psi & -\cos \theta \cos \psi & -\cos \theta \sin \psi \end{pmatrix} \begin{pmatrix} E_1 E_2(\tau) \\ E_1 N_2(\tau) \\ N_1 N_2(\tau) \\ N_1 E_2(\tau) \end{pmatrix}$$

$$(4.10)$$

I decided to first cross-correlate all possible horizontal component pairs and to rotate afterwards. In this way, less data points (d.p.) have to be stored.

In the following, the advantage of this processing order is explained by an example with three stations having recorded one daily time series each:

If there are three stations, there will be three possible stationpairs (see Formula (3.1)). It is also assumed that the stations are not in line with each other. For both processing schemes the initial situation is the same: All in one there are six horizontal component daily time series (one north and one east component time series per station). Assuming a sampling rate of 1 Hertz this equals $(6 \cdot 87600 \text{ s} \cdot 1 \text{ Hz})(\text{d.p.}) = 525600$ (data points).

If the data are rotated first, there will be two radial and two transverse daily time series per station. Altogether, $3 \text{ (stations)} \cdot 4 \frac{\text{(components)}}{\text{(station)}} \cdot 87600 \frac{\text{(d.p.)}}{\text{(component)}} = 1051200$ (d.p.) have to be stored. If the CCFs (361 d.p. each) of all possible horizontal component pairs (TT, RR, TR, RT) are calculated, the following additional data points have to be saved: $3 \text{ (stationpairs)} \cdot 4 \frac{\text{(rotated component pairs)}}{\text{(stationpair)}} \cdot 361 \frac{\text{(d.p.)}}{\text{(rotated component pair)}} = 4332$ (data points). As 87 CCFs can be computed per day the number of data points increases to: $4332 \cdot 87 = 376884$. Adding the data points of the rotated daily time series and the CCFs
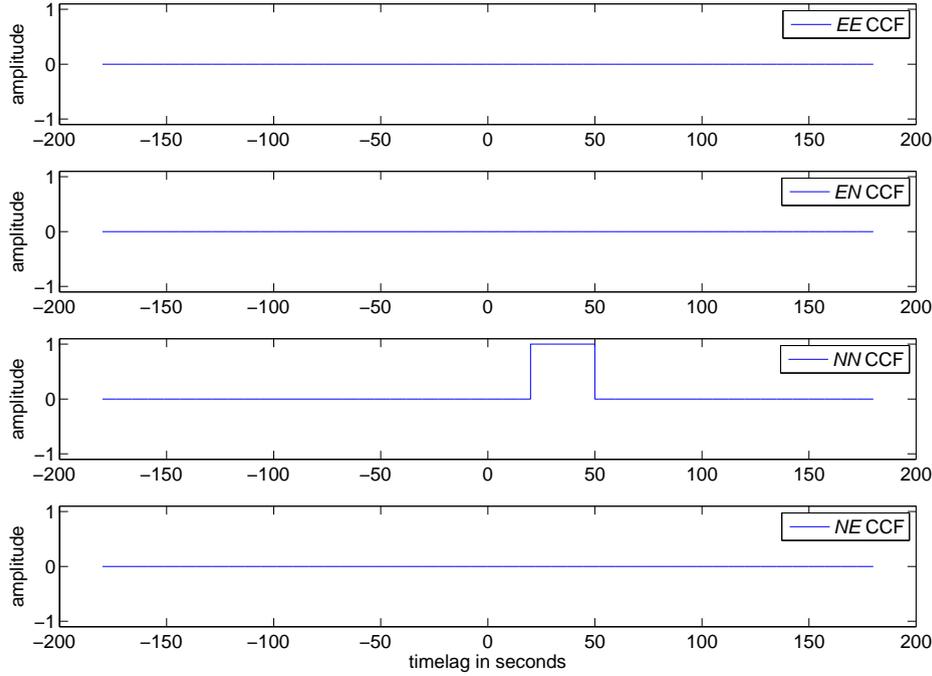
Figure 4.11.: Synthetic cross-correlations of all four possible horizontal component combinations. Only the NN CCF contains a rectangular signal at positive timelag.

one obtains

$$1051200 \text{ (d.p.)} + 376884 \text{ (d.p.)} = 1428084 \text{ (data points)}.$$

In contrast, if the horizontal component data (EE, EN, NN, NE) are first cross-correlated, the number of saved data points is:

3 (stationpairs)$\cdot$4 $\frac{\text{(component pairs)}}{\text{(stationpair)}}\cdot$87 $\frac{\text{(CCFs per day)}}{\text{(component pair)}}\cdot$361 $\frac{\text{(d.p.)}}{\text{(CCF per day)}}$ = 376884 (d.p.). The rotation of these CCFs into the TT, RR, TR and RT CCFs afterwards yields 3 (stationpairs)$\cdot$ 4 $\frac{\text{(rotated component pairs)}}{\text{(stationpair)}}\cdot$ 87 $\frac{\text{(CCFs per day)}}{\text{(rotated component pair)}}\cdot$ 361 $\frac{\text{(d.p.)}}{\text{CCF per day}}$ = 376884 data points. Adding the data points of the initial and the rotated CCFs one obtains

$$376884 \text{ (d.p.)} + 376884 \text{ (d.p.)} = 753768 \text{ (data points)}.$$

Hence, if the waveforms are first cross-correlated and then rotated only about half as many data points need to be stored.

4.4.5. Testing the Rotation Algorithm with a Synthetic Rectangular Function

Before rotating the real data, the functionality of the function XCORR_rotate_horizontal_components (see Appendix B.1) was tested. To keep it comprehensible, a rectangular function serves as test signal. Figure 4.11 shows the four initial synthetic cross-correlation functions. Only the NN cross-correlation contains a signal with an amplitude equal to one at a timelag from 20 s to 50 s. The three other CCFs (EE, EN and NE) are equal to zero. Assuming[4] an azimuth of 45° and a backazimuth of 225° the result of the rotation of the initial signals is shown in Figure 4.12. As expected the TT and the RR cross-correlations 'share' the initial signal to equal parts. The values of the TR and RT CCFs are also right

---

[4]The assumption of the backazimuth being equal to the azimuth plus 180° is not unrealistic, if the stations are located not far apart from each other.
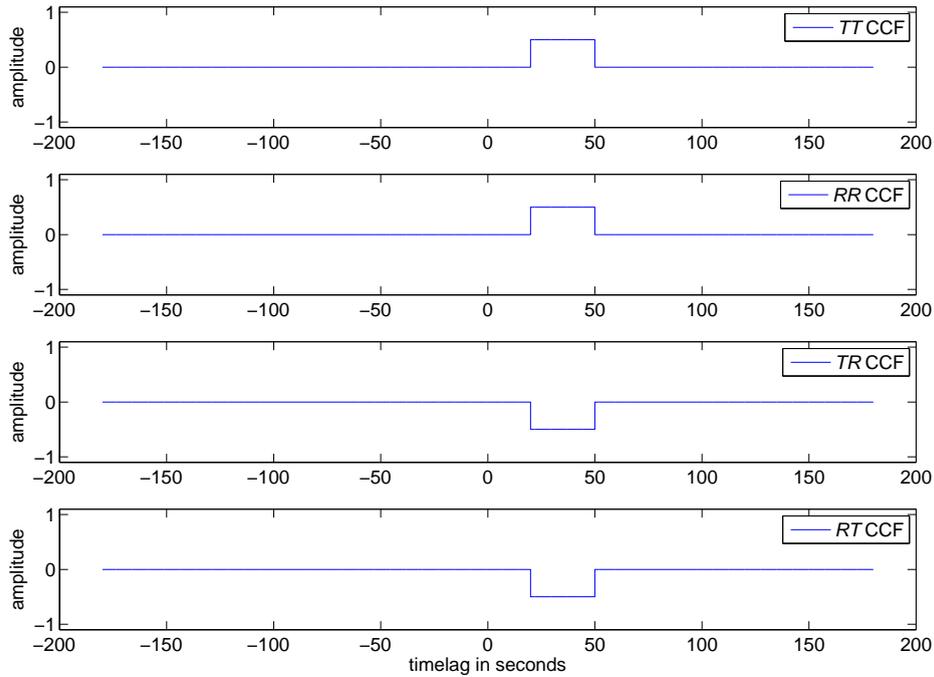
Figure 4.12.: Rotated synthetic cross-correlations. The azimuth is 45°, the backazimuth 225°.

and can be checked with Formula (4.9). Figure 4.13 shows the rotated CCFs with a given azimuth of 90° and a backazimuth of 270°. If one zooms in at the cross-correlations of the RR, TR and RT combinations one can see that the CCFs are not exactly equal to zero due to numerical noise. These uncertainties are of such a small order (between 10^{-17} and 10^{-33}) compared to the expected signal on the TT cross-correlation that they can be neglected.

4.5. Normalisation of the Cross-Correlation Functions

As already mentioned in Chapter 2.3, a normalisation of the CCFs in the frequency domain enhances the retrieval of the Green's function between pairs of stations. Before the influence of this normalisation on the CCFs is investigated, I analyse the differences between the CCFs that are calculated with one-hour and with 20-minutes long time windows. The CCFs of the two stationpairs TMO20-TMO57 and TMO53-TMO54 are chosen. The interstation distance of $\sim$ 16 kilometres between TMO20 and TMO57 is larger than the one between TMO53 and TMO54 ($\sim$ 3 kilometres). The connecting line of TMO20 and TMO57 roughly lies in a west-east direction, the connecting line of TMO53 and TMO54 in a north-south direction (see the maps in Chapter 3). I arbitrarily choose one day of data (Monday, Julian date: 101.2011) of the Z component of the four stations. The CCFs of the two stationpairs are computed using 20 minutes long time windows as well as one-hour long time windows. By stacking the CCFs (used MATLAB script: XCORRstack.m) calculated with the 20 minutes long time windows on the one hand, and the CCFs calculated with one-hour long time windows on the other hand, one obtains two one-day long CCFs, and call them CCF_20m and CCF_1h. Figures 4.14, 4.15, and 4.16 show the difference between these two one-day long CCFs of the stationpair TMO53-TMO54. For the sake of clarity the one-day long CCFs are filtered (used MATLAB script: XCORRfilter.m) in three different frequency bands: 0.1 - 1 Hertz (Figure 4.14), 1 - 5 Hertz (Figure 4.15),
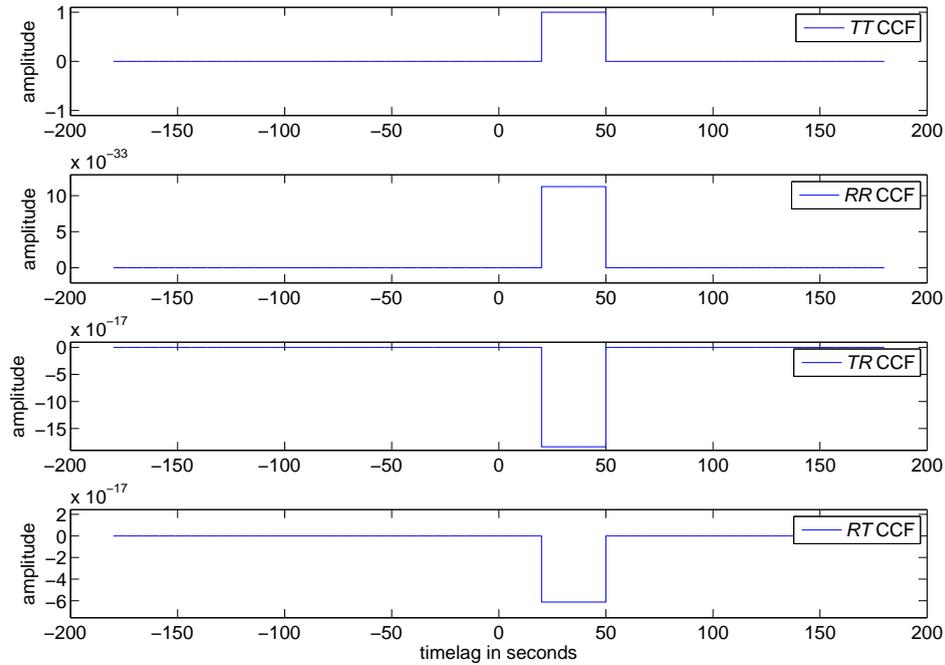
Figure 4.13.: Rotated synthetic cross-correlations. The azimuth is 90°, the backazimuth 270°. Note the different y-axis scales! Although a signal is only expected in the *TT* cross-correlation, signals in the other three cross-correlations also occur. These signals can be reduced to numerical uncertainties.

and 5 - 25 Hertz (Figure 4.16). Due to the linearity of the cross-correlation procedure it should make no difference, if a CCF is obtained by stacking the CCFs calculated with short time windows or by cross-correlating the longer time series (Bensen et al., 2007). However, in Figures 4.14, 4.15, and 4.16 there clearly is a deviation between the two differently obtained one-day long CCFs. The order of this deviation is about one order smaller than the order of the values of the two one-day long CCFs. A reason for this might be the more frequent reoccurence of data sample combinations in the cross-correlations calculated with the shorter time windows due to their overlap of 200 seconds. The results for the stationpair TMO20-TMO57 are not shown, because they are similar to the results of the stationpair TMO53-TMO54.

In the scope of this thesis, it is theoretically possible to estimate the Green's function in a frequency band between 0.1 and 50 Hertz. The lower frequency limit is set by the corner frequency of the high-pass filter in the preprocessing of the data. The sampling rate of 100 Hertz defines the upper limit. As more than two samples per cycle are required to define a frequency component in the original data, the band limiting frequency f_c - also called Nyquist frequency - is equal to:

$$f_c = \frac{1}{2\Delta t} \tag{4.11}$$

(Bendat and Piersol, 1986). Hence, with $\Delta t = 0.01$ seconds one obtains 50 Hertz as the upper frequency limit.

For a proper estimation of the Green's function in certain frequency bands (here with limits in between 0.1 Hertz and 50 Hertz) it is necessary to equalise the spectral amplitudes of the CCFs in the chosen bandwidth. Figure 4.17 shows the raw amplitude spectrum of a one-day long CCF (the stack of all 87 CCFs of Julian day 101.2011) of
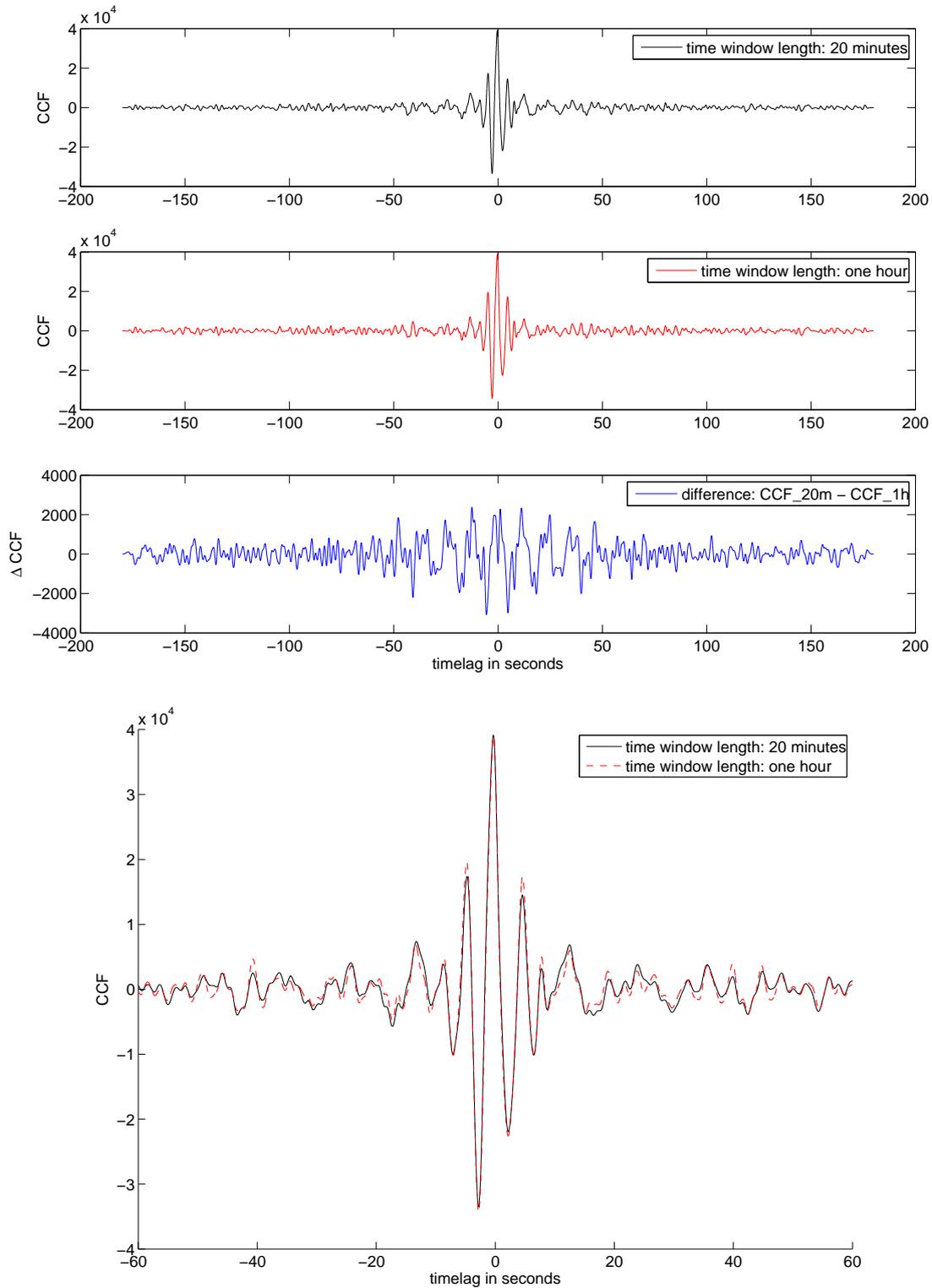
Figure 4.14.: On top in the first row, the one-day long CCF of the *ZZ* components of the stationpair TMO53-TMO54 calculated with 20 minutes long time windows, and filtered in the frequency band from 0.1 to 1 Hertz. At timelag zero a low-frequency signal clearly emerges. On top in the second row, the CCFs are computed using one-hour long time windows. The difference between the CCF_20m and CCF_1h is shown below. The difference between the CCFs is one order smaller than the order of the values of the CCFs above. At the bottom, the CCF_20m and CCF_1h are plotted upon each other. Deviations especially in the amplitudes outside of the signal range are bigger than around timelag zero.
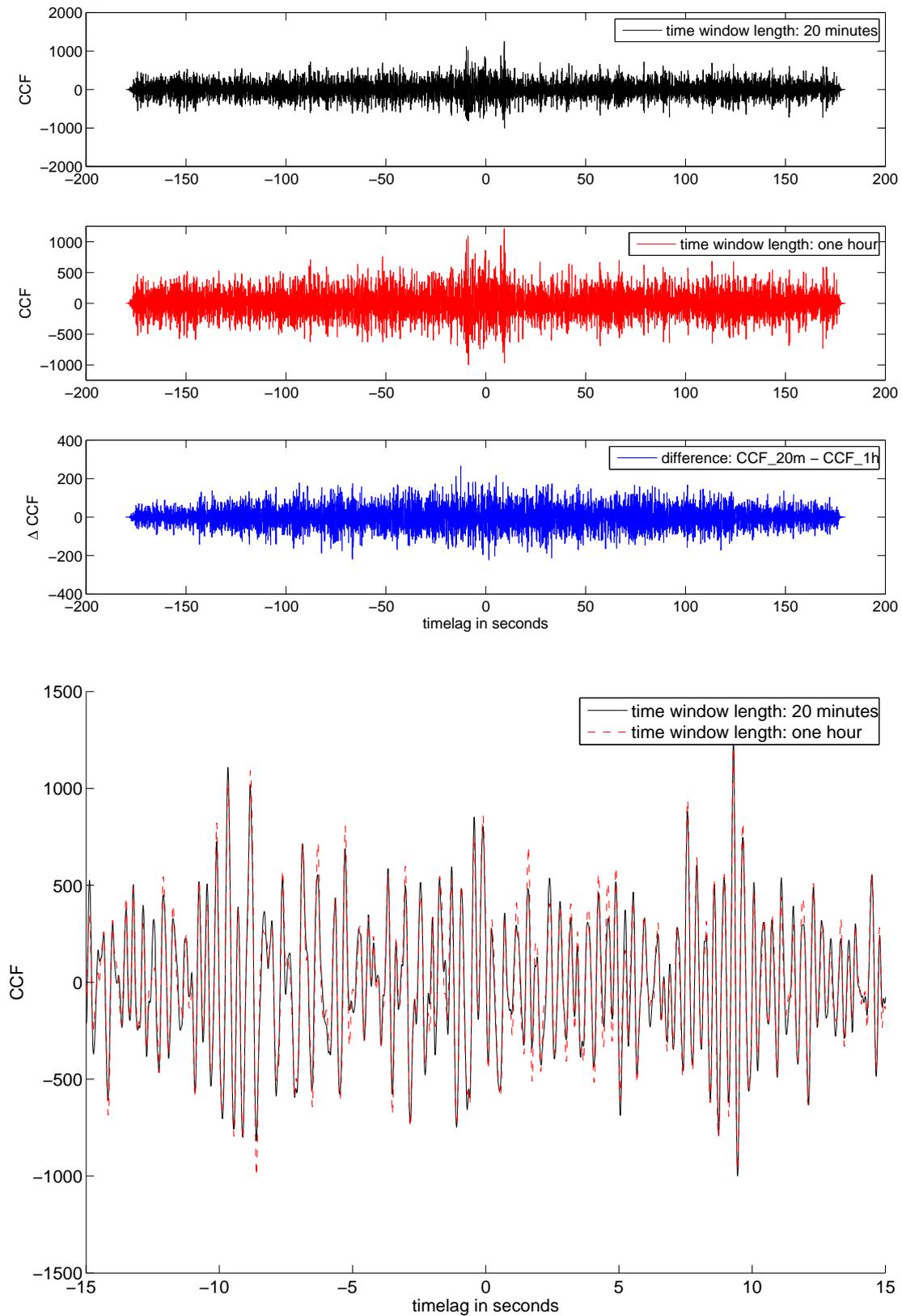
Figure 4.15.: The same as in Figure 4.14 with the CCFs of the ZZ components filtered between 1 - 5 Hertz. At the timelags of about ± 10 seconds two signals emerge. The difference between the CCF_20m and the CCF_1h is again about one order smaller than the order of the values of the CCFs above.
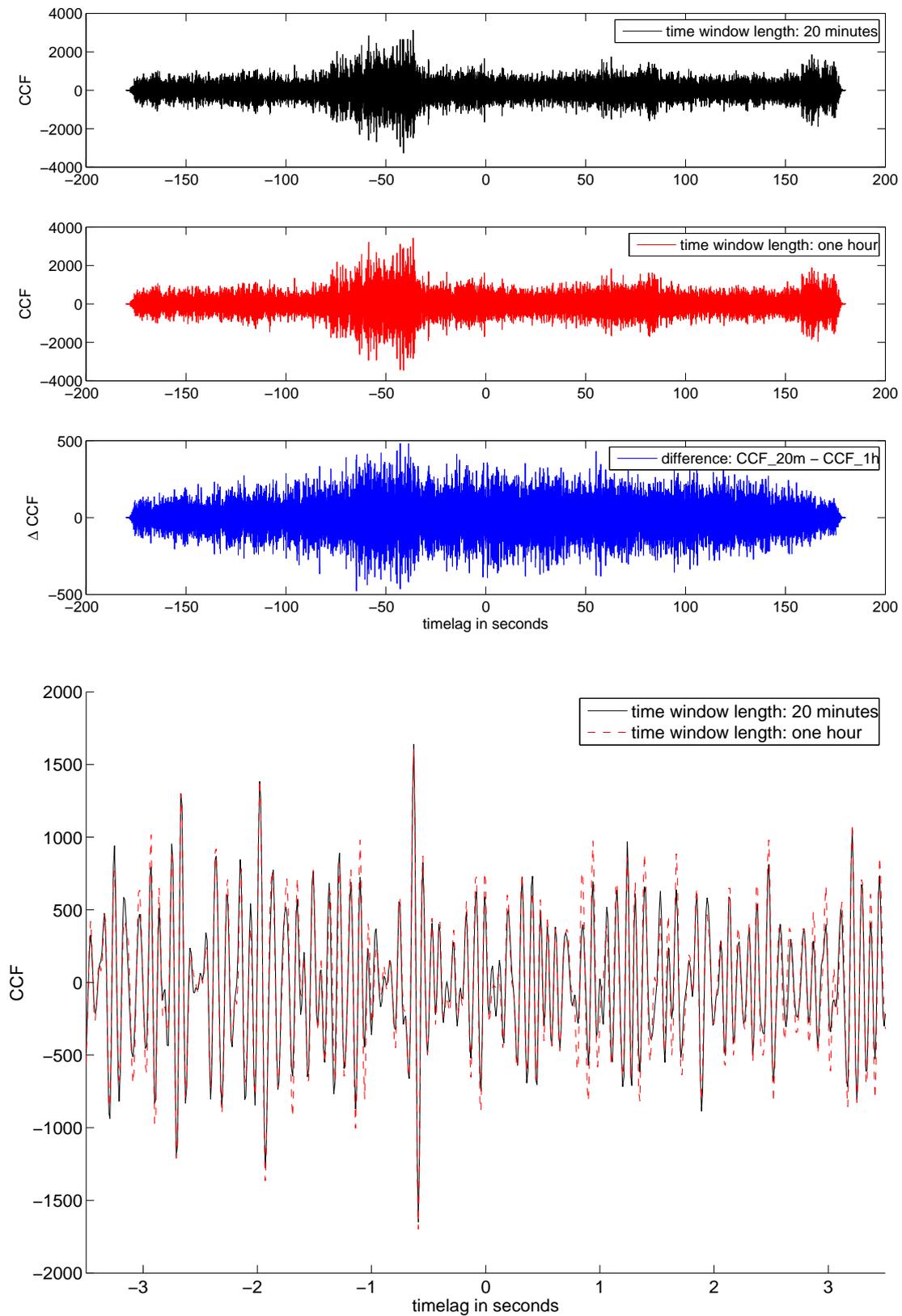
Figure 4.16.: The same as in Figure 4.14 with the CCFs of the *ZZ* components filtered between 5 - 25 Hertz. The order of the difference of the CCFs is again one order smaller than the order of the CCF values above.
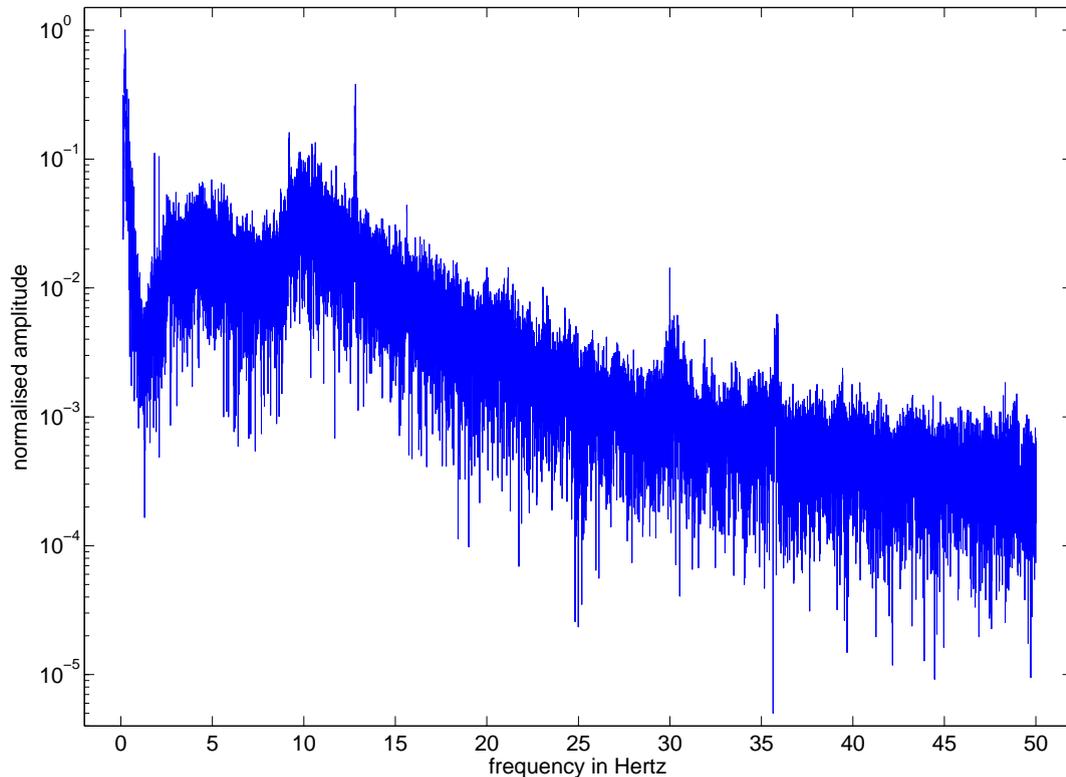
Figure 4.17.: The amplitude spectrum of a one-day long CCF (calculated with the daily time series of the Julian day 101.2011 (Monday)) of the stationpair TMO20-TMO57. At low frequencies between 0.1 Hertz and 1 Hertz, the amplitudes reach their maximum. Below 25 Hertz, there are also high amplitudes. Above 25 Hertz the amplitudes are very low (with two exceptions around 30 Hertz and around 36 Hertz). At about 13 Hertz a relatively high amplitude peak occurs.

the stationpair TMO20-TMO57. This spectrum is calculated with the MATLAB function XCORRgetFFT (see Appendix B). The spectral amplitudes are far from being equal. At very low frequencies (0.1 - 1 Hertz) the amplitudes are very high. Another amplitude peak occurs at about 13 Hertz. Between ∼ 25 Hertz and 50 Hertz the amplitudes are very small. By forcing the amplitudes of the complex spectrum in the desired frequency band to one and to zero elsewhere, one obtains an equalised amplitude spectrum. This method is proposed by Brenguier et al. (2008) and called Spectral Whitening (SW). As only a bandwidth is considered in the normalisation process, Spectral Colouring would be a better name than Spectral Whitening. Note, that this normalisation in the frequency domain implies a non-linear modification of the CCFs.

How the cross-correlations change by applying Spectral Whitening is shown in Figure 4.18. After the normalisation in the frequency domain, the low frequencies do not dominate the cross-correlation as much as before. In the example in Figure 4.18, even a signal at a lag time of -50 seconds appears that was invisible without the Spectral Whitening. The MATLAB function used for the Spectral Whitening is XCORRspectralwhitening_withlimitsinput (see Appendix B).

In many passive seismic interferometry studies the data are not only normalised in the frequency domain but also in the time domain (TD). A very common processing scheme is the normalisation of the time series in the time domain prior to the calculation of the
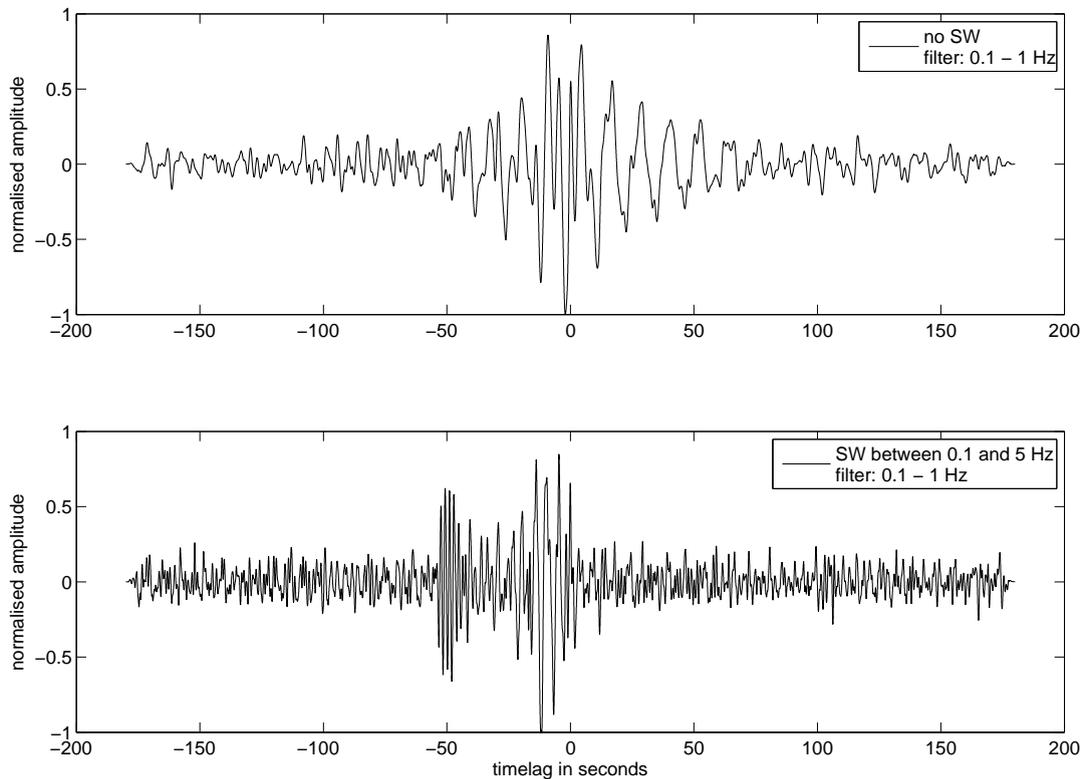
Figure 4.18.: One-day long CCF of the *ZZ* component combination of the stationpair
TMO20-TMO57 with (at the bottom) and without (on top) Spectral Whiten-
ing. The Spectral Whitening is applied to each of the 87 CCFs of the Julian
day 101.2011. Afterwards the CCFs are stacked, and then filtered between
0.1 Hertz and 1 Hertz. The same applies to the CCF shown on top except
for the Spectral Whitening. A signal clearly emerges in the normalised CCF
at a lag time of -50 seconds. Also the higher frequencies appear more clearly
in the normalised CCF. In a last step, the CCFs were both normalised to
their respective maximum.

CCFs. This procedure shall reduce the effects on the later calculated cross-correlations of, for example, earthquakes or instrumental irregularities (Bensen et al., 2007). Afterwards, these CCFs are normalised in the frequency domain. Bensen et al. (2007) suggested this processing of seismic noise waveforms in order to obtain useful CCFs. However, Groos et al. (2012) observed no systematic improvement of these CCFs which are normalised in the time and the frequency domains in comparison to those CCFs wich are only normalised in the frequency domain depending on the time window length. But Groos et al. (2012) also stated that a time domain normalisation might improve the CCFs when data sets are used not on a continental but e.g. on a local scale. Hence, a waveform preserving normalisation of the CCFs is tested with this data set (used MATLAB function: XCORRtimedomainnormalisation). Therefore, each data point of the cross-correlation is divided by the root mean square (RMS) value of the entire waveform. There are also other (non-linear) methods of normalising data in the time domain. But in the scope of this thesis they will not be discussed. Here, it is just pointed out that the used waveform preserving time domain normalisation is capable to replace the other time domain normalisation methods (Groos et al., 2012). Again, the CCFs of the stationpair TMO20-TMO57 are used. One of the CCFs, which are compared to each other (Figure 4.19), is obtained in the following procedure: Each CCF of the 87 CCFs of one day is normalised in the time domain. Afterwards, these normalised 87 CCFs are stacked to a one-day long CCF. Furthermore, a Spectral Whitening is applied to the one-day long CCF. Altogether, 60 one-day long CCFs are calculated this way (Julian days 101. - 160.2011). In a final step, these 60 normalised one-day long CCFs are stacked, so that one obtains a so-called two month CCF. For the computation of the other two month CCF only the step of the time domain normalisation is left out. Again for the sake of clarity, the CCFs are filtered between 0.1 Hertz and 1 Hertz.

