Dissertation

Magnetotelluric measurements across the southern Barberton Greenstone Belt, South Africa

vorgelegt von
Diplom-Geophysikerin

Sissy Kütter

zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
Dr. rer. nat.

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Potsdam

Potsdam, den 30. März 2015

Betreuerin: Dr. Ute Weckmann, GFZ Potsdam

Gutachter:
1. Gutachter: Prof. Dr. Michael Weber, GFZ Potsdam
2. Gutachter: Prof. Dr. Andreas Junge, Goethe-Universität Frankfurt am Main
3. Gutachter: Prof. Dr. Andreas Hördt, TU Braunschweig

Published online at the
Institutional Repository of the University of Potsdam:
URN urn:nbn:de:kobv:517-opus4-83198
http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-83198

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich meine Dissertation selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, alle Ausführungen, die anderen Schriften wörtlich oder sinngemäß entnommen wurden, kenntlich gemacht sind und die Arbeit in gleicher oder ähnlicher Fassung noch nicht Bestandteil einer Studien- oder Prüfungsleistung war.

Potsdam, den 30. März 2015 _____

Summary

The Barberton Greenstone Belt (BGB) in the northwestern part of South Africa belongs to the few well-preserved remnants of Archean crust. Over the last centuries, the BGB has been intensively studied at surface with detailed mapping of its surfacial geological units and tectonic features. Nevertheless, the deeper structure of the BGB remains poorly understood. Various tectonic evolution models have been developed based on geo-chronological and structural data. These theories are highly controversial and centre on the question whether plate tectonics - as geoscientists understand them today - was already evolving on the Early Earth or whether vertical mass movements driven by the higher temperature of the Earth in Archean times governed continent development.

To get a step closer to answering the questions regarding the internal structure and formation of the BGB, magnetotelluric (MT) field experiments were conducted as part of the German-South African research initiative Inkaba yeAfrica. Five-component MT data (three magnetic and two electric channels) were collected at ~200 sites aligned along six profiles crossing the southern part of the BGB. Tectonic features like (fossil) faults and shear zones are often mineralized and therefore can have high electrical conductivities. Hence, by obtaining an image of the conductivity distribution of the subsurface from MT measurements can provide useful information on tectonic processes.

Unfortunately, the BGB MT data set is heavily affected by man-made electromagnetic noise caused, e.g. by powerlines and electric fences. Aperiodic spikes in the magnetic and corresponding offsets in the electric field components impair the data quality particularly at periods >1 s which are required to image deep electrical structures. Application of common methods for noise reduction like delay filtering and remote reference processing, only worked well for periods <1 s. Within the framework of this thesis two new filtering approaches were developed to handle the severe noise in long period data and obtain reliable processing results. The first algorithm is based on the Wiener filter in combination with a spike detection algorithm. Comparison of data variances of a local site with those of a reference site allows the identification of disturbed time series windows for each recorded channel at the local site. Using the data of the reference site, a Wiener filter algorithm is applied to predict physically meaningful data to replace the disturbed windows. While spikes in the magnetic channels are easily recognized and replaced, steps in the electric channels are more difficult to detect depending on their offset. Therefore, I have implemented a novel approach based on time series differentiation, noise removal and subsequent integration to overcome this obstacle. A second filtering approach where spikes and steps in the time series are identified using a comparison of the short and long time average of the data was also implemented as part of my thesis. For this filtering approach the noise in the form of spikes and offsets in the data is treated by an interpolation of the affected data samples. The new developments resulted in a substantial data improvement and allowed to gain one to two decades of data (up to 10 or 100 s).

The re-processed MT data were used to image the electrical conductivity distribution of the BGB by 2D and 3D inversion. Inversion models are in good agreement with the surface geology delineating the highly resistive rocks of the BGB from surrounding more conductive geological units. Fault zones appear as conductive structures and can be traced to depths of 5 to 10 km. 2D models suggest a continuation of the faults further south across the boundary of the BGB. Based on the shallow tectonic structures (fault system) within the BGB compared to deeply rooted resistive batholiths in the area, tectonic models including both vertical mass transport and in parts present-day style plate tectonics seem to be most likely for the evolution of the BGB.

Zusammenfassung

Der Barberton Grünsteingürtel (BGB) in Südafrika gehört zu den wenigen Regionen mit noch gut erhaltener Archaischer Kruste. Seit Jahrhunderten wurde der BGB eingehend untersucht und seine geologischen und tektonischen Strukturen detailliert kartiert. Über die tiefere Struktur des BGB ist hingegen wenig bekannt. Zahlreiche Evolutionsmodelle, die auf Altersbestimmungsdaten und strukturellen Informationen beruhen wurden über die Jahre aufgestellt. Diese Theorien sind zumeist widersprüchlich. Sie konzentrieren sich im Wesentlichen auf die Frage, ob plattentektonische Prozesse bereits bei der Entwicklung der frühen Erde eine Rolle spielten oder ob vertikale Tektonik, angetrieben durch die im Archaikum höheren Temperaturen, die Bildung der heutigen Kontinente bestimmt hat.
Um neue Erkenntnisse über die interne Struktur und Entwicklungsgeschichte des BGB zu erhalten, wurden im Rahmen der Deutsch-Südafrikanischen Forschungsinitiative Inkaba yeAfrica magnetotellurische (MT) Messungen durchgeführt. Entlang von sechs Profilen, die den gesamten südlichen Teil des BGB's überdecken, wurden nahezu 200 MT-Stationen installiert. Tektonische Strukturen wie z. B. (fossile) Verwerfungszonen können erhöhte Leitfähigigkeiten haben, wenn sich leitfähige Mineralisationen innerhalb der Scherzonen gebildet haben. Durch die Abbildung der elektrischen Leitfähigkeitsverteilung des Untergrundes mit Hilfe von MT Messungen kann der Verlauf tektonischer Strukturen nachvollzogen werden, woraus Schlüsse über möglicherweise abgelaufene tektonische Prozesse gezogen werden können.
Der gesamte MT Datensatz weist starke Störeinflüsse durch künstliche elektromagnetische Signale auf, die bspw. von Stromleitungen und elektrischen Zäunen stammen. Insbesondere langperiodische Daten (>1 s) sind davon betroffen, die für die Auflösung tieferer Strukturen notwendig sind. Die Anwendung etablierter Ansätze wie Verschiebungsfiltern und der *Remote Reference*-Methode, führte zu Verbesserungen vorrangig für Perioden < 1 s. Der langperiodische Bereich ist durch impulsartige Störsignale in den magnetischen und dazugehörigen Stufen in den elektrischen Feldkomponenten geprägt. Im Rahmen dieser Arbeit wurde ein neuartiger Zeitbereichs-Filter entwickelt, welcher auf einer abgewandelten Form des Wiener Filters beruht und diese Art von Störsignalen aus den Daten entfernt. Durch den Vergleich der Datenvarianz einer lokalen Station mit der einer Referenzstation können gestörte Zeitsegmente identifiziert werden. Anschließend wird ein Wiener-Filter-Algorithmus angewendet, um für diese Segmente mithilfe der Referenzdaten physikalisch sinnvolle Zeitreihen zu berechnen, mit denen die Daten der lokalen Station ersetzt werden. Während impulsartige Störsignale in den magnetischen Datenkanälen relativ einfach erfasst werden können, ist die Detektion von Versätzen in den elektrischen Zeitreihen je nach Versatzhöhe problematischer. Um dieses Problem zu umgehen, habe ich einen Algorithmus entwickelt, bei dem die Zeitreihen differenziert, gefiltert und im letzten Schritt integriert werden. In einer zweiten von mir entwickelten Filtermethode werden die Störsignale durch den Vergleich des kurzzeitigen und des langzeitigen Datenmittelwerts ausfindig gemacht. Bei diesem Filter werden die Störsignale aus den Zeitreihen entfernt und durch eine lineare Interpolation ersetzt. Durch die beiden Filtermethoden wurde eine deutliche Verbesserung der Datenqualität bis zu 10 und teilweise 100 s erreicht.
Zur Interpretation des MT-Datensatzes wurden 2D und 3D Inversionen durchgeführt. Die so erhaltenen elektrischen Leitfähigkeitsmodelle zeigen eine gute Übereinstimmung mit den kartierten, geologischen Strukturen. Die Gesteine des BGB weisen in den Modellen hohe Widerstände auf und sind deutlich von leitfähigen benachbarten geologischen Strukturen abgegrenzt. Verwerfungszonen korrelieren mit leitfähigen Strukturen, die sich bis in eine Tiefe von 5 bis 10 km erstrecken. Eine Fortsetzung der Verwerfungszonen über die südliche Grenze des BGB wird in den 2D-Ergebnissen angedeutet. Insgesamt

zeigen die Inversionsmodelle, dass vermutlich sowohl plattentektonische als auch vertikaltektonische Prozesse bei der Entstehung des BGB eine wichtige Rolle spielten.

Contents

Nomenclature

1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
δ	Skin depth [m]
λ	Regularization parameter
B	Magnetic flux density [nT]
D	Electric displacement current [C/m^2]
E	Electric field [V/m]
H	Magnetic field [A/m]
J	Induced current density [A/m^2]
J	Jacobian or sensitivity matrix
Z	Impedance tensor [Ω]
C	Covariance matrix
d	Measured data vector
e	Data error vector
M	Convolution matrix
m	Model parameter vector
R	Rotation matrix
T	Magnetic transfer function
μ_0 or μ	Permeability of free space ($4\pi \cdot 10^{-7}$ H/m)
μ_r	Relative magnetic permeability of a medium
$\Omega(\mathbf{m})$	Model regularization term
ω	Angular frequency ($\omega = 2\pi f$) [s^{-1}]

Φ	Penalty function
ϕ	Phase [^{o}]
$\psi(t)$	Wiener filter function in time domain
$\psi(f)$	Wiener filter coefficients in frequency domain
ρ	Resistivity [Ω m]
ρ_a	Apparent resistivity [Ω m]
σ	Conductivity [S/m]
σ_α	α-trimmed standard deviation
τ	Regularization parameter
ε_0	Permittivity of free space ($\approx 8.8542 \cdot 10^{-12}$ F/m)
ε_r	Relative dielectric permittivity
$a(t)$	Data variance time series [ ]
d_{calc}	Replacement data vector
d_{orig}	Original data vector
d_{tr}	Training data vector
F	Forward operator
f	Frequency [Hz]
i	Imaginary unit ($i^2 = -1$)
k	Complex wave number
q	Charge density [C/m^3]
r	Variance ratio
t_{user}	User-defined multiplier
ts	Time series
BB	Broad Band
BGB	Barberton Greenstone Belt
BIF	Banded iron formation
DC	Direct current
DFG	German Research Foundation

EDL	Earth Data Logger
EM	Electromagnetic
EMERALD	Electro - Magnetic Equipment, Raw - data And Locations Database
Ex, Ey	Electric field components in x and y direction
FFT	Fast Fourier Transform
FFT^{-1}	Inverse Fast Fourier Transform
GIPP	Geophysical Instrument Pool Potsdam
GMS	Granite - Monzogranite - Syenitegranite
GPS	Global Positioning System
HF	High Frequency
Hx, Hy, Hz	Magnetic field components in x, y and z direction
IARWR	Intersite Activity Ratio Wiener-filter Replacement
loc	Local site
LP	Low Pass
LTA	Long time average
MOR	Mid-ocean ridge
MT	Magnetotelluric
NLCG	Non - linear conjugate gradient
OV	Overlap
ref	Reference site
RMS	Root mean square
RR	Remote reference
S.P.A.M.	Short - Period Automatic Magnetotelluric
S/N	Signal to noise ratio
SAMPLE project	South Atlantic Margin Processes and Links with onshore Evolution project
Sbx	Sensor box
STA	Short time average
SVD	Singular Value Decomposition

T	Period [s]
t	Time [s]
TE	Transverse electric mode or E - polarization
TF	Transfer function
TM	Transverse magnetic mode or B - polarization
TTG	Tonalite - Trondhjemite - Granodiorite
WL	Window length

Introduction

The Archean eon (4 - 2.5 Ga) marks the transition of the Earth from an hostile environment to a place where the earliest life forms could evolve. Unfortunately, only few remnants of the early crust have been preserved from the Archean until today to study the first stages of the evolution of our planet. One of these remnants is the Barberton Greenstone Belt (BGB) located at the border between South Africa and Swaziland (Fig. 1). Due to its good preservation, the BGB is a prime locality for studies on the Archean eon (e.g. Hall, 1918; Anhaeusser, 1969; Knoll and Barghoorn, 1977; de Ronde and de Wit, 1994).
The geological setting of the BGB and its neighbouring formations has been well studied and mapped at the surface. Hence, more or less detailed knowledge of the age, mineral content, structural properties, the surface traces of the different geologic units and outline of tectonic features is available. Unfortunately, this detailed knowledge is only limited to the surface and deeper structures of the BGB have been much discussed by geo-scientists and are a source for various highly controversial theories. These theories centre mainly on questions regarding tectonics and processes responsible for the BGB's formation and evolution during the Archean. Due to the lack of modern geophysical investigations in the Barberton area its present-day structure remains unclear.
The formation of the BGB spans over a time of ~ 490 Ma between 3.57 and 3.08 Ga ago (de Ronde and de Wit, 1994). Three geological groups can be identified within the BGB (from oldest to youngest): Onverwacht, Fig Tree and Moodies Group. Different authors (e.g. de Ronde and de Wit, 1994; Ward, 1995) suggest (based on structural studies) a tectonogenesis of the BGB similar to present-day style plate tectonics. This tectonic evolution did not occur continously but instead can be divided into different stages. Between 3.49 and 3.45 Ga submarine eruptions in a mid-ocean ridge-like (MOR) setting led to the formation of the Onverwacht Group, a thick (several kilometres) accumulation of ultramafic and mafic volcanic rocks. The second stage can be subdivided in a tectonic environment comparable to an intra-oceanic suprasubduction system (3.45 - 3.42 Ga) and an intra-arc and inter-arc setting. This led in the third stage to the amalgamation of the arc (3.23 - 3.08). The final stage (around 3.08 Ga) is described as a transition from a convergent setting to a transtensional regime and is marked by hydrothermal activities. The formation of the Fig Tree Group mostly falls before the beginning of the third stage (3.26 - 3.23 Ga) although some Fig Tree Group rocks even exhibit an age of 3.45 Ga. The siliciclastic rocks of the Moodies Group were deposited during the collisional period of the third stage (3.23 - 3.11 Ga). Due to the higher thermal gradient and the lack of stabilized crust during Archean times the mentioned authors describe the tectonic processes as similar to today and not as exact equivalents. The possibility of an oceanic setting during early Archean times with a later onset of shallow subduction and collision of small proto-continents and/or island arcs is also suggested by Lowe (1994), Foley et al. (2003) and Moyen et al. (2006). Some of these authors even speak of an early Wilson cycle (periodic opening and closing of ocean basins) in a smaller scale setting, again due to the lack of large, stabilized continental

crust in the Archean.

A major suture zone has been identified by de Ronde and de Wit (1994) - the Saddleback-Inyoka fault system (faults ③ and ④ in Fig. 1) - which is interpreted by the authors as the collision zone between two arc and trench blocks formed during stage two of the BGB formation (around 3.2 Ga) and is regarded as a further indication for early plate tectonics. In the parts where the Saddleback and the Inyoka fault are separated this suture zone has a width of several hundred metres to a couple of kilometres.

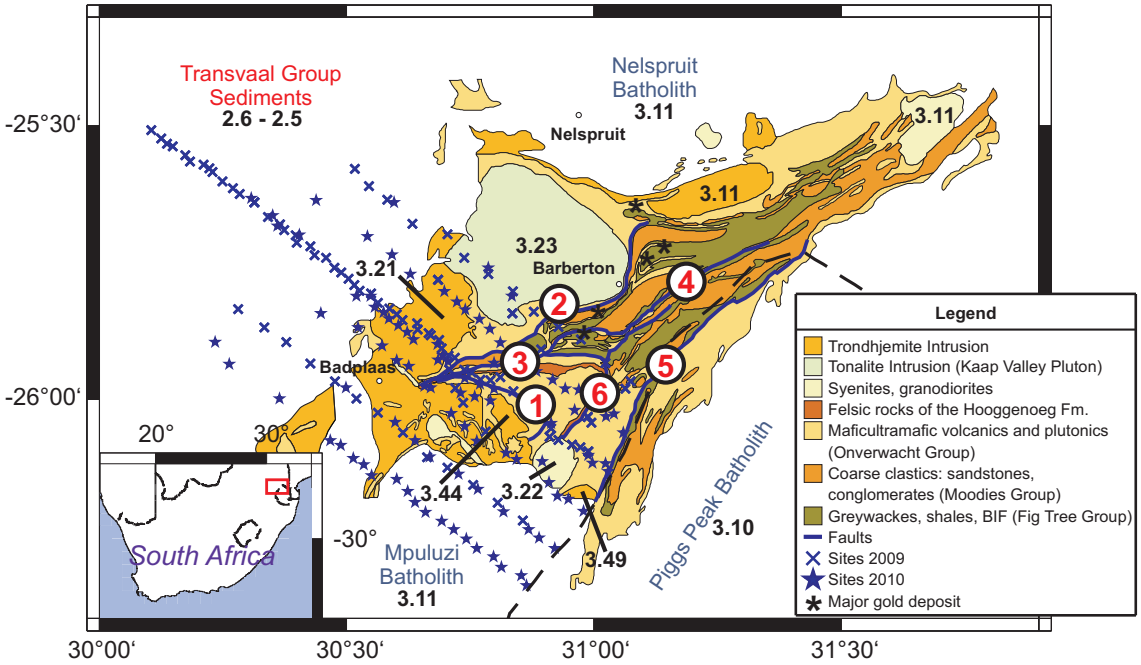


Figure 1: Geological map of the Barberton Greenstone Belt and location of the MT sites (blue markers). Faults (in blue): ① Komati, ② Moodies, ③ Inyoka, ④ Saddleback, ⑤ Maanhaar, ⑥ Kromberg (after de Ronde and de Wit, 1994 and Lowe, 1994). Black numbers in the map denote the ages (in Ga) of geological units taken from de Ronde and de Wit (1994). Inset map shows the location of the measurement area within South Africa (red rectangle, map after Schoene and Bowring, 2007 and de Ronde and de Wit, 1994).

In contrast to the theories of early plate tectonics (sometimes also referred to as horizontal tectonics) in Archean times are the theories on a crustal evolution due to vertical mass movement driven strongly by the higher temperature gradient at this time. Hamilton (1998) describes the formation of Archean crust in the following way: Originally, episodic periods of submarine volcanism led to thick layers of volcanic rocks interlayered in volcanic quiet times by sedimentary rocks. The deformation within the greenstone belts as we observe it in the BGB today marked by extensive syncline and anticline structures and fault zones were then caused by the uprising of diapiric batholiths. The batholiths were mobilized by partial melting mainly caused by the higher radiogenic heat production in Archean times. This resulted in the vertical or horizontal displacement of units of the BGB. In regions where different layers (e.g. Onverwacht and Fig Tree Group rocks) show a difference in mechanic strength this could also cause detachment and delamination (Heubeck and Lowe, 1994).

According to Hamilton (1998), early plate tectonic processes (rifting and convergence) only started around 2.0 Ga long after the formation and deformation of the BGB and was fully established as we observe plate tectonics today around 0.8 Ga.

Heubeck and Lowe (1994) states that the deformation of the BGB was not due to plate tectonics but

rather a result of crustal shortening, buckle folding and subhorizontal detachment. The Moodies Group (youngest rocks of the BGB) is here described as being deposited during an extensional period above or adjacent to thrust-faulted rocks of the Onverwacht and Fig Tree Group. Heubeck and Lowe (1994) also question the interpretation of the Saddleback-Inyoka fault system as a suture zone.
Other authors such as Moyen et al. (2006), Moyen et al. (2007) and Van Kranendonk et al. (2007a) describe the formation of Archean greenstone belts as a mixture of the two tectonic concepts with a dominance of vertical tectonics at first and a subsequent transition to horizontal tectonics with subduction and accretion processes around 3.2 - 3.1 Ga.
Along the majority of the boundary of the BGB but especially at the western and southern margins, Tonalite - Trondhjemite - Granodiorite (TTG) intrusive bodies can be found. The TTG intrusions were either the result of vertical tectonics (Hamilton, 1998) or of primitive plate tectonic processes as described by Foley et al. (2003). In general, these bodies increase in size with their decreasing age and are linked to the stages of the BGB formation (de Ronde and de Wit, 1994). Due to a higher heat flow and hence, a higher melting rate at the MORs, the oceanic crust was much thicker in the Archean than today (25 km at about 3.5 Ga) and could not be subducted as a whole (Foley et al., 2003). During this time, melts necessary for the generation of TTG complexes could only be formed in small amounts by melting processes in the upper oceanic crust after the delamination of the lowermost ultramafic layer and subsequent exposure of the upper oceanic crust to the higher temperatures of the primitive mantle of the Early Earth (Foley et al., 2003). Later on (3.3 to 3.0 Ga) due to the gradual cooling of the Earth, the oceanic crust became thinner and could be subducted as a whole. The subduction of the in parts hydrothermally altered topmost layer allowed a larger generation of TTG melts. This process could mark the onset of early plate tectonics around again 3.2 Ga.
To answer the questions regarding the onset of plate tectonics, in-depth knowledge of the subsurface structure of greenstone belts as provided by geophysical studies is necessary. Unfortunatelly, especially at the BGB hardly any geophysical measurements were carried out. Based on gravity and DC electrical data (de Beer et al., 1988) collected at the central and northern part of the BGB, rocks of the Onverwacht Group have a thickness of 5 to 6 km and are in part overlain by Fig Tree and Moodies Group sedimentary rocks which are only several hundred metres thick. The subsurface structure and extent of the extensive fault system which reflects the tectonic history of the BGB was not investigated.
As part of the presented work, two Magnetotelluric (MT) field experiments were carried out by the GFZ Potsdam in 2009 and 2010 which focus on the structure of the fault system and BGB rocks in its southern part and the possibility of a southward continuation of the faults within the geologic units adjacent to the BGB. This project is a part of the German-South African research initiative Inkaba yeAfrica and provided the possibility to set-up an array of $\sim$ 200 MT sites aligned along six profiles (Fig. 1). The collected five-component MT data (three magnetic and two electric channels) was measured in a period range between 10^{-3} and 10^{3} s.
The MT method is especially suitable to investigate an area with fossil fault zones such as the BGB as it is able to image the Earth's deep interior over the entire crust down to the upper mantle. In general, the mainly mafic to ultramafic rocks and the neighbouring batholiths provide an electrically highly resistive background. In contrast, the fossil faults, parts of the Fig Tree Group and Transvaal Group sediments to the northwest of the BGB can be expected to appear as conductive structures in conductivity images based on the MT data. In the case of fossil faults, the expected high conductivities are linked to ore mineralizations (especially gold) along the shear planes of the faults (Hall, 1918; Dirks et al., 2013). For

the probable suture zone of the Saddleback-Inyoka fault system, graphite could attribute to high conductivities as well as it is often associated in an orogenesis (e.g. Ritter et al., 2005; Weckmann et al., 2007). Rocks of the Fig Tree Group contain in part banded iron formations (BIF) and ferruginous shales (de Ronde and de Wit, 1994; Ward, 1995) which can cause a localized rather shallow increase in conductivity.

Unfortunatelly for a MT survey, the Barberton area is a highly populated region with a dense network of powerlines, electric fences of the numerous farms and game reserves, a gas pipeline which uses anti-corrosion currents, a DC railway line and mines. All these man-made facilities can be viewed as electromagnetic (EM) sources which superpose the natural EM signals and disturb the MT measurements. In case of the Barberton MT data, often the entire period range is affected and cannot be directly used for further analysis and interpretation. Since MT studies encounter more and more the difficulties of man-made EM noise present in all inhabitated areas, numerous methods for noise removal and for filtering MT data have been developed over the years (e.g. Gamble et al., (1979); Junge, 1996; Egbert (1997); Oettinger et al., 2001; Weckmann et al., 2005). One of the most established and widely used methods is remote referencing (RR, Gamble et al., (1979)) which exploits the close correlation of magnetic data over long distances ($\sim$ 100 km) by integrating noise-free magnetic data of a remote site (at least several kilometres of distance to the noisy MT site) within the data processing algorithm. Another commonly used method is the delay filter which is especially adapted to remove EM noise with a fixed frequency (e.g. from powerlines). These latter two approaches often help to clean MT data in a wide period range. In a lot of cases, the filtering or processing methods used for MT data focus on cleaning the magnetic data and neglect EM noise present in the electric channels (e.g. RR method). If the electric data is heavily interspersed with noise and contributes significantly to the distortion of the MT transfer functions, the electric channels need to be filtered as well. This can be observed in the BGB MT data set presented in this work where both, magnetic and electric, components exhibit severe EM noise. Moreover, data cleaning methods can often not be applied generally and need to be customized for the specific EM noise dominant in the investigation area. For this work, due to the severe effect of EM noise on the BGB data, the classical methods (RR and delay filter) were not sufficient to improve the data quality and advanced filtering methods need to be investigated and adapted to the Barberton data.

The MT time series are heavily interspersed with spikes in the magnetic channels and correlated steps in the electric channels. This type of noise mainly distorts long period MT data necessary for resolving deep subsurface structures. Hence, a majority of the presented work focuses on the understanding and removal of this noise type. This is further supported by synthetic noise tests with noise-free MT data.

The work by Kappler (2012) serves as a foundation for the presented filtering method since the author showed promising results in the detection and removal of spikes. The method operates in the time domain using MT data of a remote site for the spike detection. In order to avoid gaps in the time series after the removal of the spikes, data is predicted using an adapted Wiener filter. This approach is expanded within this work especially regarding the detection of steps in the electric channels.

Removing EM noise and filtering the BGB MT data is vital since only data with a good signal to noise ratio (S/N) can be used for obtaining reliable models of the conductivity distribution of the subsurface. Existing 2D (Rodi and Mackie, 2001, implemented in the software WinGlink) and 3D (Egbert and Kelbert, 2012; Meqbel, 2009; implemented in ModEM) inversion algorithms are used in this study. Due to the much smaller model size in 2D (when compared to 3D models) a finer discretization along a profile is possible allowing for the resolution of smaller scale structures such as narrow fault zones. Difficulties

may arise when the data exhibits deviations from 2D structures and violates the assumptions made for 2D studies. This is in part true for the BGB MT data. Although the vast part of the geological structures at the BGB such as the major faults have a northeast strike which is perpendicular to the MT profiles, large parts of the southern BGB are affected by NW-SE trending shear zones (de Ronde and de Wit, 1994). This complicates the 2D analysis since it violates the basic 2D assumption of two different orthogonal current systems and might result in model artefacts. Nevertheless, the 2D inversions give a basic understanding of the data set and an overview of the conductive and resistive structures of the BGB and its surrounding geological units.

3D inversions allow for random strike directions of the subsurface and an 'all in one' analysis of the BGB MT array data set. This is especially helpful to delineate geological structures between the profiles and to spatially limitate them. In 3D modelling with small cells in order to incorporate very detailed structures is not possible for a large data set due to computational limitations. Hence, a combined interpretation of 2D and 3D results considering their characteristics is necessary to get a step closer to understanding the subsurface of the BGB and the tectonic processes in the Archean.

Chapter 1

Magnetotelluric measurements across the Barberton Greenstone belt

The formation of the crust during the early stages of the Earth's evolution is a much discussed topic among geoscientists. One of the key aspects is the character of the tectonic processes which controlled the formation of the first stable crust and protocontinents. Many scientists argue that although especially the thermal framework was different than today, plate tectonics similar to what we presently observe already operated in the Archean (e.g. de Ronde and de Wit, 1994; Ward, 1995; Moyen et al., 2006). In contrast, many authors favour an evolution scheme based on vertical tectonics and mass movement driven by the higher temperature gradient in the Archean and a, hence, fast convection cycle of crustal formation and remelting processes (e.g. Hamilton, 1998; van Kranendonk et al., 2009). Unfortunately, only few places remain where the evolution of the Early Earth can be investigated. One of these places is the Barberton Greenstone Belt (BGB) in South Africa. It has been well-preserved from its formation between 3.57 and 3.08 Ga ago (de Ronde and de Wit, 1994) until today.

The surface structure and geology of the BGB is well known from a number of geological studies (e.g. Hall, 1918; Anhaeusser, 1969; de Ronde and de Wit, 1994). However, due to a lack of modern deep reaching geophysical studies, there is only little information on the deeper architecture of the BGB. Imaging the complex subsurface and tectonic features is essential for understanding the geological history of the BGB and is a major aim of this thesis.

In the framework of the research initiative Inkaba yeAfrica two high resolution magnetotelluric (MT) field experiments were carried out in 2009 and 2010 across the southern part of the BGB. The MT method allows to resolve the electrical conductivity distribution down to great depths (to the lithosphere) and can give an overview over large investigation areas. For the BGB, conductive mineralizations can be expected along the shear planes of the extensive fault system (Fig. 1.2) caused either by graphite or sulfide enrichments (Ritter et al., 2005). Even only 1% of graphite coating a host rock can increase the conductivity to 1 - 10 S/m (Ritter et al., 2005 and references therein). Due to shearing processes, the graphite is usually interconnected (Jödicke et al., 2004). The conductive zones of the fault system contrast with the high resistivities of the surrounding mafic to ultramafic rocks of the BGB. For details on the MT method see app. A.

Prior to a further analysis of the BGB MT data set by 2D and 3D inversion methods, understanding the properties and information contained in the MT data is essential. Therefore, this chapter aims to give a first overview of the complex BGB MT data set.

1.1 MT measurement setup at the BGB

As part of the German-South African geoscientific research initiative Inkaba yeAfrica, two high-resolution MT field experiments were carried out at the Barberton Greenstone Belt in South Africa. The measurements took place in April/May 2009 (marked by crosses in Fig. 1.1) and July/August 2010 (marked by stars in Fig. 1.1). The field experiments were funded by the GFZ Potsdam and the German Research Foundation (DFG, WE 2938/8-1). Five-component MT data was collected at 193 sites for three to four days in the period range between 10^{-3} and 10^{3} s. The sites are aligned along six profiles with a length between 80 and 110 km. Site spacing along the profiles varies between 2, 5 and 10 km. The distance between the profiles is ~ 10 km.

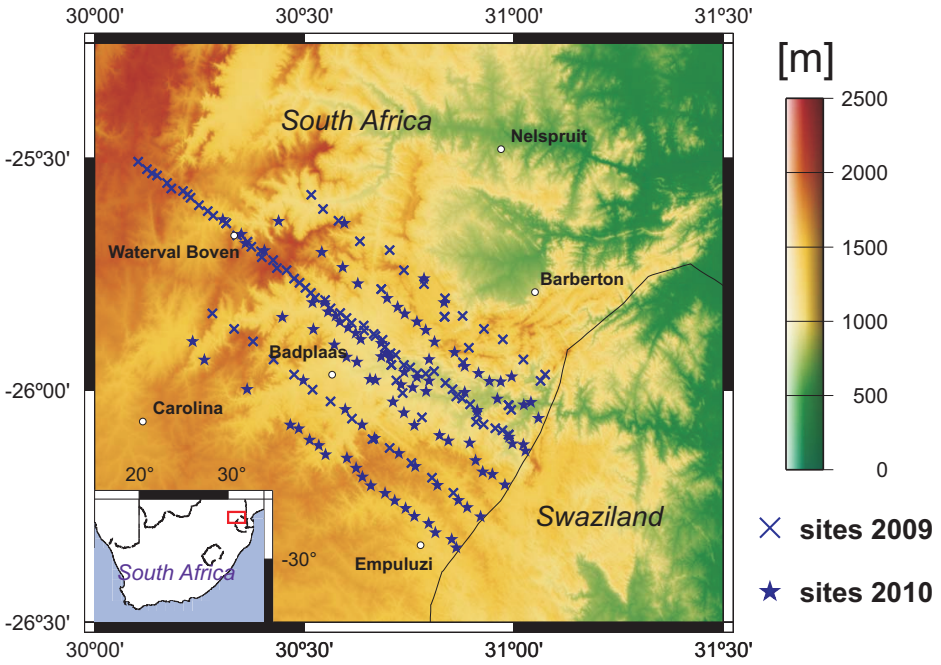


Figure 1.1: Topographic map of the Barberton Greenstone Belt and positions of the MT stations (blue markers). Inset map shows the location of the measurement area within South Africa (red rectangle).