Concluding, for the following analysis of the CCFs one has to be aware of the differences between the vertical and the horizontal component CCFs due to the different time window lengths that were used for the calculation of the CCFs. The processing step (3) consists of a normalisation of the CCFs only in the frequency domain, because a preceding normalisation in the time domain does not change the CCFs significantly. The following processing scheme should provide useful cross-correlations: All CCFs obtained by cross-correlating the time series fragments of a daily time series are normalised in the frequency domain. The bandwidth, in which the amplitudes of the complex spectrum are set to one, should be at least as big as the corner frequencies of the bandpass filter applied afterwards. Then, these CCFs of one day are stacked, so that one obtains one-day long CCFs. Further stacking and filtering is optional and depends on the application. Due to the linearity of the stacking process, it does not matter if the CCFs are stacked before or after the filtering. If horizontal and vertical CCFs are compared to each other the processing scheme above should be slightly changed for the vertical CCFs. If four consecutively CCFs (four times 20 minutes yields 80 minutes) of one day are stacked, and only then spectral whitened, the difference is smaller between the one-day long CCFs calculated with 20 minutes and with one-hour long time series. This fact is emphasised in Figure 4.20.
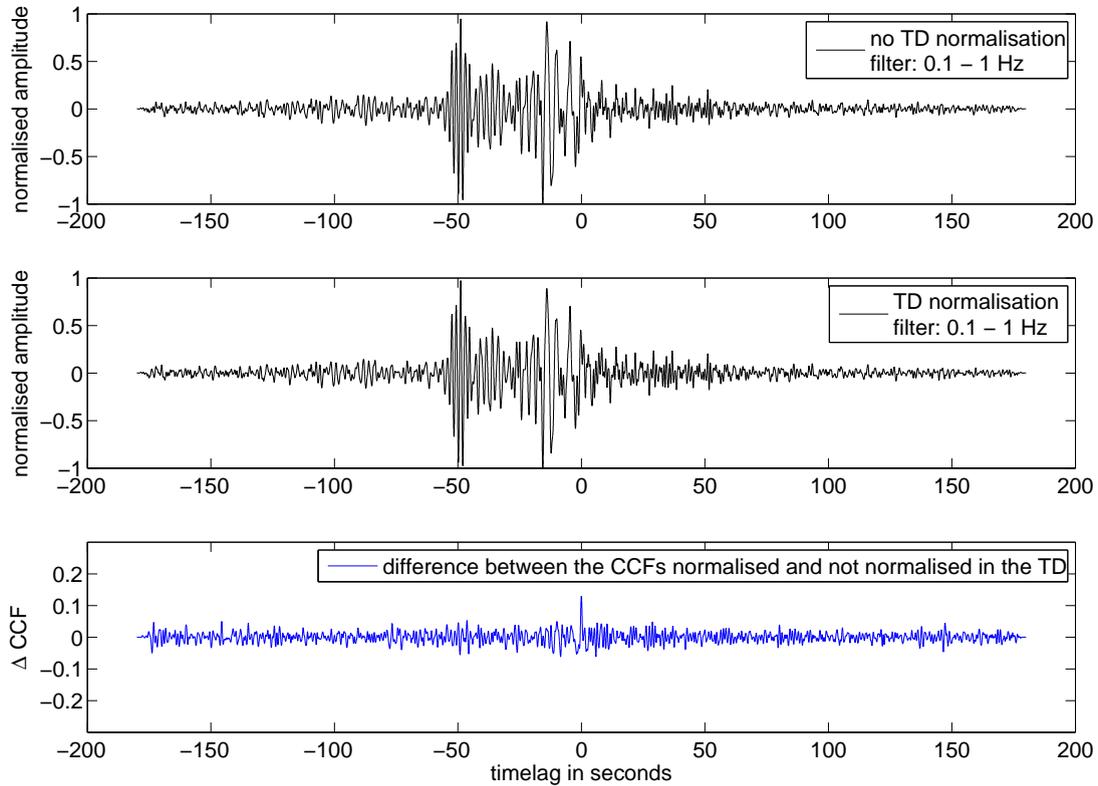
Figure 4.19.: Two-month long CCFs of the ZZ component combination of the stationpair
TMO20-TMO57 with (in the middle) and without (on top) a normalisation
in the time domain. Spectral Whitening between 0.1 Hertz and 2 Hertz is
applied in both cases to the one-day long CCFs. Afterwards, 60 one-day long
CCFs are stacked. In a last step, the CCFs were filtered between 0.1 Hertz
and 1 Hertz, and both normalised to their respective maximum to obtain the
two month CCFs shown above. At the bottom, the difference of the different
processing schemes is depicted.
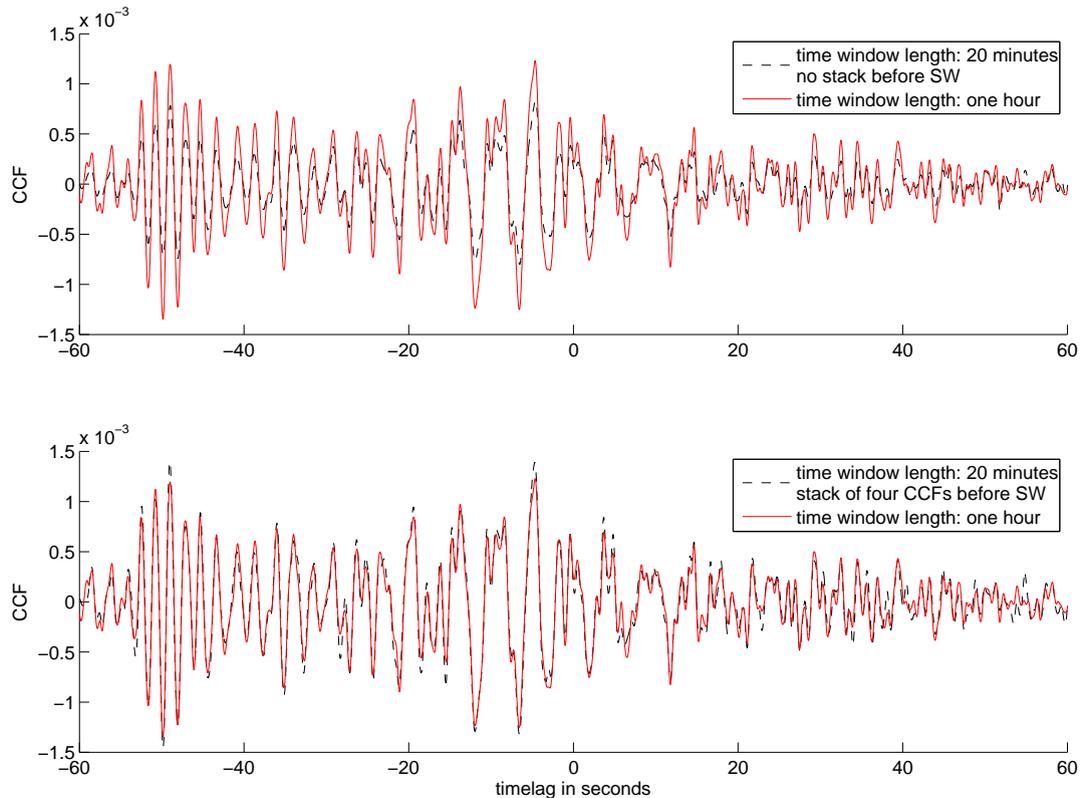
Figure 4.20.: One-day long CCFs of the stationpair TMO20-TMO57 of Julian day 100.2011 (Sunday). On top: The comparison of the one-day long CCF calculated with one-hour long time windows and the one-day long CCF calculated with 20 minutes long time windows. At first a Spectral Whitening (0.1 Hertz to 20 Hertz) is applied to all the obtained CCFs calculated with the time series fragments. In the case of 20 minutes segments, the number of normalised CCFs is 87, in the case of one hour segments it is 25. Afterwards, the 87 CCFs are stacked and filtered between 0.1 Hertz and 1 Hertz (black dashed line), and the 25 CCFs are also stacked and filtered in the same frequency band (red curve). At the bottom: The same as above except of one additional processing step applied to the CCFs calculated with 20 minutes long time windows. Here, four consecutively CCFs of the raw 87 CCFs are stacked, and then normalised in the frequency domain. This way, 21 normalised CCFs are obtained, and stacked and filtered afterwards. With the additional stacking step, the CCFs calculated with one hour and with 20 minutes long time windows look more the same due to the amplification of the signals of the black dashed curve.

5. Results

The first part of this chapter (Section 5.1) deals with the analysis of the properties of the cross-correlations, calculated with the TIMO2 data set. It is investigated if the cross-correlations are usable for passive imaging (Subsections 5.1.1 and 5.1.2) or passive monitoring techniques (Subsection 5.1.3). In the second part of this chapter (Section 5.2) noise sources which dominate the cross-correlations are analysed. At the end of every subsection the obtained results are summarised and discussed.

5.1. Properties of the Cross-Correlations

In this section the cross-correlations are investigated with respect to signals. In the first subsection (Section 5.1.1) the similarities and differences between all the CCFs of the year 2012 of the ZZ, RR, and TT components are discussed. Afterwards, more detailed information on the observed coherent signals in the CCFs are given by the analysis of the CCFs of three different stationpairs. The symmetry relative to timelag zero and the polarisations of the waveforms in the cross-correlograms of the ZZ, RR, and TT components are investigated. The third part of this section (Subsection 5.1.3) deals with the temporal stability of signals in selected cross-correlograms.

5.1.1. Similarities and Differences between the Vertical and Horizontal Cross-Correlations

The CCFs of the ZZ, RR, and TT components are investigated with respect to coherent signals. In this context, coherent signal means the coherent noise from common sources, which is recorded by both stations of a cross-correlation pair (Lin et al., 2008).
A very common method for depicting cross-correlograms of many stationpairs is the distance plot. In this sort of plots the cross-correlograms are normalised to their respective maximum and plotted against the corresponding interstation distance. Figure 5.1 exemplarily shows such a distance plot of the CCFs of the ZZ components. The CCFs of the available data of the year 2012 (Julian days 1 to 305) are obtained by stacking all the one-day long CCFs (see Section 4.5). No band-pass filter is applied to the CCFs. Thus, Figure 5.1 shows the cross-correlograms with frequencies above 0.1 Hertz (0.1 Hertz is the corner frequency of the high-pass filter used in the preprocessing of the time series, see Section 4.1). Various wavegroups occur in the cross-correlograms in Figure 5.1. Moreover, an asymmetry is observed between the causal and acausal parts of the cross-correlograms, as low-frequency signals dominate the acausal part of the cross-correlograms.
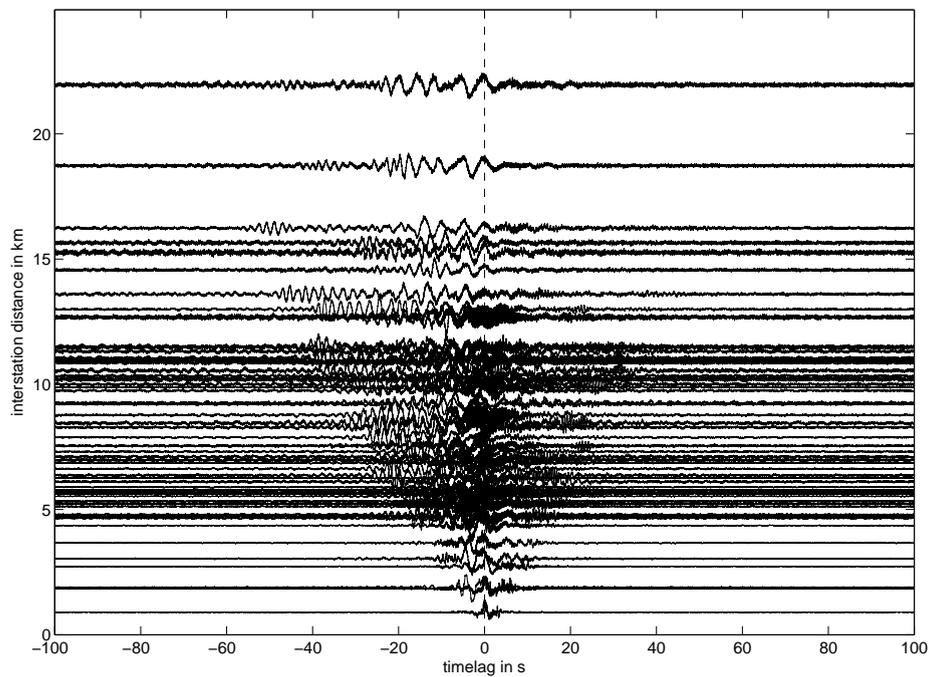
Figure 5.1.: The distance plot of all available data of 2012 of the ZZ components. The
CCFs are not filtered. Therefore, the cross-correlograms with frequencies
above 0.1 Hertz (preprocessing) are depicted. An asymmetry between the
signals in the causal and in the acausal part is observed. Especially at in-
terstation distances from about 4.5 kilometres to about 12 kilometres the
cross-correlograms lie very close to each other.

Table 5.1.: The nine frequency ranges used for the detailed analysis of the cross-correlation functions.

frequency range	lower corner frequency in Hertz	upper corner frequency in Hertz
a)	0.1	0.2
b)	0.2	0.4
c)	0.4	0.8
d)	0.8	1.6
e)	1.6	3.2
f)	3.2	6.4
g)	6.4	12.8
h)	12.8	25.6
i)	20	45

Due to the broad frequency range and several stationpairs having almost the same interstation distances, the plot in Figure 5.1 is not appropriate for a further investigation of the signals in the CCFs. Hence, I decided in a first step to divide the frequency band between 0.1 Hertz and 50 Hertz, which can theoretically be analysed (see Section 4.5), into nine different, smaller frequency bands comprising one octave (except for the last frequency range, see Table 5.1). In a second step, the interstation distances are split into four ranges: 0 - 4.5 kilometres, 4.5 - 8.5 kilometres, 8.5 - 12.5 kilometres, and 12.5 - 22.5 kilometres. The distribution of the interstation distances is given by the histogram in Section 3.2. This way, nine times four distance plots for each component combination (ZZ, RR and TT) are obtained. The MATLAB scripts used for the distance plots are: XCORR_distanceplot.m; XCORR_distanceplot_section.m; XCORR_distanceplot_stacked_CCFs.m (see Appendix B.2).

Dependent on the frequency range and the interstation distance the following properties of the cross-correlograms are enquired: Do the cross-correlograms contain coherent signals? If they do, will the dominant signals occur at timelag zero, only in the causal or only in the acausal or in both parts of the cross-correlograms? If signals occur in the cross-correlograms at a timelag of 0 seconds, the signals will arrive at both stations of a cross-correlation pair at the same time. If there are signals, for example, in the acausal part, they will arrive at the western station first, and at the eastern station afterwards. The reason for this is the chosen order of the input variables for the MATLAB xcorr-function: I always defined the time series of the western station as the first input variable and the time series of the eastern station as the second input variable (see Section 4.3).

It starts with the investigation of the CCFs filtered in the lowest frequency band between 0.1 Hertz and 0.2 Hertz. All the cross-correlograms of each component combination, i.e. ZZ, RR, and TT, contain a dominant signal either at timelag zero or in the acausal part. Figures 5.2, 5.3, and 5.4 show the distance plots of the ZZ, the RR, and the TT components. In general, the signal is slightly shifted to the acausal part of the cross-correlograms with increasing interstation distance. Moreover, the signal range in the cross-correlograms of the TT components extends to higher negative timelag values in comparison to the signal ranges occurring on the ZZ and RR cross-correlograms. For example, the signal in the cross-correlogram of the TT components of the stationpair with the largest interstation distance of $\sim$ 22 kilometres lasts from about 0 seconds to -35 seconds, while on the ZZ component it lasts from about 10 seconds to -20 seconds, and on the RR component only from 0 seconds to -20 seconds.

The common criterion that the interstation distance d has to be larger than three times the wavelength λ so that the far field approximation is satisfied (Lin et al., 2008), is not
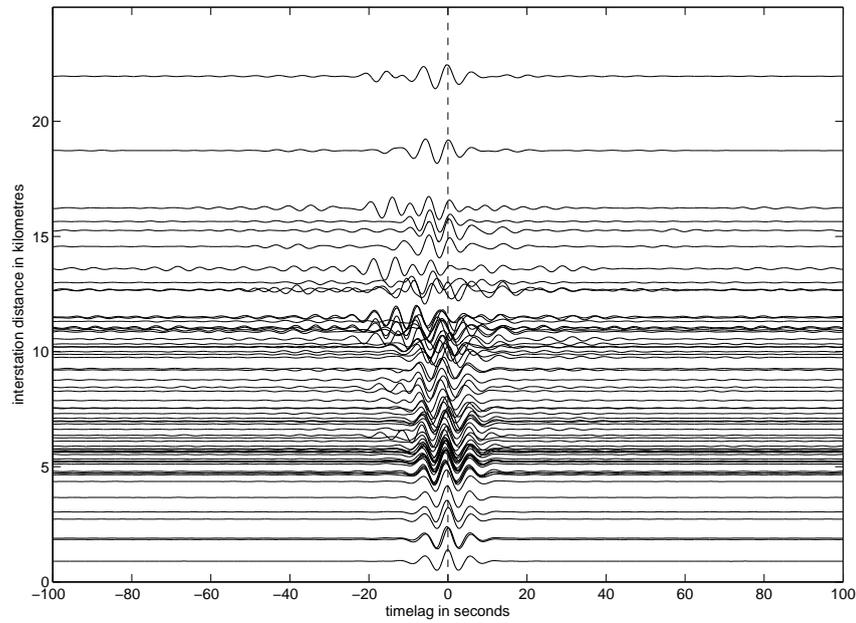
Figure 5.2.: The distance plot of the available data of the year 2012 of the ZZ components. The CCFs are filtered in the frequency band from 0.1 Hertz to 0.2 Hertz. There is a dominant signal in all cross-correlograms.
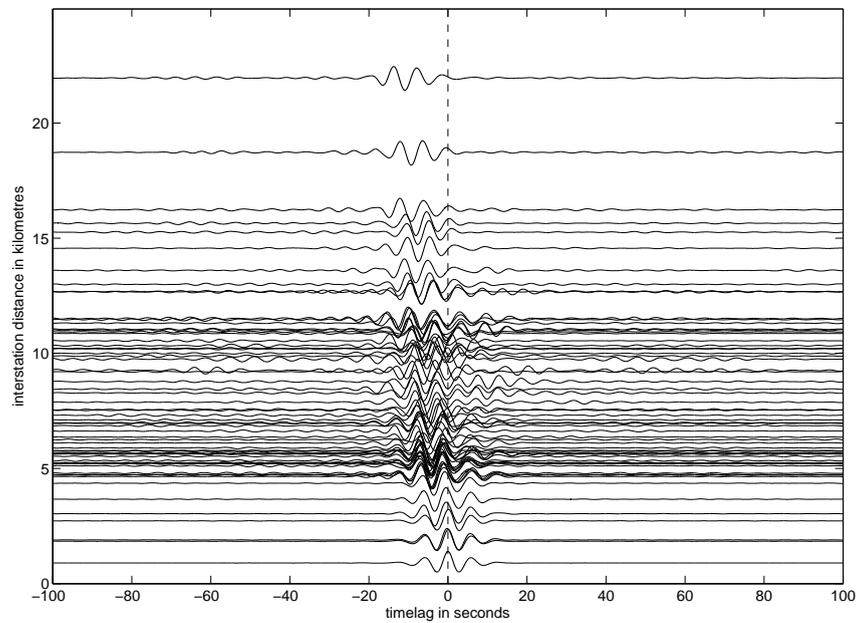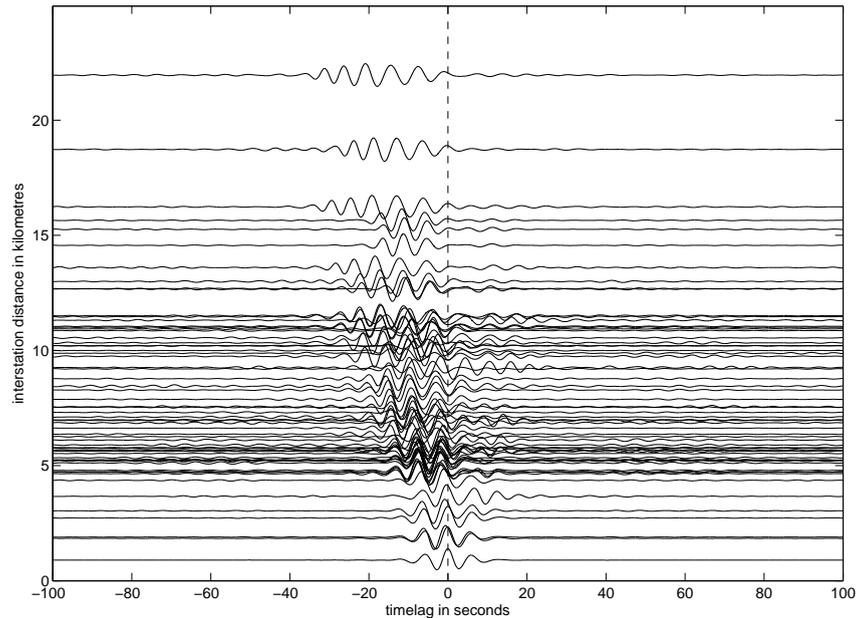

Figure 5.3.: The distance plot of all available data (RR components) of the year 2012. The CCFs are filtered between 0.1 Hertz and 0.2 Hertz. Very similar to the cross-correlograms of the ZZ components the signals are slightly shifted to the acausal part of the cross-correlogram with increasing interstation distances. Signals in the acausal part of a cross-correlogram arrive at first at the western station of a stationpair, and secondly at the eastern station.

Figure 5.4.: The distance plot of all available data (TT components) of the year 2012. The CCFs are filtered between 0.1 Hertz and 0.2 Hertz. In many cross-correlograms, the acausal part is dominated by high-amplitude signals.

fulfilled by the data set in the frequency range from 0.1 Hertz to 0.2 Hertz. Thus, based on the largest interstation distance of $\sim$ 22 kilometres waves of a maximum wavelength of 7.3 kilometres can properly be analysed. In turn, with the frequency of $f = 0.1$ Hertz and the maximum wavelength $\lambda = 7.3$ kilometers, the velocity v of the wave yields:

$$v = f \cdot \lambda = 0.1 \cdot 7.3 \ \frac{\text{km}}{\text{s}} = 730 \ \frac{\text{m}}{\text{s}}. \tag{5.1}$$

With the frequency of 0.2 Hertz, on the other hand, the wave velocity should not exceed 1470 metres per second to satisfy the criterion $d > 3 \cdot \lambda$. Under the assumption of surface waves the propagating velocities can be estimated to values between 4 kilometres per second and 2 kilometres per second in the bandwidth from 0.1 Hertz to 0.2 Hertz (e.g. Groos and Ritter, 2009; Poli et al., 2013). These estimated values are much higher than the maximum velocities which can be analysed in this frequency band with the TIMO2 cross-correlations.

The asymmetric occurrence of the signals in the cross-correlograms filtered from 0.1 Hertz to 0.2 Hertz implies a non-uniform distribution of noise sources. If the Fresnel zones or Fresnel volumes of every stationpair were covered by noise sources, the cross-correlograms would contain symmetric signals (see Section 2.3). Here, the signals of many cross-correlograms occur mainly in the acausal part. Thus, I conclude that the source(s) for these signals is (are) located closer to the western stations than to the eastern stations. A further analysis of the origin of the noise sources in this frequency range is described at the beginning of Section 5.2.

In the frequency bands b), and c) (0.2 Hertz to 0.4 Hertz, and 0.4 Hertz to 0.8 Hertz) high-amplitude signals also emerge in the causal part of several cross-correlograms. But still, the signal portion in the acausal part of the cross-correlograms is higher than the one in the causal part. Especially the cross-correlograms of stationpairs with very long interstation distances do not contain high-amplitude signals in the causal part. These facts apply to the cross-correlograms of all three component combinations. More information
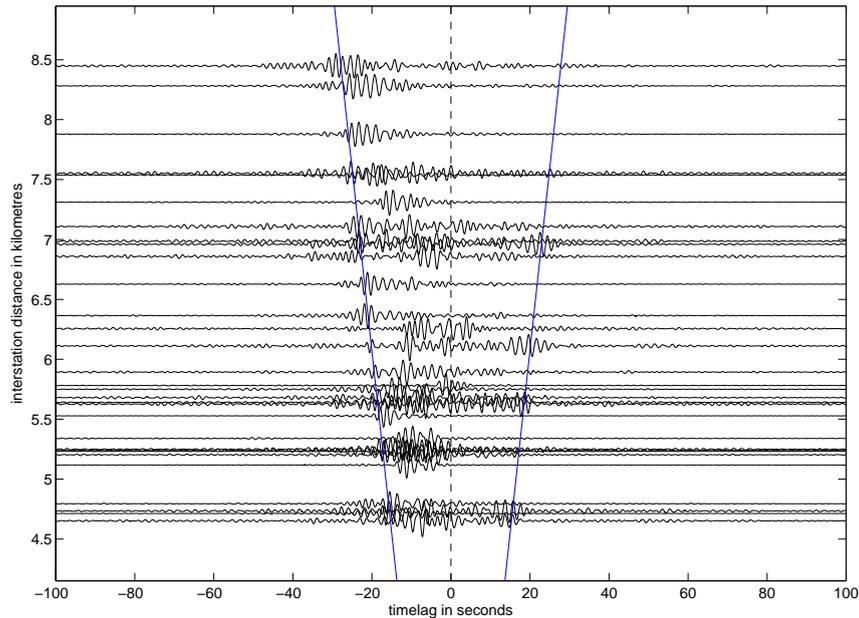
Figure 5.5.: The distance plot of the available data of the year 2012 of the ZZ components. The CCFs are filtered in the frequency band from 0.4 Hertz to 0.8 Hertz. Only the interstation distance interval from 4.5 kilometres to 8.5 kilometres is depicted. In some cross-correlograms there are signals not only in the acausal but also in the causal part. Signals in the causal part belong to waves propagating in a roughly east-west direction. Signals in the acausal part belong to waves propagating in a roughly west-east direction. The blue lines mark the arrival times of signals propagating with an apparent velocity of 300 metres per second.

on the origin of the noise sources, which contribute to the dominant signals in the cross-correlograms, are given in Section 5.2. Examples of the cross-correlograms of the ZZ, RR, and TT components filtered between 0.4 Hertz and 0.8 Hertz are shown in Figures 5.5, 5.6, and 5.7. Only the cross-correlograms in the interstation distance interval from 4.5 kilometres to 8.5 kilometres are depicted. In each plot two blue lines mark an apparent velocity of 300 metres per second. Note that the apparent velocity is not equal to the actual propagation velocity of the signal. This is only the case, if the signal propagates along the connecting line of a stationpair. The apparent velocity is just the quotient of the interstation distance of a cross-correlation pair and the time difference of the arrival times of a signal at the two stations.

In all three distance plots (Figures 5.5, 5.6, and 5.7) high-amplitude signals mainly occur at the inner region that is limited by the 300 metres per second velocity lines. In the following, I will call this region the signal area. In contrast to the ZZ and RR cross-correlograms the signals of the TT cross-correlograms appear more symmetrically. In these cross-correlograms a more consistent arrival of relatively narrow wavelets can be observed at the left and right sides of the signal area. The same observation applies to the TT cross-correlograms in the interstation distance intervals of 0.5 kilometres and 4.5 kilometres, and of 8.5 kilometers and 12.5 kilometres.

In frequency band d) (0.8 - 1.6 Hertz) the averaged portion of signals is nearly the same in the causal and in the acausal parts of the cross-correlograms of all three component combinations. Figures 5.8, 5.9, 5.10, and 5.11 show the cross-correlograms of the TT components in all four interstation distance intervals. Again, blue lines mark the arrival times
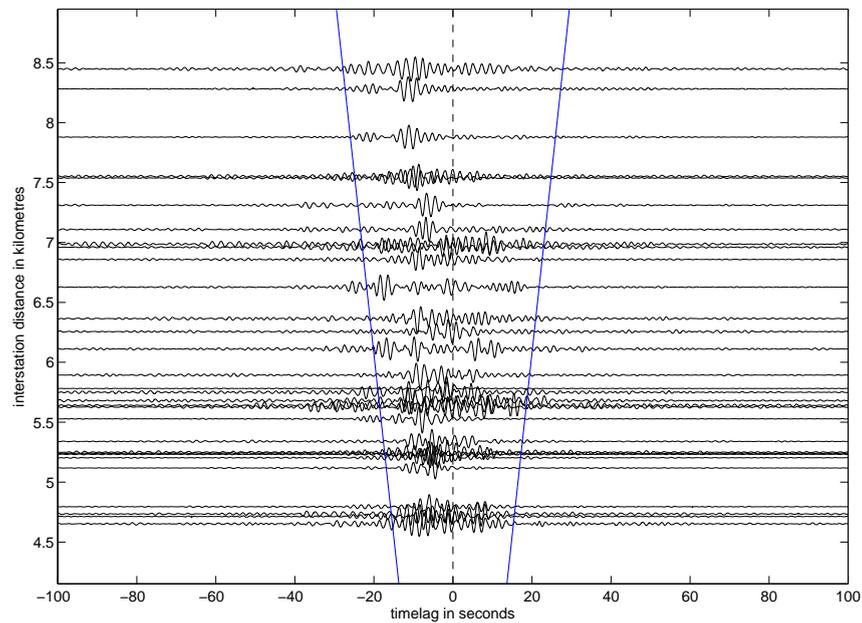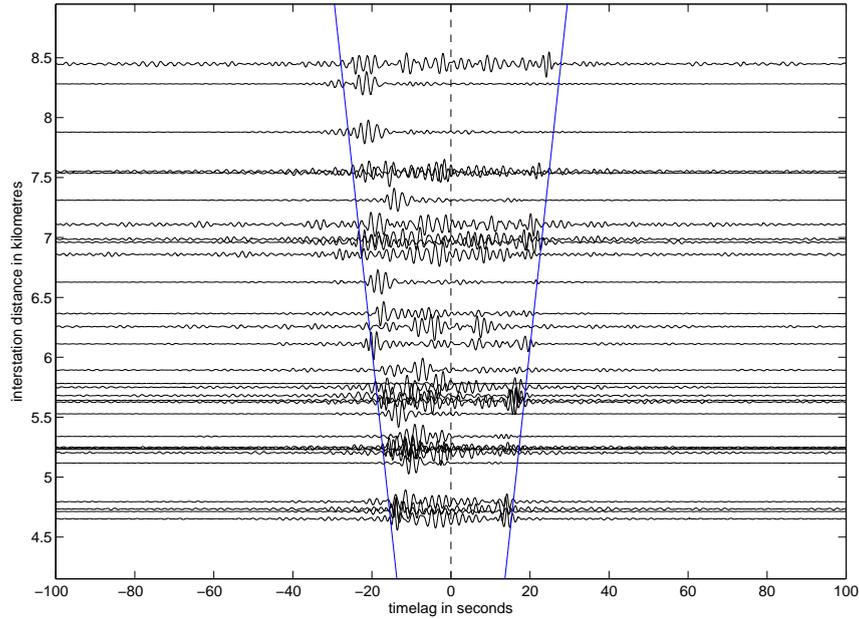
Figure 5.6.: The distance plot of the available data of the year 2012 of the RR components. The CCFs are filtered in the frequency band from 0.4 Hertz to 0.8 Hertz. Only the interstation distance interval from 4.5 kilometres to 8.5 kilometres is depicted. In some cross-correlograms, there are signals not only in the acausal but also in the causal part. Compared to the signals depicted in the distance plot of the ZZ components (Figure 5.5) the high-amplitude signals in the RR cross-correlograms arrive earlier on average.

Figure 5.7.: The distance plot of the available data of the year 2012 of the TT components. The CCFs are filtered in the frequency band from 0.4 Hertz to 0.8 Hertz. Only the interstation distance interval from 4.5 kilometres to 8.5 kilometres is depicted. In many cross-correlograms, there are signals not only in the acausal but also in the causal part. A form of symmetry of some signals in several cross-correlograms can be observed.

of signals travelling with a velocity of 300 metres per second. Several cross-correlograms with interstation distances smaller than about eleven kilometres contain symmetric signals.

In the frequency band from 0.8 Hertz to 1.6 Hertz the symmetric signals of some TT cross-correlograms emerge even more clearly than in the frequency band from 0.4 Hertz to 0.8 Hertz. Not only the cross-correlograms of the TT, but also the cross-correlograms of the ZZ components contain symmetric signals in the frequency range of 0.8 Hertz and 1.6 Hertz. The ZZ cross-correlograms between 4.5 kilometres and 8.5 kilometres are shown in Figure 5.12. For example, the ZZ cross-correlogram with an interstation distance of about 8.5 kilometres contains a symmetric signal at ± 24 seconds.

Compared to the TT cross-correlograms there are more superimposed signals in the signal area of the ZZ cross-correlograms. Hence, the determination of symmetric signals in the ZZ cross-correlograms is not as clear as in the TT cross-correlograms.

The CCFs of the RR components behave similar to those of the ZZ components. At first sight, no symmetric signal can be identified in the RR cross-correlograms filtered between 0.8 Hertz and 1.6 Hertz (see Figure 5.13).

The cross-correlograms filtered between 1.6 Hertz and 3.2 Hertz of all three component combinations show that at very large interstation distances a clear differentiation between coherent noise (signal) and incoherent noise (noise) is not possible anymore. An example is given in Figure 5.14 with the cross-correlograms of the ZZ components.