As shown in Fig. 1.2, the MT sites cover the southern part of the greenstone belt and the transition to the igneous intrusive rocks of the Mpuluzi Batholith towards the south and the Transvaal Group sediments northwest of the greenstone belt. This site layout was chosen to investigate a possible southwards continuation of the suture and fault system of the BGB (blue lines in Fig. 1.2) and to delineate the geological units of the BGB from the surrounding rocks.

The MT stations were set up as described in detail in app. A.3. Metronix induction coils and non-polarizable Ag/AgCl electrodes were used to measure the three magnetic and two electric field components, respectively. For data recording, three different types of data loggers were used: S.P.A.M. MkIII and MkIV (**S**hort-**P**eriod **A**utomatic **M**agnetotelluric instrument, 24 bit system, 50 kHz to DC; Dawes 1990, Ritter and Dawes 1992, Ritter 1995, Nascimento 1997) and EDL (**E**arth **D**ata **L**ogger, 6-channel 24 bit recording system, 500 Hz to DC). The equipment was provided by the Geophysical Instrument Pool Potsdam (GIPP). For the recording of the high frequency (HF) and broad band (BB) MT data different settings are used which are summarized in table 1.1.

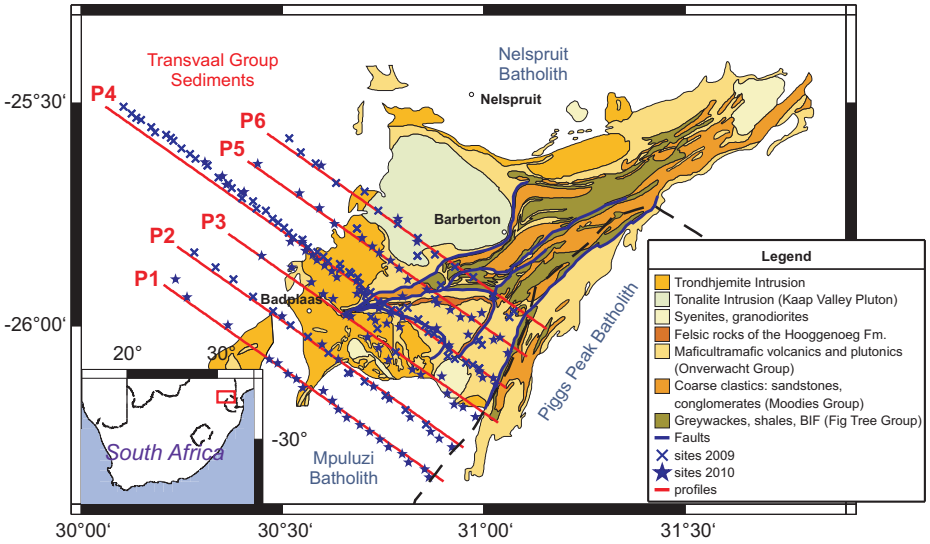


Figure 1.2: Geological map of the Barberton Greenstone Belt (BGB) and positions of the MT sites (blue markers). Inset map shows the location of the measurement area within South Africa (red rectangle, map after Schoene and Bowring, 2007 and de Ronde and de Wit, 1994).

Table 1.1: Setup for recording EM fields at the BGB for high frequency (HF) and broad band (BB) measurements. With the used recording systems measurements can be carried out in different bands. For a better recording of the higher bands (LP1, LP2, ...), LP0 is set to a high sampling frequency but is not recorded.

	Band	Sample Freqency	
HF	LP0	25 000 Hz	not recorded
	LP1	5 000 Hz	recorded for ~ 30 min
BB	LP0	2 500 Hz	not recorded
	LP1	500 Hz	recorded for ~ 3 - 4 days
	LP2	50 Hz	recorded for ~ 3 - 4 days

1.2 Overview of the BGB MT data set

After standard processing as outlined in app. A.4, the MT data can be visualized in pseudosections to obtain a first overview (Fig. B.1 and B.2 in app. B). For the shown data a declination angle of 18.25 W is used for data processing (Magnetic declination calculator from the National Geophysical Data Center, NGDC, last retrieved Feb. 8, 2015).

In general, broad band MT data is available in the period range between 10^{-2} to 10^{3} s. Gaps in the data are either due to rare technical problems (e.g. recording system failure, power breakdown or damage of the MT instruments) or more often, due to low data quality caused by EM noise. In the latter case, no MT transfer functions (TF) can be obtained by standard processing (see app. A.4). EM noise can also be observed in the Barberton data when the TF are viewed separately for each site (e.g. Fig. 1.3 and 1.4).

Severe EM noise in the Barberton area is caused by several sources: a DC railway line, a gas pipeline with anti-corrosion currents, the power supply network of settlements and heavy load/duty devices of mines (see map Fig. 1.3). Towards the western part of the measurement area less settlements can be found due to an increase of farm land. Normally, this would be an improvement, however, these farms

often use electric fences. Not shown in the map are game reserves which can be found throughout the Barberton region. They deploy electric fences along their boundaries as well.
EM fields generated by EM noise sources are often much larger than the natural EM fields. Hence, the natural signal is superposed and the MT recordings are dominated by the artificial EM fields (see chap. 2). Furthermore, the MT measurements are located in the near field of the noise source which violates the assumption of plane-polarized source fields in MT. In the presence of various different noise sources, their signals interfere with each other (Szarka, 1988). Especially, if the emitted noise is of a random type or different sources emit similar noise signals, a differentiation of which source caused a specific disturbance in the data can be difficult. In years with a generally low activity of the natural EM fields as in 2009, the signal to noise ratio is further deteriorated.
Most of the apparent resistivity curves (e.g. Fig. 1.3 and 1.4) have values between 10^2 and $10^6\,\Omega m$ (only considering their noise-free parts), reflecting the generally high resistivity of the subsurface in this area. On the one hand, this facilitates the penetration of the natural EM fields to larger investigation depths, but on the other hand it enables the EM noise to travel far as well affecting large numbers of MT sites simultaneously. EM noise, e.g. in the form of sudden impulses or spikes can easily be measured as far as 70 - 100 km away from its origin in a high resistive subsurface (Szarka, 1988 and references therein). During the MT measurements at the BGB in 2009, a suitable location for a reference site was searched in a $\sim$ 300 km radius from the measurement area. Even in this distance coherent noise was observed in the time series. Setting up a reference site which is even further away from the target area is not possible because the maintenance (e.g. battery supply) of the reference site becomes logistically impractical.
Since EM noise is dominant throughout the area, the TF of nearly all 193 MT sites are in general heavily influenced by it. To gain a better impression of the data set as a whole and the influence of EM noise on MT data, I grouped the sites according to the distortion of the ρ_a and φ curves. The map in Fig. 1.3 shows the location of the main EM noise sources and MT sites colour-coded according to their group. The groups are divided as follows.

- **Type 1: 69 sites (blue)**
  EM noise affects the data starting from around 1 s. It is expressed as a (i) steep increase (with a $\sim 45^{\,o}$ slope) of ρ_a with an accompanied decrease of φ towards zero; (ii) increase of ρ_a with a step at $\sim$ 1 s and decrease of φ or (iii) scatter of the data points.

- **Type 2: 12 sites (green)**
  ρ_a curves exhibit a (strong) increase over (almost) the whole period range while φ drops to zero or stays between 0 and 45^o.

- **Type 3: 18 sites (red)**
  ρ_a exhibits a steep drop while φ increases till $\sim$ 1 s. This effect can either be caused by EM noise or by a complex structured subsurface.
  If the data points for periods above 1 s are not omitted by the robust processing due to severe noise, ρ_a increases steeply while φ decreases. This behaviour of the MT TF is typical for MT sites which are located in the near field of an EM noise source.

- **Type 4: 24 sites (yellow)**
  Both ρ_a and φ are scattered over (almost) the whole period range. For 20 of these sites, this is

caused by EM noise. At the other 4 sites, the deterioration of the data is caused by hardware errors (e.g broken cables, low battery voltage).

Not all sites can be grouped to the described types. At 17 sites, xy and yx components are assigned to two different types (Fig. 1.3 mixed type, black marker). This might be caused by different EM noise sources dominant in either the xy or yx component. The remaining 53 sites could not be assigned to a specific type (Fig. 1.3 unsorted, white marker). In general, these unsorted sites show an erratic behaviour more or less throughout the entire period range (e.g. site 399 in Fig. 1.3).
MT sites of type 1 are in general located in the southeastern part of the measurement area where settlements and farms can be found. Type 2 and 4 MT sites occur close to the DC railway tracks or the gas pipeline. Sites of type 3 form mainly two clusters which are further away from the DC railway line and the gas pipeline. This can be an indicator that the steep decrease of ρ_a and increase of φ is caused by geological structures instead of EM noise. Unsorted sites can mainly be found in the northeastern part of the measurement area. This part has an especially complex geological structure which could cause geological noise (see sec. 2) and be the reason for the erratic behaviour of the TF of the MT sites located there.
MT data distorted by EM noise cannot be used for further analysis by 2D and 3D inversion methods (see chap. 3). If corrupted data would be used the inversion would either result in falsified models or would not even be possible. Hence, noisy data needs to be masked as done in Fig. B.3 and B.4 (in app. B). This leads to a very limited usable data set for which only investigation of the very shallow parts of the BGB would be feasible.
Since the EM noise is superposed on the natural EM signal, filtering of the data can be used to recover the natural signal. This is especially important for longer period data (> 1 s) which is essential for imaging deeper structures (see app. A.2). In order to apply filtering methods succesfully, knowledge about the present EM noise and its properties is necessary. Therefore, looking not only at the distortions of the TF but also analyzing the time series is essential.
Exemplarily, time series are shown for sites 005 and 353 (Fig. 1.4). Mainly two types of noise can be observed in the BGB data: (i) harmonic oscillating noise (Fig. 1.5) and (ii) randomly occuring spikes and steps (Fig. 1.6). Time harmonic noise is characterized by a base frequency and often an amplitude increase at the base frequency and its multiples as well (spectra in Fig. 1.5). For the BGB data set harmonic noise with a frequency of 50 Hz is dominant affecting periods below 0.1 s.
Random noise can occur in different shapes within the time series (Fig. 1.6). In general, spikes occuring in one or more of the magnetic channels are accompanied by steps in the electric channels. Spikes can either be localized in time, affecting only a couple of neighbouring samples or form a longer succession distorting extended amounts of data. Steps appear usually not as simple offset between one sample and the next but are rather rugged over several samples including often minor spikes as well (see zoom-in in Fig. 1.6).
A further analysis especially of random noise and the removal of EM noise from MT data is shown and discussed in detail in the next chapter.

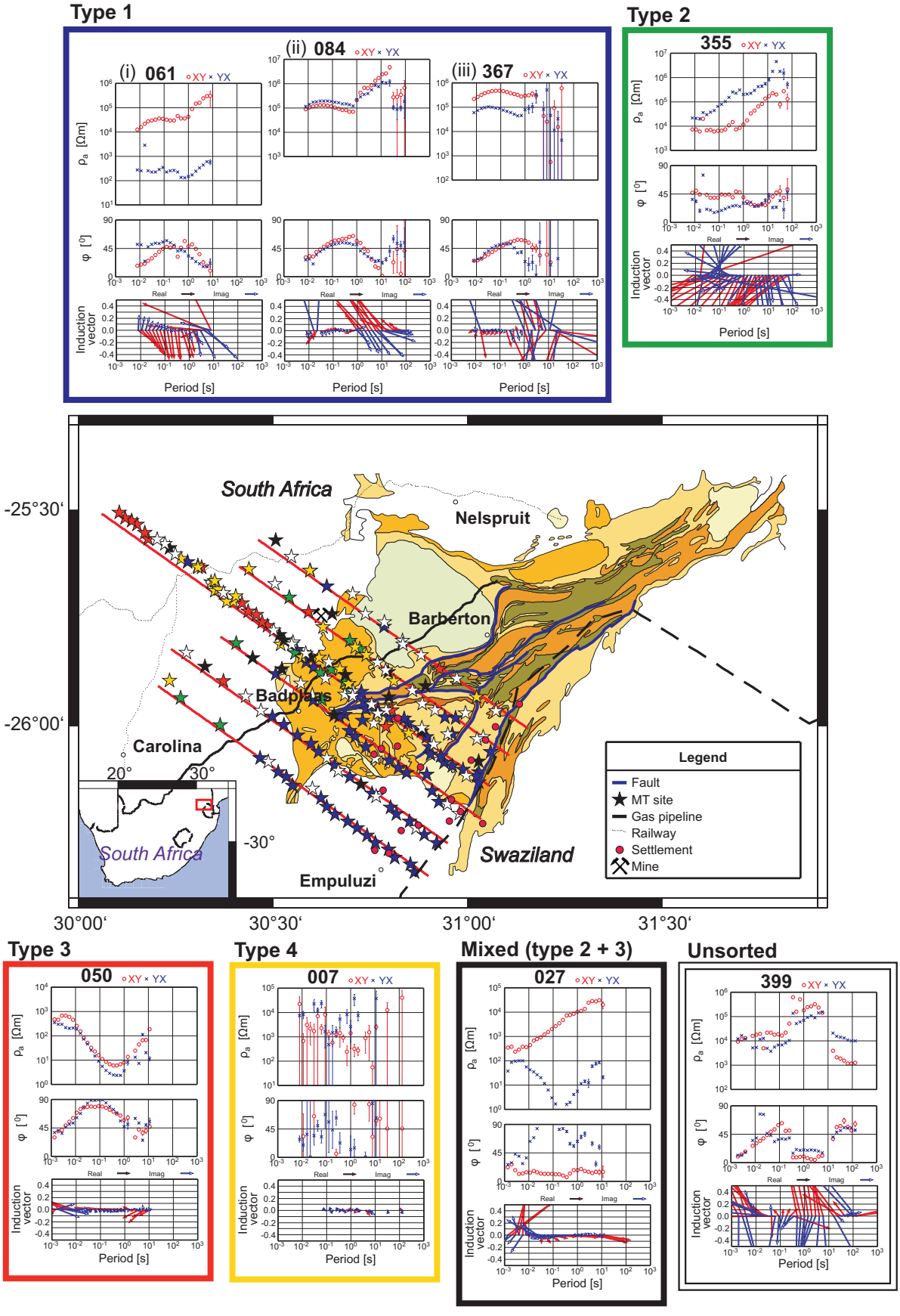


Figure 1.3: Classification of noise types (see text for description) as observed in the BGB MT data set, site markers in map (center) and transfer function examples (top and bottom) are colour-coded (e.g. blue for type 1, map after Schoene and Bowring, 2007 and de Ronde and de Wit, 1994). Locations of major EM noise sources are marked in the map.

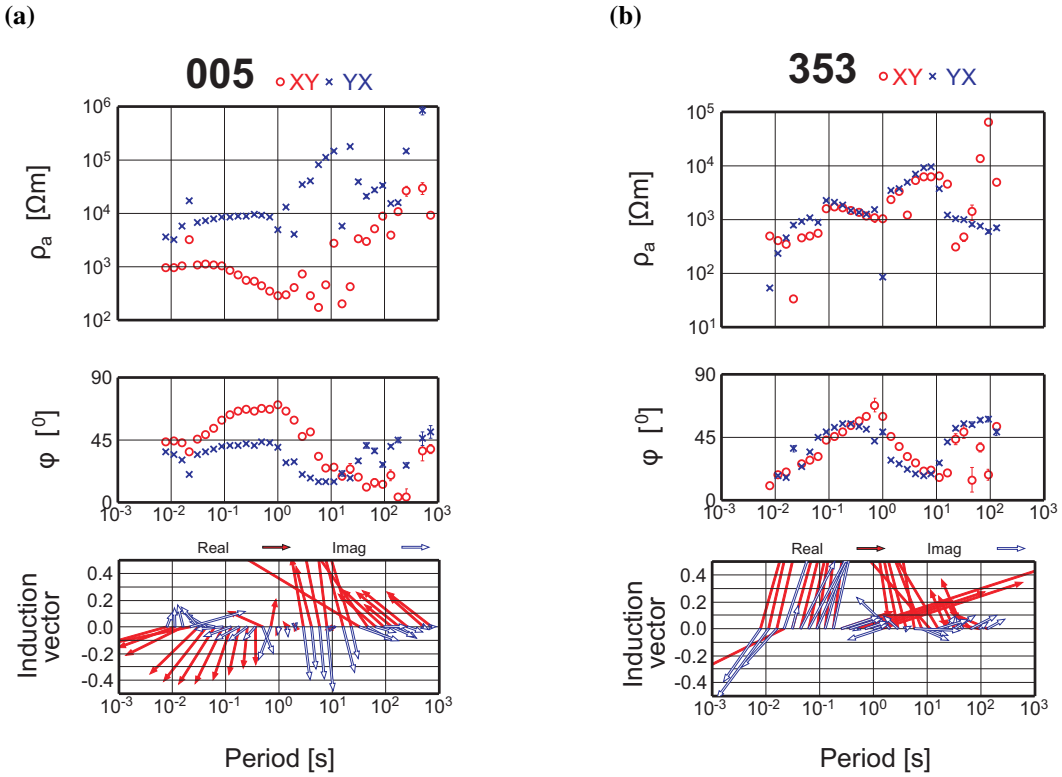


Figure 1.4: Sites 005 (type 1) and 353 (unsorted): Results from single site processing (for site 005 see also Fig. 2.61 and 2.62).

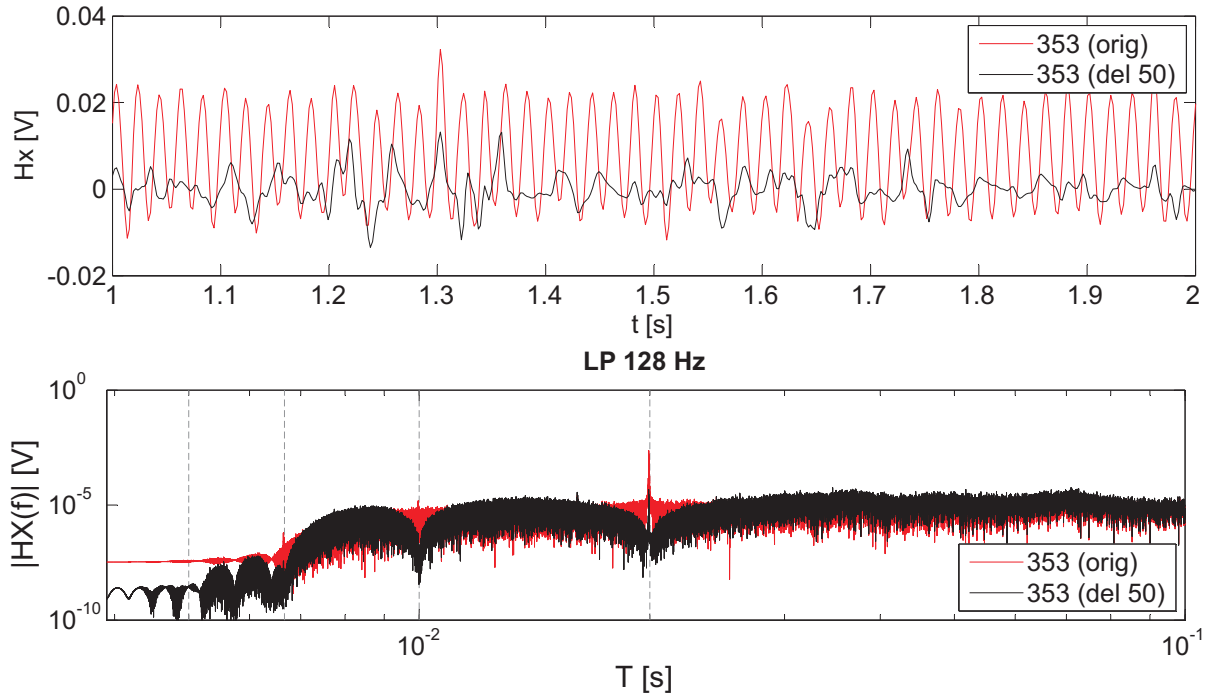


Figure 1.5: Comparison of time series (top) and spectra (bottom) of site 353 with (red) and without (black) harmonic oscillating noise with a frequency of 50 Hz (LP 128 Hz, 262 144 samples for spectra calculations). This type of noise can usually be observed in all five EM field components simultaneously.

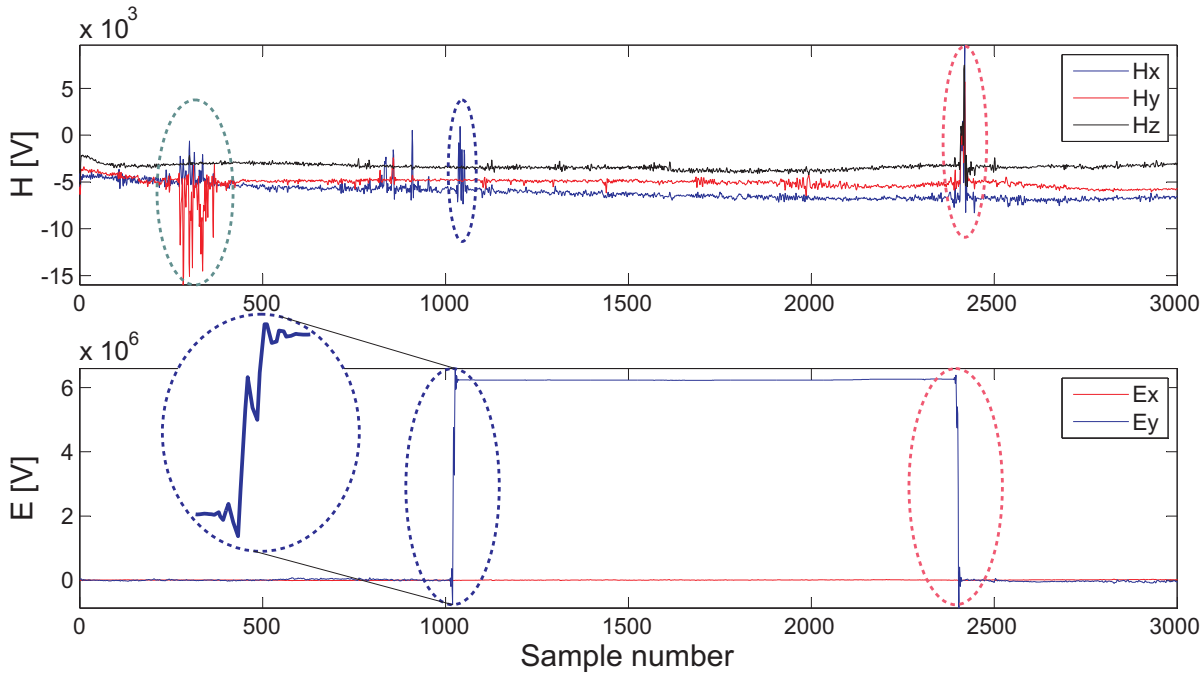


Figure 1.6: Comparison of time series of site 005 (LP 128 s, length ~ 27 h). Different manifestations of noise can be observed: Succession of high amplitude spikes only in the horizontal magnetic field components (~ 110 samples ≈ 1 h, green ellipse), spikes in the Hx component (~ 30 samples ≈ 16 min) corresponding to a step in Ey (~ 20 samples ≈ 11 min, blue ellipses) and a spike observable in all magnetic components with a corresponding step only in one electric component (red ellipses).

Chapter 2

Data analysis and filtering

As shown in the previous chapter, the magnetotelluric (MT) data from the Barberton greenstone belt (BGB) is severely affected by cultural electromagnetic (EM) noise and exhibits an overal poor data quality. Prior to any further analysis, the data needs to be filtered to remove as much of the noise as possible and to obtain a data set with a sufficient quality. To achieve that, the EM noise observed in the data needs to be investigated to understand its characteristics and to find proper filtering and processing methods for cleaning the data.
For the analysis of the observed noise, tests with synthetically distorted data are conducted (sec. 2.1). Synthetic studies have the advantage that they allow full control over the added noise to see exactly how and in which period range the data is altered. Arbitrarily occuring, high-amplitude spikes and steps (irregular noise, Szarka, 1988) similar to the ones observed in the BGB time series are synthetically generated and added to nearly noise-free data (sec. 2.1). A systematic variation of the synthetic noise regarding, e.g. the percentage of added noise or the noise amplitude, provides a qualitative overview of the resulting corruption of MT transfer functions (TF).
A second noise form often observed in the Barberton data are periodic noise signals generated by power lines ($f = 50\,\mathrm{Hz}$, regular noise, Szarka, 1988). The frequency of such signals is determined by a peak in the power spectrum at the frequency of the noise signal.
In general, every EM signal which is not plane polarized and, hence, in the near field of a MT site can be considered as EM noise. The distance between the noise source and the receiver must be larger than three times the skin depth (app. A.2 eq. A.23) to obtain far field conditions (i.e. plane polarized signals). Recorded time series always consist of the wanted natural signal and EM noise. Szarka (1988) and Junge (1996) give an overview of EM noise types and their removal from the data. The following noise types can be observed:

- Instrumental noise which can be caused by strong temperature changes that mainly affect induction coils can be avoided by burying the instruments sufficiently deep.
- Environmental noise caused by sensor motion due to weather conditions, animals or seismic activity can be minimized by a careful site setup and burial of the cables and instruments.
- Geological noise (Bahr, 1991) due to near surface small-scale inhomogeneities can lead to galvanic distortion or static shift in the data. Steep topography in the measurement area is another source for geological noise.

- Cultural noise generated by either active or passive sources. The term passive refers to noise sources which induce secondary fields, e.g. grounded pipelines and electric fences (ground leakage currents can disturb data in a wide period range). Active sources are, e.g. electric networks in mines and anti-corrosion systems (to create a potential difference between the pipeline and the soil).

Cultural EM noise is the main source for the distortion of the BGB MT data and its analysis and removal is the main topic of this chapter.
One possible source for the observed impulse-like noise is the DC railway line which runs through the investigation area (Fig. 1.3). Ground return currents do not follow the path of the overhead wires of the railway line, but instead these traction currents cause pulsations and distort the measured EM fields (Szarka, 1988).
Sometimes cultural EM noise is not removed but instead used for the investigation of the conductivity structure in the measurement area (e.g. Brasse and Junge, 1984; Adam et al., 1989; Qian and Pedersen, 1991). In order to do that, the source field and source mechanisms need to be well understood. In areas (e.g. at the BGB) with several EM noise sources (operating at the same time), a systematic investigation of one specific noise type is hardly feasible since the various noise signals are superposed and interfere with each other. Hence, using the observed noise for investigation purposes was not attempted in this thesis.
Considering the growing level of cultural noise in areas of geoscientific interest, its removal from the actual MT signal is essential (e.g. Larsen et al., 1995; Oettinger et al., 2001; Uchida et al., 2005). Depending on the power of the noise source, the noise signal might even easily exceed the power of the natural signal (Ritter et al., 1998) which makes the removal of the noise extremely difficult.
Proposed ways to erase noise in the data are based on the following concepts (Junge, 1996).

- Sophisticated (robust) stacking and filtering procedures (e.g. Hattingh, 1989),
- visual or automatic noise detection, subsequent removal and data replacement using a data prediction scheme (e.g. Fontes et al., 1988; Kappler, 2012),
- analysis of the relationship of time series from different sites, investigation of their coherency, usage of multistation data (e.g. Gamble et al., 1979; Egbert, 1997) and
- statistical approaches/robust processing schemes, application of weighting schemes according to the statistical analysis of the data (e.g. Larsen, 1989; Larsen et al., 1996).

A number of the mentioned papers fall into more than one of the stated categories and nearly all methods use in some ways robust statistics.
Due to the increasing amount of data (MT data sets with a few hundred different sites are not uncommon anymore), a visual inspection of all measured time series is not feasible. Hence, automated ways to detect and improve the data quality are preferable.
Filtering or processing approaches for removing EM noise from MT data can be divided into two main groups based either on time series (time domain, e.g. Fontes et al., 1988; Hattingh, 1989; Egbert and Booker, 1992) or spectral (frequency domain, e.g. Weckmann et al., 2005) analysis. The two well-established methods, remote referencing (Gamble et al., 1979) and delay filtering, which are shortly described in section 2.2, are frequency domain approaches. For the BGB MT data set, both methods

proved to be most effective for periods $< 1 - 10$ s and in the case of the delay filter for removing noise signals with a specific frequency content (power line signals). In the presence of the described arbitrary, high-amplitude noise in the magnetic and electric channels, however, these approaches fail to improve the data. The arbitrary character (i.e. no fixed frequency) of this noise type makes the application of the delay filter impossible. The commonly observed high-offset steps in the electric channels are not addressed by the remote reference method which solely focuses on the magnetic data. Hence, a significant amount of noise remains in the MT data which most severely affects long periods (T >1 s) due to the lower amount of data in this range (limited recording time for each MT site).
Since the arbitrary spikes and steps are clearly visible within the time series, a time domain based approach for their detection seems to be most appropriate. Within the framework of this thesis, the *reference site filter* introduced by Kappler (2012) is further developed and tested (sec. 2.3.1). This method can be viewed as a hybrid filtering approach: Noise detection takes place in the time domain whereas the replacement of noisy data utilizes the frequency domain. The newly developed STA/LTA filter (sec. 2.3.2) is a solely time domain based approach. These two filters are especially designed for spike and step detection and removal. The results and experience gained by the synthetic noise tests serve as additional input for the filter development and for refining the filter parameter settings. For the synthetically distorted data, a comparison of the filtered to the original time series is possible and is a further asset to the understanding of filter and data processing methods.

2.1 Characterization of synthetic noise and its effect on MT data

To systematically study the effect a certain type of noise can have on MT transfer functions, studies where synthetic noise is added to (nearly) noise-free real-world data are conducted. The MT data from the Barberton Greenstone Belt (BGB) exhibits dominant spikes and steps in the magnetic and electric time series, respectively, for nearly all sites. Therefore, the synthetic noise studies will focus on this specific noise type.
Since no site of the BGB MT data set has a quality sufficient for the proposed tests, MT data from two sites in the Skeleton Coast Park in Namibia is taken (locations marked in Fig. 2.1). The northwestern part of Namibia is characterized by a complex geological setting. This is reflected in the MT data as well which has significant main diagonal components (Fig. 2.5) indicating a 3D-structured subsurface. The results of the synthetic noise tests are mainly shown for the off diagonal elements since they have a smooth variation over the entire period range and a comparison between the original and the distorted data are easier to plot for these components.
Data was recorded for four days in October 2011 within the SAMPLE project (**S**outh **A**tlantic **M**argin **P**rocesses and **L**inks with onshore **E**volution). Fig. 2.2 shows the original transfer functions of the two sites. Spikes and steps are added to the time series of site 105. Site 103 serves as a reference site in the later processing and filtering procedure (sec. 2.2 and 2.3.1).
In the following tests, the synthetic noise aims to be similar to the observed spikes and steps. Therefore, different approaches for generating and adding noise are tested: Spikes are either added additively (sec. 2.1.1) or multiplicatively (sec. 2.1.2) to the time series, whereas steps are only added in an additive manner. Moreover, the effect of spikes and steps with a rather regular amplitude and step offset compared to a random approach is investigated. The influence of the occurance rate and amplitude of the synthetic noise on the data quality is of particular interest.