Furthermore, many cross-correlograms filtered between 1.6 Hertz to 3.2 Hertz contain periodically reoccurring 'signal packets'. These 'signal packets' look like beats. Figure 5.15 shows the ZZ cross-correlogram of 2012 of stationpair TMO20-TMO58 (interstation distance $\sim$ 10 kilometres) filtered in the frequency range from 1.6 Hertz to 3.2 Hertz. On the one hand, broad periodically reoccurring 'signal packets' are observed (top of Figure
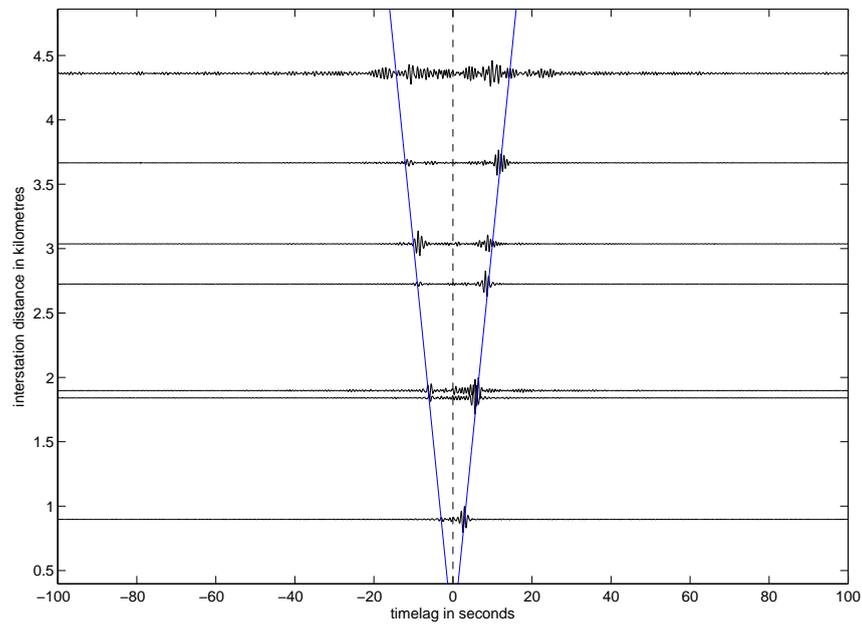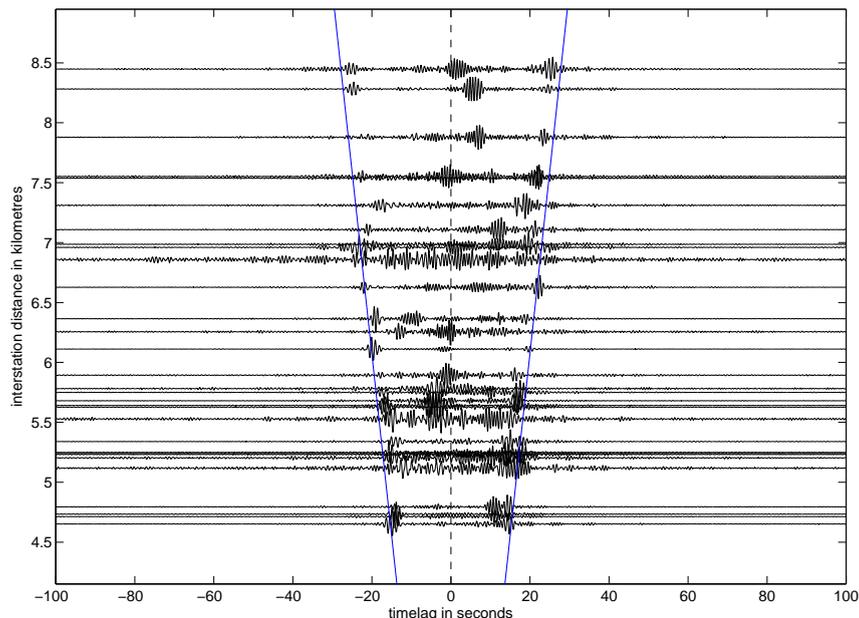
Figure 5.8.: The distance plot of the available data of the year 2012 of the TT components. The CCFs are filtered in the frequency band from 0.8 Hertz to 1.6 Hertz. Only the interstation distance interval from 0.5 kilometres to 4.5 kilometres is depicted. Even though the amplitudes of the signals in the causal and the acausal parts of one cross-correlogram are not equalised, a symmetry, for example of the signals recorded by the stationpair located about 3 kilometres apart, can clearly be observed.

Figure 5.9.: The distance plot of the available data of the year 2012 of the TT components. The CCFs are filtered in the frequency band from 0.8 Hertz to 1.6 Hertz. Only the interstation distance interval from 4.5 kilometres to 8.5 kilometres is depicted. Many of the cross-correlograms feature symmetric signals, which travel between the pairs of stations with an apparent velocity slightly higher than 300 metres per second.
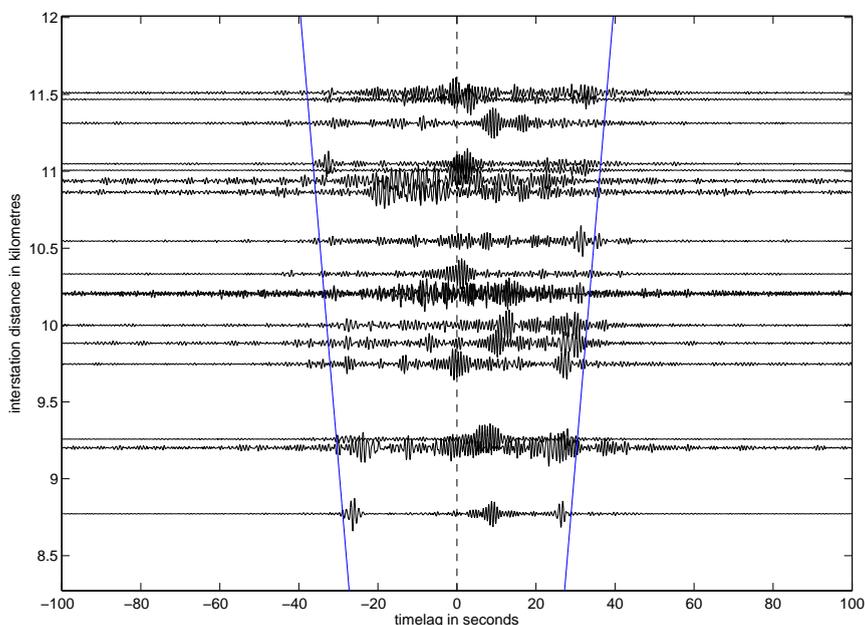

Figure 5.10.: The distance plot of the available data of the year 2012 of the TT components. The CCFs are filtered in the frequency band from 0.8 Hertz to 1.6 Hertz. Only the interstation distance interval from 8.5 kilometres to 12.5 kilometres is depicted. As in Figures 5.8, and 5.9 symmetric signals can be observed in several cross-correlograms.
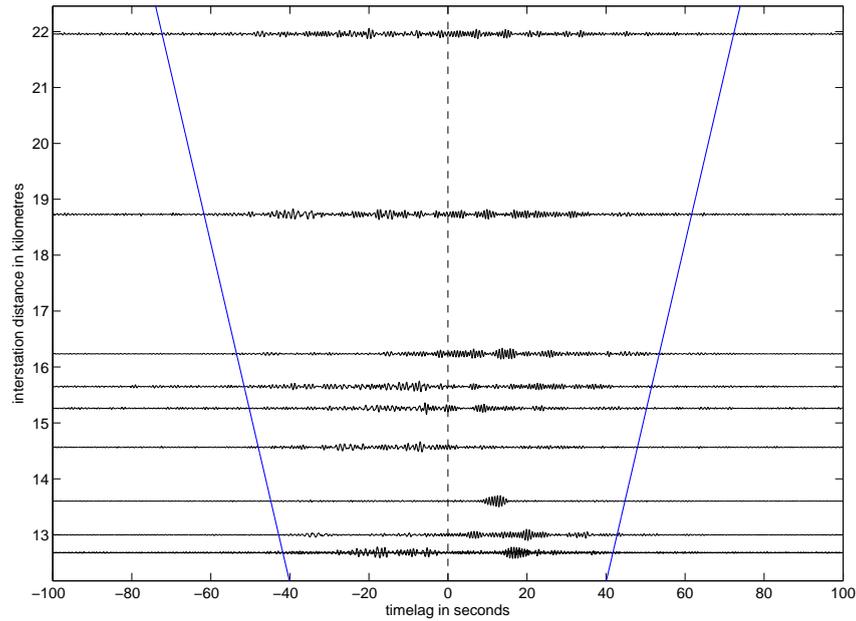
Figure 5.11.: The distance plot of the available data of the year 2012 of the TT components. The CCFs are filtered in the frequency band from 0.8 Hertz to 1.6 Hertz. Only the interstation distance interval from 12.5 kilometres to 22.5 kilometres is depicted. Signals occur in the causal part as well as in the acausal part of the cross-correlograms. However, symmetric signals are not identified.

5.15). The highest amplitude values of these 'packets' occur about every 50 seconds in the cross-correlogram, e.g. at lag times of 50 seconds, 100 seconds, 150 seconds, and so on. On the other hand, smaller periodically reoccurring 'signal packets' are also observed (bottom of Figure 5.15). These wave packets extend over a timelag period of circa 2 - 3 seconds.

With increasing interstation distance the amplitude of the beat-like 'signals' also grows. A reason for this might be a decreasing signal amplitude ($\rightarrow$ decreasing SNR, if the beat-like 'signals' are considered as noise) in the cross-correlograms of stationpairs located farther apart. Two distance plots of the TT components are exemplarily depicted in Figures 5.16, and 5.17. Due to the bias by the beat-like 'signals' the cross-correlograms of this frequency band are not further discussed.

In the cross-correlograms filtered between 3.2 Hertz and 6.4 Hertz a clear differentiation between signals and noise is not possible anymore except for cross-correlograms of stations being very close to each other. In the signal area the amplitudes are higher compared to the amplitudes of the waveforms outside of the signal area. Though, in many cases an exact arrival time of coherent signals in the cross-correlograms cannot be determined. Figure 5.18 illustrates how the coherent noise vanishes in the incoherent noise.

In the frequency band from 6.4 Hertz to 12.8 Hertz the SNR further decreases. Even at the shortest interstation distances the arrival times of signals cannot be clearly defined anymore.

All the CCFs filtered in the two highest frequency bands (12.8 Hertz to 25.6 Hertz, and 20 Hertz to 45 Hertz) are more or less biased by beat-like 'signals'. As these periodically occurring 'signals' look differently compared to those in Figure 5.17, another plot is depicted in Figure 5.19.
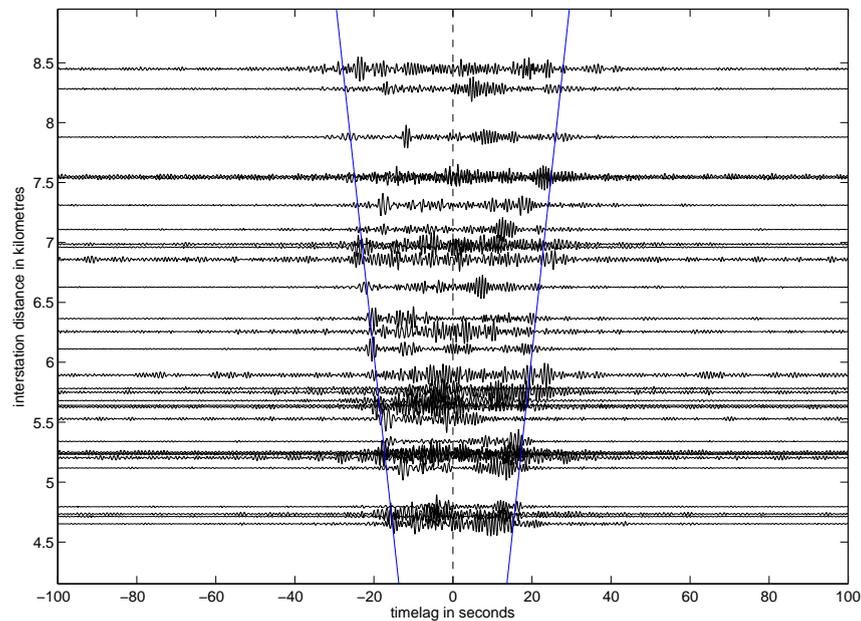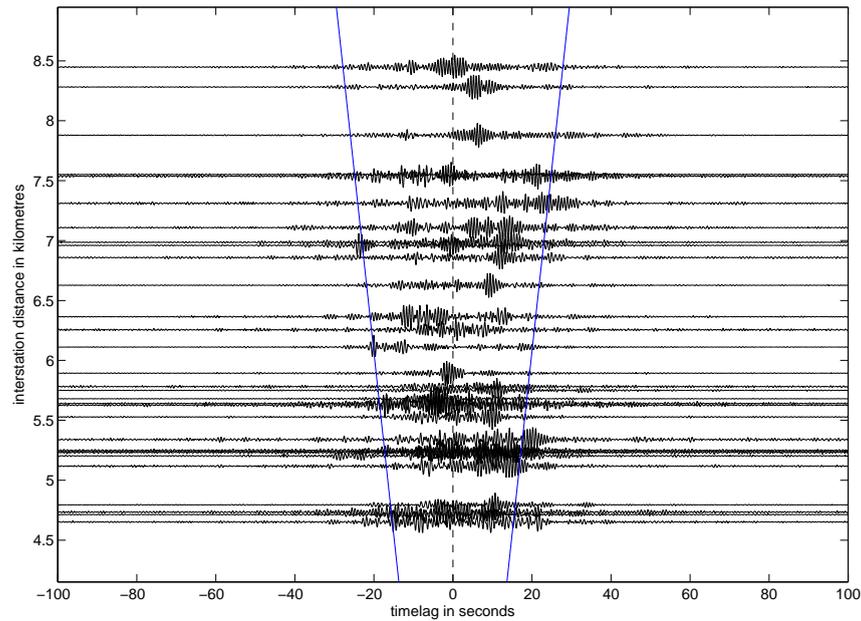
Figure 5.12.: The distance plot of the available data of the year 2012 of the ZZ components. The CCFs are filtered in the frequency band from 0.8 Hertz to 1.6 Hertz. Only the interstation distance interval from 4.5 kilometres to 8.5 kilometres is depicted. The signal area, which lies in between the two blue 300 metres per second velocity lines, looks 'noisier' than the signal area of the distance plot of the TT components depicted in Figure 5.9. In some ZZ cross-correlograms many signals interfere with each other, so that no exact arrival time of a signal can be determined.

Figure 5.13.: The distance plot of the available data of the year 2012 of the RR components. The CCFs are filtered in the frequency band from 0.8 Hertz to 1.6 Hertz. Only the interstation distance interval from 4.5 kilometres to 8.5 kilometres is depicted. The signal area between the two blue lines looks about as noisy as the signal area in the distance plot of the ZZ components (Figure 5.12).
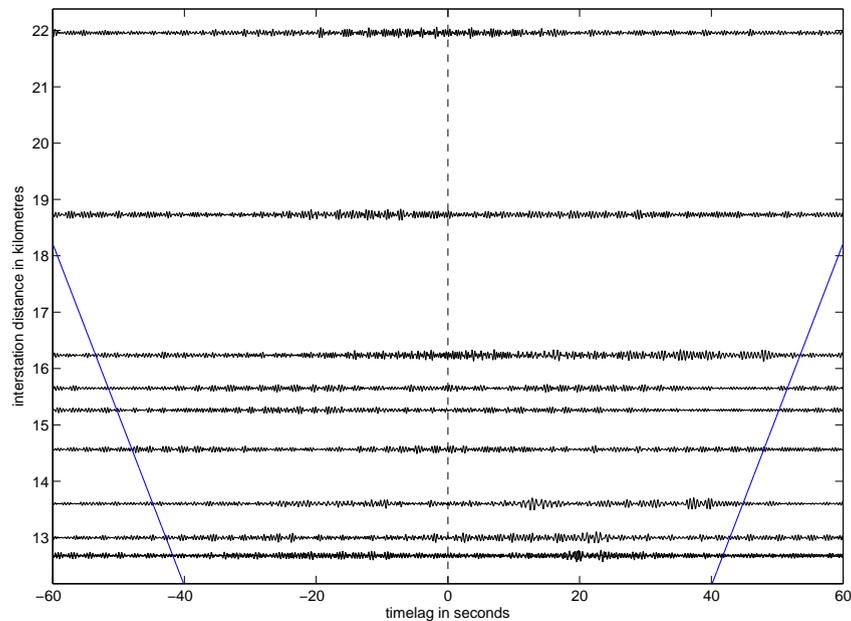

Figure 5.14.: The distance plot of the available data of the year 2012 of the ZZ components. The CCFs are filtered in the frequency band from 1.6 Hertz to 3.2 Hertz. Only the cross-correlograms with the largest interstation distances are depicted. Note the different timelag limits compared to the plots above. For a better orientation the two blue 300 metres per second velocity lines are plotted again. Signals with frequencies between 1.6 Hertz and 3.2 Hertz disappear at these large interstation distances.
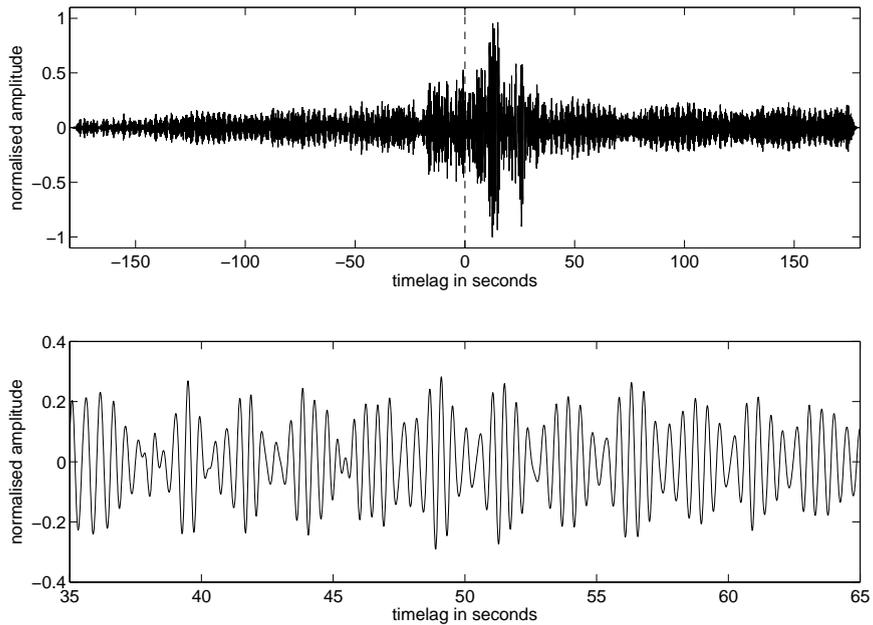
Figure 5.15.: Top: The cross-correlogram of stationpair TMO20-TMO58 of the year 2012 of the ZZ components. The CCF is filtered in the frequency band from 1.6 Hertz to 3.2 Hertz. Periodically reoccurring (about every 50 seconds) 'signal packets' are observed. Bottom: Zoom of the waveform above between 35 seconds and 65 seconds. The cross-correlogram also contains smaller 'wave packets', which reach their maximum circa every 2 - 3 seconds.
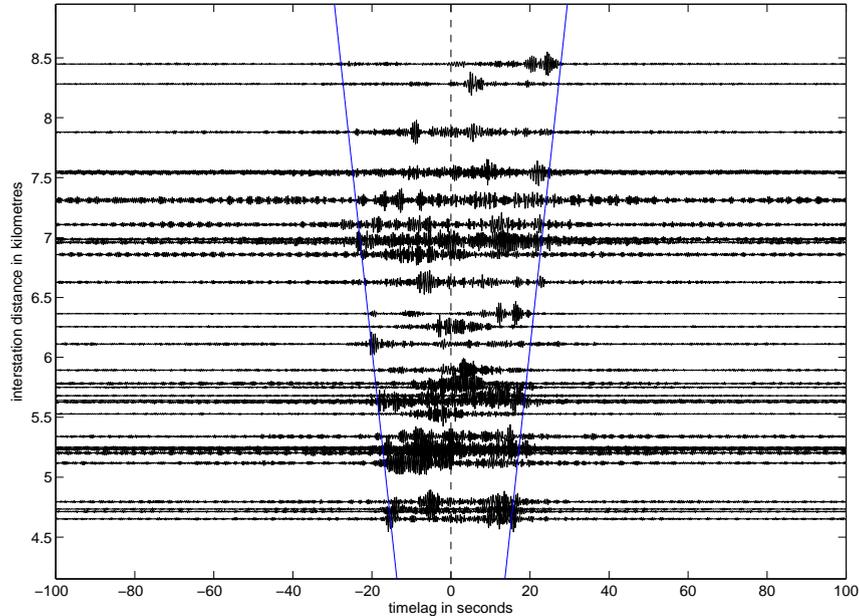

Figure 5.16.: The distance plot of the available data of the year 2012 of the TT components. The CCFs are filtered in the frequency band from 1.6 Hertz to 3.2 Hertz. Only the interstation distance interval from 4.5 kilometres to 8.5 kilometres is depicted. A symmetric signal in the cross-correlogram at the bottom emerges at a timelag of ± 16 seconds. Beat-like 'signals' are already observed, but not as clearly visible as in the cross-correlograms of stationpairs with larger interstation distances (see Figure 5.17).
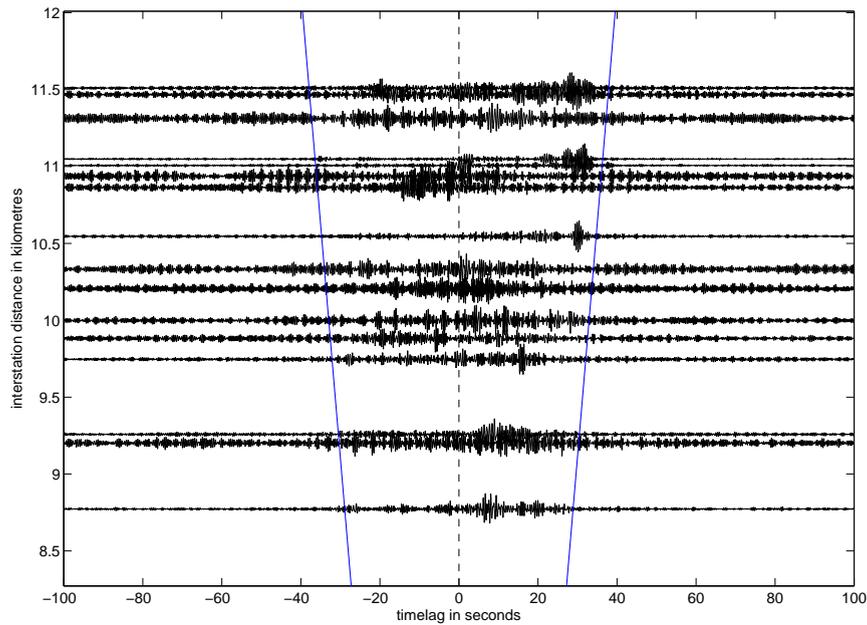
Figure 5.17.: The distance plot of the available data of the year 2012 of the TT components. The CCFs are filtered in the frequency band from 1.6 Hertz to 3.2 Hertz. Only the interstation distance interval from 8.5 kilometres to 12.5 kilometres is depicted. Here, beat-like 'signals' emerge more clearly than in the cross-correlograms of stations with smaller interstation distances.
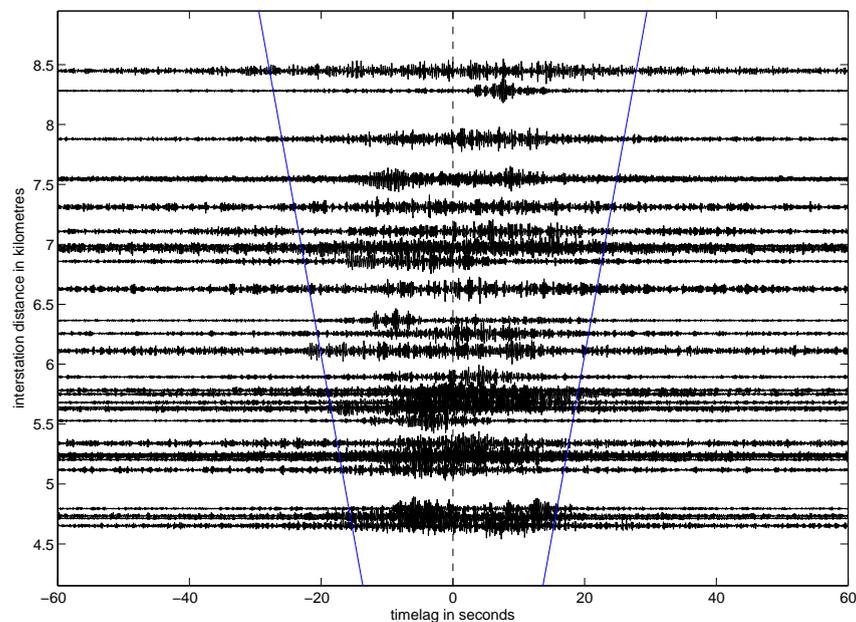

Figure 5.18.: The cross-correlograms (RR components) with interstation distances between 4.5 to 8.5 kilometres are shown. The CCFs are filtered in the frequency band from 3.2 Hertz to 6.4 Hertz. In between the two blue 300 metres per second velocity lines the signal amplitude is higher than outside of this lag interval. However, the SNR is too high for clearly defining signal phases. Here, the timelag limits are changed from ± 100 seconds to ± 60 seconds.
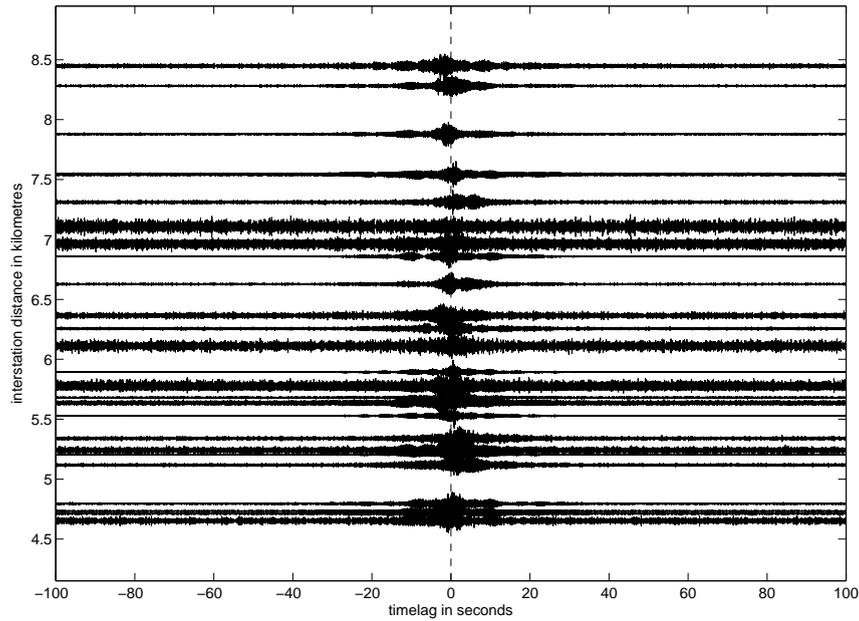
Figure 5.19.: The cross-correlograms (ZZ components) with interstation distances between 4.5 to 8.5 kilometres are shown. The CCFs are filtered between 12.8 Hertz and 25.6 Hertz. Almost all of the cross-correlograms contain beat-like 'signals'.

Summarising the observations above, the following is concluded:

1. The signal to noise ratio of the cross-correlograms decreases with increasing interstation distances and increasing frequency.

2. At very large distances (>12.5 kilometres) coherent signals disappear in the cross-correlograms filtered between 1.6 Hertz and 3.2 Hertz.

3. Between 3.2 Hertz and 6.4 Hertz clear arrival times of signals in cross-correlograms of closer (4.5 kilometres to 12.5 kilometres) stationpairs can seldomly be determined.

4. Above 6.4 Hertz the SNR of the cross-correlograms of very close (<4.5 kilometres) pairs of stations is also too low for a clear differentiation between signal and noise.

From this follows that low-frequency signals travel more coherently over longer distances than high-frequency signals.

In the lower frequency bands a) to c) (from 0.1 Hertz to 0.8 Hertz) the main portion of the signals occurs in the acausal part of the cross-correlograms. This is equivalent to a roughly west-east propagation direction of the low-frequency noise field. Above about 0.8 Hertz the averaged portion of signals in the acausal and causal parts equalises. Thus, the distribution of noise sources is not the same in different frequency bands.

In some frequency ranges the cross-correlograms are biased by beat-like amplitude modulations (see the example in Figure 5.15, where the cross-correlogram of stationpair TMO20-TMO58 is depicted). These are the frequency ranges e) from 1.6 Hertz to 3.2 Hertz, h) from 12.8 Hertz to 25.6 Hertz, and i) from 20 Hertz to 45 Hertz. It is considered that these 'beats' are formed in the cross-correlation process. Under the assumption that two sinusoidal signals of nearly the same frequency are recorded by two stations, the cross-correlogram of these two waveforms will contain beat-like 'signals'. In Appendix C synthetic data examples of the cross-correlations of two sine functions are studied. Moreover, the spectrograms of one day of the vertical ambient noise between 0.8 Hertz and 4 Hertz

recorded by the stations TMO20 and TMO58 are depicted in Appendix C. In these spectrograms periodic signals with distinct frequencies are observed over the whole day. How these periodic signals contribute to the beat-like 'signals' in the cross-correlations should be further investigated, so that one might even be able to eliminate the 'beats' from the CCFs.

In comparison to the cross-correlograms of the TT component, the cross-correlograms of the other two component combinations (ZZ, and RR) appear to be less symmetric. It is also conspicuous that the signal areas of the ZZ and the RR distance plots look 'noisier' than the signal area of the TT distance plots. In turn, this might be the reason why symmetric signals do not emerge as clearly out of the cross-correlograms of these two component pairs as they emerge out of the TT cross-correlograms.

Most of the signals of all three component cross-correlograms travel with apparent velocities higher than about 300 metres per second. Especially the ZZ and TT cross-correlograms contain many signals with very high apparent velocities. These signals might originate from very dominant noise sources located outside of the Fresnel zones (see Section 2.3) of the respective stationpair.

In summary, the cross-correlations filtered in the frequency band from 0.8 Hertz and 1.6 Hertz have the most balanced relation between signals in the causal and in the acausal parts of the cross-correlations. Moreover, they feature the best SNR, and they are also not perturbed by beat-like 'signals'. At frequencies above 0.8 Hertz the criterion $d > 3 \cdot \lambda$ is also satisfies by the main part of the used stationpairs. (For example, if a wave velocity of maximal 600 meters per second is assumed at a frequency of 0.8 Hertz, the criterion is fulfilled by all stationpairs with an interstation distance higher than 2250 metres.) In the following section, it is further analysed if the cross-correlations filtered in this preferred frequency range are usable for noise-based imaging techniques.

5.1.2. More Detailed Information on the CCFs filtered between 0.8 Hertz and 1.6 Hertz of Three Selected Stationpairs

For a further, more detailed investigation of the signals in the cross-correlograms I choose to analyse the ZZ, RR, and TT CCFs of three stationpairs: TMO53-TMO54, TMO22-TMO57, and TMO20-TMO61. Stations TMO53 and TMO54 are ~ 3.0 kilometres, stations TMO22 and TMO57 are ~ 8.4 kilometres, and stations TMO20 and TMO61 are ~ 5.3 kilometres apart from each other. The connecting line of stationpair TMO53-TMO54 lies in a roughly north-south direction, the connecting line of TMO22-TMO57 lies in a roughly northwest-southeast direction, and the connecting line of TMO20-TMO61 lies in a roughly west-east direction (see the map in Figure 5.20). Hence, with the choice of these stationpairs different interstation distances as well as different stationpair azimuths are covered. For the three pairs of stations the following questions are studied and answered:

1. Are the signals in the ZZ, RR, and TT cross-correlograms symmetric to lag time zero? Furthermore, do the CCFs converge to the Green's function?

2. Is it possible to identify what type of seismic waves the signals in the cross-correlograms consist of?

The latter question can be answered by determining the propagation velocity of the waves and by generating polarisation diagrams (see Section 2.1).

With respect to question 1 I only choose this frequency band where the best balance between signals in the causal and acausal parts of the cross-correlograms is observed. Following the results of the previous section (Section 5.1.1) I merely use the CCFs (of the year 2012) filtered between 0.8 Hertz and 1.6 Hertz.

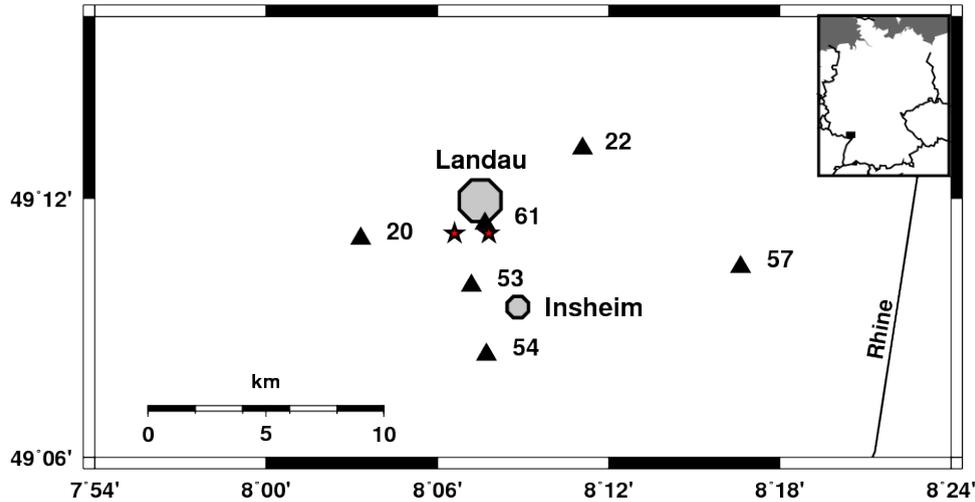To be able to detect symmetric comparatively high-amplitude signals more easily, not only

Figure 5.20.: This map shows the stations of the three cross-correlation pairs used for the further investigation of the signals in the respective cross-correlograms: TMO53-TMO54, TMO20-TMO61, and TMO22-TMO57. The smallest interstation distance has stationpair TMO53-TMO54 ($\sim$ 3.0 kilometres), the largest interstation distance has stationpair TMO22-TMO57 ($\sim$ 8.4 kilometres). The stations connecting lines vary between a roughly north-south (TMO53-TMO54), a northwest-southeast direction (TMO22-TMO57), and a roughly west-east direction (TMO20-TMO61). The red stars denote the landing points of the two boreholes of the geothermal power plant in Landau.

the cross-correlograms are shown, but also the causal part and the time-flipped acausal part of the CCF are plotted above each other. Moreover, the causal SNR (SNR_c) and the acausal SNR (SNR_{ac}) are each calculated as the ratio between the peak signal in a signal time window and the highest amplitude in a noise window. The noise window is always chosen behind the signal window ('trailing noise'; Bensen et al., 2007). A table in Appendix D lists the lag times of the determined signal and noise windows, the highest amplitude values within these windows, and the resulting SNR values for the three stationpairs.