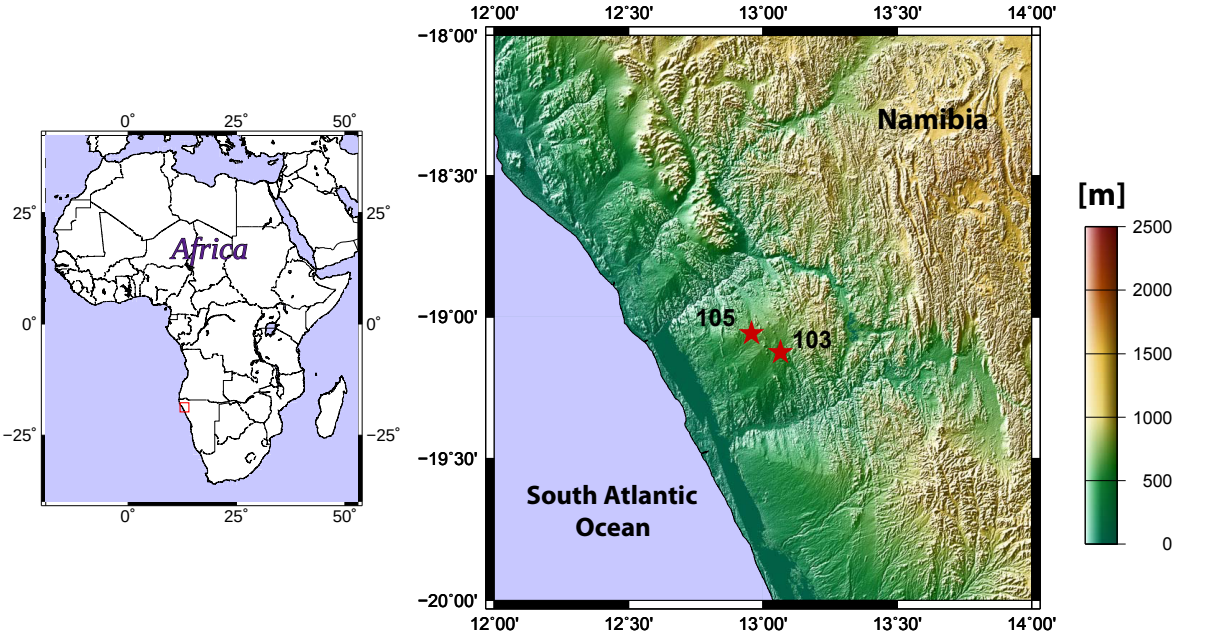


Figure 2.1: Location of sites 103 and 105 (right map, map section marked in red in the left map) used for the synthetic noise studies.

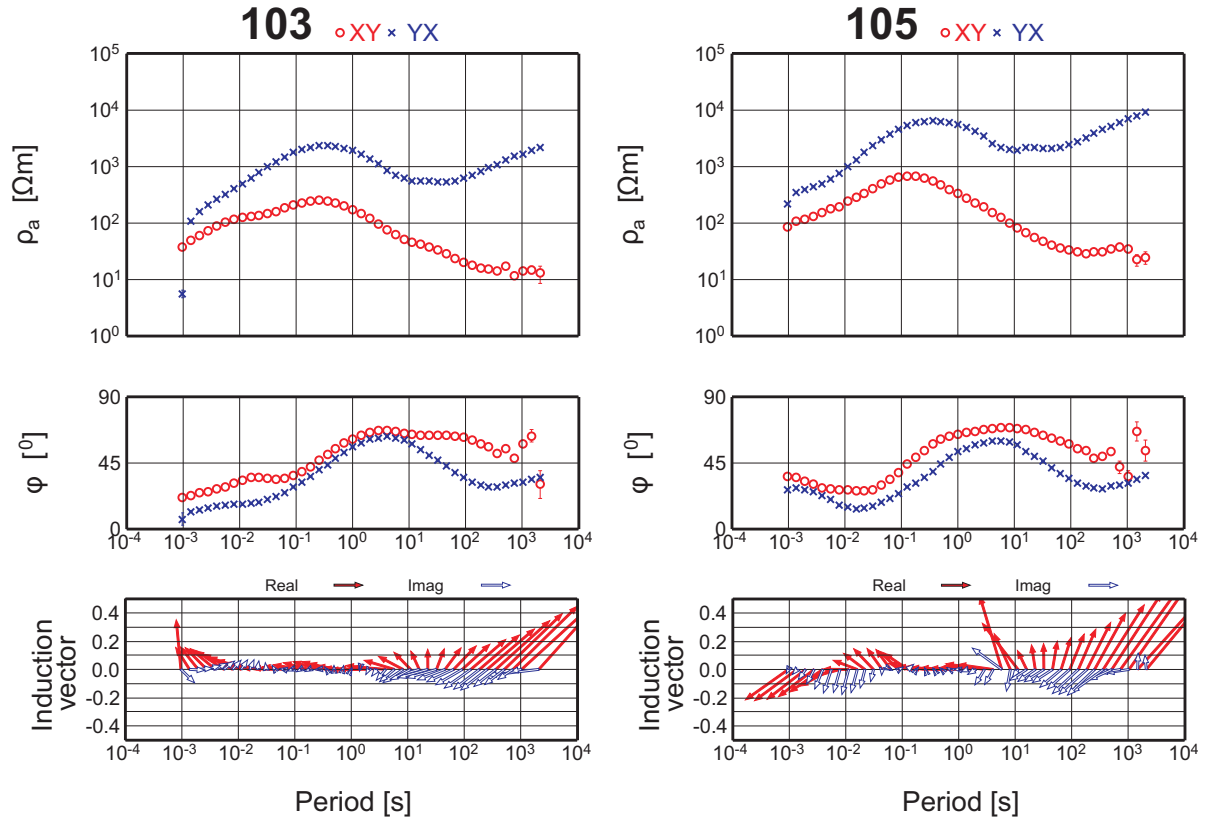


Figure 2.2: Apparent resistivity ρ_a (top), phase φ (middle) and induction vectors (bottom) of sites 103 and 105 (see Fig. 2.1 for their position) for the off diagonal components of the impedance tensor. The main diagonal components are shown exemplarily for site 105 in Fig. 2.5.

2.1.1 Additive noise

In this subsection, only spikes are added to the time series additively, hence, in the following this noise type is called *additive noise*. This approach is conform with various filtering and processing schemes where the data ts_{obs} is considered to be a combination of the natural EM signal ts_{nat} and the EM noise ts_{noise}: $ts_{obs} = ts_{nat} + ts_{noise}$ (e.g. Ritter et al., 1998; Larsen et al., 1996).
Initially, a defined percentage of samples to which noise should be added or substracted is randomly selected. The noise is calculated using three different values (2.1):

$$ts_{noise} = ts_{orig} \pm \underbrace{\overline{ts} \cdot n_{add} \cdot m_{user}}_{\text{additive noise}} \tag{2.1}$$

- $\overline{ts}$ is the mean value of the entire time series and serves to adjust the noise to the general level of the time series,
- n_{add} is the multiplier for the additive noise itself and is randomly selected from the interval $[0, v_{max})$ where v_{max} is the upper user-defined bound of the interval and
- m_{user} is a user-defined fixed value to change the general amplitude of the added spikes.

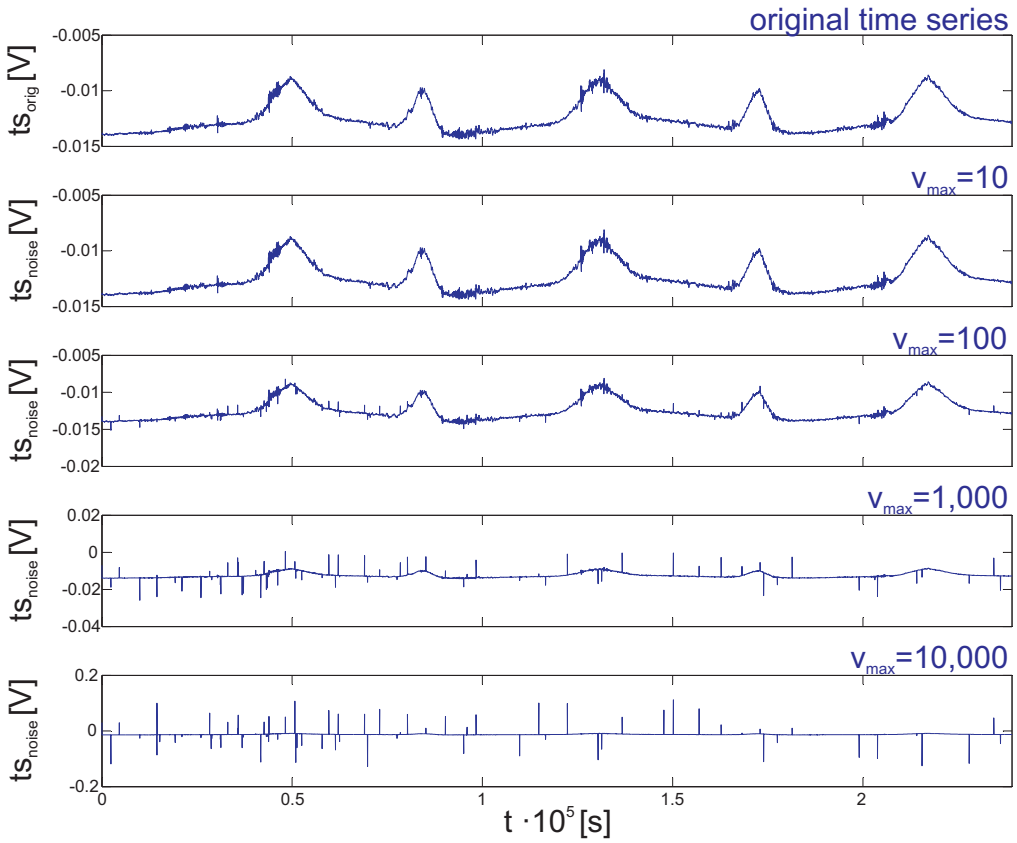


Figure 2.3: Examples for time series (≈ 3 days, channel Hx, site 105) with *additive noise*. Original time series ts_{orig} (top panel) shows long period periodically occuring variations of the natural signal and only very few spikes which might be caused by EM noise; time series corrupted with 1% noise and v_{max} varying between 10 and 10,000 (ts_{noise}, $m_{user} = 0.001$, second to bottom panel).

Figures 2.3 and 2.4 illustrate the *additive noise* exemplarily for channel Hx of site 105 with varying values for v_{max} and m_{user}. Both values influence the amplitude of the added noise (cf. the changing ordinate scales). For small values ($v_{max} \leq 100$, $m_{user} < 0.1$) the noise is hardly discernable and the shape of the curve is still determined by the time variations of the natural EM fields, whereas for larger values the additive noise dominates the amplitude range and shape of the time series.

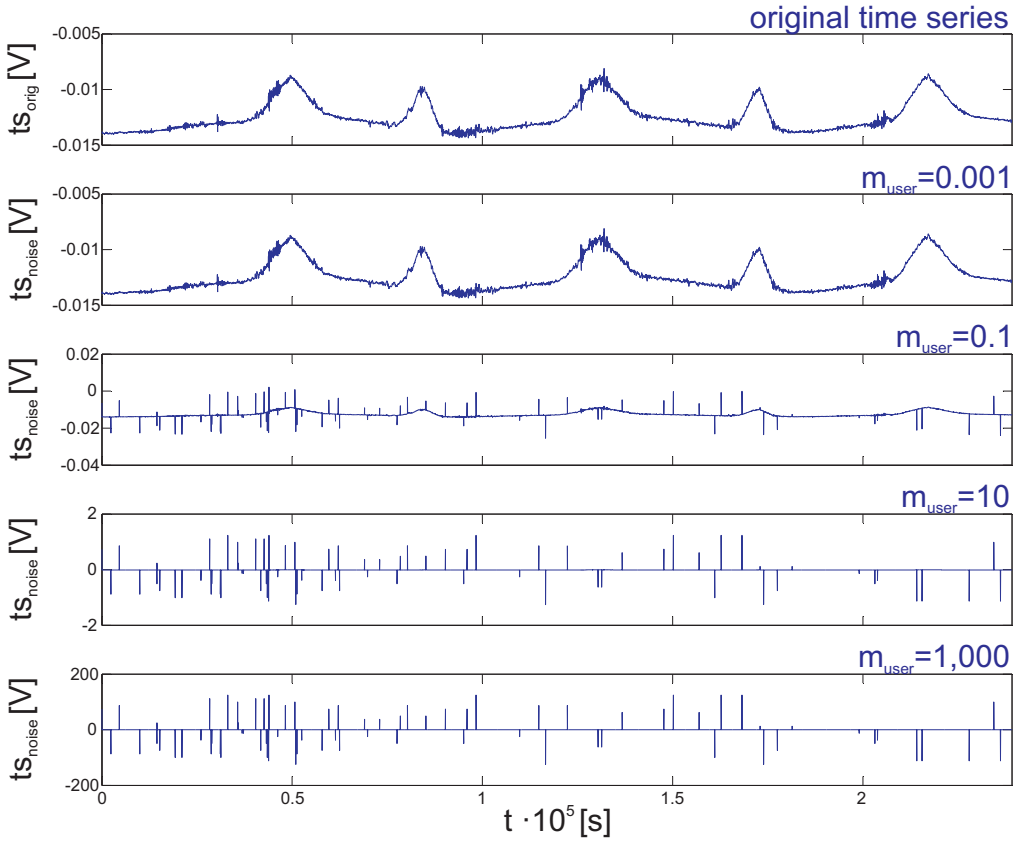


Figure 2.4: Examples for time series (≈ 3 days, channel Hx, site 105) with *additive noise*. Original time series ts_{orig} (top panel) shows long period periodically occuring variations of the natural signal and only very few spikes which might be caused by EM noise; time series corrupted with 1% noise and m_{user} varying between 0.001 and 1,000 (ts_{noise}, $v_{max} = 10$, second to bottom panel). m_{user} increases the amplitude of the spikes more than an increase of the interval v_{max} for n_{add} (cf. with Fig. 2.3).

Figure 2.6 comprises the resulting MT transfer functions for varying v_{max} and m_{user} values with 1% noise added to channel Hx. An interesting aspect is that for some curves the *additive noise* seems to affect certain period ranges more than others. Some data points even got rejected by the robust processing or are shifted to levels out of the shown range, whereas the directly neighbouring points hardly deviate from their original values (Fig. 2.6, black solid ellipses). This can be explained on the one side by the varying effect the noise has on the different period ranges and on the other side by the amount of windows the added noise actually affects in the different bands (Tab. 2.1). The first aspect is best illustrated by the amplitude spectra with and without noise (Fig. 2.7). Higher frequency bands (LP 16 Hz till LP 8192 Hz, see Tab. A.1 and Fig. A.2) are only altered towards the boundaries of the bands, whereas the main part of the spectrum is basically as it was before. This explains the deviation of the data points especially pronounced in the transition from one period band to the next.

For the second aspect, it is important to note that the time series for the different bands are divided

into windows for the calculation of the transfer functions. The longer the period for which the transfer function is calculated the longer must also be the window (cf. sec. A.4). Depending on the distribution of the noise more or less time windows are affected by it as summarized in table 2.1. As a general rule, high percentages are reflected in low data quality. One exception is band LP 128 Hz. Although to more than half of the windows noise is added the ρ_a and φ curves show the same behaviour as before (T between 10^{-2} and 10^{-1} s). This can be attributed to the large amount of data for these periods. Therefore, a sufficient number of noise-free windows remains to calculate stable ρ_a and φ values. For longer periods where there are already only few time windows due to time constraints during recording, adding noise is more severe. Therefore, longer periods are the first to be affected by randomly occuring high-amplitude noise. Sometimes a sufficient number of windows with good data quality can still be found for calculating the transfer functions for the shorter periods within a certain period range, since shorter time windows are used (cf. Fig. 2.6, black solid ellipses). The longer the period and therefore, the longer the time windows, the harder it gets to find sufficiently clean windows. Hence, for longer periods the transfer functions are either scattered or get rejected by the robust processing.

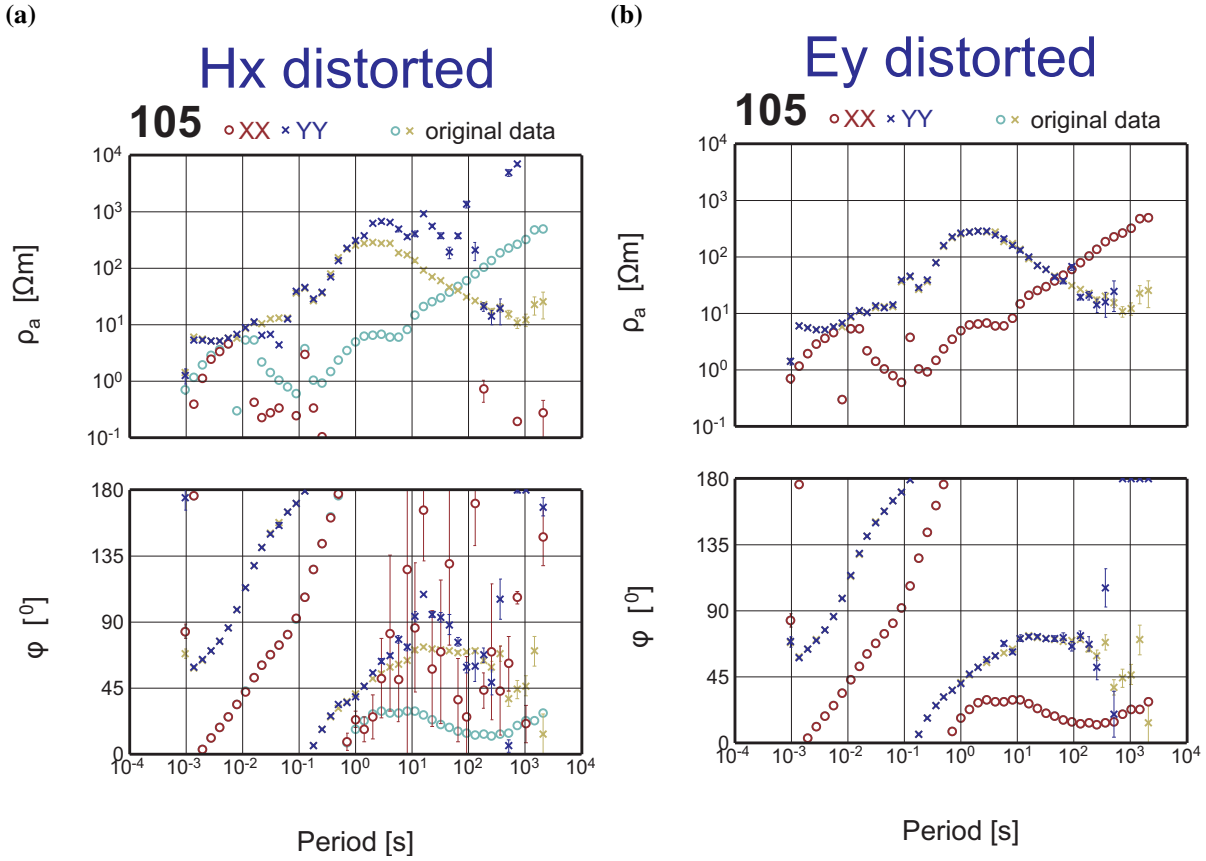


Figure 2.5: Apparent resistivity ρ_a and phase φ for 1% arbitrary noise in Hx **(a)** ($v_{max} = 10,000$, $m_{user} = 0.001$) and Ey **(b)** ($v_{max} = 10,000$, $m_{user} = 0.001$). Original curves for the main diagonal elements are shown in the background.

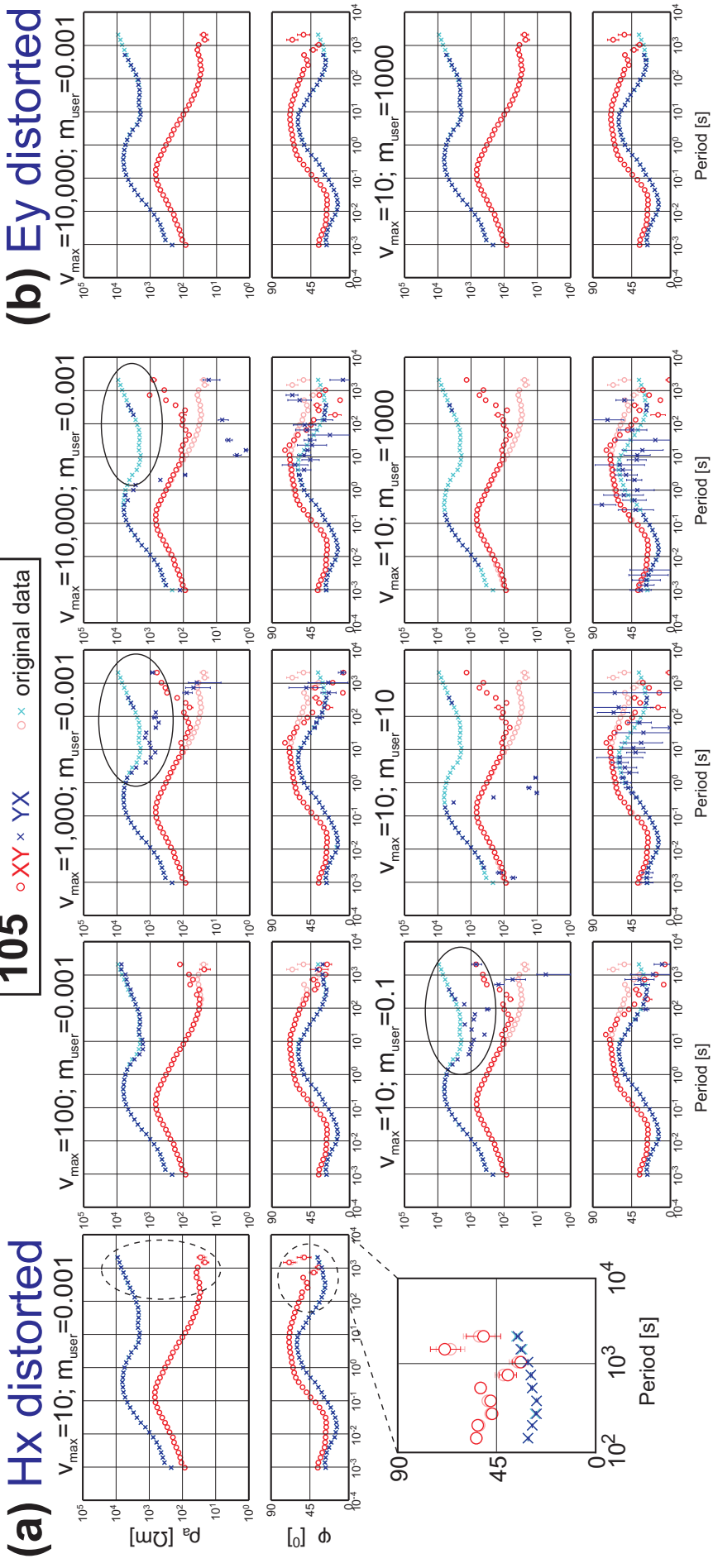


Figure 2.6: Apparent resistivity ρ_a and phase φ obtained from standard processing of synthetically corrupted data: **(a)** 1% noise is added to channel Hx using varying values for ν_{max} and m_{user}. Original curves (cf. Fig. 2.2) are shown in the background in lighter colours. For ν_{max} = 10 and m_{user} = 0.001, the results show a negligible degradation of the data (dashed ellipses and magnified detail of the phase). The yx component is distorted as well (black solid ellipses) since the Hx component is included in the calculations (see text and sec. A.1). **(b)** For comparison, processing results where channel Ey is distorted with 1% noise are shown as well.

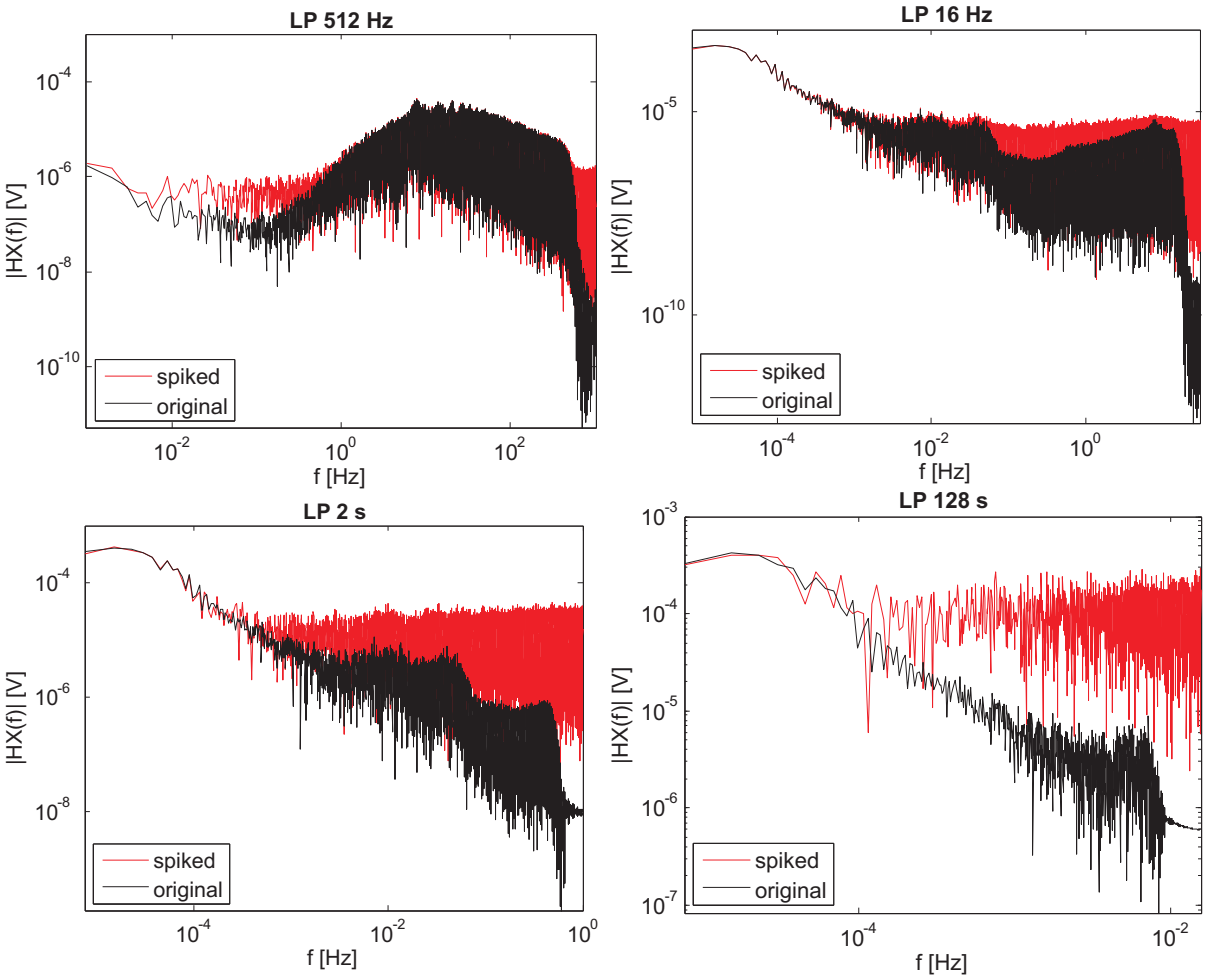


Figure 2.7: Amplitude spectrum of channel Hx, original (black) and with 1% arbitrary additive noise added (red, $v_{max} = 10,000$, $m_{user} = 0.001$). No. of samples (stacked windows) used for spectra calculations: LP 512 Hz: 2,097,152 (16,384 windows); LP 16 Hz: 8,388,608 (65,536 windows); LP 2 s: 261,144 (2,040 windows); LP 128 s: 4,096 (32 windows) (for filtering details see Tab. A.1 and Fig. A.2).

Table 2.1: Percentage of windows affected by synthetic spikes (additive noise) for the different low pass (LP) filtered bands (window length (WL) 128 or 256; no overlap; 1% of the samples of either channel Hx or channel Ey distorted, $v_{max} = 10,000$, $m_{user} = 0.001$, for filtering details see Tab. A.1 and Fig. A.2)).

WL		LP (Hz)						LP (s)		
		8192	2048	512	128	16	2	2	16	128
Hx	128	4 %	3 %	10 %	54 %	3 %	17 %	46 %	63 %	69 %
	256	5 %	3 %	11 %	68 %	3 %	20 %	57 %	81 %	87 %
Ey	128	4 %	3 %	10 %	54 %	3 %	17 %	46 %	63 %	69 %
	256	5 %	3 %	11 %	68 %	3 %	20 %	57 %	81 %	87 %

In order to investigate the influence noise in the electric channels has on the processing results, one E channel is corrupted as well. If channel Ey is distorted by the same noise (additive noise in the form of spikes) that is used for Hx their noisy spectra look similar (cf. Fig. 2.8 and 2.7) and the percentage of windows affected by noise in the different bands is the same as for channel Hx. However, the transfer functions differ only marginally from the original undistorted curves (Fig. 2.6, **(b)**) showing

that mostly the longest periods are affected and only slight deviations at some of the transitions from one period band to the next can be observed. In contrast to distorting one magnetic channel, distorting one electric channel affects only one of the off diagonal components of the impedance tensor (here it is the yx component). This behaviour is due to the way the transfer functions are calculated (see sec. A.1). In 3D, the horizontal components of the magnetic field influence both Z_{xy} and Z_{yx}, whereas the electric field influences only one of them. Looking at the main diagonal elements without and with noise distortion shows a comparable picture (Fig. 2.5). Here the Hx distortion again affects both components, whereas Ey distortion only influences the yy component.

In general, electric fields are weaker (smaller in amplitude) than the magnetic field. Since the noise is also multiplied by the mean value of the time series $\overline{ts}$ the noise in the electric channels is effectively smaller than in the magnetic channels. Hence, *additive noise* as described in this section added to the electric channels is less severe than noise in the magnetic channels which can be observed in real-world data as well. However, depending on the proximity of a noise source, noise in the electric channels can also be much larger in its amplitude. In the worst case, the noise could completely dominate the time series, masking the small variations of the natural EM fields. The time series would then be more or less useless for processing and further analysis because such noise could never be removed effectively. Fortunately, real-world noise in E channels is seldom this severe.

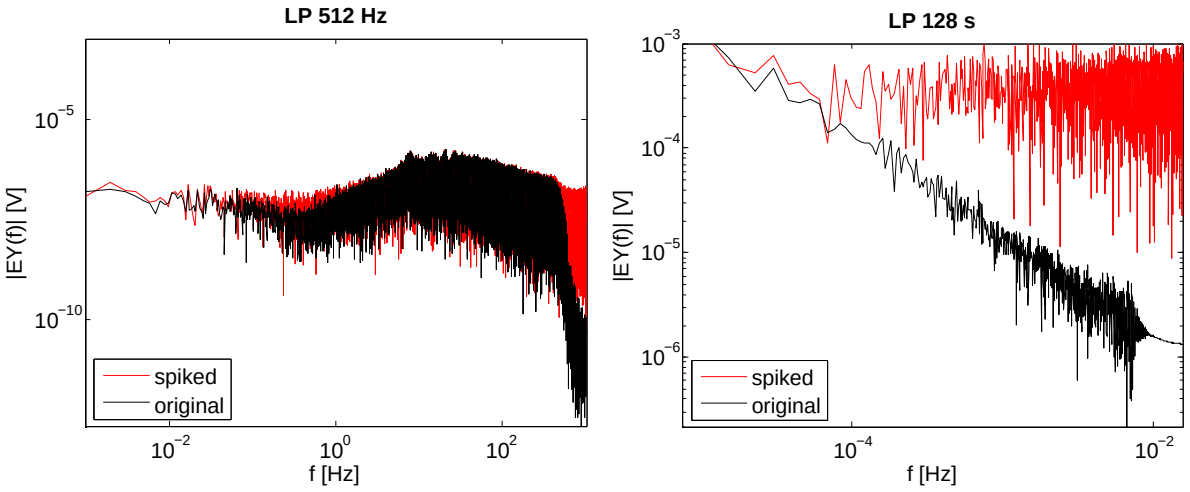


Figure 2.8: Amplitude spectrum of channel Ey, original (black) and with 1% arbitrary noise added (red, v_{max} = 10,000, m_{user} = 0.001). No. of samples for spectra calculations: LP 512 Hz: 2,097,152; LP 128 s: 4,096 (for filtering details see Tab. A.1 and Fig. A.2).