Figures 5.21, 5.22, and 5.23 show the 2012 cross-correlograms of the ZZ, RR, and TT components filtered between 0.8 Hertz and 1.6 Hertz of stationpair TMO53-TMO54. The best signal to noise ratio (SNR) of the three cross-correlograms of stationpair TMO53-TMO54 (top of Figures 5.21, 5.22, and 5.23) is observed on the TT components with an estimated $SNR_c \approx 12$, and an estimated $SNR_{ac} \approx 61$. The SNR values of the RR components are equal to about 3 (causal) and to about 5 (acausal). On ZZ the SNR_c equals circa 2, and the SNR_{ac} is about 13. The TT waveforms of the causal and acausal parts between (-)7.5 seconds and (-)10 seconds are similar in relation to phase and frequency (see the bottom of Figure 5.23). Only the amplitudes of the signal in the acausal part are higher than the amplitudes in the causal part. Assuming a signal arrival time of about 8.5 seconds the signals propagate with an apparent velocity of about 360 metres per second between the two stations. Compared to the TT components the ZZ components are noisier. The cross-correlogram on top of Figure 5.21 does not clearly reveal symmetric signals. Though, plotting the waveforms of the acausal and causal parts above each other exposes signals between (-)8.5 seconds and (-)11.5 seconds identical in phase and frequency. Only the amplitude of the two signals differs (Figure 5.21). The RR cross-correlogram is the noisiest. Clear signal arrival times do not emerge. Furthermore, symmetric signals propagating with velocities higher than 200 metres per second are not observed (Figure 5.22).
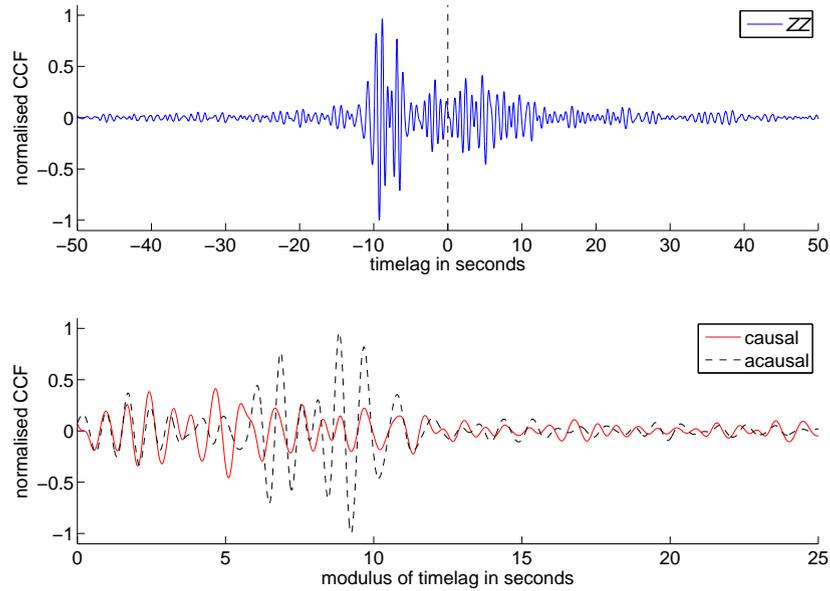
Figure 5.21.: On top: The ZZ cross-correlogram of the year 2012 of stationpair TMO53-TMO54 filtered between 0.8 Hertz and 1.6 Hertz, and normalised to its maximum. At the bottom: The causal part (red line) and the flipped acausal part (black dashed line) of the cross-correlogram shown above are plotted against the modulus of the timelag. Between the lag times of 1 second and 2.5 seconds, and between 8.5 seconds and 11.5 seconds the waveform in the causal part is very similar to the waveform in the acausal part. In the latter case only the amplitude of the causal waveform is smaller.
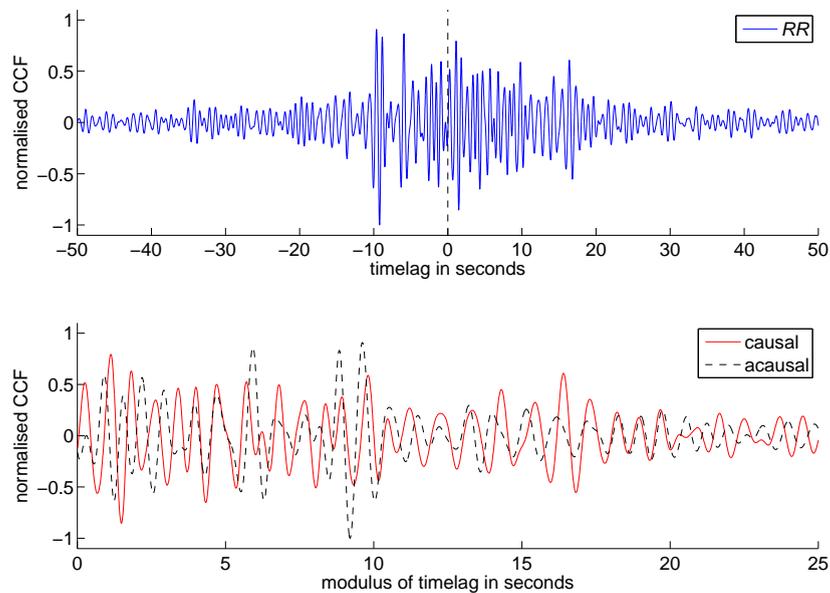

Figure 5.22.: On top: The RR cross-correlogram of the year 2012 of stationpair TMO53-TMO54 filtered between 0.8 Hertz and 1.6 Hertz, and normalised to its maximum. At the bottom: The causal part (red line) and the flipped acausal part (black dashed line) of the cross-correlogram shown above are plotted against the modulus of the timelag. The highest waveform similarity is observed around the timelags of 4 seconds and of 19 seconds.
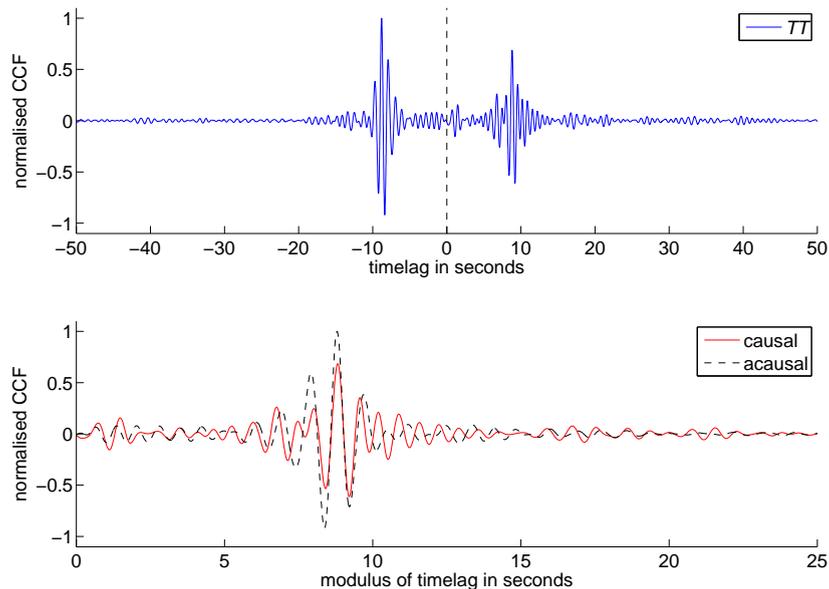
Figure 5.23.: On top: The TT cross-correlogram of the year 2012 of stationpair TMO53-TMO54 filtered between 0.8 Hertz and 1.6 Hertz, and normalised to its maximum. At the bottom: The causal part (red line) and the flipped acausal part (black dashed line) of the cross-correlogram shown above are plotted against the modulus of the timelag. Between 8 seconds and 10 seconds the waveforms in the causal and in the acausal parts are very similar to each other.

The symmetry analysis of signals in the ZZ, RR, and TT cross-correlograms of the year 2012 of stationpair TMO20-TMO61 (Figures 5.24, 5.25, and 5.26) shows results comparable to those obtained with stationpair TMO53-TMO54. There are nearly identical waveforms with respect to phase and frequency, but not to amplitude in the causal and acausal parts of the ZZ cross-correlogram at lag times between (-)14 seconds and (-)17 seconds (see bottom of Figure 5.24), and in the TT cross-correlograms at lag times between (-)13 seconds and (-)16 seconds (see bottom of Figure 5.26). With an arrival time of the signal at about 14 seconds an apparent velocity of about 380 metres per second is obtained. The RR components are again very noisy ($\text{SNR}_\text{c} \approx 4$, $\text{SNR}_\text{ac} \approx 2$), especially compared to TT ($\text{SNR}_\text{c} \approx 8$, $\text{SNR}_\text{ac} \approx 6$). However, between (-)13.5 seconds and (-)14.5 seconds the RR waveforms of the causal and acausal parts are also slightly similar to each other (Figure 5.25).

Very nearly symmetric signals are observed on all three component combinations of the cross-correlograms of stationpair TMO22-TMO57. Again, the most distinct symmetry and the best SNR is obtained on the TT components between lag times of 23.5 seconds and 32 seconds (TT: $\text{SNR}_\text{c} \approx 6$, $\text{SNR}_\text{ac} \approx 3$; RR: $\text{SNR}_\text{c} \approx 3$, $\text{SNR}_\text{ac} \approx 2$; ZZ: $\text{SNR}_\text{c} \approx 2$, $\text{SNR}_\text{ac} \approx 4$). Assuming an arrival time of the signals at 23.5 seconds the apparent velocity is equal to about 360 metres per second.

In the following the polarisations of the nearly symmetric waves identified in the cross-correlograms above are analysed. Under the assumption that the CCF converges to the Green's function, one expects, for example, for Rayleigh waves an elliptical polarisation of the wave in the RR-ZZ-plane (see Sections 2.1 and 4.4). However, if the CCF does not converge to the Green's function, other particle movements might be observed, and in turn, they might also provide usable information on the signals.

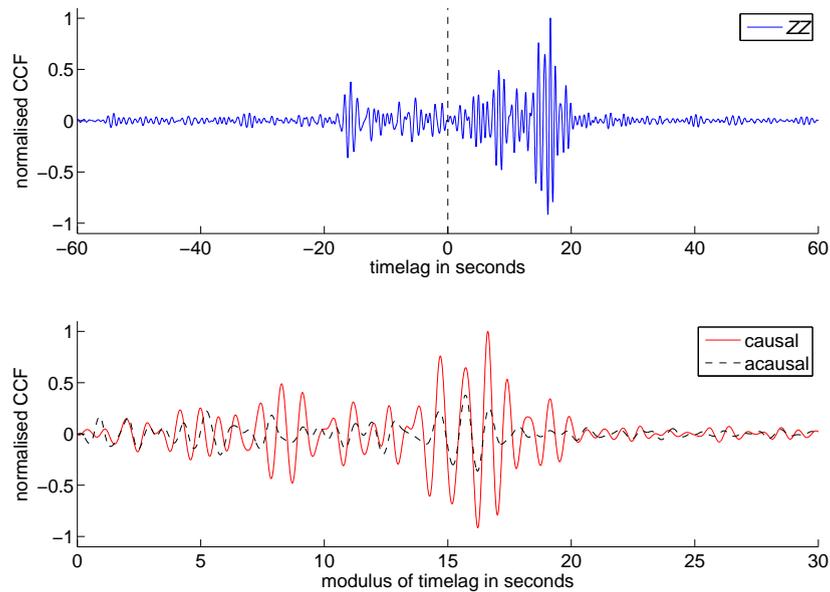The time windows which are chosen for the calculation of the polarisations of the observed

Figure 5.24.: On top: The ZZ cross-correlograms of the year 2012 of stationpair TMO20-TMO61 filtered between 0.8 Hertz and 1.6 Hertz. At the bottom: The causal part (red line) and the flipped acausal part (black dashed line) of the cross-correlogram shown above are plotted against the modulus of the timelag. The waveforms of the causal and acausal parts coincide best at lag times between 14 seconds and 17 seconds.
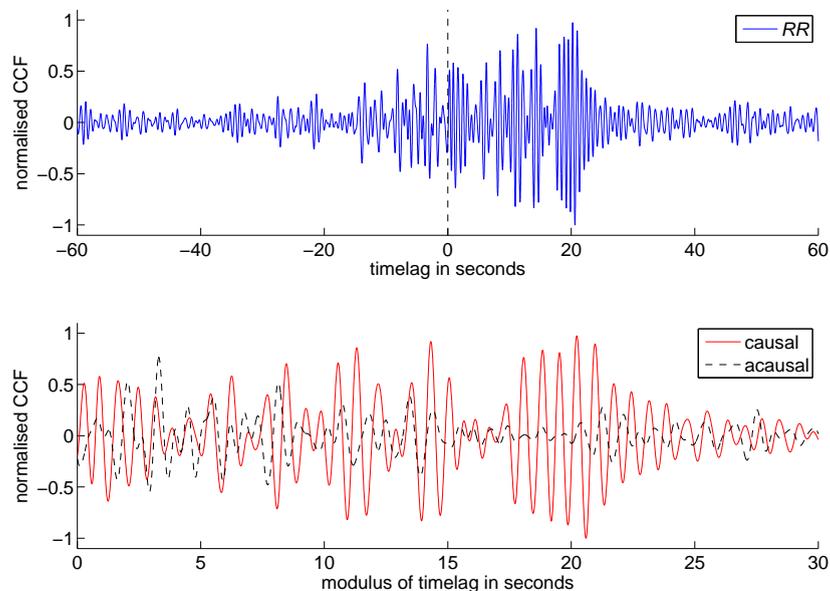

Figure 5.25.: On top: The RR cross-correlograms of the year 2012 of stationpair TMO20-TMO61 filtered between 0.8 Hertz and 1.6 Hertz. No clear arrival times of signals are identifiable. Bottom: The causal part (red line) and the flipped acausal part (black dashed line) of the cross-correlogram shown above are plotted against the modulus of the timelag. Around 14 seconds the causal and acausal waveforms look nearly the same taking the amplitudes of the two signals not into account.
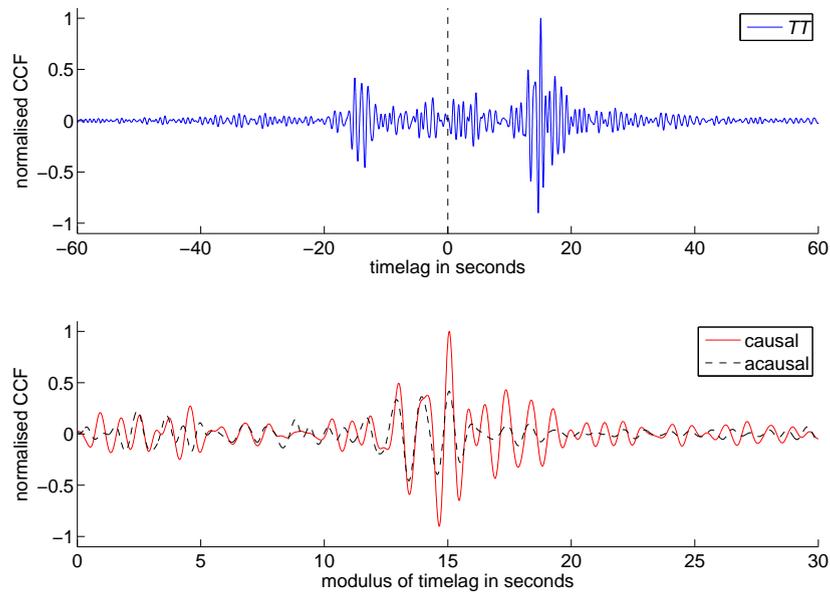
Figure 5.26.: On top: The *TT* cross-correlograms of the year 2012 of stationpair TMO20-TMO61 filtered between 0.8 Hertz and 1.6 Hertz. Bottom: The causal part (red line) and the flipped acausal part (black dashed line) of the cross-correlogram shown above are plotted against the modulus of the timelag. A very distinct agreement between the waveforms of the causal and acausal parts is obtained between lag times of 13 seconds and 16 seconds.
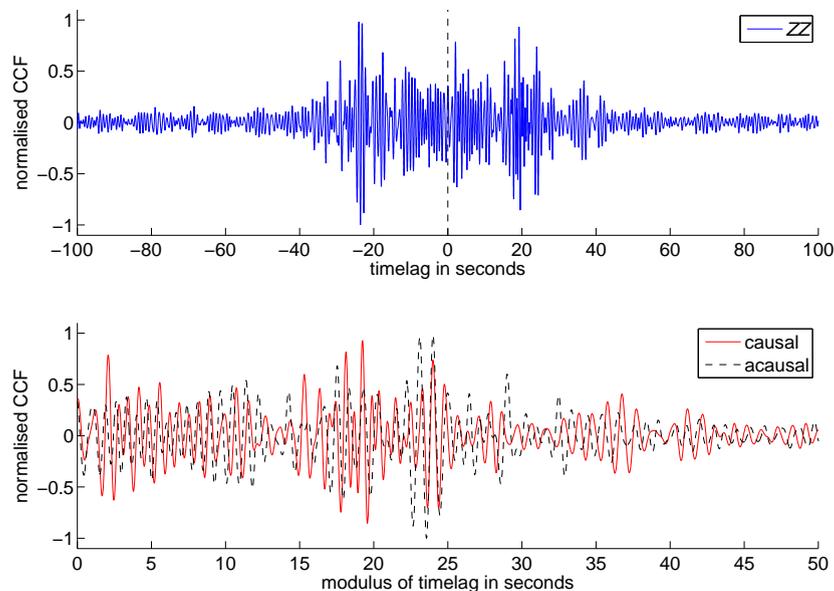

Figure 5.27.: On top: The *ZZ* cross-correlograms of the year 2012 of stationpair TMO22-TMO57 filtered between 0.8 Hertz and 1.6 Hertz. Bottom: The causal part (red line) and the flipped acausal part (black dashed line) of the cross-correlogram shown above are plotted against the modulus of the timelag. The highest coincidence between the waveforms of the causal and acausal parts is obtained at timelags around 24 seconds.
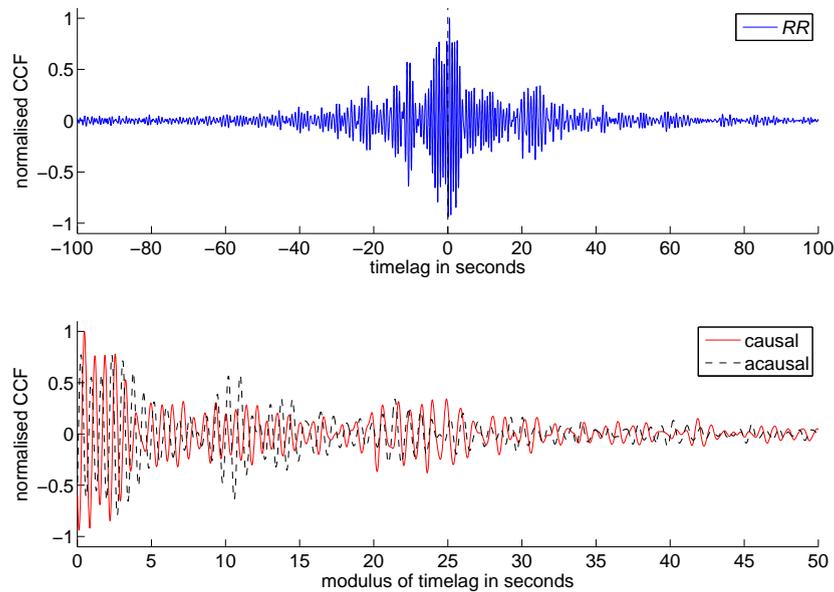
Figure 5.28.: On top: The *RR* cross-correlograms of the year 2012 of stationpair TMO22-TMO57 filtered between 0.8 Hertz and 1.6 Hertz. Bottom: The causal part (red line) and the flipped acausal part (black dashed line) of the cross-correlogram shown above are plotted against the modulus of the timelag. Around a lag time of about 23 seconds the waveforms of the causal and acausal parts agree best with each other.

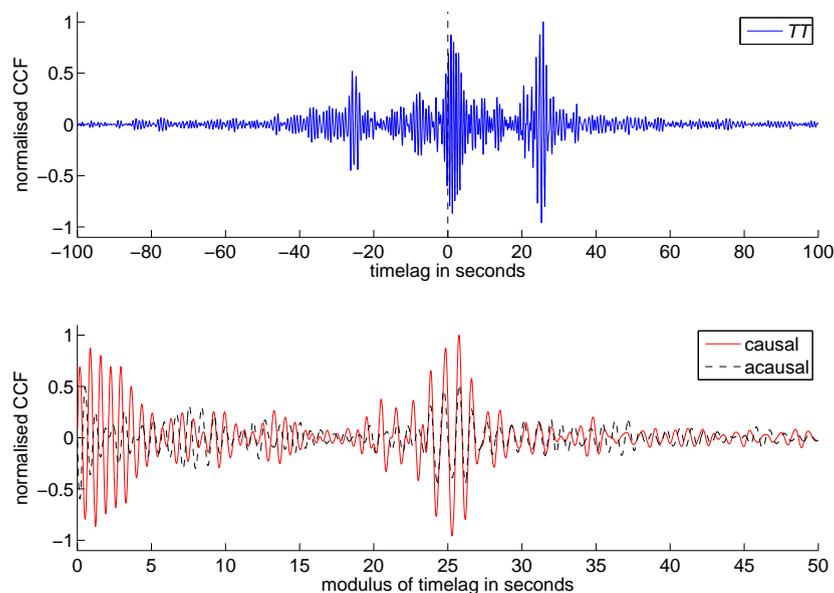Figure 5.29.: On top: The *TT* cross-correlograms of the year 2012 of stationpair TMO22-TMO57 filtered between 0.8 Hertz and 1.6 Hertz. Bottom: The causal part (red line) and the flipped acausal part (black dashed line) of the cross-correlogram shown above are plotted against the modulus of the timelag. Between 23.5 seconds and 32 seconds the waveforms of the causal and acausal parts are very similar to each other.

Table 5.2.: The analysed stationpairs listed with the respective interstation distance, with the time windows chosen for the calculation of the polarisation diagrams, and with the maximum and minimum apparent velocities, which result from dividing the interstation distance by the time window limits. Due to the analysis of symmetric signals only the modulus of the time window limits is given.

Stationpair	interstation distance in kilometres	time window limits in seconds	apparent velocity in metres per second
TMO53-TMO54	3.04	7.5 - 10	405 - 304
TMO20-TMO61	5.34	13 - 17	411 - 314
TMO22-TMO57	8.45	22 - 26	384 - 325

waves are listed in Table 5.2 for the three investigated stationpairs. Within these time windows not only the amplitudes of the ZZ and RR cross-correlograms but also the amplitudes of the ZZ and TT cross-correlograms are plotted against each other. For these plots the MATLAB script XCORR_plot_polarisation.m was developed.

The three Figures 5.30, 5.31, and 5.32 – one figure per stationpair – consist of six plots each. The left column depicts the acausal part, the right column the causal part. In the first row of each figure, the waveforms of the ZZ (black dashed line), RR (red line), and TT (blue line) components within the defined time window are shown. In the mid-row the polarisation diagrams of the waves on the RR and ZZ components are plotted. In the third row, one can see the polarisation diagrams of the waves on the TT and ZZ components.

The polarisation diagrams for stationpair TMO53-TMO54 are shown in Figure 5.30. In the chosen time window an elliptic movement of the wave is observed in the RR-ZZ-plane. This is evidence for the wave being a Rayleigh wave. In the TT-ZZ-plane the energy of the wave is higher on the TT components than on the ZZ components. Actually, one can also observe a slightly elliptical polarisation of the wave in the TT-ZZ-plane in the acausal part of the cross-correlogram.

Figure 5.31 shows the polarisation diagrams for stationpair TMO20-TMO61. In the analysed time window of the cross-correlogram of this stationpair no standard type of wave can be clearly identified. In the RR-ZZ-plane the main energy of the wave propagates along the ZZ components.

The polarisation diagrams for stationpair TMO22-TMO57 are depicted in Figure 5.32. A slightly elliptical polarisation of the wave can be identified in the RR-ZZ-plane, and also for the acausal part in the TT-ZZ-plane. However, this ellipticity is not as obvious as the ellipticity of the waves propagating between stations TMO53 and TMO54 (see mid-row of Figure 5.30). The wave observed at timelags between 22 seconds and 26 seconds has higher amplitudes on the TT components than on the ZZ components (at the bottom left).

With the above obtained results the questions asked at the beginning of this section are answered in the following way:

A symmetry between the waveforms of the causal and the acausal parts of the cross-correlograms of all three component combinations (ZZ, RR, and TT) is observed. Within a certain time window (see Table 5.2) the waveforms of the causal and acausal parts are very nearly identical with respect to phase and frequency. Only the amplitudes of the signals noticeably deviate from each other. Due to the higher signal to noise ratio the symmetric signals emerge more clearly out of the TT cross-correlograms than out of the ZZ or RR cross-correlograms. The values of the estimated SNRs of the TT components vary between ~ 3 and ~ 61, the SNRs of ZZ range within ~ 1 and ~ 13, and the SNR values of the RR cross-correlograms vary from ~ 2 to ~ 5. Therefore, using the symmetry
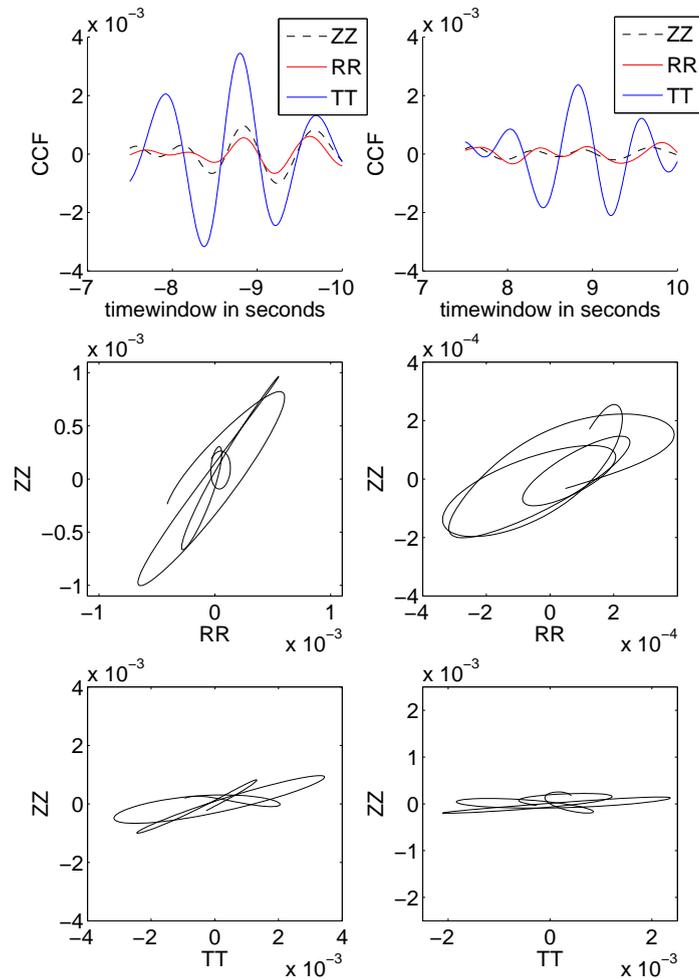
Figure 5.30.: Polarisations in the RR-ZZ-plane and in the TT-ZZ-plane of the waveforms between the negative lag times of -7.5 seconds and -10 seconds (lefthand side) and the waveforms between the positive lag times of 7.5 seconds and 10 seconds (righthand side) of stationpair TMO53-TMO54. The three cross-correlograms of the year 2012 are filtered between 0.8 Hertz and 1.6 Hertz. They are not normalised to their respective maximum. The maximum amplitude values of the whole cross-correlograms are of the order between 10^{-3} and 10^{-4}. Hence, the amplitudes in the two chosen time windows above are very small. In the RR-ZZ-plane the waves in the acausal part as well as in the causal part move elliptically (mid-row). In the TT-ZZ-plane most of the wave energy concentrates on the TT components (last row). The waveform in the acausal part also reveals a slightly elliptical polarisation in the TT-ZZ-plane (at the bottom left).
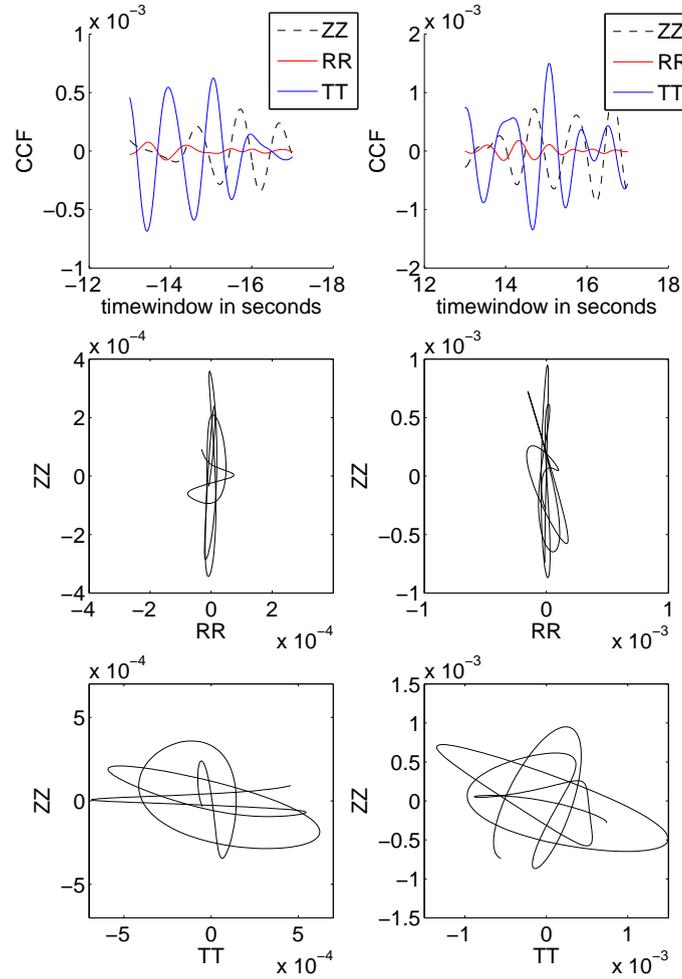
Figure 5.31.: Polarisations in the RR-ZZ-plane and in the TT-ZZ-plane of the waveforms between the negative lag times of -13 seconds and -17 seconds (lefthand side) and the waveforms between the positive lag times of 13 seconds and 17 seconds (righthand side) of stationpair TMO20-TMO61. The three cross-correlograms of the year 2012 are filtered between 0.8 Hertz and 1.6 Hertz. They are not normalised to their respective maximum. The maximum amplitude values of the whole cross-correlograms are of the order between 10^{-3} and 10^{-4}. Hence, the amplitudes in the two chosen time windows above are very small. In the RR-ZZ-plane most of the energy is on ZZ (mid-row). This might be an indicator for a P-wave. However, it is not further investigated in the scope of this thesis. In the TT-ZZ-plane the particle movements of the wave in the causal part as well as in the acausal part show no distinct behaviour.
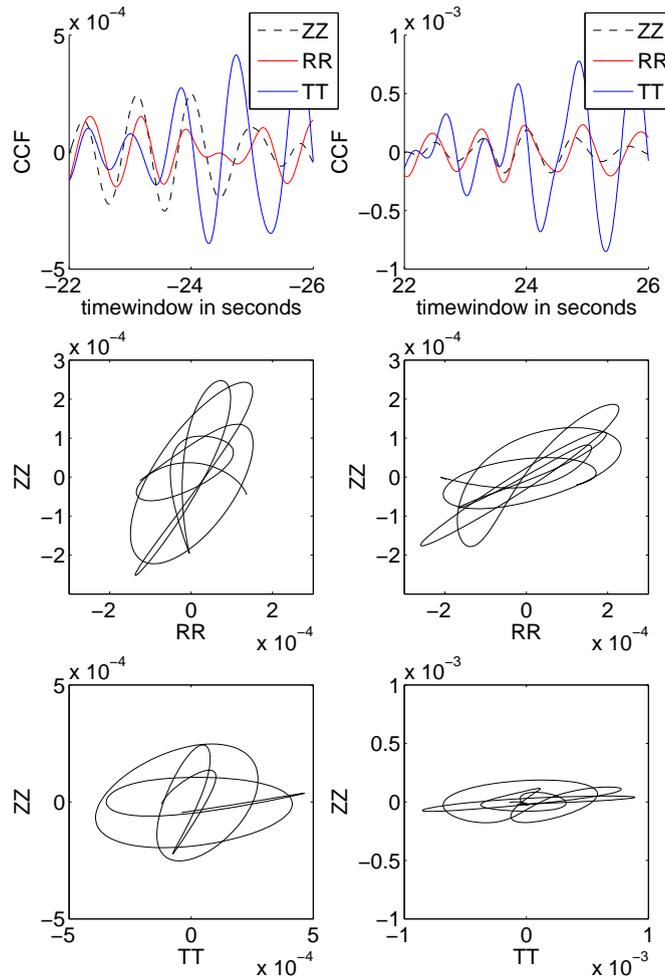
Figure 5.32.: Polarisations in the RR-ZZ-plane and in the TT-ZZ-plane of the waveforms between the negative lag times of -22 seconds and -26 seconds (lefthand side) and the waveforms between the positive lag times of 22 seconds and 26 seconds (righthand side) of stationpair TMO22-TMO57. The three cross-correlograms of the year 2012 are filtered between 0.8 Hertz and 1.6 Hertz. They are not normalised to their respective maximum. The maximum amplitude values of the whole cross-correlograms are of the order of 10^{-4}. With exception of the polarisation diagram at the bottom right all particle movements have a more or less elliptical shape (mid-row and at the bottom left). In the TT-ZZ-plane the energy of the wave in the causal part is much higher in TT direction than in ZZ direction (at the bottom right).

of a CCF and a high signal to noise ratio as indications for a CCF being an estimate of the Green's function (see numerical two-dimensional example in Section 2.3), the CCFs of the TT components rather converge to Green's functions than the CCFs of the ZZ or RR components. Under the assumption that the TT cross-correlograms contain Love waves (this assumption is discussed later in this section), especially a Love wave ambient noise tomography might yield reliable results.

If the signals in the ZZ and RR cross-correlograms mainly consist of Rayleigh waves, is not very clear. The polarisation diagrams in the RR-ZZ-plane (mid-rows of Figures 5.30, 5.31, and 5.32) do not always show a distinct ellipticity as it is expected for Rayleigh waves. There are many reasons why no clear elliptical shape is observed in the RR-ZZ-plane:

- The signal does not consist of a Rayleigh wave.

- The Rayleigh wave does not propagate along the connecting line of a stationpair or rather along the radial axis of the two stations. Thus, the particle movement will not only be observed in the RR-ZZ-plane, but also in the TT-ZZ-plane. This might also explain, why a slight ellipticity is sometimes observed in the TT-ZZ-planes (third rows in Figures 5.30, 5.31, and 5.32).

- A higher mode present in the Rayleigh wave interferes with the fundamental mode (Jay et al., 2012).

- Due to the interference of many waves, which might originate from various sources and which propagate between the stations with similar apparent velocities, the signal caused by only one wave is perturbed.

The determined velocities between about 300 metres per second and 410 metres per second of the more or less symmetric waves in the CCFs (Table 5.2) most likely correspond to surface wave velocities. (Here, it is assumed that the apparent velocities are the actual velocities of the waves.) At a geologically similar setting in the Lower Rhine Embayment near to the city of Cologne, Köhler et al. (2007) estimated in the frequency band from 0.8 Hertz to 1.6 Hertz phase velocities between 1.6 kilometres per second and 500 metres per second for Rayleigh waves, and phase velocities between 700 metres per second and 400 metres per second for Love waves. These values are a little bit higher than our obtained results. Unfortunately, for the area around Landau there is only a one-dimensional model for the P- and S-wave velocities (v_P and v_S). In the uppermost layer the velocity of the S-wave is equal to 490 metres per second. Under the assumption that the Rayleigh wave velocity can be expressed in terms of the S-wave velocity by $\sim 0.9 \cdot v_S$ (Lay and Wallace, 1995), this yields an approximate Rayleigh wave velocity of 440 metres per second for the area around Landau. Though, due to all the approximations this value should be handled with care! Nonetheless, this velocity estimate is compared to the determined velocities of the signals in the cross-correlogram of stationpair TMO53-TMO54. In this cross-correlogram the most distinct ellipticity is observed in the RR-ZZ-plane, and hence, the signal corresponds most likely to a Rayleigh wave. The maximum velocity equals 405 metres per second (see Table 5.2), and therefore, this value is smaller than the estimated 440 metres per second.

On the TT components one expects to observe Love waves. The obtained velocities (Table 5.2) do not object this assumption. As the signal to noise ratio is higher in the TT cross-correlations than in the ZZ and RR cross-correlations, it is concluded that the Love waves produce clearer signals in the cross-correlations than the Rayleigh waves. On the one hand, this might be due to a better noise source distribution for Love waves. On the other hand, Rayleigh waves originate from the couplings between P- and vertical polarised S-waves, while Love waves originate from horizontal polarised S-waves only (Müller, 1973). These different couplings of waves might cause the differences between the observations on the RR/ZZ and TT cross-correlations. Another indication of the preponderance of the

assumed Love waves are the polarisations of the waves in the TT-ZZ-plane. The diagrams show that the energy of the symmetric waves is higher on the TT components than on the ZZ components[1] (third row of Figures 5.30, 5.31, and 5.32).

A dominance of Love waves in the ambient seismic noise field is also observed by others, like, for example, Köhler et al. (2007), who analysed the noise field near to the city of Cologne using the three-component modified spatial autocorrelation method (MSPAC). Jay et al. (2012) state that, opposed to the Love waves, the peaks of the Rayleigh waves do not display a clear linear moveout. One of the explanations they give, is that the ambient noise wave field at Uturuncu Volcano might be dominated by Love waves. Moreover, they calculate only a Love and not a Rayleigh wave ambient noise tomography.

Furthermore, the observations of Jay et al. (2012) and Behm and Snieder (2013) concerning the moveout of the signals in the ZZ, RR, and TT cross-correlograms are in agreement with our observations. In both studies the linear moveout of the signals in the TT cross-correlograms is clearer than the moveout on the ZZ and RR components (see Section 5.1.1).

The main origin of the Love waves, which are observed by Behm and Snieder (2013), can be traced back to the traffic activity in the surroundings of the station deployment. Already in the 1950's, Aki (1957), who can be denoted as the father of the spatial autocorrelation method (SPAC), studied the traffic induced seismic noise at Hongu, Tokyo. He observed a strong polarisation of the horizontal components in a direction perpendicular to the direction of wave propagation, which shows that the waves are of Love type. Thus, the traffic in and around Landau might also be a source for the Love waves observed in this study.

By the analysis of three (out of 65) stationpairs with different azimuths and interstation distances I gave an approximately representative insight into the properties of the signals occurring in the cross-correlograms filtered between 0.8 Hertz and 1.6 Hertz. However, especially the symmetry of the cross-correlograms in the analysed frequency band might not apply to all the data.

5.1.3. Temporal Stability of the Signals in the Cross-Correlations

In this section the temporal stability of the signals in the cross-correlations is analysed. This is insofar interesting as continuous signals are required for noise-based – also called passive – monitoring (Hadziioannou et al. (2009); Section 2.3). Relatively stable interferograms already emerge by stacking the CCFs of only one day duration. Hence, for the detection of continuous signals the following type of plot is created: At first, the amplitudes of the used one-day long cross-correlograms are denoted in colours of the rainbow. This means, for each day one obtains a topview of a one-day long cross-correlogram with the positive amplitudes coloured in red and the negative amplitudes coloured in blue. In a next step, a rectangle is line by line filled with all the chronologically sorted one-day cross-correlograms. Due to the colouring, this type of plot is called colourplot in the following. The MATLAB script, which was developed for the creation of these plots, is called colorplot.m (Appendix B.2).

Due to very similar results between the different stationpairs I focus here on the stationpair TMO53-TMO54 (see map and details given in Section 5.1.2). The CCFs of this stationpair are filtered in three different frequency bands: from 0.1 Hertz to 0.4 Hertz, from 0.4 Hertz to 0.8 Hertz, and from 0.8 Hertz to 1.6 Hertz. Higher frequency ranges are not analysed due to the observed beat-like 'signals', and due to the decreasing signal to noise ratio (see

---

[1]However, one must keep in mind that the CCFs used in this study are normalised in the frequency domain (see Section 4.5). This processing method modifies the CCFs non-linearly, so that the amplitude information must be handled with care. Another important fact concerning the amplitudes of the cross-correlograms of the vertical and the horizontal components was already mentioned at the end of Section 4.5: Due to the different time window lengths, used for the calculation of the ZZ CCFs and the horizontal CCFs, differences in amplitude, caused by the different data processings, cannot be excluded.
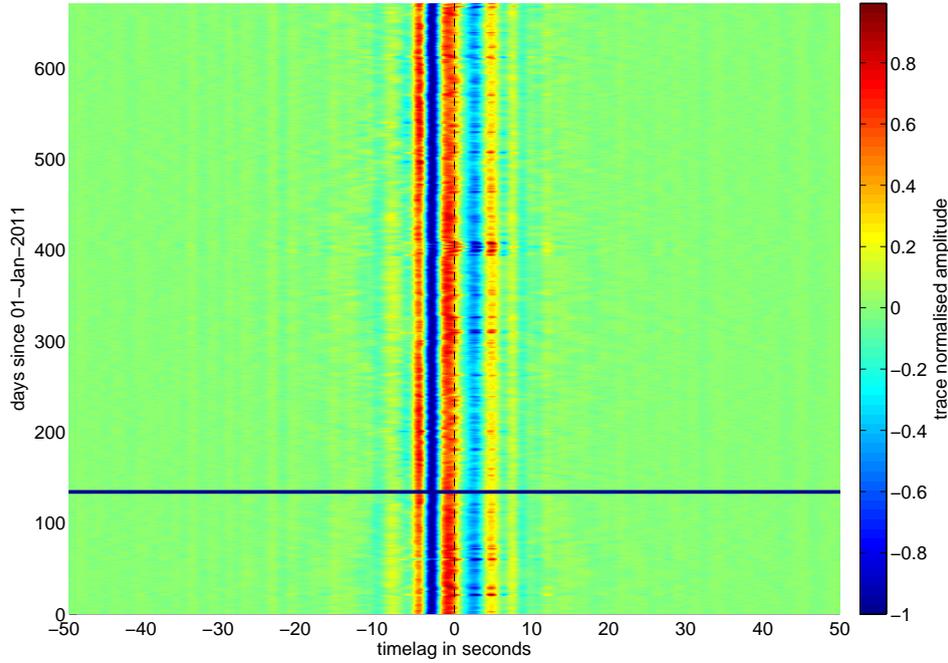
Figure 5.33.: Colourplot with the available one-day long ZZ cross-correlograms of the years 2011 and 2012 of stationpair TMO53-TMO54. The CCFs are filtered between 0.1 Hertz and 0.4 Hertz. The dark blue lines mark those days where no data is available. Temporally very stable signals occur between lag times of -10 seconds and 10 seconds. The amplitudes of the signals are higher in the acausal than in the causal part.

Section 5.1.1). All available ZZ, RR, and TT CCFs of the years 2011 and 2012 are used.

5.1.3.1. 0.1 Hertz to 0.4 Hertz

In Figure 5.33 the colourplot with the available one-day long ZZ cross-correlograms of the years 2011 and 2012 of stationpair TMO53-TMO54 are depicted. The CCFs are filtered between 0.1 Hertz and 0.4 Hertz. Note, that each trace is normalised to its own maximum, and that the dark blue lines mark those days where no data is available. Figures 5.34, 5.35 show the same but with the cross-correlograms of the RR and TT components. Temporally very stable signals are observed in all three colourplots between lag times of -10 seconds and +10 seconds. While the colourplots of the ZZ and RR components look very similar to each other, the TT cross-correlograms exhibit strong signals also in the causal part. Especially the codas of the cross-correlograms of the horizontal components also contain faintly visible signals which propagate with almost the same apparent velocity over the whole time period (e.g. in Figure 5.35 at a timelag of about -30 seconds).

5.1.3.2. 0.4 Hertz to 0.8 Hertz

Figures 5.36, 5.37, and 5.38 show the colourplots with the available one-day long ZZ, RR, and TT cross-correlograms filtered between 0.4 Hertz and 0.8 Hertz. Signals stable over the whole time period are observed on all three components. In the colourplot of the ZZ components the strongest signals occur between 0 seconds and about -7 seconds (Figure 5.36). At the same lag times continuous signals are also observed on the RR components with comparatively high amplitudes (Fig. 5.37) and on the TT components with relatively
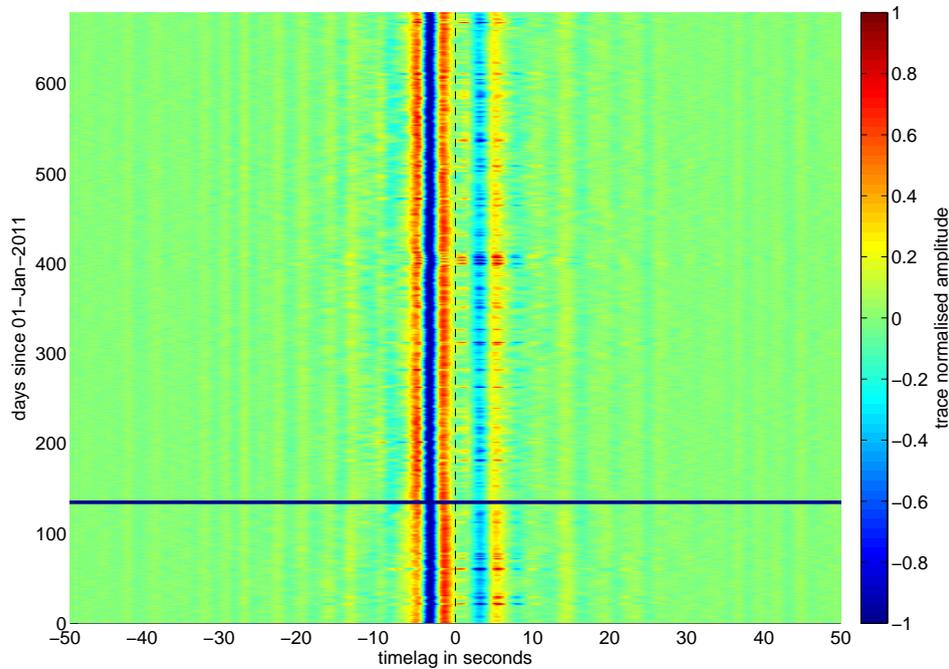
Figure 5.34.: Colourplot with the available one-day long RR cross-correlograms of the years 2011 and 2012 of stationpair TMO53-TMO54. The CCFs are filtered between 0.1 Hertz and 0.4 Hertz. The observed continuous signals are very similar to those in the ZZ cross-correlograms (Figure 5.33).
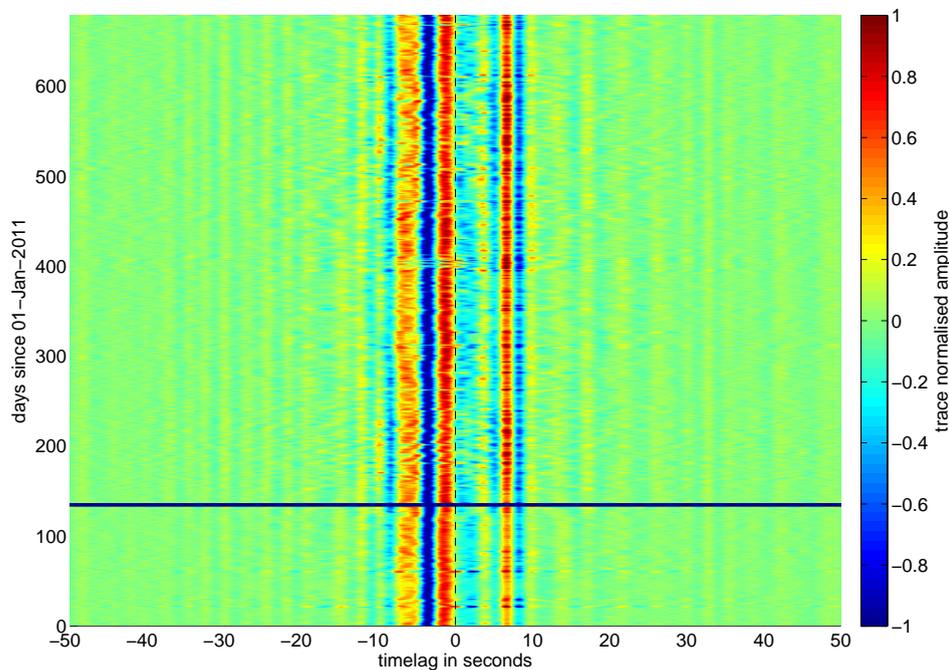

Figure 5.35.: Colourplot with the available one-day long TT cross-correlograms of the years 2011 and 2012 of stationpair TMO53-TMO54. The CCFs are filtered between 0.1 Hertz and 0.4 Hertz. The amplitudes of the signals in the causal and acausal parts are almost equal.
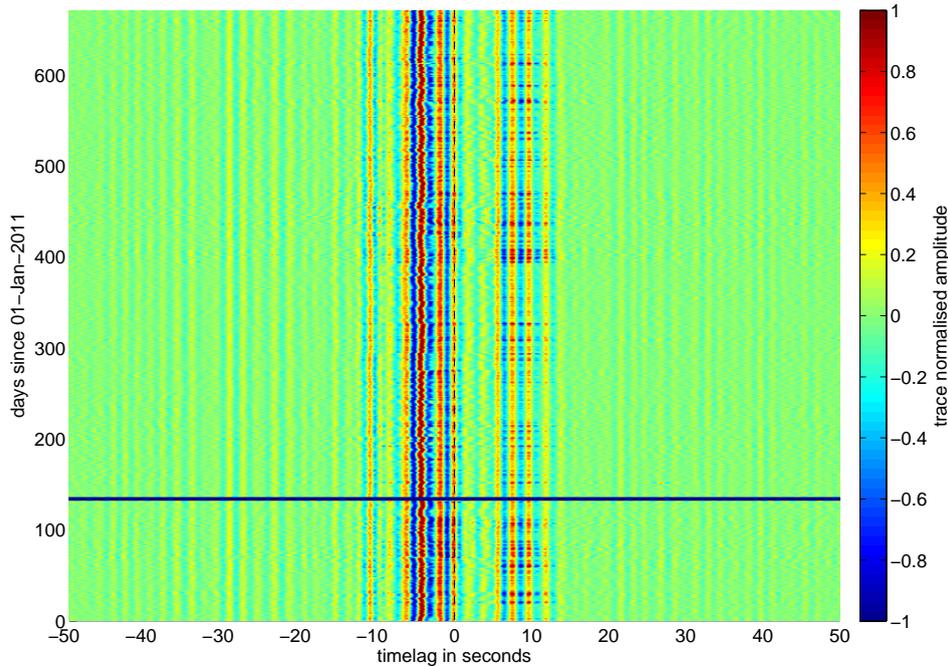
Figure 5.36.: Colourplot with the available one-day long ZZ cross-correlograms of the years 2011 and 2012 of stationpair TMO53-TMO54. The CCFs are filtered between 0.4 Hertz and 0.8 Hertz. Temporally stable signals occur between ± 12 seconds and around -29 seconds in the codas of the cross-correlograms.

low amplitudes 5.38. While the causal parts of the ZZ and the TT cross-correlograms contain strong signals, there are only high-amplitude signals in the acausal part of the RR cross-correlograms. Furthermore, the cross-correlograms of all three components contain stable low-amplitude signals in the coda.

5.1.3.3. 0.8 Hertz to 1.6 Hertz

Finally, the colourplots of the cross-correlograms of the ZZ, RR, and TT components filtered between 0.8 Hertz and 1.6 Hertz are depicted in Figures 5.39, 5.40, and 5.41. Again, temporally stable signals are observed on all three components over the whole time period. At larger timelags ($>|\pm 15|$ seconds) relatively stable signals are also detectable in the codas. As already mentioned in Section 5.1.2, the RR components are very noisy in this frequency range. The SNR of the TT cross-correlograms, on the other hand, is very high, and signals, symmetric to lag time zero, clearly emerge. In the colourplot of the TT cross-correlograms a weekly cycle of the strengths of the signal amplitudes in the acausal part is observed (Figure 5.42).

For example, the signals in the one-day long cross-correlograms of the days 443, 449, 450, 456, and 457 (since 1 January 2011) have a smaller amplitude than the cross-correlograms in between these days. Converting these 'day numbers' into dates yields: 443 $\widehat{=}$ 18 March 2012 (Sunday), 449 $\widehat{=}$ 24 March 2012 (Saturday), 450 $\widehat{=}$ 25 March 2012 (Sunday), and so on. Figure 5.43 also shows the colourplot of the TT cross-correlograms filtered between 0.8 Hertz and 1.6 Hertz but now with another normalisation of the cross-correlograms: Instead of normalising each trace to its own maximum, the absolute maximum of all the cross-correlograms of 2011 and 2012 is determined and each trace is divided by this absolute maximum. This way, one obtains a temporal resolution of the variation of the signals' amplitudes over the whole time period. Compared to the colourplot with the trace
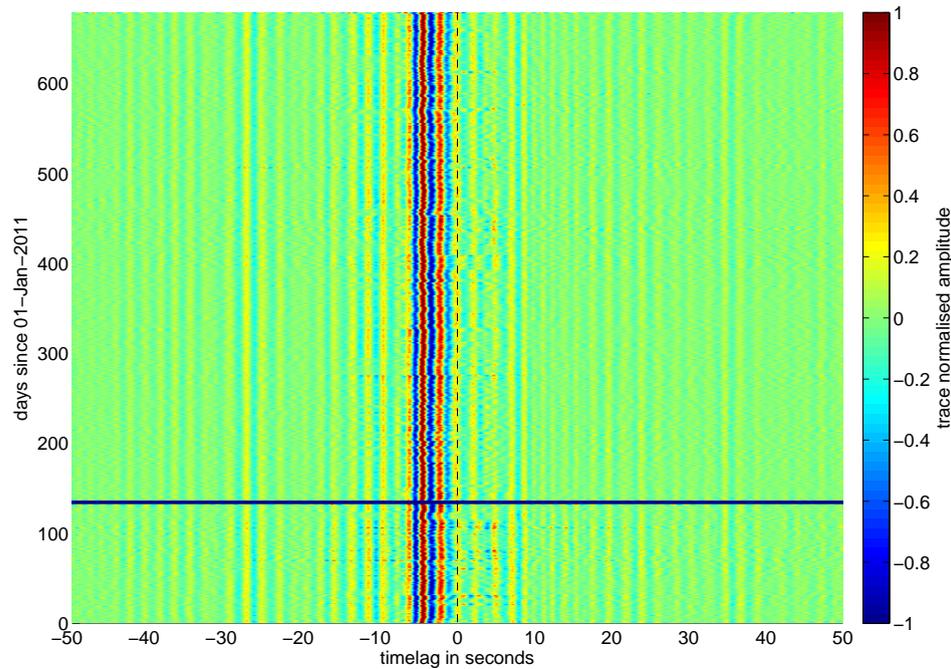
Figure 5.37.: Colourplot with the available one-day long RR cross-correlograms of the years 2011 and 2012 of stationpair TMO53-TMO54. The CCFs are filtered between 0.4 Hertz and 0.8 Hertz. A very strong and stable signal is observed in the acausal part between 0 seconds and -7 seconds. Moreover, there are continuous signals around -37 seconds and around 37 seconds.
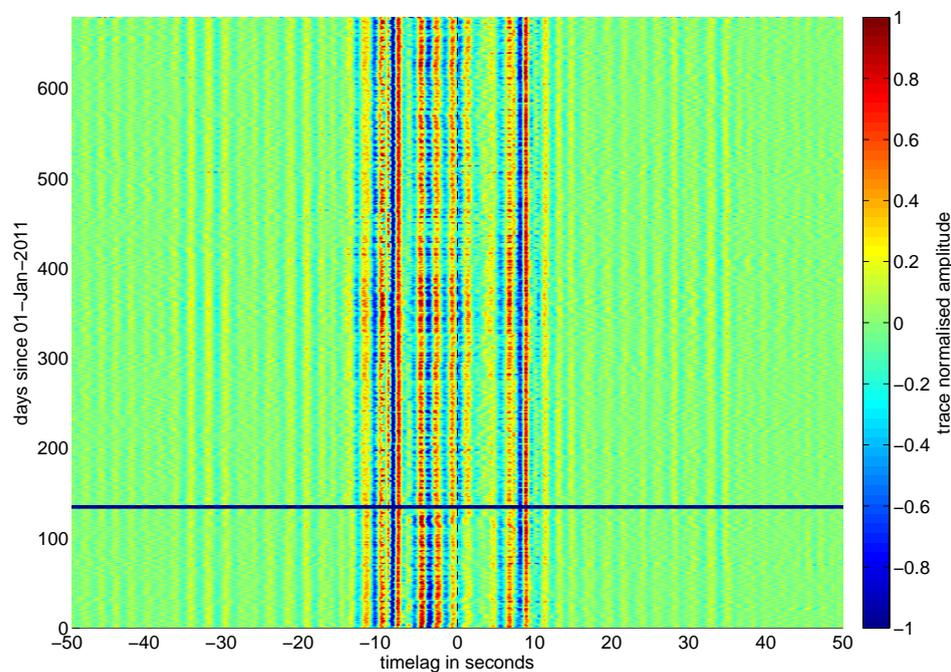

Figure 5.38.: Colourplot with the available one-day long TT cross-correlograms of the years 2011 and 2012 of stationpair TMO53-TMO54. The CCFs are filtered between 0.4 Hertz and 0.8 Hertz. Strong and continuous signals emerge around the lag times of -10 seconds and 10 seconds. Like in Figures 5.36 and 5.37 the codas contain stable signals.
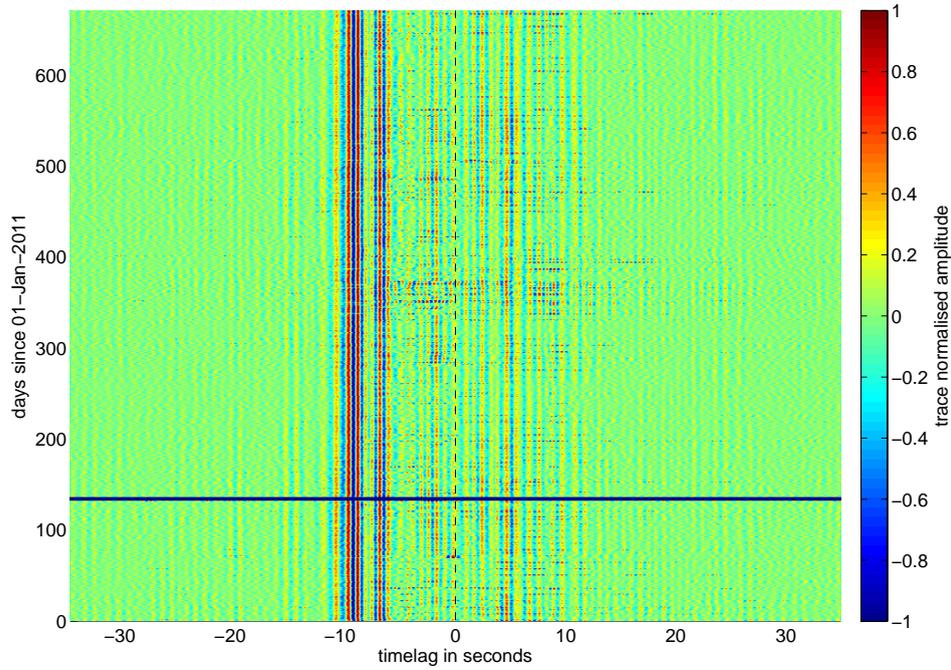
Figure 5.39.: Colourplot with the available one-day long ZZ cross-correlograms of the years 2011 and 2012 of stationpair TMO53-TMO54. The CCFs are filtered between 0.8 Hertz and 1.6 Hertz. Compared to the lower frequency ranges the noise level is higher. But still one can observe continuous signals even in the coda of the cross-correlograms.

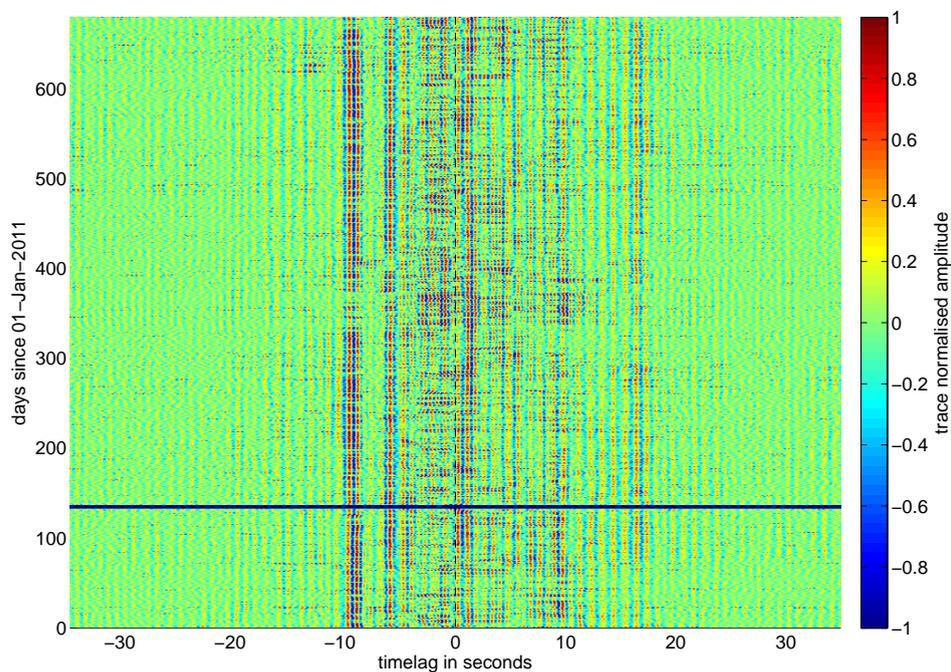Figure 5.40.: Colourplot with the available one-day long RR cross-correlograms of the years 2011 and 2012 of stationpair TMO53-TMO54. The CCFs are filtered between 0.8 Hertz and 1.6 Hertz. Many signals interfere between -20 seconds and 20 seconds.
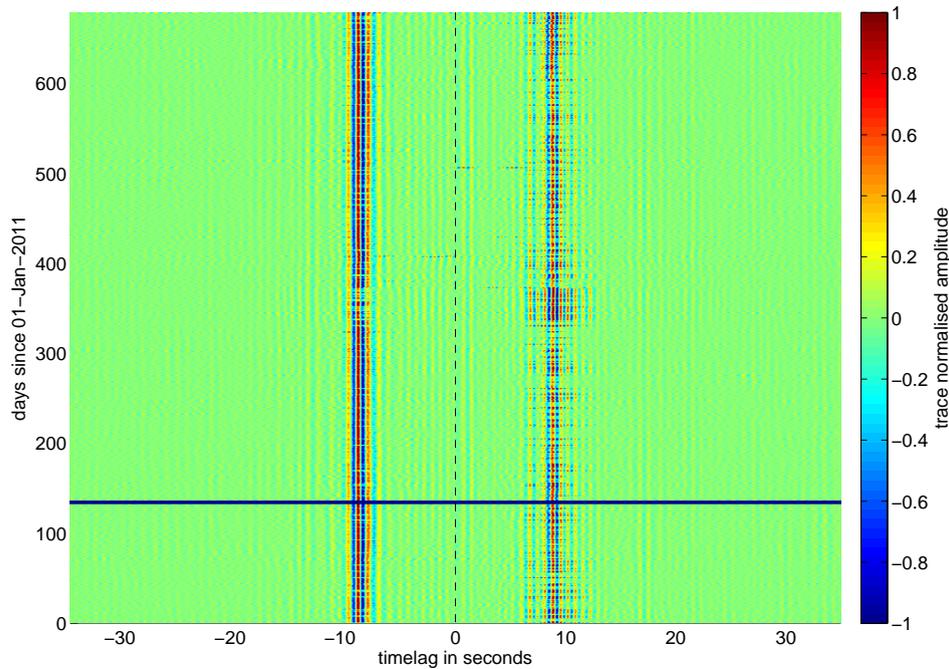
Figure 5.41.: Colourplot with the available one-day long TT cross-correlograms of the years 2011 and 2012 of stationpair TMO53-TMO54. The CCFs are filtered between 0.8 Hertz and 1.6 Hertz. Clearly, two temporally stable signals, symmetric to timelag zero, emerge in the cross-correlograms. Around 34 seconds continuous signals are detectable.
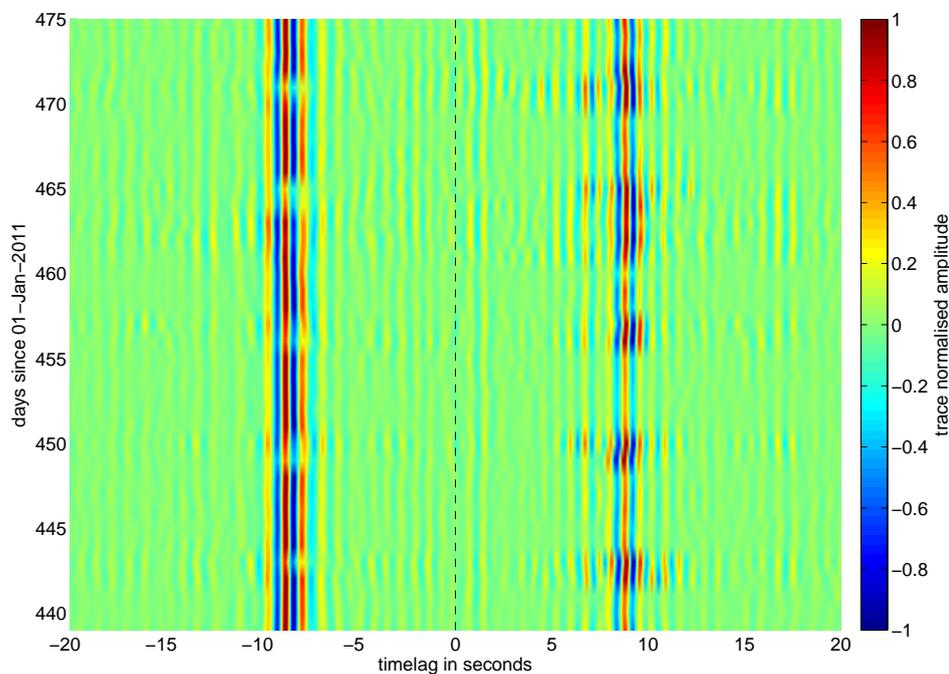

Figure 5.42.: Zoom of Figure 5.41. The one-day long TT cross-correlograms of stationpair TMO53-TMO54 between 14 March 2012 (Wednesday) and 19 April 2012 (Thursday) are depicted. The CCFs are filtered between 0.8 Hertz and 1.6 Hertz. A weekly variation of the signal amplitude in the acausal part is clearly identifiable.
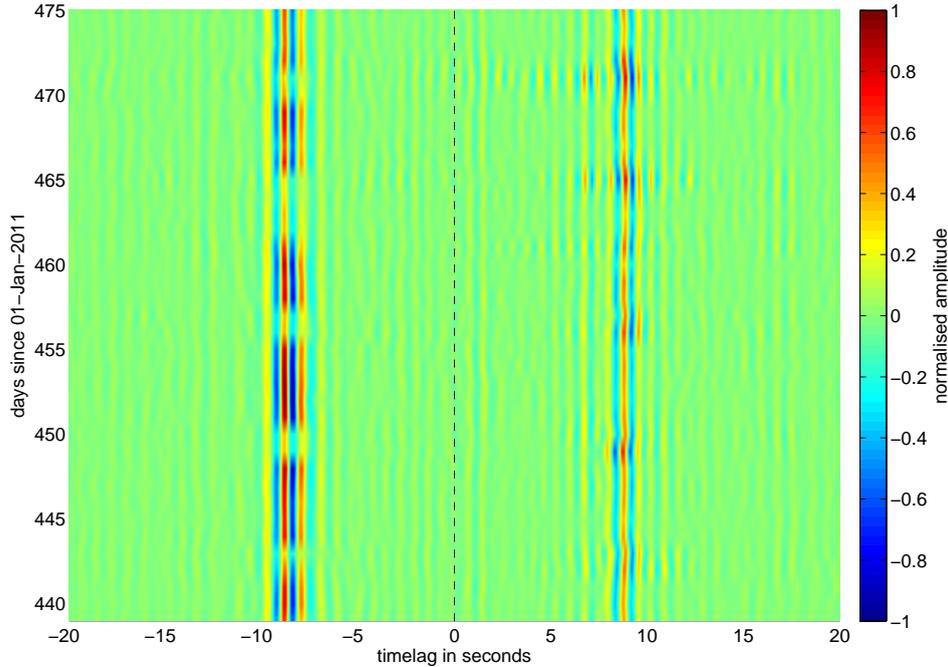
Figure 5.43.: The same colourplot as in Figure 5.41 except for the normalisation of the traces of the TT components (frequency range: 0.8 Hertz to 1.6 Hertz): Every trace is normalised to the absolute maximum of all the available CCFs of the years 2011 and 2012. Variations of the signal amplitude in the acausal part are clearly identifiable. Low amplitudes are connected to weekends and Easter 2012 (days 462-465).

normalised amplitudes (Figure 5.42) the signals in the causal part of the cross-correlograms in Figure 5.43 only slightly vary in amplitude over time. This is an indicator that mainly the signal strengths in the acausal part vary weekly. Figure 5.43 also reveals Easter 2012. The days 462 to 465 (since 1 January 2011) coincide with Good Friday, Holy Saturday, Easter Sunday, and Easter Monday. These specific temporal variations are an indicator for a man-made noise source. In the ZZ and RR cross-correlograms such clear workday-/weekend-variations are not identifiable. This might be due to the comparatively low SNR on ZZ and RR. It is also possible that different noise sources emit signals, which, on the one hand, are only observed on the TT components, and on the other hand, are only observed on the RR or ZZ components.

In summary, a temporal stability of the signals in the ZZ, RR, and TT cross-correlations is observed over a long time period and over the whole bandwidth between 0.1 Hertz and 1.6 Hertz. In the coda of the cross-correlograms temporally stable signals of very low amplitude are also detectable. Thus, the basic requirements for monitoring are fulfilled by the TIMO2 data set (see Section 2.3). The delay time between the arrivals of the observed coda waves at different time periods (for example, at winter and at summer) might be estimated in future work. In a further step, these delay times might be reducible to temporal variations of the medium, through which the waves of the coda propagated.

Moreover, colourplots reveal temporal variations of the signal amplitudes. The colourplot in Figure 5.42 of stationpair TMO53-TMO54, for example, shows workday- and weekend-dependent signal strengths. In turn, anthropogenic noise sources (see Section 2.2) might be identifiable, if one looks closer at the signal variations in colourplots. Future work should also investigate temporal signal variations between night- and day-times. This might reveal additional information on the origin of the seismic noise field around Landau.

5.2. Localisation of Noise Sources

For a better understanding of the origin of the noise sources which contribute to the signals in the CCFs filtered in bandwidths with frequencies between 0.1 Hertz and 1.6 Hertz two different approaches for localising these sources (Horstmann, 2010) are explained and evaluated with the TIMO2 data set in the following.

In the one case, a noise source far away from the network is assumed, and the direction from which the planar wavefront of this noise source enters the network is determined (see Section 5.2.1). In the other case, a point-like noise source inside of the network is considered, and the approximate location of this point source is calculated (see Section 5.2.2).

At the end of Section 5.2.2 a short summary and a discussion of the results are given.