2.1.2 Multiplicative noise

In this section, spikes are generated by multiplying the original time series at a selected percentage of samples with a noise value (see eq. 2.2 and 2.4). This multiplication approach allows for a wider amplitude range of the spikes and they are easier to generate and handle than the *additive noise* described in the previous section. Instead of spikes, steps are added to the electric channels in this synthetic noise generating approach (Fig. 2.9). The occurance of noise as spikes in the magnetic and steps in the electric channels can frequently be observed in time series with real-world noise (see Fig. 1.6).

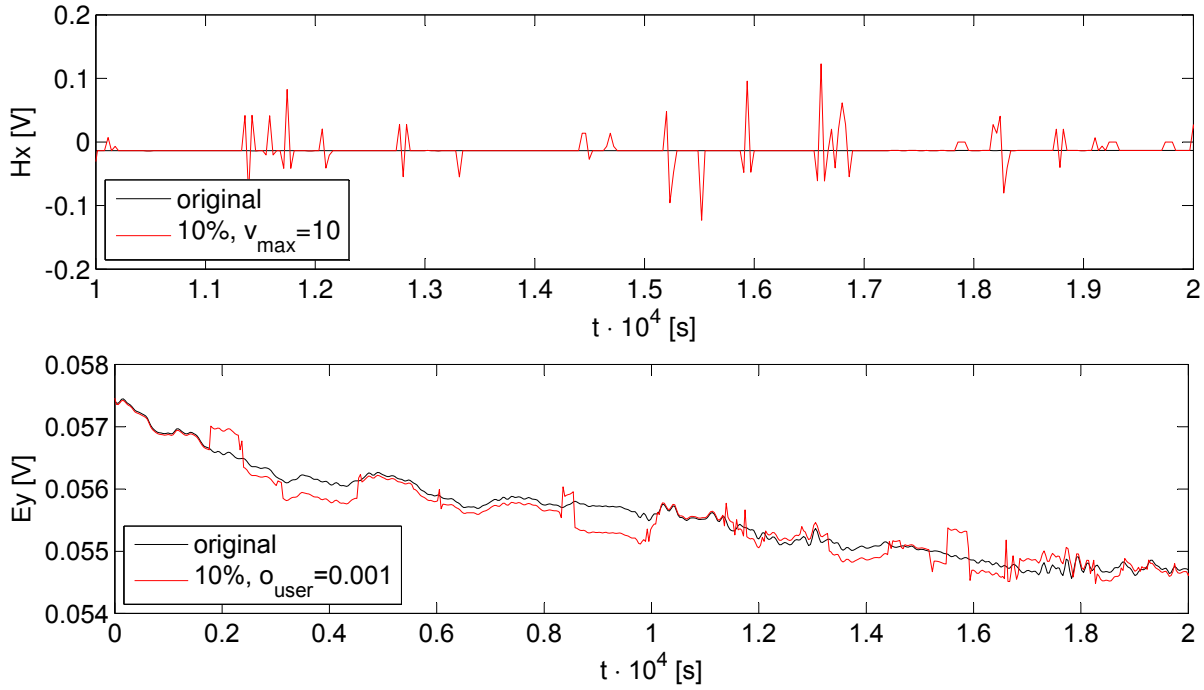


Figure 2.9: Example of time series with synthetic spikes in a magnetic (top, band LP 128 s, ~ 3 h) and synthetic steps in an electric channel (bottom, band LP 128 s, ~ 6 h). Synthetic spikes can distort either single samples or several consecutive samples. The synthetic steps are not generated simply as offsets between two data levels within a time series but can consist of several consecutive offsets and/or spikes.

Two slightly different ways of introducing spikes and steps to magnetic and electric time series are presented in this section. In the following, they are called *regular* and *arbitrary spikes and steps* and can be described mathematically as follows.

Regular spikes and steps

$$ts_{noise} = ts_{orig} \cdot n_{mult} \qquad \text{(spikes)} \tag{2.2}$$

$$ts_{noise} = ts_{orig} \pm \underbrace{\overline{ts} \cdot o_{user}}_{\text{step offset}} \qquad \text{(steps)} \tag{2.3}$$

- n_{mult} is a fixed user-defined value, it is randomly decided by the algorithm from spike to spike whether the value itself is used or its negative,
- o_{user} is a fixed user-defined multiplier to regulate the offset of the steps and
- $\overline{ts}$ is the mean value of the time series.

Arbitrary spikes and steps

$$ts_{noise} = ts_{orig} \cdot n_{mult} \qquad \text{(spikes)} \tag{2.4}$$

$$ts_{noise} = ts_{orig} \pm \underbrace{\overline{ts} \cdot n_{mult} \cdot o_{user}}_{\text{step offset}} \qquad \text{(steps)} \tag{2.5}$$

- $n_{mult} \in [0, v_{max})$ is a multiplier, randomly selected from the user-defined interval,
- o_{user} is a fixed user-defined multiplier to regulate the offset of the steps and

- $\overline{ts}$ is the mean value of the time series.

For adding spikes to the magnetic channels, at first, only a third of the user-defined percentage of samples of a given time series are randomly selected. These sample positions mark the central peak of each spike and the value of the time series is multiplied by n_{mult} (cf. equations 2.2 and 2.4). To avoid double selection of samples for spiking, already chosen samples get excluded from the further selection process. In the next step, the samples directly before and after the peak are multiplied with $-0.5 \cdot n_{mult}$ which results in spikes distorting at least three consecutive samples. Wider spikes (affecting more than one sample) can also be observed in 'real-world' noisy data (see Fig. 1.6).
In general, this procedure should lead to the set percentage, however, if neighbouring samples got selected in the beginning, the actual percentage of affected samples is lower. In this case, the remaining percentage of samples is randomly selected from the remaining noise-free samples, resulting in a few single sample spikes.
If both magnetic and electric channels get distorted, the same sample positions that are selected for spikes are used to add steps to the electric channels. Starting from a selected sample (all the way to the end of the time series), a constant offset value is added to the time series (cf. equations 2.3 and 2.5). At the next selected sample a new offset is calculated using again equation 2.3 or 2.5 and so on. Since usually three consecutive samples are marked the steps are rugged and are not simply boundaries between two levels. This can be observed for 'real' steps as well (see Fig. 1.6). In general, the step size or offset can be arbitrarily large. However, if the mean value of the time series is much smaller than the step offset, numerical and digitilization limitations (when writing the data into a file) destroy the data since the small variations of the time series get lost due to the large steps (Fig. 2.10). Since intervals where no noise is added are compared, the time series should ideally be identical. However, this is only the case for $o_{user} < 10$ for the presented example. Table 2.2 summarizes the absolute differences between the original and the artificially noisy time series. The mentioned behaviour of the time series in the presence of huge steps is comparable to the saturation of a channel during recording when the dynamic range of the analogue-to-digital converter is not sufficient to record also the largest fluctuations (probably due to noise) in the electric or magnetic fields.
The magnitude of the step offsets is reflected in the spectra as well (Fig. 2.11). Whereas for small steps ($o_{user} = 0.001$) the amplitude spectrum is still in the range of the original, the spectrum gets shifted to larger values with increasing step size.

Table 2.2: Relative difference between original time series of channel Ey and time series where 10% noise in the form of steps was added with varying o_{user} values (arbitrary steps: $n_{mult} \in [0, 10)$).

o_{user}	regular steps	arbitrary steps
0.001	0	0
0.01	0	0
0.1	0	0
10	0 . . . 6.3 %	0 . . . 0.2 %
1000	0.2 . . . 100 %	6.3 . . . 100 %

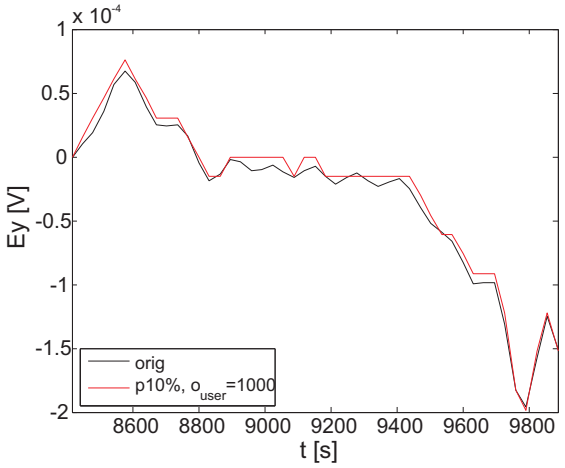


Figure 2.10: Comparison between original time series of channel Ey and with 10% regular steps added ($o_{user} = 1000$) in a clean part of the time series (band LP 128 s, ~ 25 min).

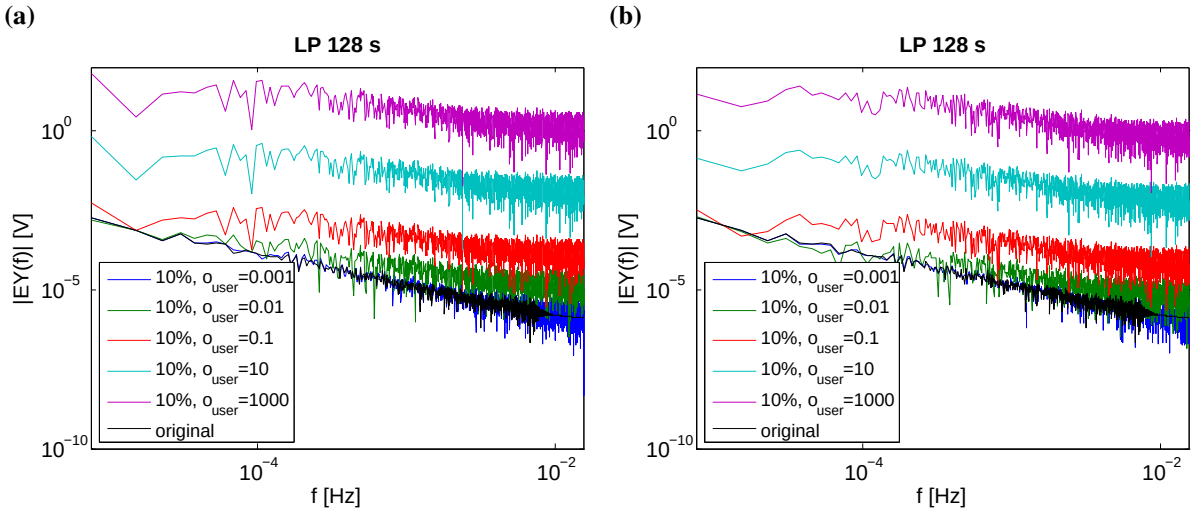


Figure 2.11: Comparison of amplitude spectra of channel Ey with 10% noise (distorting over half of the windows) added in the form of regular **(a)** and arbitrary steps **(b)** ($n_{mult} \in [0, 10)$) for varying o_{user} values. No. of samples for spectra calculations: LP 128 s: 4,096 (32 windows) (for filtering details see Tab. A.1).

The main difference between *regular* and *arbitrary* noise (Fig. 2.12 and 2.13) is the definition of the multiplier n_{mult} which can be fixed or arbitrary. In the arbitrary case, this multiplier is used for the offset calculation for the steps as well. Figures 2.12 and 2.13 show examples of time series with 10% of the data influenced by spikes or steps. If n_{mult} is fixed the amplitudes of the spikes and the step offsets follow the shape of the curve and the noise looks artificial (middle panels). The added spikes and steps appear to be closer to real noise if n_{mult} is random (bottom panels).

Comparing the transfer functions (Fig. 2.14 and 2.15), the results with *arbitrary noise* show a higher degradation of the data quality than with *regular noise*. For *arbitrary noise*, the effect is also visible starting from shorter periods (around 10^{-1} s or sometimes even shorter). Hence, not only the amplitude of noise seems to matter but also the amplitude distribution. Again, noise in the magnetic channel is affecting the quality of the transfer functions more severely than noise in the electric channels.

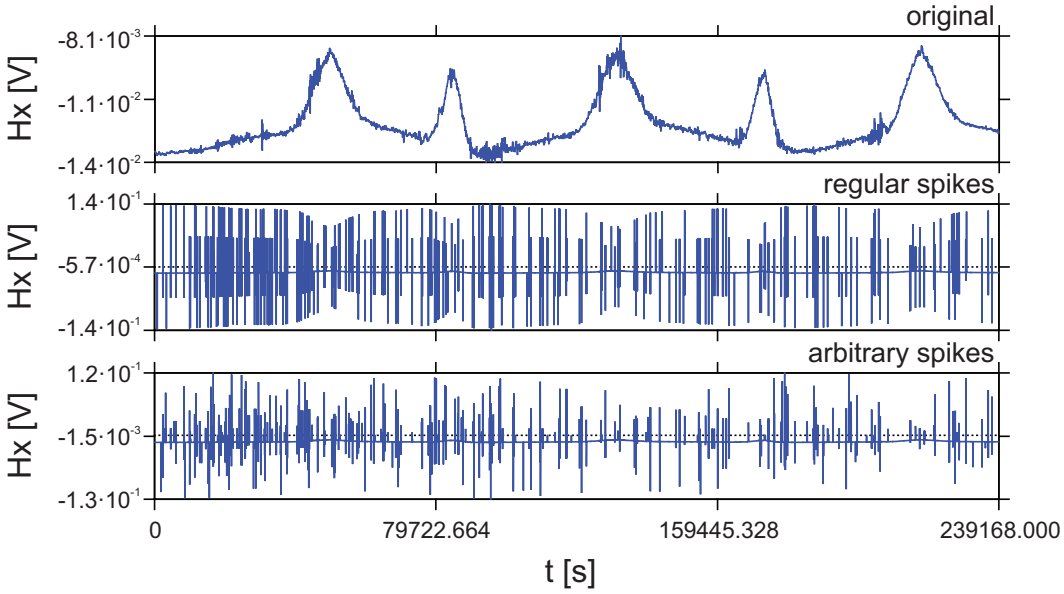


Figure 2.12: Original time series of channel Hx (top), regular spiked (middle, $n_{mult} = 10$) and arbitrarily spiked (bottom, $n_{mult} \in [0, 10)$) time series ($\approx$ 3 days).

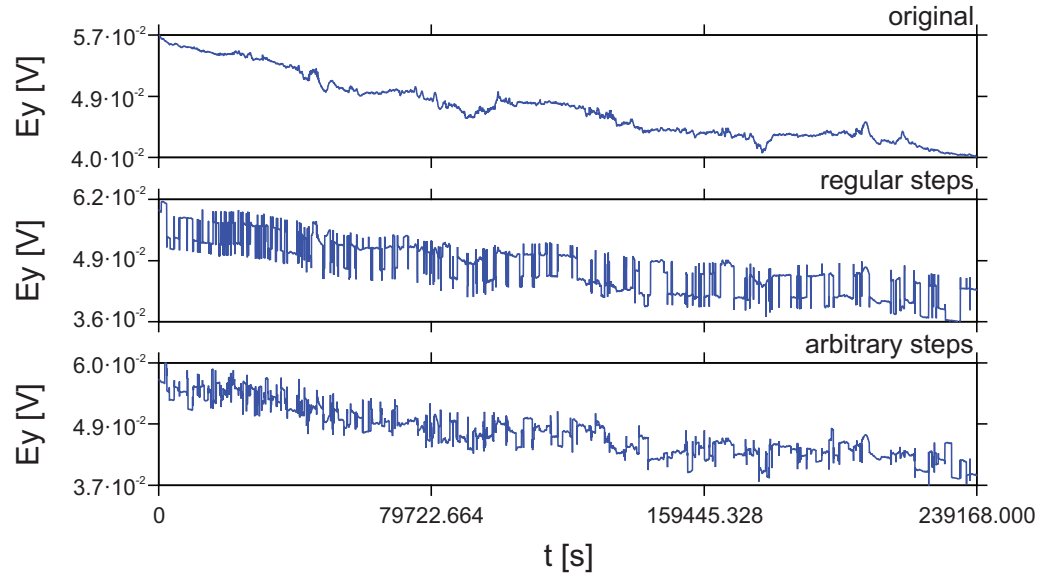


Figure 2.13: Original time series of channel Ey (top), with regular steps (middle, $o_{user} = 0.01$) and arbitrary steps (bottom, $o_{user} = 0.01$,$n_{mult} \in [0, 10)$, $\approx$ 3 days).