5.2.1. Planar Wave Incidence from outside of the Station Network

In the frequency bands between 0.1 Hertz and 0.2 Hertz, between 0.2 Hertz and 0.4 Hertz, and between 0.4 Hertz and 0.8 Hertz the portion of high-amplitude signals in the cross-correlograms of all three component combinations (ZZ, RR, and TT) is higher in the acausal part than in the causal part. If one assumes a planar wavefront, propagating across the network as the source for these signals, the following can be done to verify this assumption:

At first, a measure of the direction of the incoming wavefront needs to be defined. Hence, I use the backazimuth, the angle measured clockwise from north to the normal of the wavefront. With a fixed angle of incidence the so-called effective distance the wave has to propagate between two stations of a cross-correlation pair can be calculated. For example, if the wavefront travels across the network in a west-east direction, the angle will be equal to 270°. Furthermore, if the two stations' connecting line also lies in west-east direction (this is equivalent to a stationpair azimuth of 270°), the effective distance will exactly be the interstation distance of the stationpair. On the other hand, if the connecting line of a pair of stations lies in north-south direction, the effective distance, the wavefront has to cover to arrive at the two stations, will be equal to zero. Thus, if the direction of the incoming wavefront is a good estimate for the real data, the signals in the cross-correlograms of stationpairs with a short effective distance will arrive at an earlier time than the signals in the cross-correlograms of stationpairs with a large effective distance. Plotting the cross-correlograms against the respective effective distances will then reveal a moveout – an increase of the arrival times of the wave with increasing effective distances. Actually, these plots are comparable with the distance plots in Section 5.1.1 with the cross-correlograms just being reorganised. The names of the used MATLAB scripts are XCORR_distanceplot_effective_incidentangle.m and XCORR_distanceplot_effective_incidentangle_stacked_CCFs.m (see Appendix B.2).

5.2.1.1. 0.1 Hertz to 0.2 Hertz

First, the cross-correlograms filtered between 0.1 Hertz and 0.2 Hertz are analysed. Therefore, the angle of the incident wavefront is varied from 180° to 360° in increments of 10°. Figure 5.44 shows the best result for the cross-correlograms of the ZZ components. Except for very few cross-correlograms the signals show a clear moveout, when an incident angle of $(300 \pm 15)°$ is assumed. For a better understanding of the error estimate on the incident angle, Appendix E shows five effective distance plots using the same cross-correlograms and five different consecutive incident angles. Within $\pm 15°$ it is not clear which incident angle results in the best moveout of the signals in the effective distance plots.

The effective distance plot of the RR cross-correlograms of the year 2012 (also filtered between 0.1 Hertz and 0.2 Hertz) are depicted in Figure 5.45. The best result is obtained
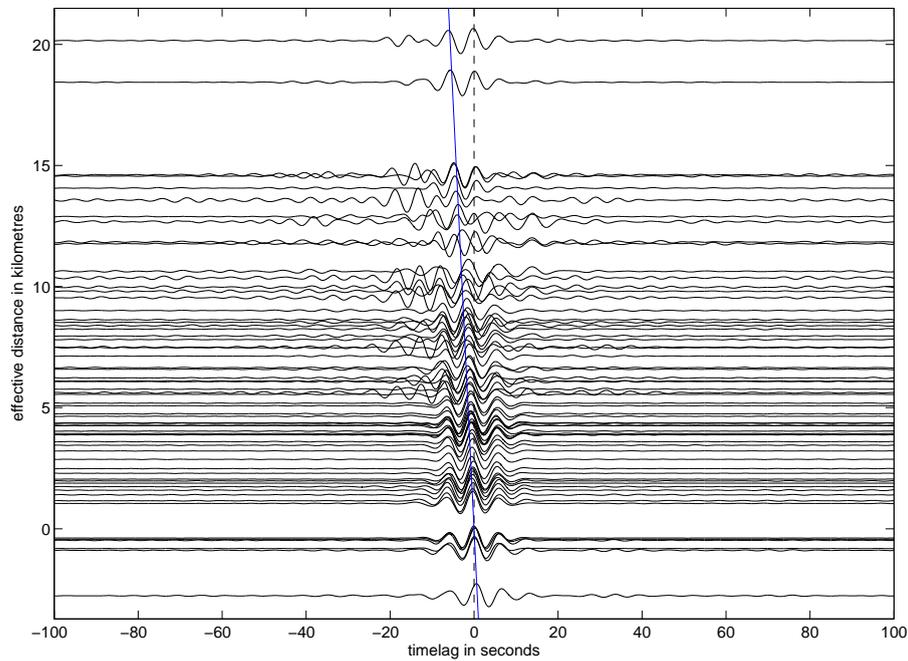
Figure 5.44.: The cross-correlograms of the ZZ components filtered between 0.1 Hertz and 0.2 Hertz are plotted against the respective effective distances which result from an incident angle of the wavefront of $300°$. A clear moveout is observed. The blue line indicates an (apparent) velocity of 3500 metres per second. At effective distances larger than about 9 kilometres the signals do not align as well as at smaller distances. Cross-correlograms depicted at negative effective distances will be obtained if the order of the arrivals of the wavefront at two stations of a stationpair is reciprocal to the order of the arrivals of two stations whose cross-correlograms are depicted at positive effective distances.
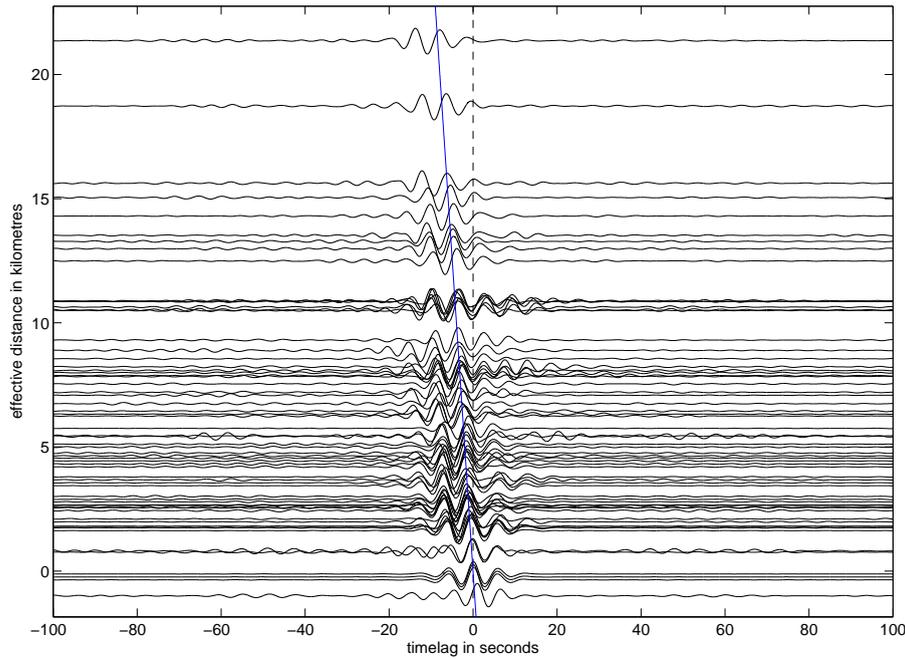
Figure 5.45.: The effective distance plot of the RR components. All available data of the year 2012 are used, and filtered between 0.1 Hertz to 0.2 Hertz. The angle of the incident wavefront equals $290°$. The blue lines indicates an (apparent) velocity of 2500 metres per second.

by assuming an incident angle of the wavefront of $(290 \pm 15)°$.

Moreover, Figure 5.46 shows the effective distance plot of the TT components. The incident angle that delivers the best result is equal to $(280 \pm 15)°$. An alignment of the signals as good as in Figures 5.44 and 5.45 is not observed.

On the one hand, it is important to keep in mind, that the interstation distance of the stations are too small for satisfying the far field approximation in this frequency range (see Subsection 5.1.1). On the other hand, one has to be aware of the fact that the cross-correlograms depicted in Figure 5.45 were rotated into the direction of the connecting line of the respective stations of each cross-correlation pair, and that those in Figure 5.46 were rotated into the direction perpendicular to the connecting line. Thus, although the stations are obviously not uniformly surrounded by noise sources, it is surprising that the signals of the rotated cross-correlograms show such a clear moveout. A possible explanation might be the unequal azimuthal distribution of the pairs of stations. Most connecting lines lie in a roughly west-east direction, which coincides with the angle of the incident wave ($\sim 290°$). In turn, this means that many RR components point into directions which roughly agree with the direction the wave travels along. In Figure 5.47 a polar area diagram illustrates the number of azimuths and backazimuths per $10°$ interval for the stationpairs of the TIMO2-network in 2012.

5.2.1.2. 0.2 Hertz to 0.4 Hertz

The same procedure as above (Subsection 5.2.1.1) is carried out with the cross-correlograms filtered between 0.2 Hertz and 0.4 Hertz. Figures 5.48, 5.49, and 5.50 show the best results of the effective distance plots for the three different component combinations ZZ, RR, and TT. Due to the increasing number of signals the moveouts are not as clear as in the lowest frequency band from 0.1 Hertz to 0.2 Hertz. The angle of the incident wavefront seems to
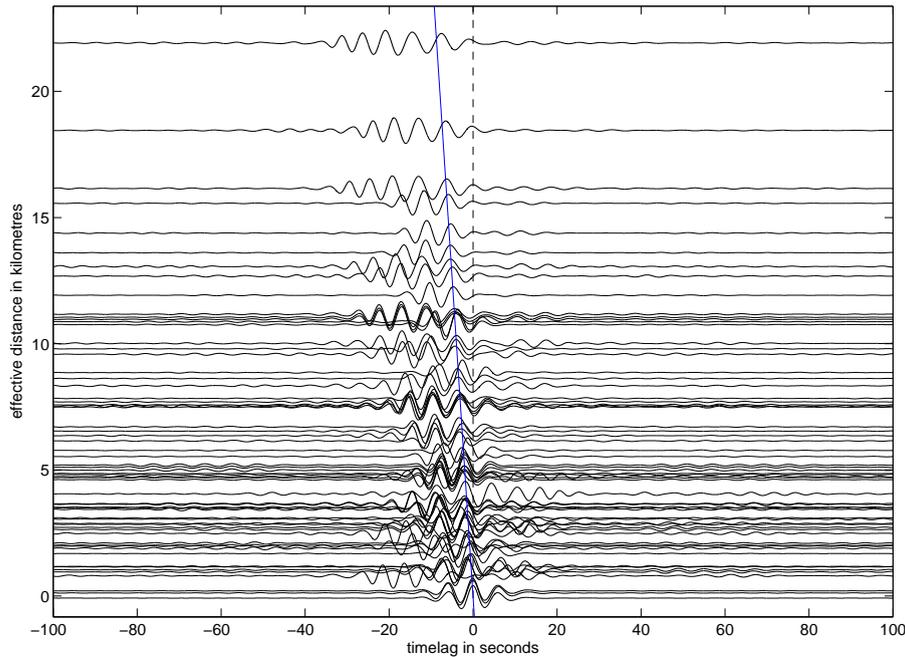
Figure 5.46.: The effective distance plot of the TT components. All available data of the year 2012 are used, and filtered between 0.1 Hertz to 0.2 Hertz. The best moveout is observed for an angle of 280°. The blue line indicates an (apparent) velocity of 2500 metres per second.

be frequency independent (in the bandwidth from 0.1 Hertz to 0.4 Hertz) for the signals observed on the ZZ components. For the RR components the angle only slightly changes from at first $(290 \pm 15)°$ to $(300 \pm 15)°$ at frequencies between 0.2 Hertz and 0.4 Hertz. The moveout in the effective distance plot of the TT components looks worse than the ones of the ZZ and RR components. The best result for TT is obtained with an incident angle of the wavefront of $(280 \pm 15)°$. Again, it is important to keep in mind that the horizontal CCFs are rotated by an angle according to their stationpair azimuth and not by an angle which corresponds to the angle of the incoming wavefront.

5.2.1.3. 0.4 Hertz to 0.8 Hertz

In the frequency range from 0.4 Hertz to 0.8 Hertz a very clear moveout cannot be observed anymore in the effective distance plots of all three component combinations. An example of the effective distanceplot of the ZZ components with a vaguely perceptible moveout is given by Figure 5.51.
Thus, all the signals in the cross-correlograms filtered in the frequency band between 0.4 Hertz and 0.8 Hertz cannot be explained anymore by a plane wavefront propagating across the network. As already mentioned above, one (or more) point source(s) inside of the network might then be the reason for the occurrence of these signals with frequencies higher than about 0.4 Hertz. This case is analysed in the next section.

5.2.2. Point Sources inside of the Station Network

A method for localising a point source in a given area is the so-called migration analysis (after Horstmann, 2010). For this analysis, a theoretical grid of evenly distributed hypothetical point sources is superimposed on the area of interest. At first, a uniform velocity v of the waves emitted by the hypothetical point sources is defined. Then, the
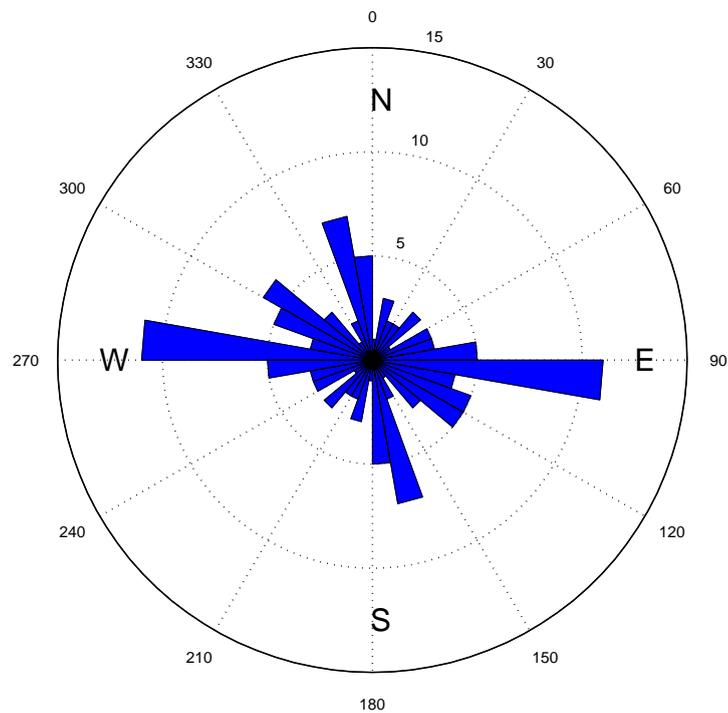
Figure 5.47.: If one defines the azimuth merely at the western station of a stationpair, one will obtain azimuth values between 0° and 180° only. Furthermore, if the backazimuth is only measured at the eastern station of a stationpair, the backazimuth values will vary between 180° and 360°. In the polar area diagram above the azimuthal distribution of the pairs of stations of the TIMO2-network in 2012 is depicted on the righthand side of the circle (0° - 180°). On the lefthand side of the circle (180° - 360°) the backazimuthal distribution is shown. The azimuthal (and also the backazimuthal) coverage is not balanced. Many connecting lines between pairs of stations lie in roughly west-east direction. Due to the local character of the station network the azimuthal and backazimuthal distributions are almost symmetric.

Figure 5.48.: The ZZ cross-correlograms of the year 2012 filtered between 0.2 Hertz to 0.4 Hertz are plotted against their effective distances assuming a planar wave incidence with an angle of $300°$. The blue line marks a velocity of 1500 metres per second.

Figure 5.49.: The RR cross-correlograms of the year 2012 filtered between 0.2 Hertz to 0.4 Hertz are plotted against their effective distances assuming a planar wave incidence with an angle of $300°$. The blue line marks a velocity of 1500 metres per second.
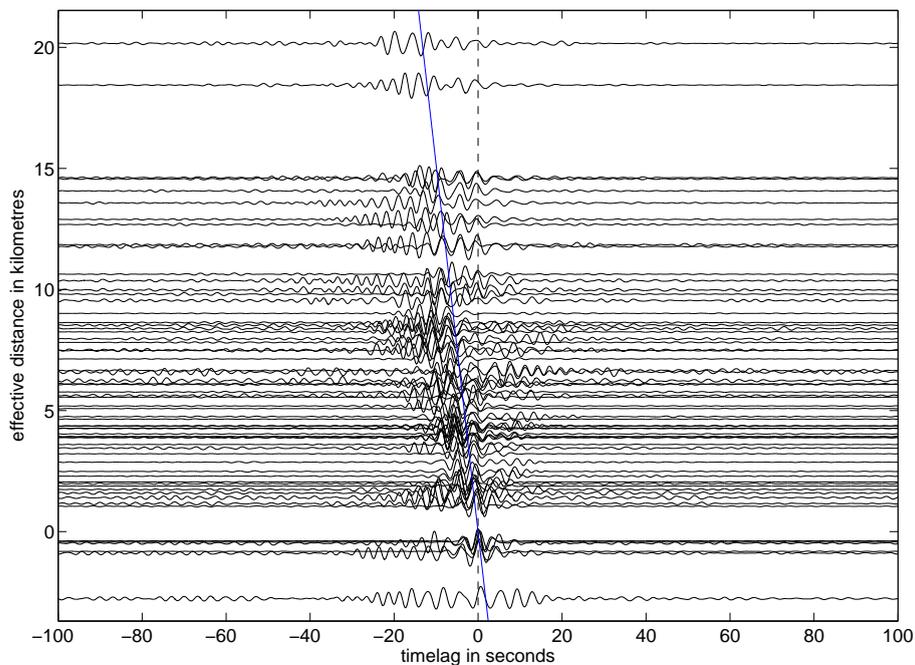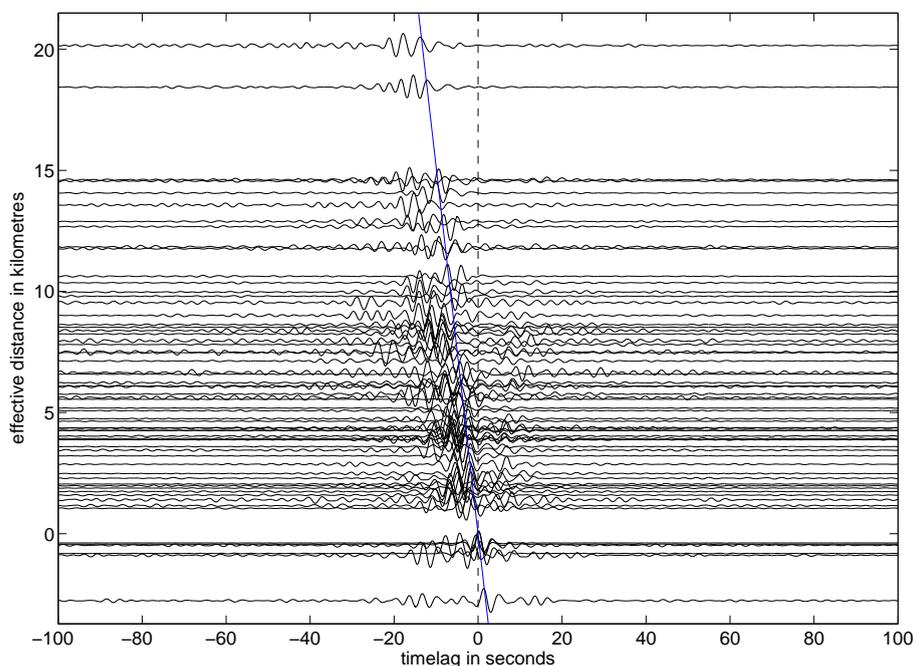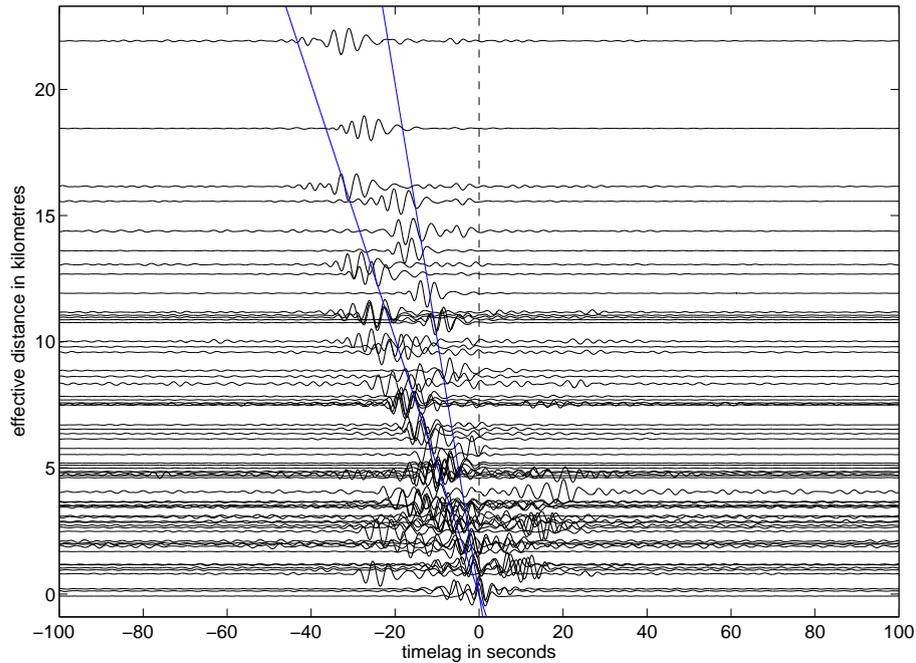
Figure 5.50.: The TT cross-correlograms of the year 2012 filtered between 0.2 Hertz to 0.4 Hertz are plotted against their effective distances assuming a planar wave incidence with an angle of 280°. The blue line with the higher slope marks a velocity of 1000 metres, the blue line with the smaller slope indicates a velocity of 500 metres per second.

following scheme is performed for each grid point at x_i (with the number of grid points i): The traveltimes t_{st} from the point source to the n seismic stations ($st = 1, 2, ..., n$) located at x_{st} are computed: $t_{st} = \frac{|x_i - x_{st}|}{v}$. Afterwards, for each possible combination of two stations ($P(n)$) the difference of the two respective traveltimes is calculated. This difference corresponds to that lag time in the CCF at which the hypothetical signal would occur. To be able to compare the real data with the hypothetical arrival times of one point source, the envelope (modulus of the Hilbert-transformed cross-correlation) of each CCF is computed and normalised to one. Then, the amplitude values at the hypothetical arrival times in the CCFs of all stationpairs are added. If the grid point is located at the real point source and if the pre-defined velocity is the actual velocity of the signal, the sum of the amplitudes will reach its maximum value. On the other hand, if the defined velocity and the hypothetical point source location do not coincide with the real source, a very low amplitude value sum will be obtained. Hence, the sum of the amplitude values is a measure of the similarity between hypothetical arrival times of one point source and the real data. This value is called semblance value. In a perfect setup, the maximum semblance value is equal to the number of stationpairs $P(n)$.

The semblance values of each grid point are determined one after another (migrated) following the scheme above. In the end, one obtains a spatial distribution of the semblance values. A criterion for a good localisation of a point source is a relatively small area with relatively high semblance values.

The MATLAB function used for the migration analysis is called Migrationsanalyse (see Appendix B.2), and was written by Tobias Horstmann. He also successfully tested the functionality of this MATLAB script with a synthetic data set (Horstmann, 2010).

Furthermore, Ma et al. (2013), for example, use the same approach successfully: They are

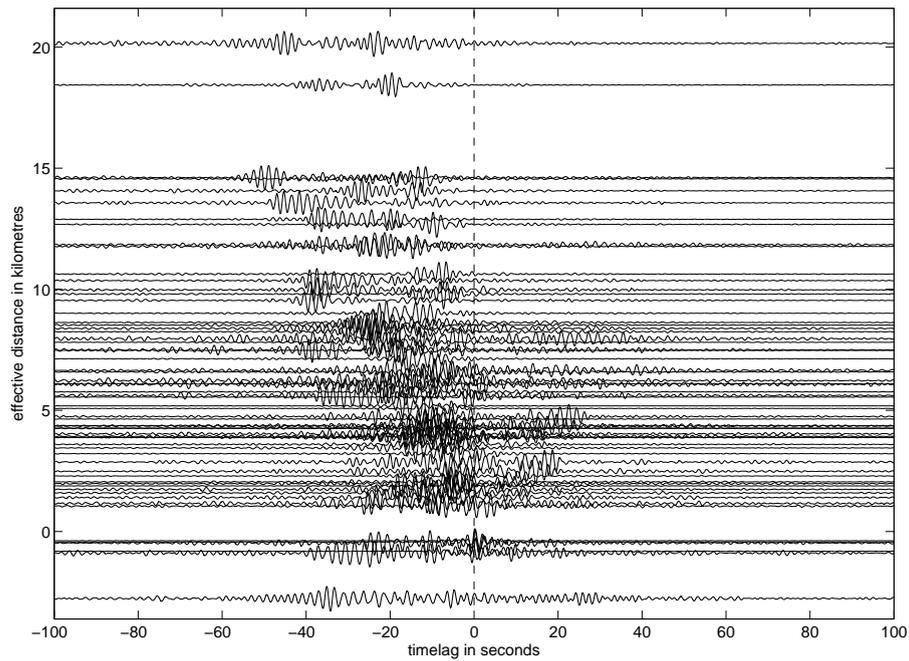Figure 5.51.: The *ZZ* cross-correlograms of the year 2012 filtered between 0.4 Hertz to 0.8
Hertz are plotted against their effective distances assuming a planar wave
incidence with an angle of 280°. Only a slight moveout is observed. There-
fore, most of the occurring signals cannot only be explained by a wavefront
propagating across the network. There must be other, additional sources for
the signals in the cross-correlograms.

able to locate a passive scatterer in Southern Peru from ambient noise cross-correlations. For our migration analyses a grid is defined between the latitudes of 49.05° and 49.35°, and the longitudes of 7.75° and 8.40°. The grid spacing is about 250 metres in west-east and also in north-south direction ((49.35° − 49.05°)/130 yields about 250 metres and (8.40° − 7.75°)/190 also yields about 250 metres).

The migration analyses are computed with the CCFs filtered in two different frequency bands: 0.4 Hertz to 0.8 Hertz and 0.8 Hertz to 1.6 Hertz. Due to the preponderance of signals in the acausal part of the CCFs in the frequency range from 0.4 Hertz to 0.8 Hertz, a dominant source for these signals is assumed in the western part of the network. Thus, I choose to use the CCFs of the year 2012 for the migration analyses, because in this year the station TMO65 was installed in the very west of the TIMO2-network. In turn, this station constellation guarantees a better coverage of the western area than the network constellation without TMO65 (see the map in Section 3.2). Under the assumption of temporally stable signals (exemplarily for one stationpair see Section 5.1.3) I use the CCFs of all available stationpairs in 2012, even though some of the twelve used stations recorded data not at the same time. The only stations which have no recording time in common are the stations TMO50 and TMO65. Therefore, the number of CCFs used for the migration analysis equals 65 instead of 66 ($12 \cdot 11 \cdot \frac{1}{2} = 66$). The additionally used MATLAB script is zusatz_migrationsanalyse_filter.m.

5.2.2.1. 0.4 Hertz to 0.8 Hertz

In the frequency range from 0.4 Hertz to 0.8 Hertz fourteen different migration analyses are calculated each with another predefined signal velocity emitted by the hypothetical point sources. Velocities are varied between a minimum of 200 metres per second and a maximum of 2.1 kilometres per second. Between 300 metres per second and 700 metres per second the highest semblance between hypothetical arrival times and the real data is obtained. The concordance is of 55% on average. At lower (< 200 metres per second) and higher (> 700 metres per second) velocities the semblance values decrease to about 35%. The best result of the migration analyses will be obtained, if one assumes a velocity of 600 metres per second. In this case, the migration analysis yields a relatively small area with high semblance values (shown in yellow colour) west-north-west of the station network (see Figure 5.52). This yellow area coincides with the village of Albersweiler and a small region around this village including the Albersweiler Quarry. Around this yellow area there are also regions with relatively high semblance values (shown in orange colour) which extend on the one hand to the south-east of station TMO65 and on the other hand to the far north-west of the network.

Based on the result of the migration analysis shown in Figure 5.52 two additional seismic stations were deployed in the western part of the TIMO2-network to improve the localisation of the noise sources. The one of the stations was installed south-east to station TMO65 and the other north-east to TMO65. More information on the stations' set-up are given in Section 3.2. So far, the data of the two new stations could not have been used for a further investigation of the noise sources, because not enough data were available until the end of June 2013 due to technical problems (defect of the power supply).

Using the geographic coordinates of the point with the highest semblance value obtained by the migration analysis depicted in Figure 5.52 it is possible to plot the cross-correlograms against their respective effective distances related to the source location (Horstmann, 2010). This is very similar to the effective distance plots introduced in Section 5.2.1, where a planar wave incidence was assumed. The effective distance d between the point source located at x_S and two stations of a cross-correlation pair located at x_1 and x_2 can
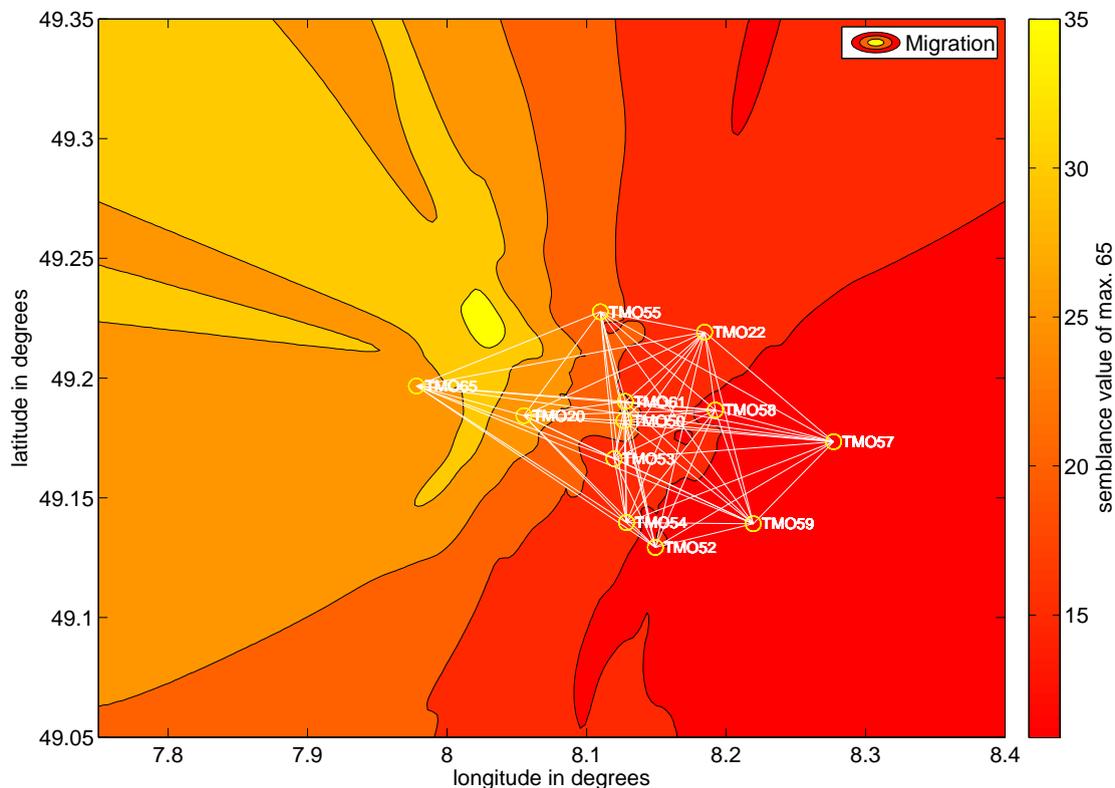
Figure 5.52.: The best result of the migration analysis with the ZZ CCFs of the year 2012
            filtered between 0.4 Hertz and 0.8 Hertz. The predefined velocity is equal to
            600 metres per second. The highest achieved semblance value of maximal 65
            is 36.9. This corresponds to a concordance of 57% between the hypothetical
            arrival times and the real data. The point with the highest semblance value
            is located at 49.2277° N and 8.0237° E.

be formulated by:

$$d = (t_2 - t_1) \cdot v = \left(\frac{|x_2 - x_S|}{v} - \frac{|x_1 - x_S|}{v} \right) \cdot v = |x_2 - x_S| - |x_1 - x_S| \qquad (5.2)$$

with the velocity v of the signal emitted by the source and the traveltime t_1 between the source and the first station, and the traveltime t_2 between the source and the second station. If the modulus of the distance between the first station and the source is higher than the modulus of the distance between the second station and the source ($|x_1 - x_S| > |x_2 - x_S|$), the effective distance d will be negative. I always choose the western station of a stationpair as the first station and the eastern station of a stationpair as the second station.

The used MATLAB scripts are: workspace_saving_calc_distplot_pointsource.m, XCORR_distanceplot_effective_pointsource.m, and XCORR_distanceplot_effective_pointsource_section.m.

Using the geographic coordinates 49.2277° N and 8.0237° E, one obtains the effective distance plot in Figure 5.53. As already mentioned above, cross-correlograms depicted at negative effective distances will be obtained if the order of the arrivals of the wave at two stations of a stationpair is reciprocal to the order of the arrivals of two stations whose cross-correlograms are depicted at positive effective distances. The black dashed line indicates the wave's propagation velocity of 600 metres per second, which yielded the best result for the migration analyses.

Many high-amplitude signals lie on the black dashed velocity line. However, in some cross-correlograms the amplitude of the signal propagating with 600 metres per second is (much) smaller than the amplitude of signals at other timelags. The origin of the latter signals cannot be explained by the point source located at 49.2277° N and 8.0237° E. Moreover, this explains why the maximum achieved semblance value of the migration analysis depicted in Figure 5.52 is only equal to 57% and not higher.

5.2.2.2. 0.8 Hertz to 1.6 Hertz

In the frequency range from 0.8 Hertz to 1.6 Hertz nine different migration analyses are computed. Beginning with the smallest velocity equal to 200 meters per second the values are increased by 50 meters per second to the highest velocity of 600 metres per seconds. The highest semblance value is reached in the migration analysis with a predefined velocity of 350 metres per second. The result is depicted in Figure 5.54.

As before, the cross-correlograms of the year 2012 (this time filtered between 0.8 Hertz and 1.6 Hertz) are plotted against their respective effective distance assuming a point source in the south-east of the station network at 49.1677° N and 8.2050° E (Figure 5.55).

Comparable to the effective distance plots in Figure 5.53, the signals with the highest amplitude of each cross-correlogram do not all lie on the black dashed line indicating a velocity of 350 metres per second. The sources for the signals which cannot be explained by a point source located at 49.1677° N and 8.2050° E are unidentified.

However, the yellow area with the highest semblance values in Figure 5.54 coincides with the location of a windpark. It is the Windpark Offenbach an der Queich with three wind turbines. If one assumes that at larger distances away from the windpark the signals originating from this windpark are dominated by other noise sources, and hence, do not include station TMO65 (which is farthest apart from the windpark) in the migration analysis, one will obtain an even higher concordance of 60% between the hypothetical arrival times and the real data. The result of the migration analysis without the stationpairs including TMO65 is shown in Figure 5.56. The result without TMO65 is better with respect to the percentage concordance between the hypothetical arrival times and the real data. The location of the area with the highest semblance values, on the other hand, does not change

Figure 5.53.: Upper left: The effective distance plot with the cross-correlograms of the year 2012 filtered between 0.4 Hertz and 0.8 Hertz. The location of the point source is at 49.2277° N and 8.0237° E. The other diagrams show sections of the effective distance plot in the upper left corner (upper right: -3 kilometres to 3 kilometres; lower left: 3 kilometres to 9 kilometres; lower right: 9 kilometres to 15 kilometres). The black dashed lines indicate a propagation velocity of 600 metres per second. The points of intersection of this line and each cross-correlogram are equal to that timelags at which the amplitude information was taken for the migration analyses.

Figure 5.54.: The best result of the migration analysis with the ZZ CCFs of the year 2012 filtered between 0.8 Hertz to 1.6 Hertz. The predefined velocity is equal to 350 metres per second. The highest achieved semblance value of maximal 65 is 34.6. This corresponds to a concordance of 53% between the hypothetical arrival times and the real data. The point with the highest semblance value is located at 49.1677° N and 8.2050° E.
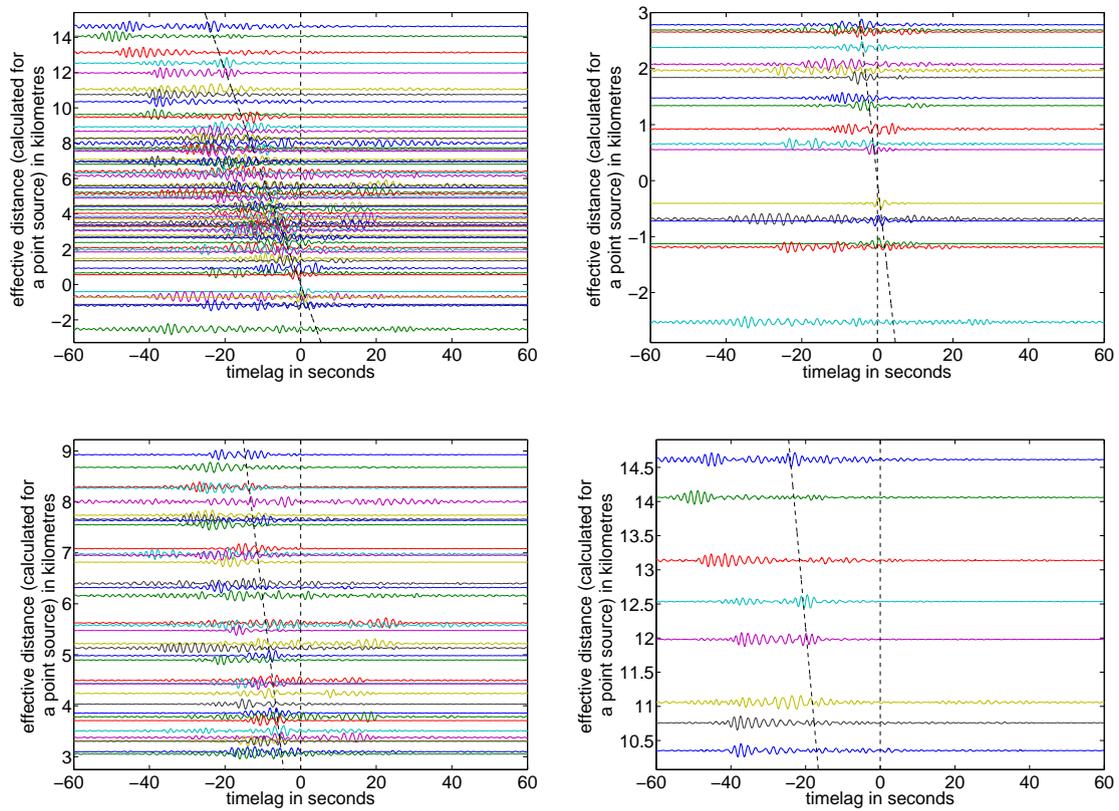
Figure 5.55.: Upper left: The effective distance plot with the cross-correlograms of the year 2012 filtered between 0.8 Hertz and 1.6 Hertz. The location of the located point source is at 49.1677° N and 8.2050° E. For the sake of clarity, the other diagrams show enlarged sections of the effective distance plot in the upper left corner (upper right: 0.5 kilometres to -1.5 kilometres; lower left: -2.7 kilometres to -5 kilometres; lower right: -10 kilometres to -12 kilometres). The black dashed lines indicate a propagation velocity of 350 metres per second. If the effective distances are positive, the source will be closer to the first (western) station than to the second (eastern) station of a cross-correlation pair. On the other hand, if the first station is further away from the source than the second station, the effective distances will have negative values (see Formula (5.2)).
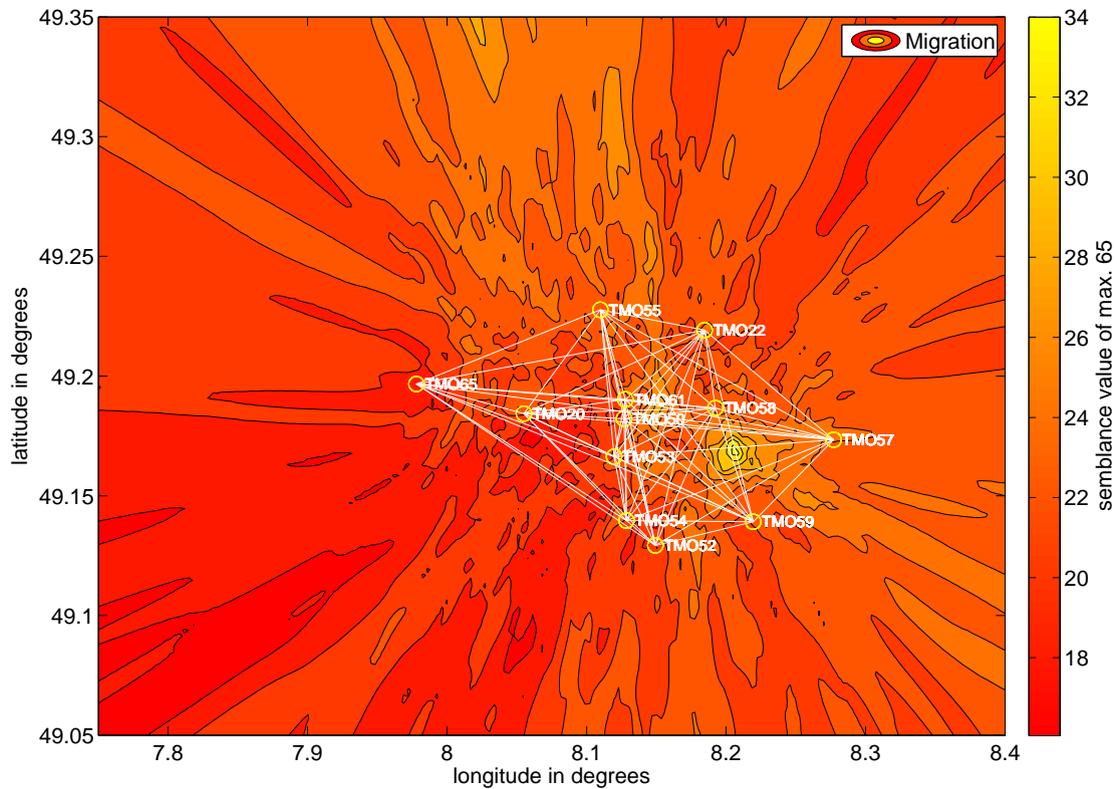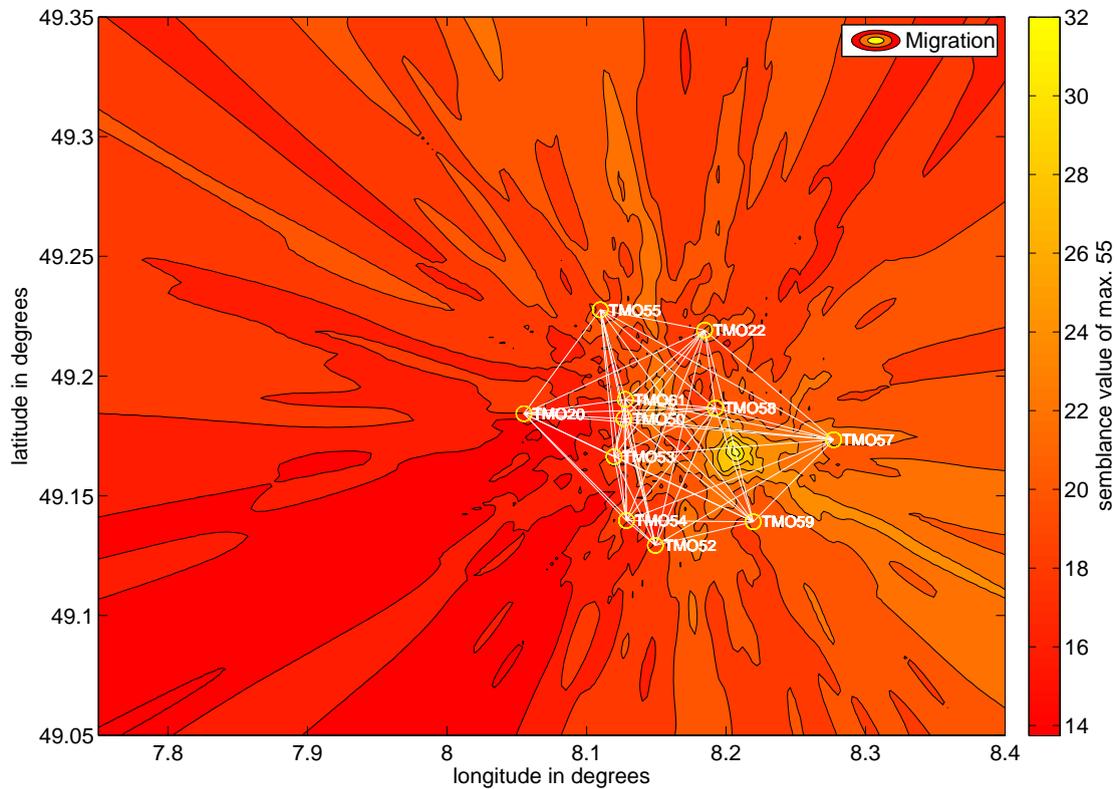
Figure 5.56.: The best result of the migration analysis with the *ZZ* CCFs of the year 2012 filtered between 0.8 Hertz to 1.6 Hertz. The stationpairs with station TMO65 are excluded from the analyses. The predefined velocity is equal to 350 metres per second. The highest achieved semblance value of maximal 55 is 33. This corresponds to a concordance of 60% between the hypothetical arrival times and the real data. The point with the highest semblance value is located at 49.1677° N and 8.2084° E.
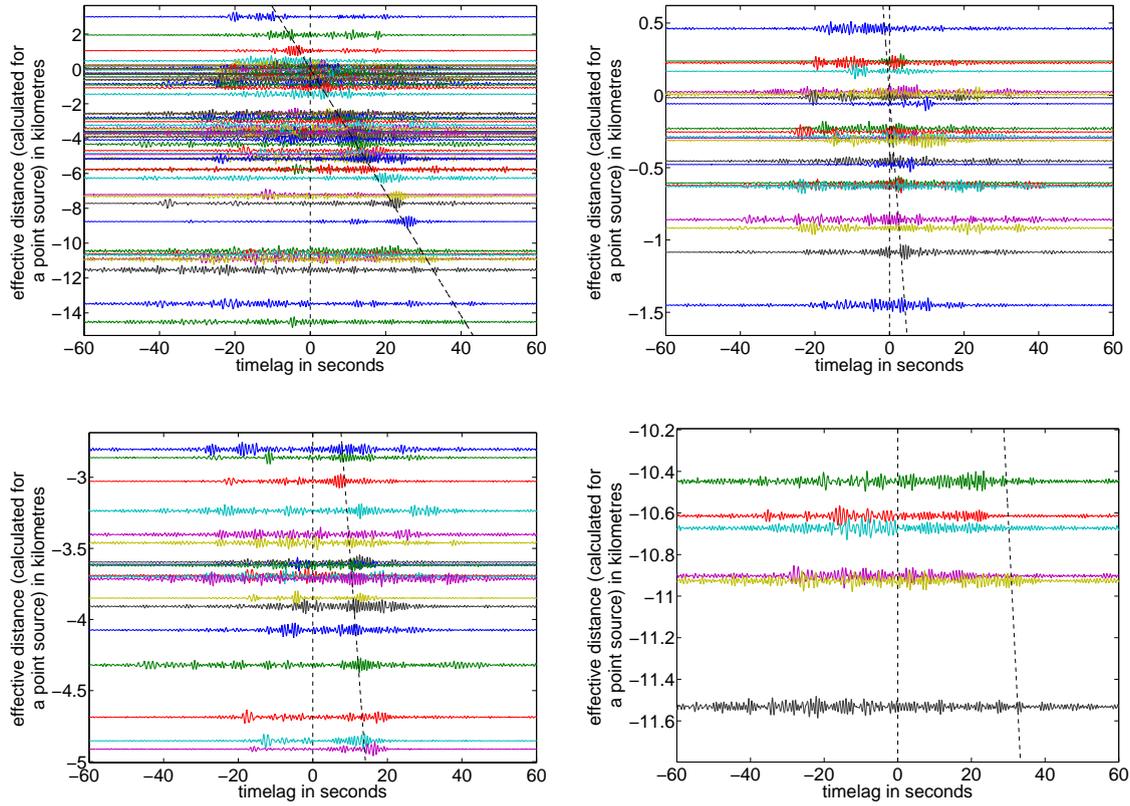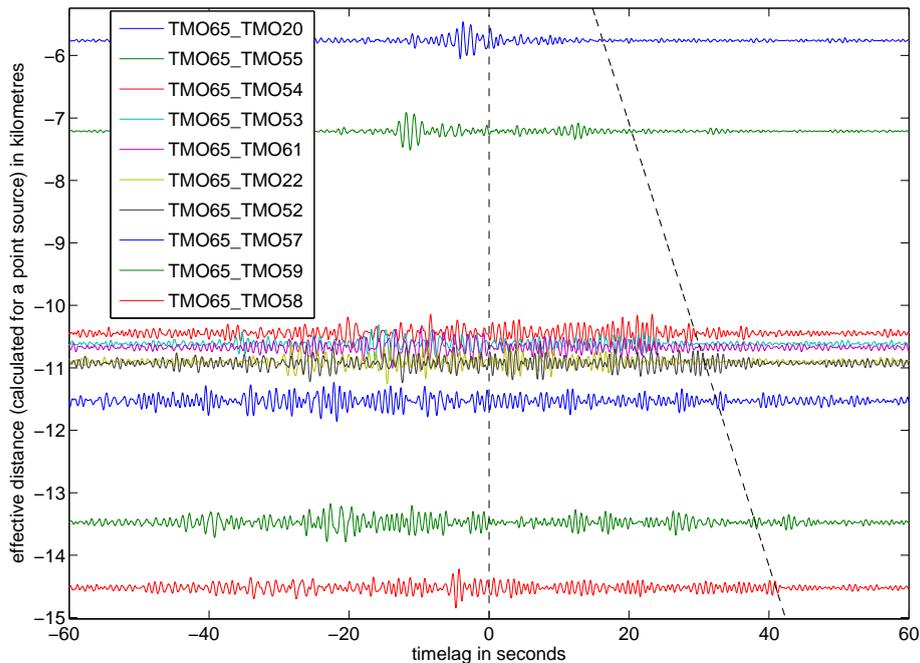
Figure 5.57.: The effective distance plot only with the cross-correlograms of the station-
pairs with station TMO65. The assumed point source is located at 49.1677°
N and 8.2050° E. The black dashed line indicates a velocity of 350 metres
per second. In many cross-correlograms relatively high-amplitude signals lie
on this line. The effective distances ($d = |x_2 - x_S| - |x_1 - x_S|$) are negative,
because the distance between the first station (here: TMO65) and the source
is larger than the distance between the second station and the source (see
the explanation for the formula of the effective distance in Section 5.2.2.1).

significantly, if station TMO65 is excluded from the analysis.

Figure 5.57 shows only the cross-correlograms of the stationpairs with station TMO65
plotted against their respective effective distances. At the intersection points of the 350
metres per second line and the cross-correlograms relatively high-amplitude signals occur.
Therefore, our assumption of the signals emitted by the windpark being too weak to be
detected at station TMO65 might be invalid.

In the following the information given in this section and the section before (Section 5.2.1)
are summarised and interpreted:

In the lowest frequency band from 0.1 Hertz to 0.2 Hertz the assumption of a planar wave
incidence is justified: A clear moveout is observed in the effective distance plots (see Fig-
ures 5.44, 5.45, and 5.46). In this frequency band the most likely origin for the signals
are the ocean-generated microseisms. Regarding the wave's incidence angle of 300° the
Atlantic Ocean and the North Sea around the British Isles including the British Channel
are most probably the origin of these low-frequency noise signals. Friedrich et al. (1998)
identified this region as a source for secondary microseisms using the Gräfenberg array.

In the bandwidth from 0.2 Hertz to 0.4 Hertz the ocean-generated microseisms interfere
with signals of other (unidentified) noise sources (see Figures 5.48, 5.49, and 5.50). But
still the signals originating from the ocean-generated microseisms dominate the cross-
correlograms.

In the frequency band from 0.4 Hertz to 0.8 Hertz other noise sources than the ocean-
generated microseisms start to dominate the CCFs. However, an influence of the CCFs
by the ocean generated microseisms is still detectable. In the effective distance plot of the

ZZ CCFs filtered between 0.4 Hertz and 0.8 Hertz (see Figure 5.51) a slight moveout can still be observed, if the angle of the incident wave is equal to 280°. This angle only slightly differs from the angle of the incoming wave that was determined in the lowest frequency band (see above). The results of the migration analyses with the CCFs filtered between 0.4 Hertz and 0.8 Hertz also show that the ocean-generated microseisms still slightly influence the CCFs (Figure 5.52). An area with relatively high semblance values (coloured in orange) extends to the far north-west of the TIMO2-network, which is again the same direction that points to the seas between the British Isles and France. It is still indeterminate, if there is a dominant noise source in the region with the highest semblance values. Nonetheless, using the data of the two additional installed stations should yield a more distinct result.

At frequencies between 0.8 Hertz and 1.6 Hertz another noise source dominates the CCFs. This time it is a point-like source in the south-east of the station network. By the migration analyses with CCFs filtered between 0.8 Hertz and 1.6 Hertz I localise a small area, where the Windpark Offenbach an der Queich is located (Figures 5.54 and 5.56).

Principally, wind turbines generate seismic and acoustic waves in two ways: On the one hand, the vibrations of the wind turbine, which is strongly coupled to the ground by massive concrete foundations, are transmitted into the subsurface surroundings through these foundations. On the other hand, the rotating blades of the wind turbine cause low-frequency acoustic signals, which also might acoustically couple to the ground (Styles et al., 2005). The distances over which signals created by wind turbines can be measured depends on many factors like, for example, the height of the tower, the size of the foundation and so on. Styles et al. (2005) state that at distances >10 kilometres wind turbine induced signals are still detectable. Furthermore, the frequencies of some of the wind turbine induced signals are related to overtones of the blade-passing frequency of the wind turbine. Styles et al. (2005) conclude among other things that seismic signals between 0.5 Hertz and 5 Hertz are generated by wind turbines. Hence, so far there are no objections to the Windpark Offenbach an der Queich being the origin of the dominant noise in the frequency band between 0.8 Hertz and 1.6 Hertz. Of course, future work should establish a more solid connection between the temporal occurrence of the signals in the CCFs and the wind strengths varying over time, to make sure that the windpark is the actual source.

6. Summary and Outlook

It was tested, whether stable interferograms could be obtained with the TIMO2 data set (TIMO: Deep structure of the Central Upper Rhine Graben). As this was the case, the ambient noise cross-correlations were further investigated with respect to coherent signals. It was examined, if these signals are applicable, on the one hand, for determining the seismic velocity distribution in the subsurface (imaging), and on the other hand, for deriving temporal variations of the medium (monitoring). Properties of the seismic noise wavefield around Landau were also analysed in this context.

To get an overview of the properties of the vertical (ZZ) and horizontal (RR and TT) cross-correlations, they were filtered in nine different frequency bands (Section 5.1.1). All bandwidths lie within the frequency range between 0.1 Hertz and 45 Hertz. Furthermore, the cross-correlograms of the year 2012 of 65 stationpairs were plotted against the interstation distances. From these plots it could be inferred that the signal to noise ratio (SNR) of the signals in the cross-correlograms generally decreases with increasing interstation distances and with increasing frequencies. Hence, low-frequency signals propagate over longer distances more coherently. At lower frequencies (about from 0.1 Hertz to 0.8 Hertz) an asymmetry between signals in the acausal and causal parts of the cross-correlation functions is observed on all three component combinations ZZ, RR, and TT. In a first step, it was concluded that the noise sources are not uniformly distributed, and that the waves propagate along a roughly west-east direction across the network. In a further step, a localisation of the noise sources for these asymmetric signals was conducted (see below). At higher frequencies (>0.8 Hertz) the portion of signals in the causal and acausal part equalises. Compared to the ZZ and RR cross-correlograms, the TT cross-correlograms are characterised by a higher SNR, and symmetric signals emerge more clearly. It turned out that the frequency range from 0.8 Hertz to 1.6 Hertz meets the requirements best for a further investigation of the signals with respect to the applicability of imaging techniques. In this bandwidth, signals occur at negative timelags as well as at positive timelags, the SNR is comparatively high, and the cross-correlations are not perturbed by beat-like 'signals' (Section 5.1.1).

It is assumed that those beat-like amplitude modulations are formed in the cross-correlation process. For example, if two sinusoidal signals of nearly the same frequency are (independently) recorded by two stations of a cross-correlation pair, the cross-correlogram will contain 'beats' (Appendix C). Further investigation is necessary for an even better understanding of the origin of these beat-like 'signals', so that one might even be able to eliminate them from the cross-correlograms.

In the preferred frequency range from 0.8 Hertz to 1.6 Hertz the cross-correlograms of the

year 2012 of three stationpairs with different interstation distances and different azimuths were analysed in more detail (Section 5.1.2). A symmetry in relation to phase and frequency between the signals in the causal and acausal parts within a certain timelag window is observed on all three components (with one exception). Confirming the observations above, the signals in the TT cross-correlograms emerge more clearly than the signals on the ZZ and RR components. Hence, the TT cross-correlations more likely converge to the Green's function than the ZZ or RR cross-correlations.

For the identification of the wave types I determined the propagation velocities, and I generated polarisation diagrams. The propagation velocity of the signals varies from about 300 metres per second to about 410 metres per second (see Table 5.2). These velocities most likely correspond to Rayleigh wave or Love wave velocities. Slightly higher values for the Love and Rayleigh wave velocities were obtained by Köhler et al. (2007) at a geological similar setting. Polarisation diagrams in the RR-ZZ-plane did not always reveal the expected ellipticity of the particle movement of Rayleigh waves. In the TT-ZZ-plane most of the wave energy concentrates on the TT components. Thus, there are no objections that the signals on the TT components correspond to Love waves. Hence, a Love wave tomography using the TT cross-correlations might yield reliable results for the area around Landau. Why Rayleigh waves do not emerge as clearly out of the cross-correlograms might be due to different reasons: On the one hand the noise distribution around the used stations might not be suited as well for observations of Rayleigh waves as for Love waves. On the other hand, the physics, which explain the occurrence of Rayleigh waves differ from those of Love waves. Rayleigh waves are formed by couplings of P- and vertically polarised S-waves, Love waves result from interaction of horizontally polarised S-waves. The finding that signals emerge more clearly on the TT cross-correlograms than on the RR cross-correlograms was also observed by others, for example, by Lin et al. (2008), Jay et al. (2012) and Behm and Snieder (2013).

In a next step, the temporal stability of signals in the cross-correlograms was analysed (Section 5.1.3). It was exemplarily (stationpair TMO53-TMO54) shown that between 0.1 Hertz and 1.6 Hertz temporally stable signals occur over the whole time period of almost two years (2011 and 2012). Not only ballistic waves could be identified but also continuous signals in the coda of the cross-correlograms, which most likely correspond to scattered and reflected waves. In this context, I also detected weekly variations in the signal amplitudes in the cross-correlations. These variations are related to human activities (more quiet days on weekends and during Easter 2012, for example). Thus, further analysis of the temporal variations of the signals might provide more detailed information on the noise sources. For example, day-/nighttime variations – also indicating man-made noise sources – might be detectable as well.

As the TIMO2 data set meets the requirements for a possible detection of temporal variations of the seismic velocities, a first monitoring test was carried out with stationpair TMO20-TMO61 using the stretching technique. With the stretching technique the relative delay time between two coda phases (of different time periods) is determined as the factor by which the time axis of the one waveform has to be stretched or compressed to obtain the highest correlation with the other waveform (Sens-Schönfelder and Wegler, 2006). It turned out that the results for the delay times react very sensitively to changes of the parameters, like, amongst others, the time window length in which the codas of the two cross-correlograms are compared to each other, or the choice of the reference trace(s) to which the other traces should be compared. Due to the huge number of possible parameter settings, a statistical approach might be the only practicable way to obtain reliable results.

As already mentioned above, an asymmetry relative to timelag zero is observed in the cross-correlograms filtered between 0.1 Hertz and 0.8 Hertz. A planar wave incidence from the north-west of the TIMO2-network explains the signals in the cross-correlograms in the frequency range from 0.1 Hertz to 0.4 Hertz (Section 5.2.1). Regarding the bandwidth of the

signals and the direction of the incoming wavefront, the sources are most likely the ocean-generated microseisms originating from the Atlantic Ocean and the North Sea around the British Isles including the British Channel. In the frequency band from 0.4 Hertz to 0.8 Hertz the ocean-generated microseisms still influence the cross-correlograms but also another source very close to the western part of the network might cause the asymmetric signals. By evaluating the data of the two additionally set up stations TMO66 and TMO67 (map Section 3.2) this open question might be answered.

In the frequency range between 0.8 Hertz and 1.6 Hertz a migration analysis was also calculated (Section 5.2.2). The highest semblance value is obtained in a small region in the south-east of the station network around Landau. This area coincides with the windpark of Offenbach an der Queich consisting of three wind turbines. Future work should investigate, if the windpark is the actual source of the dominant signals in the cross-correlations. A temporal relation between the occurrence of the signals in the cross-correlograms and the wind strengths varying over time might be observable.

Finally, future work might establish a connection between activities at the geothermal reservoirs and temporal variations of the seismic velocities. This will be possible if the waves occurring in the coda of the cross-correlograms reach into depths where the geothermal reservoirs are located. It might also be possible to extract ballistic body waves out of the TIMO2-data set by using passive seismic interferometry. Body waves are not confined to the shallow depth like the surface waves. Therefore, using body waves imaging and monitoring in larger depths might be possible around Landau.

Bibliography

Keiiti Aki. Space and time spectra of stationary stochastic waves, with special reference to microtremors. *Bull. Earthquake Res. Inst. Tokyo Univ.*, 35:415–457, 1957.

Keiiti Aki and Bernard Chouet. Origin of coda waves: source, attenuation, and scattering effects. *Journal of Geophysical Research*, 80(23):3322–3342, 1975.

Keiiti Aki and Paul G. Richards. *Quantitative seismology*. Univ. Science Books, Sausalito, Calif., 2. ed., corr. print. edition, 2009. ISBN 978-1-891389-63-4.

M. Behm and R. Snieder. Love waves from local traffic noise interferometry. *The Leading Edge*, 32(6):628–632, 2013.

Julius S. Bendat and Allan G. Piersol. *Random data: analysis and measurement procedures*. A Wiley-interscience publication. Wiley, New York [u.a.], 2. ed. (rev. and expanded) edition, 1986. ISBN 0-471-04000-2.

G. D. Bensen, M. H. Ritzwoller, M. P. Barmin, A. L. Levshin, F. Lin, M. P. Moschetti, N. M. Shapiro, and Y. Yang. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical Journal International*, 169(3):1239–1260, 2007.

Sylvette Bonnefoy-Claudet, Fabrice Cotton, and Pierre-Yves Bard. The nature of noise wavefield and its applications for site effects studies: A literature review. *Earth-Science Reviews*, 79(3):205–227, 2006.

Florent Brenguier, Nikolai M. Shapiro, Michel Campillo, Alexandre Nercessian, and Valérie Ferrazzini. 3-D surface wave tomography of the Piton de la Fournaise volcano using seismic noise correlations. *Geophysical research letters*, 34(L02305): doi:10.1029/2006GL028586, 2007.

Florent Brenguier, Nikolai M. Shapiro, Michel Campillo, Valérie Ferrazzini, Zacharie Duputel, Olivier Coutant, and Alexandre Nercessian. Towards forecasting volcanic eruptions using seismic noise. *Nature Geoscience*, 1(2):126–130, 2008.

Michel Campillo and Anne Paul. Long-range correlations in the diffuse seismic coda. *Science*, 299(5606):547–549, 2003.

Andrew Curtis, Peter Gerstoft, Haruo Sato, Roel Snieder, and Kees Wapenaar. Seismic interferometry - turning noise into signal. *The Leading Edge*, 25(9):1082–1092, 2006.

Arnaud Derode, Eric Larose, Michel Campillo, and Mathias Fink. How to estimate the Green's function of a heterogeneous medium between two passive sensors? Application to acoustic waves. *Applied Physics Letters*, 83(15):3054–3056, 2003.

A. Friedrich, F. Krüger, and K. Klinge. Ocean-generated microseismic noise located with the Gräfenberg array. *Journal of Seismology*, 2(1):47–64, 1998.

J. C. Groos. *Broadband Seismic Noise: Classification and Green's Function Estimation.* PhD thesis, Karlsruher Institut für Technologie, Fakultät für Physik, 2010.

J. C. Groos and J. R. R. Ritter. Time domain classification and quantification of seismic noise in an urban environment. *Geophysical Journal International*, 179(2):1213–1231, 2009.

J. C. Groos, S. Bussat, and J. R. R. Ritter. Performance of different processing schemes in seismic noise cross-correlations. *Geophysical Journal International*, 188(2):498–512, 2012.

J. C. Groos, R. Fritschen, and J. R. R. Ritter. Untersuchung induzierter Erdbeben hinsichtlich ihrer Spürbarkeit und eventueller Schadenswirkung anhand der DIN 4150. *Bauingenieur*, 9:in press, 2013.

Céline Hadziioannou, Eric Larose, Olivier Coutant, Philippe Roux, and Michel Campillo. Stability of monitoring weak changes in multiply scattering media with ambient noise correlation: Laboratory experiments. *The Journal of the Acoustical Society of America*, 125:3688, 2009.

M. Hobiger, U. Wegler, K. Shiomi, and H. Nakahara. Coseismic and postseismic elastic wave velocity variations caused by the 2008 Iwate-Miyagi Nairiku earthquake, Japan. *Journal of Geophysical Research: Solid Earth (1978–2012)*, 117(B09313): doi:10.1029/2012JB009402, 2012.

Tobias Horstmann. Interferometrische Analyse seismischer Hintergrundsignale. Master's thesis, Karlsruher Institut für Technologie, Fakultät für Physik, 2010.

Jennifer A. Jay, Matthew E. Pritchard, Michael E. West, Douglas Christensen, Matthew Haney, Estela Minaya, Mayel Sunagua, Stephen R. McNutt, and Mario Zabala. Shallow seismicity, triggered seismicity, and ambient noise tomography at the long-dormant Uturuncu Volcano, Bolivia. *Bulletin of volcanology*, 74(4):817–837, 2012.

Andreas Köhler, Matthias Ohrnberger, Frank Scherbaum, Marc Wathelet, and Cecile Cornou. Assessing the reliability of the modified three-component spatial autocorrelation technique. *Geophysical Journal International*, 168(2):779–796, 2007.

Werner Lauterborn, Thomas Kurz, and Martin Wiesenfeldt. *Kohärente Optik: Grundlagen für Physiker und Ingenieure; mit 73 Aufgaben und vollständigen Lösungen.* Springer, Berlin, 1993. ISBN 3-540-56769-0.

Thorne Lay and Terry C. Wallace. *Modern global seismology.* International geophysics series ; 58. Academic Pr., San Diego [u.a.], 1995. ISBN 0-12-732870-X.

F. C. Lin, M. P. Moschetti, and M. H. Ritzwoller. Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps. *Geophysical Journal International*, 173(1):281–298, 2008.

Yiran Ma, Robert W. Clayton, Victor C. Tsai, and Zhongwen Zhan. Locating a scatterer in the active volcanic area of Southern Peru from ambient noise cross-correlation. *Geophysical Journal International*, 192(3):1332–1341, 2013.

Guénolé Mainsant, Eric Larose, Cornelia Brönnimann, Denis Jongmans, Clément Michoud, and Michel Jaboyedoff. Ambient seismic noise monitoring of a clay landslide: Toward failure prediction. *Journal of Geophysical Research: Earth Surface (2003–2012)*, 117 (F01030):doi:10.1029/2011JF002159, 2012.

Gerhard Müller. *Theorie elastischer Wellen*. Universität Frankfurt, Institut für Meteorologie und Geophysik, 1973.

M. Picozzi, S. Parolai, D. Bindi, and A. Strollo. Characterization of shallow geology by high-frequency seismic noise tomography. *Geophysical Journal International*, 176(1): 164–174, 2009.

Katrin Plenkers, Joachim R. R. Ritter, and Marion Schindler. Low signal-to-noise event detection based on waveform stacking and cross-correlation: application to a stimulation experiment. *Journal of seismology*, 17(1):27–49, 2013.

A. Plesinger, M. Hellweg, and D. Seidl. Interactive high-resolution polarization analysis of broad-band seismograms. *J. Geophys*, 59:129–139, 1986.

P. Poli, H. A. Pedersen, M. Campillo, et al. Noise directivity and group velocity tomography in a region with small velocity contrasts: the northern Baltic shield. *Geophysical Journal International*, 192(1):413–424, 2013.

Philippe Roux and Mathias Fink. Green's function estimation using secondary sources in a shallow water environment. *The Journal of the Acoustical Society of America*, 113: 1406, 2003.

Christoph Sens-Schönfelder. Synchronizing seismic networks with ambient noise. *Geophysical Journal International*, 174(3):966–970, 2008.

Christoph Sens-Schönfelder and U. Wegler. Passive image interferometry and seasonal variations of seismic velocities at Merapi Volcano, Indonesia. *Geophysical Research Letters*, 33(L21302):doi:10.1029/2006GL027797, 2006.

Roel Snieder. Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase. *Physical Review E*, 69(4):046610, 2004.

Roel Snieder and Kees Wapenaar. Imaging with ambient noise. *Physics Today*, 63:44, 2010.

Peter Styles, I. Stimpson, S. Toon, R. England, and M. Wright. Microseismic and infrasound monitoring of low frequency noise and vibrations from windfarms. *Recommendations on the Siting of Windfarms in the Vicinity of Eskdalemuir, Scotland, Report to MOD/FTI/BWEA*, Keele University:125pp, 2005.

Kees Wapenaar. Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation. *Physical Review Letters*, 93(25):254301, 2004.

Kees Wapenaar, Jacob Fokkema, and Roel Snieder. Retrieving the Green's function in an open system by cross correlation: A comparison of approaches (l). *The Journal of the Acoustical Society of America*, 118:2783, 2005.

Kees Wapenaar, Deyan Draganov, and Johan Robertsson. Introduction to the supplement on seismic interferometry. *Geophysics*, 71(4):SI1–SI4, 2006.

Kees Wapenaar, Deyan Draganov, Roel Snieder, Xander Campman, and Arie Verdel. Tutorial on seismic interferometry: Part 1 - Basic principles and applications. *Geophysics*, 75(5):75A195–75A209, 2010.

Richard Weaver and Oleg Lobkis. On the emergence of the Green's function in the correlations of a diffuse field: pulse-echo using thermal phonons. *Ultrasonics*, 40(1):435–439, 2002.

Richard L. Weaver and Oleg I. Lobkis. Ultrasonics without a source: Thermal fluctuation correlations at MHz frequencies. *Physical Review Letters*, 87(13):134301, 2001.

David Wilson, Joseph Leon, Richard Aster, James Ni, John Schlue, Steve Grand, Steve Semken, Scott Baldridge, and Wei Gao. Broadband seismic background noise at temporary seismic stations observed on a regional scale in the southwestern United States. *Bulletin of the Seismological Society of America*, 92(8):3335–3342, 2002.

Appendix

A. The Cross-Correlation Function

The analytical cross-correlation function (CCF) of two time series $A(t)$ and $B(t)$ is defined as:

$$\widehat{R}_{AB}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{+T/2} A(t)B(t+\tau)dt. \tag{6.1}$$

with the time window length T and the lag time τ, by which $B(t)$ is shifted against $A(t)$. A second definition of the analytical CCF, which is also used, is:

$$\widehat{R}_{AB}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{+T/2} A(t+\tau)B(t)dt. \tag{6.2}$$

In this case, the time series $A(t)$ is shifted against $B(t)$. As the MATLAB function xcorr is based on the second definition, the following text refers to this formula only.

Due to the digital seismic data acquisition there are no continuous recordings $A(t)$ and $B(t)$ but sampled digital time series. Assuming two digital time series of finite duration with N elements A_n and B_n, which are sampled at equally spaced time intervals Δt, the digital cross-correlation can be expressed by:

$$\widehat{R}_{AB}(r) = \frac{1}{N - |r|} \sum_{n=1}^{N-r} A_{n+r}B_n \tag{6.3}$$

with the discrete timelag $r = 0, 1, 2, ..., m$, with the maximum timelag $m < N$, and with $\tau = r \cdot \Delta t$. Due to the finite length of the time series missing values are replaced by zeros. On account of this replacement by zeros the CCF, which is calculated using Formula 6.3, is also called linear digital CCF. By dividing the sum in Formula 6.3 by $\frac{1}{N-|r|}$ and not only by $\frac{1}{N}$ one obtains an 'unbiased' instead of a 'biased' CCF. Hence, with the calculation of the 'unbiased' CCF it is considered that the number of summands equal to zero increases with increasing r.

The above given definitions apply to the calculation of the CCF in the time domain. In MATLAB, however, the linear CCF is not calculated in the time domain but in the frequency domain using a Fast Fourier Transform approach. For further information please refer to Bendat and Piersol (1986).

B. MATLAB Functions

The most important MATLAB functions, which were used in the scope of this thesis, are listed and briefly explained in the following. The functions denoted by a star (*) are part of the Karlsruhe Seismology Processing (KaSP) toolbox. The functions denoted by two stars (**) were changed/improved or developed by myself.

B.1. Functions Used for the Data Processing

KABBA_TMO_NCC_preprocessing_station_reduced.m*
Raw data are preprocessed and saved by this function.

XCORRprocessingSLWDW_NOfilter.m, cc_SLWDW_preparation.m**, and CCF_new.m****
Data are prepared by the two functions CCF_new.m and cc_SLWDW_preparation.m for function XCORRprocessingSLWDW_NOfilter.m which implements the cross-correlation of the time series fragments without applying a filter or normalisation methods to the cross-correlation functions.

XCORR_rotate_horizontal_components.m, data_preparation_for_XCORR_rotate_horizontal_components.m**, rotated_CCF_calculation.m****
The horizontal component cross-correlation functions are rotated by XCORR_rotate_horizontal_components.m. The two functions data_preparation_for_XCORR_rotate_horizontal_components.m and rotated_CCF_calculation.m prepare the horizontal component cross-correlation functions for the rotation.

XCORRspectralwhitening_withlimitsinput.m**
The normalisation in the frequency domain is implemented by this function. Limits for the range where the amplitudes of the complex spectrum are set to one must be given.

XCORRtimedomainnormalisation.m*
A waveform preserving normalisation of the cross-correlation functions in the time domain was tested with this function.

XCORRstack.m*
This function stacks the given number of cross-correlations.

XCORRfilter.m*
High-pass, band-pass, and low-pass filters can be applied to the cross-correlation functions using XCORRfilter.m.

XCORRgetFFT.m*
This function calculates the fast fourier transform of the cross-correlation function.

B.2. Functions Used for the Analysis of the Data

XCORR_distanceplot.m*, XCORR_distanceplot_section.m**, and XCORR_distanceplot_stacked_CCFs.m**
The function XCORR_distanceplot.m plots the cross-correlograms against their respective interstation distances. The function XCORR_distanceplot_section.m plots only those cross-correlograms within a given interstation distance interval. XCORR_distanceplot_stacked_CCFs.m prepares the data for the two former functions.

XCORRgetcausalpart.m*
This function time-flips the acausal part of a cross-correlation function.

XCORR_plot_polarisation.m**
This function plots the polarisation diagrams of waveform segments of a cross-correlogram.

colorplot.m**
This function colours the amplitudes of the cross-correlograms and plots them chronologically sorted in a matrix line by line.

XCORR_distanceplot_effective_incidentangle.m*, and XCORR_distanceplot_effective_incidentangle_stacked_CCFs.m**
The function XCORR_distanceplot_effective_incidentangle.m plots the cross-correlograms against their effective distances assuming a planar wavefront propagating across the station network. XCORR_distanceplot_effective_incidentangle_stacked_CCFs.m prepares the data for the former function.

Migrationsanalyse.m*, zusatz_migrationsanalyse_filter.m**
zusatz_migrationsanalyse_filter.m prepares the data for the function Migrationsanalyse.m. The latter function can be used for the localisation of a point source inside of a station network.

XCORR_distanceplot_effective_pointsource.m*, XCORR_distanceplot_effective_pointsource_section.m**, and workspace_saving_calc_distplot_pointsource.m**
These functions are used for the plot of the cross-correlograms against their effective distances if a point source within the station network is assumed.

shorttermspectrogram.m*
This function calculates the spectrogram of a short time series.

test_stretching.m**, TEST_XCORRstretching.m**
These functions calculate the stretching coefficient by which the first of two synthetic sine functions has to be stretched or compressed to obtain the second sine function.

timedependence_stretchcoef.m**, XCORRstretching.m**
With these two functions the stretching technique can be applied to the cross-correlations of the TIMO2 data set. At the end, the stretching coefficients including error bars are plotted against the time.

C. Beats in the Cross-Correlations

This appendix gives an idea of how 'beats' can emerge in cross-correlations. Therefore, using two sine functions four different synthetic examples are analysed (see below). Moreover, two spectrograms of the ambient noise recorded by two TIMO2 stations are investigated with respect to periodic signals, which might cause the 'beats' in the real data.

(1) Two phase-shifted sine functions of the same frequency are cross-correlated.

(2) Two phase-shifted sine functions with slightly different frequencies (2.5 Hertz and 3 Hertz) are cross-correlated.

(3) A beat – as it is usually defined – is generated by summing a sine function with a frequency of 2.5 Hertz and a sine function with a slightly higher frequency of 3 Hertz. The obtained signal is cross-correlated with a sine function of a frequency of 2.5 Hertz.

(4) Two phase-shifted beats are cross-correlated with each other.

Case (1) is depicted in Figure C.1. The cross-correlation of two phase-shifted sine functions of the same frequency (here: 2.5 Hertz) yields a sine function of a frequency of 2.5 Hertz.
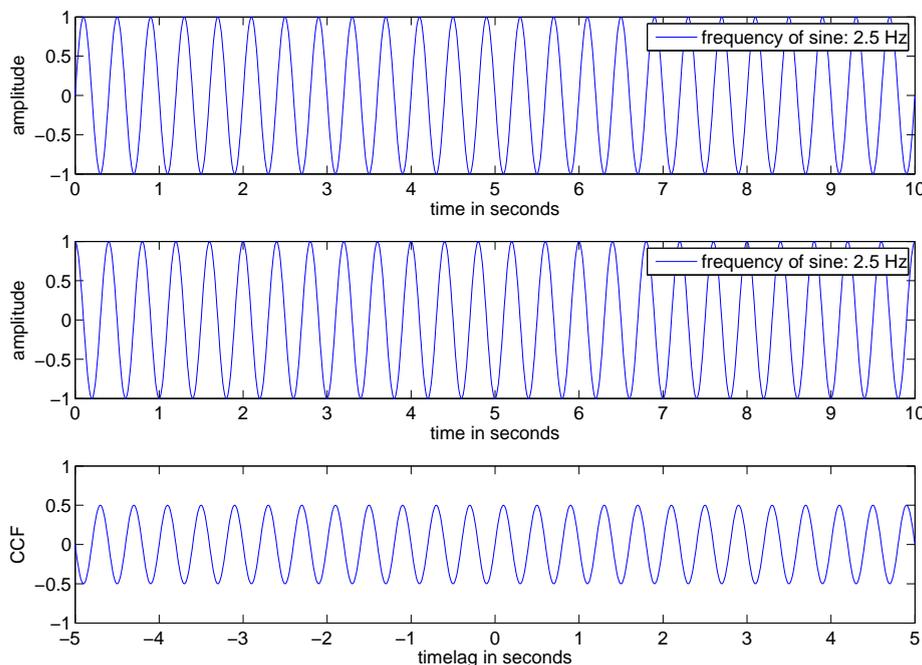

Figure C.1.: Top: Sine function of a frequency of 2.5 Hertz. Middle: Phase-shifted (compared to above) sine function of a frequency of 2.5 Hertz. Bottom: Cross-correlogram of the two waves shown above. A sine function of the same frequency as the waves above is obtained.

In Figure C.2 Case (2) is shown. Two sine functions of slightly different frequencies are cross-correlated. The result are 'beats'. At the depicted lag times of -3 seconds, -1 second, 1 second, and 3 seconds the similarity between the two initial signals is the highest.
Case (3) is illustrated in Figure C.3. A sine function of a frequency of 2.5 Hertz is cross-correlated with a common beat (the sum of two sine functions with slightly different frequencies). The cross-correlogram yields a sine-function of a frequency of 2.5 Hertz.
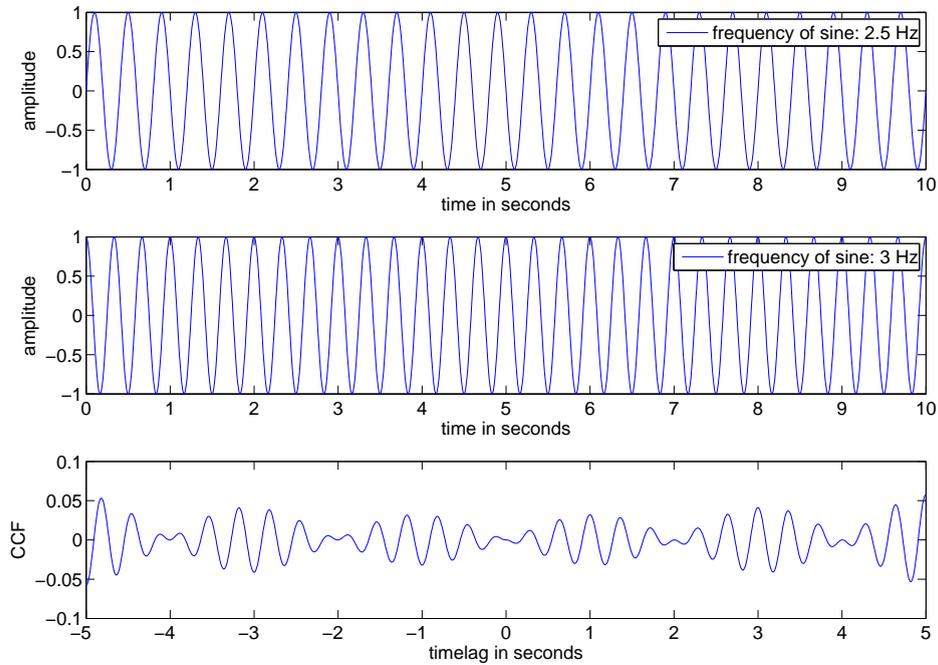
Figure C.2.: Top: Sine function of a frequency of 2.5 Hertz. Middle: Phase-shifted (compared to above) sine function of a frequency of 3 Hertz. Bottom: The cross-correlogram of the two functions shown above. Periodically occurring 'signal packets' are observed.
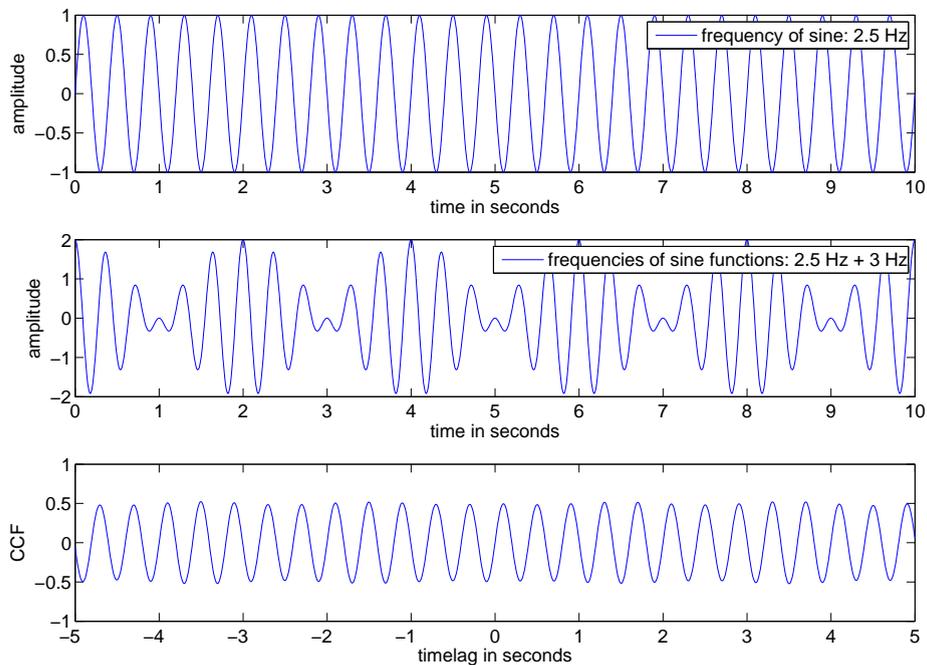

Figure C.3.: Top: Sine function of a frequency of 2.5 Hertz. Middle: Signal calculated by the sum of two sine functions with frequencies of 2.5 Hertz and 3 Hertz. Bottom: Cross-correlation of the signals shown above. A sine function of a frequency of 2.5 Hertz is obtained.
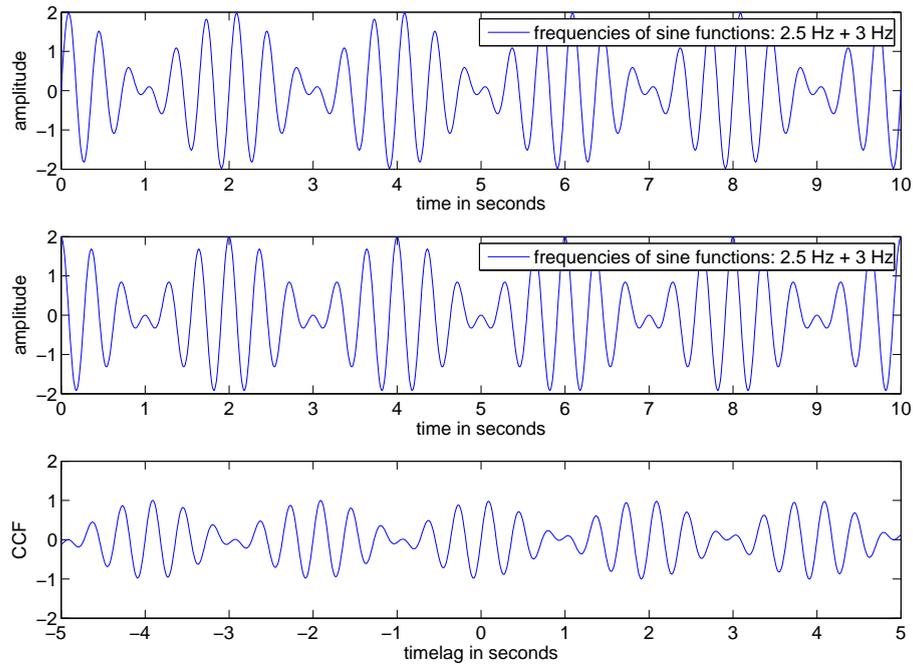
Figure C.4.: Top: The sum of two sine functions with frequencies of 2.5 Hertz and 3 Hertz. Middle: The same signal as above but phase-shifted. Bottom: The cross-correlation of the the signals shown above. 'Beats' are obtained.
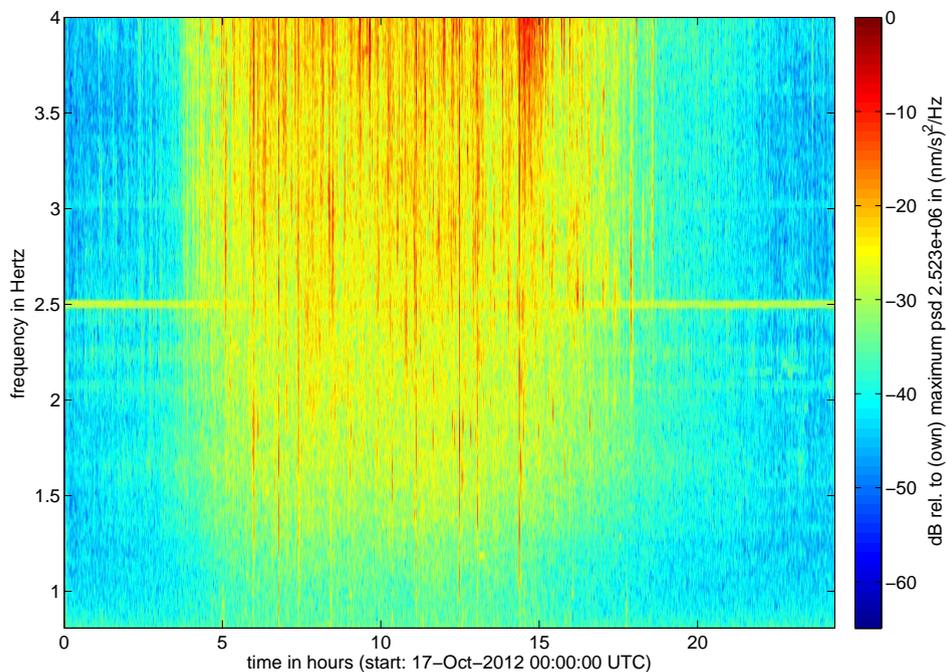

Figure C.5.: The spectrogram of the vertical component seismic noise of station TMO20 during one day in October 2012. The frequencies between 0.8 Hertz and 4 Hertz are depicted. Signals with distinct frequencies occurring over the whole day correspond to horizontal lines in the spectrogram. At 2.5 Hertz narrow-band signals are observed over the whole day.

The last case, Case (4), is depicted in Figure C.4. Two phase-shifted signals consisting of common beats (sum of the sine functions with frequencies of 2.5 Hertz and 3 Hertz) are cross-correlated. The result yields beat-like 'signals' of the same frequency as the beat frequencies of the initial signals.

Summarising the observations above, 'beats' occur in cross-correlograms either by cross-correlating two sine functions of slightly different frequencies or by cross-correlating two signals which already consist of beats (as commonly defined). A mathematical relation between the frequencies of the 'beats' in the cross-correlations and the frequencies of the initial signals should be established.

Figures C.5 and C.6 show the spectrograms of the vertical ambient noise of one day recorded by two stations of the TIMO2-network. In both spectrograms horizontal lines of increased power spectral density (psd) are observed over the whole time period at distinct frequencies. These signals correspond most likely to sinusoidal-type seismic waves originating from rotating machinery (Groos and Ritter, 2009).
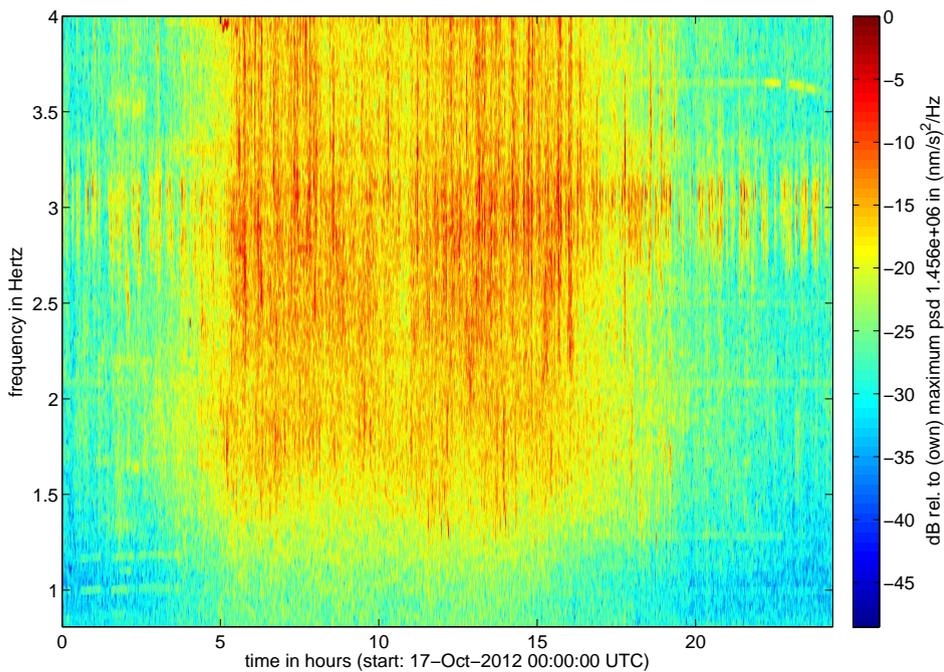

Figure C.6.: The spectrogram of the vertical component seismic noise of station TMO58 during one day in October 2012. The frequencies between 0.8 Hertz and 4 Hertz are depicted. High-amplitude transient signals occur especially at daytime. Due to these transients periodic signals at distinct frequencies are not as clearly observed as in Figure C.5. Nonetheless, at 2.1 Hertz and 2.5 Hertz, for example, horizontal lines are slightly visible.

How these periodic signals contribute to the 'beats' in the cross-correlations filtered between 1.6 Hertz and 3.2 Hertz (see the example in Section 5.1.1 depicted in Figure 5.15) is not fully understood yet.

D. Determination of the Signal to Noise Ratio

Table D.1.: The signal to noise ratio (SNR) is determined in the cross-correlograms (year 2012, filtered from 0.8 - 1.6 Hertz) of all three component combinations (comp. combi.) of the three stationpairs TMO53-TMO54, TMO20-TMO61, and TMO22-TMO57. The highest amplitude value in an individually chosen signal window is divided by the highest amplitude value in a noise window. The noise window trails the end of the signal window by about 10 seconds. The time window (TW) limits in the causal (c) and acausal (ac) parts are chosen symmetrically respective to lag time zero.

	Comp. Combi.	Part of the CCF	Peaks in Signal Window	Peaks in Noise Window	SNR (Rounded)
TMO53-TMO54 TW Limits			7.5 s - 10 s	20 s - 25 s	
	ZZ	c	0.25	0.11	2
		ac	1	0.08	13
	RR	c	0.59	0.22	3
		ac	1	0.21	5
	TT	c	0.69	0.06	12
		ac	1	0.02	61
TMO20-TMO61 TW Limits			13 s - 17 s	25 s - 30 s	
	ZZ	c	1	0.08	13
		ac	0.38	0.43	1
	RR	c	0.92	0.25	4
		ac	0.40	0.25	2
	TT	c	1	0.12	8
		ac	0.46	0.07	6
TMO22-TMO57 TW Limits			22 s - 26 s	35 s - 40 s	
	ZZ	c	0.74	0.41	2
		ac	1	0.27	4
	RR	c	0.38	0.12	3
		ac	0.22	0.12	2
	TT	c	1	0.16	6
		ac	0.52	0.17	3

E. Effective Distance Plots - Error Estimate

The following five figures (Figures E.7, E.8, E.9, E.10, and E.11) show the effective distance plots with the ZZ cross-correlograms of the year 2012 assuming a planar wave incidence from outside of the TIMO2-network. The CCFs are filtered between 0.2 Hertz and 0.4 Hertz. The incident angle of the wavefront is varied by increments of $10°$ from $280°$ to $320°$. A distinct linear moveout of the signals in the cross-correlograms is observed assuming an incident angle of $300°$ (Figure E.9). However, under the assumption of the two incident angles of $280°$ and $310°$ a moveout almost as clear as in the effective distance plot in Figure E.9 is obtained. In Figures E.7 and E.11, the obtained results do not show such a distinct moveout as the other three effective distance plots. Hence, the uncertainty on the incident angle is estimated to $±15°$.
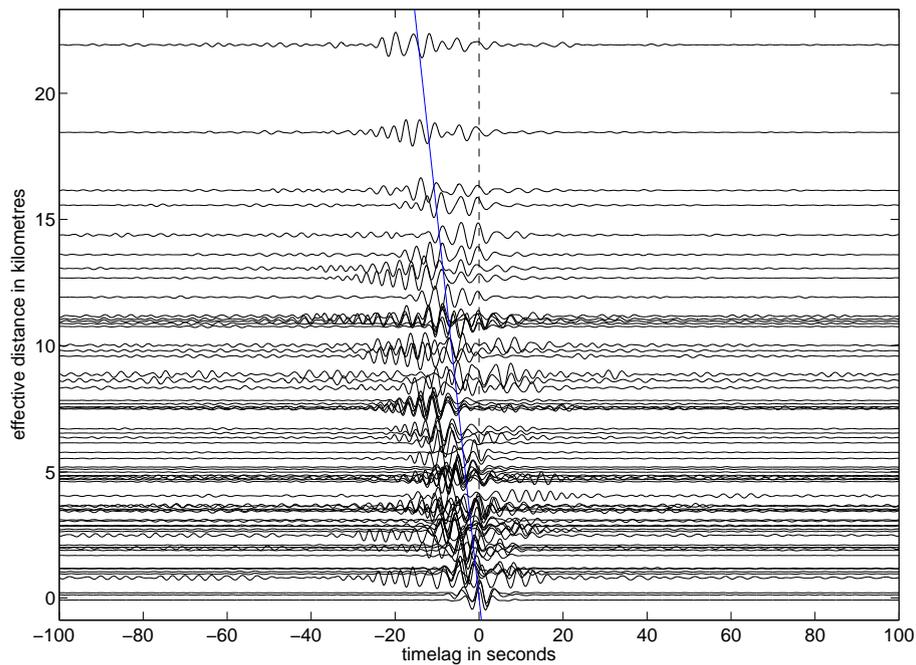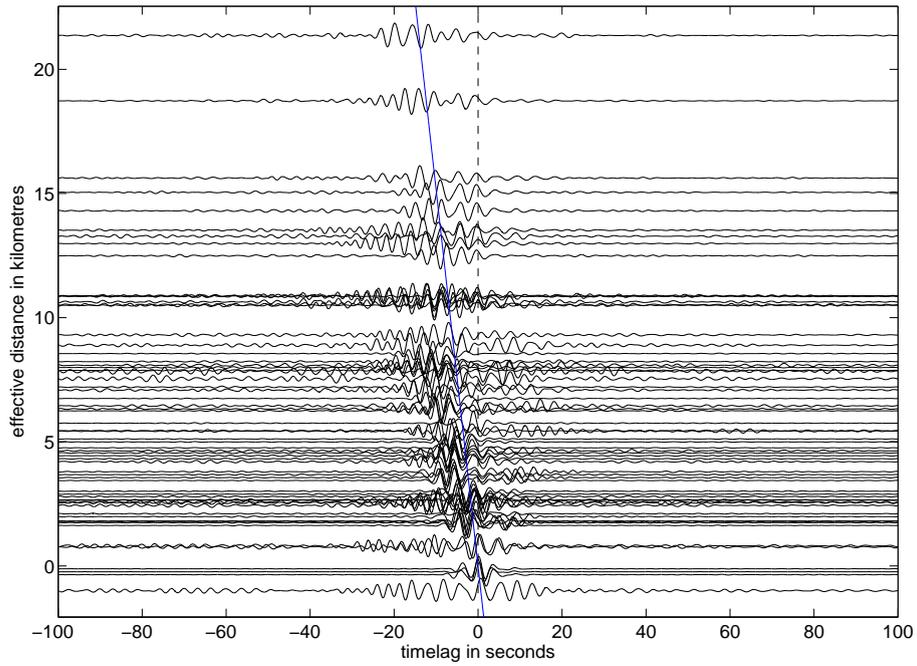

Figure E.7.: The ZZ cross-correlograms of the year 2012 filtered between 0.2 Hertz and 0.4 Hertz are plotted against their effective distances assuming a planar wave incidence with an angle of $280°$. The blue line marks a velocity of 1500 metres per second.
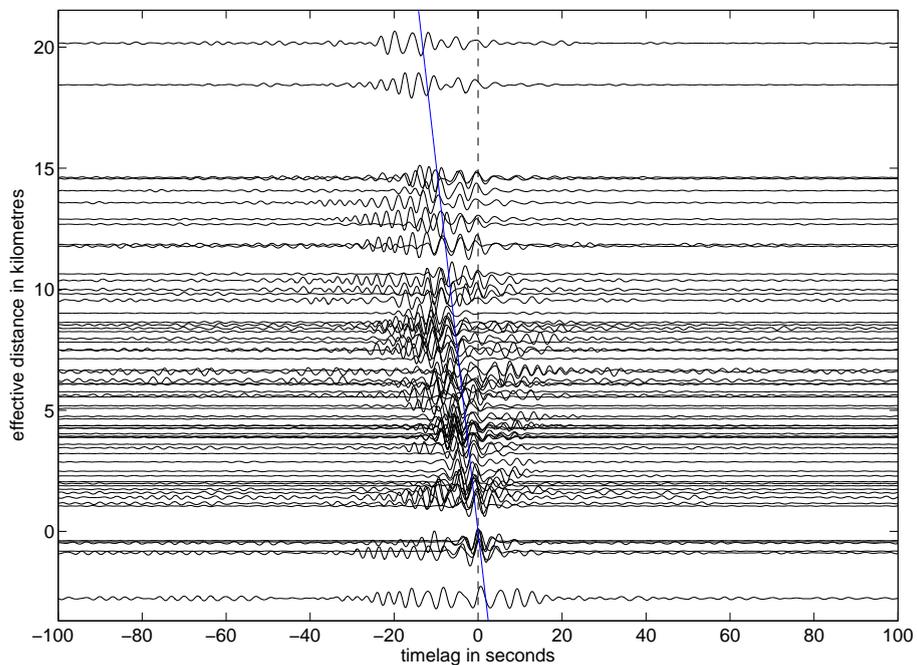
Figure E.8.: The ZZ cross-correlograms of the year 2012 filtered between 0.2 Hertz and 0.4 Hertz are plotted against their effective distances assuming a planar wave incidence with an angle of $290°$. The blue line marks a velocity of 1500 metres per second.
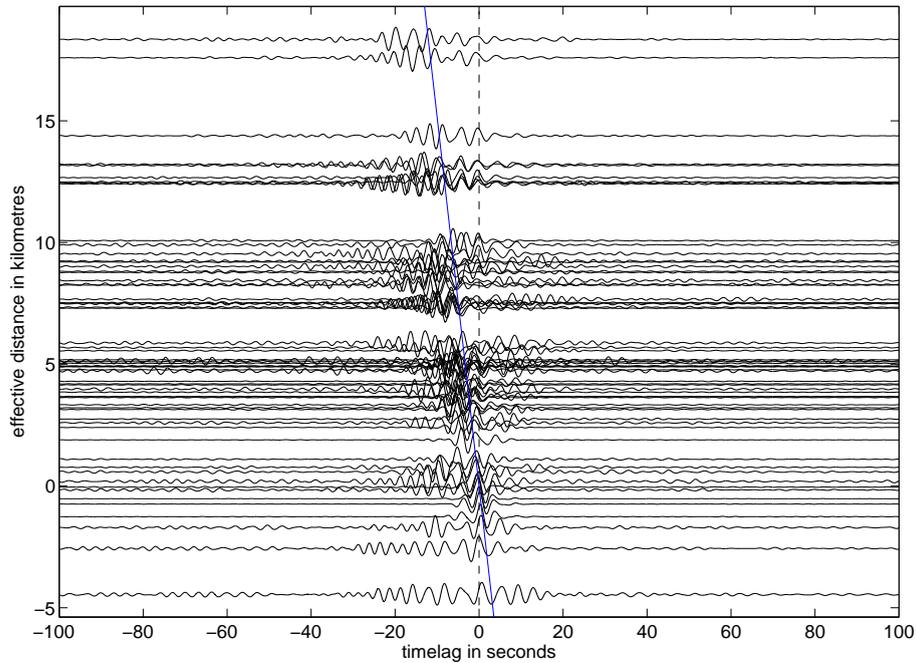

Figure E.9.: The ZZ cross-correlograms of the year 2012 filtered between 0.2 Hertz and 0.4 Hertz are plotted against their effective distances assuming a planar wave incidence with an angle of $300°$. The blue line marks a velocity of 1500 metres per second.

Figure E.10.: The *ZZ* cross-correlograms of the year 2012 filtered between 0.2 Hertz and 0.4 Hertz are plotted against their effective distances assuming a planar wave incidence with an angle of $310°$. The blue line marks a velocity of 1500 metres per second.
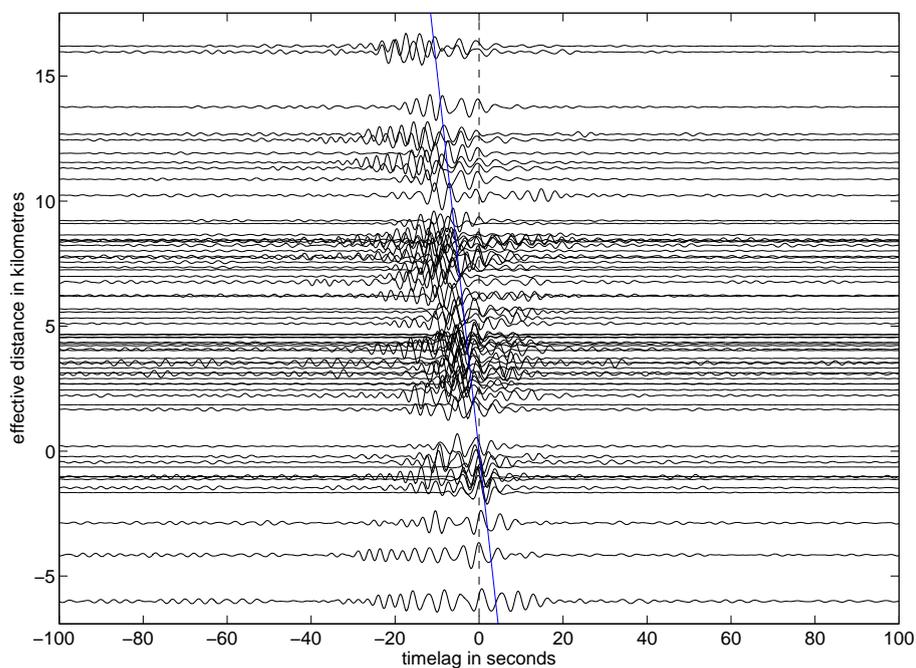

Figure E.11.: The *ZZ* cross-correlograms of the year 2012 filtered between 0.2 Hertz and 0.4 Hertz are plotted against their effective distances assuming a planar wave incidence with an angle of $320°$. The blue line marks a velocity of 1500 metres per second.