Increasing either the percentage of noise added to the time series or increasing the amplitudes of the spikes (with higher v_{max} values) leads to a more severe impact on the transfer functions (Fig. 2.16). Depending on the dimensionality of the subsurface, the resulting xy- or yx-component ρ_a and φ curves are comparably distorted if either Hy or Hx are noisy.

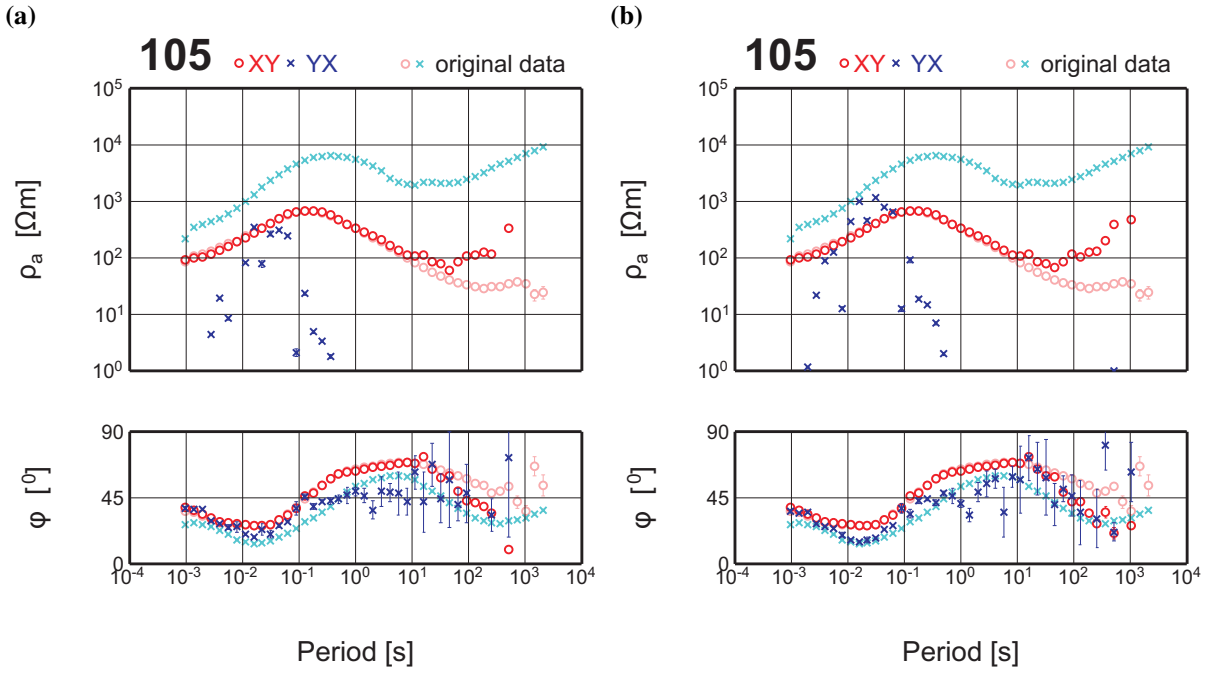


Figure 2.14: Apparent resistivity ρ_a and phase φ for 10% regular **(a)** ($n_{mult} = 10$) and arbitrary spikes **(b)** ($n_{mult} \in [0, 10)$) in channel Hx. For regular spikes ρ_a goes down to 10^{-3} Ωm (not shown here). Original curves are shown in the background in lighter colours.

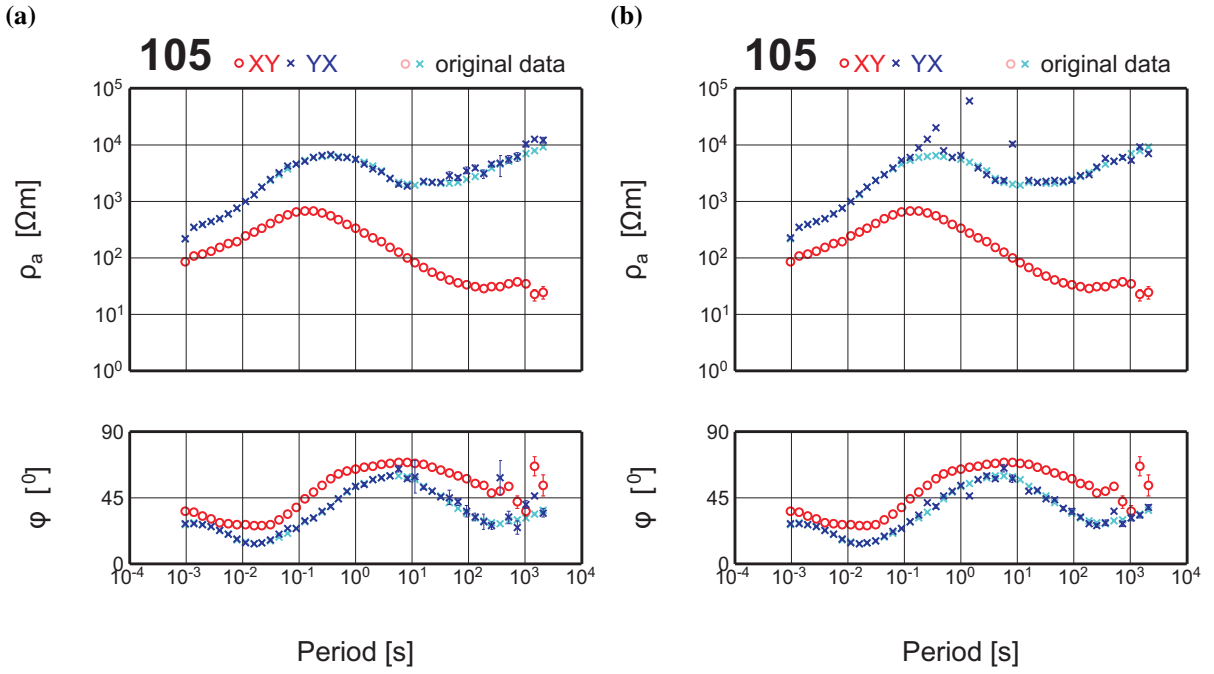


Figure 2.15: Apparent resistivity ρ_a and phase φ for 10% regular **(a)** ($o_{user} = 0.001$) and arbitrary steps **(b)** ($o_{user} = 0.001$, $n_{mult} \in [0, 10)$) in channel Ey. Original curves are shown in the background in lighter colours.

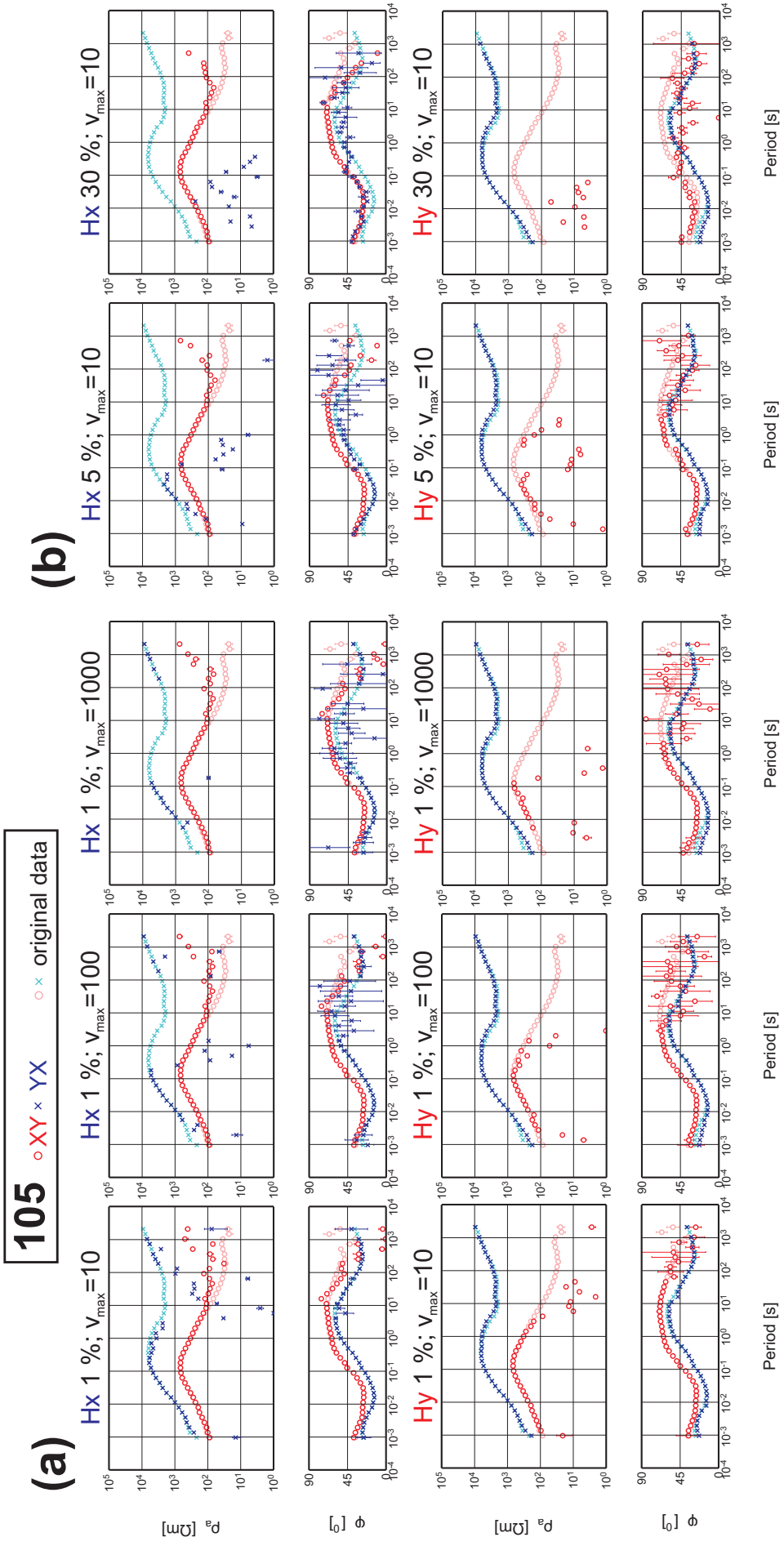


Figure 2.16: Apparent resistivity ρ_a and phase φ for varying values of v_{max} (**a**) and varying percentages of arbitrary spikes (**b**) in channel Hx or Hy. Original curves are always shown in the background in lighter colours.

Adding noise to more than just one channel decreases the data quality further (Fig. 2.17). As expected from the single channel examples, if the magnetic channel of a component is already affected by noise, adding noise to its complementary electric channel does not significantly alter the transfer functions. Important to note is, that the phase is, in general, less affected by noise and can therefore, be considered as more reliable (for $T < 10^{-1}$ s). Similar observations can be made in the presence of geological noise in form of distinct changes in the topography between sites of an array where the phase remains more stable as well (Fischer et al., 1992; Kütter, 2009).

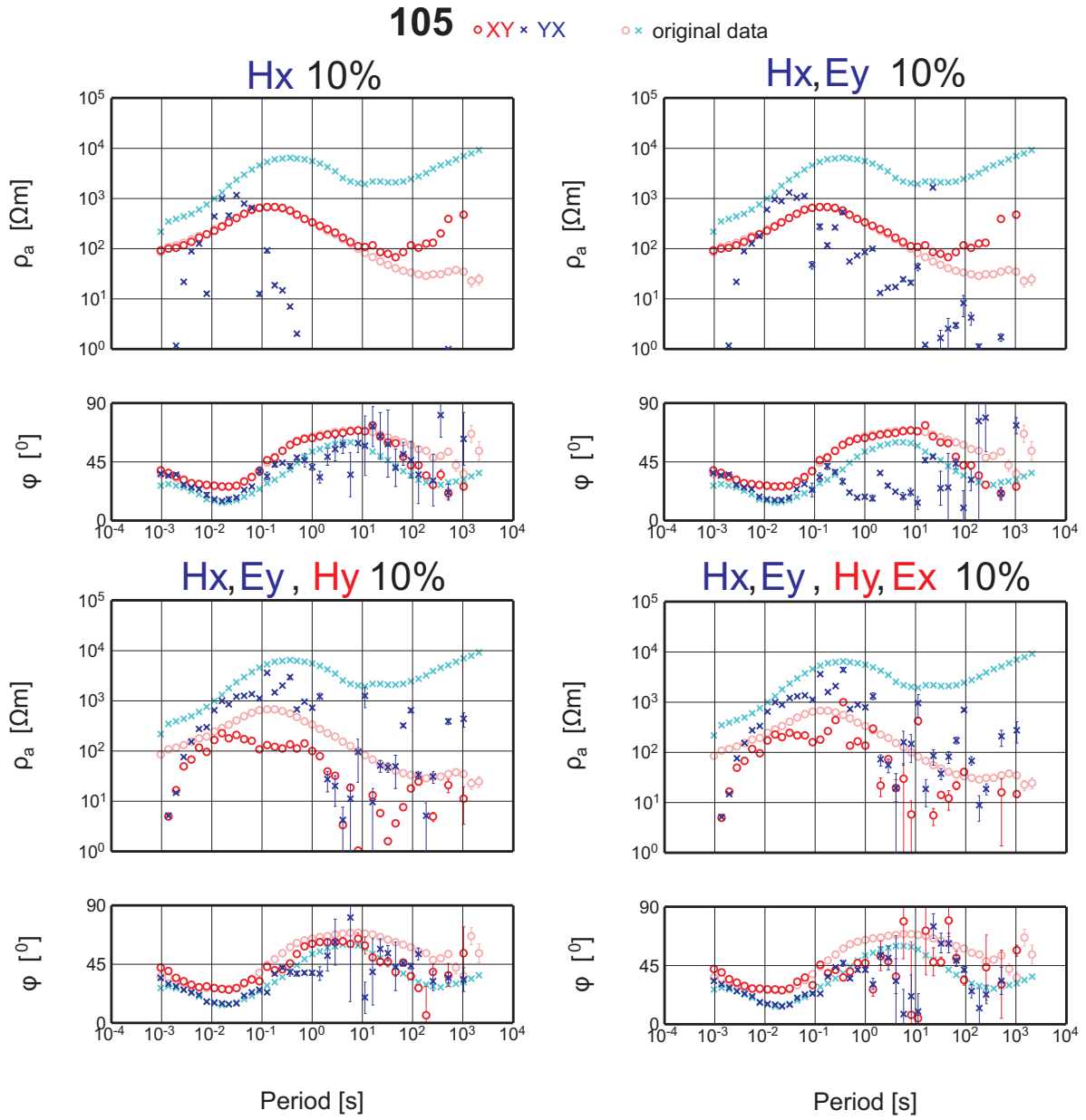


Figure 2.17: Apparent resistivity ρ_a and phase φ for 10% arbitrary spikes and steps in different magnetic and electric channels ($n_{mult} \in [0, 10)$, $o_{user} = 0.001$). Original curves are shown in the background in lighter colours.

One of the most often applied operators on MT data is the rotation matrix (eq. A.10). It is for example used for rotating the data according to the declination at the measurement point during the recording or for rotating the data in the geo-electric strike direction (sec. 3.3.1). Obviously, not only the clean data is rotated but also the present noise. If only one magnetic channel is synthetically spiked, a

rotation distributes these spikes, depending mainly on the rotation angle and data dimensionality, more or less evenly to both horizontal channels. Observing noise in more than one channel is also more in agreement with 'real' noise since rarely a noise source can be found which is aligned exactly in one of the measurement directions.
A few examples of the distorted data (either Hx or Hy) are rotated anticlockwise (towards west) by a declination angle of 10.5^o (Fig. 2.18). Rotation is applied to the data before the bandpass filtering of the time series (cf. sec. A.4). In the yx component (Hx with 1 or 10% noise, cf. Fig. 2.14 **(b)**) a bump-like feature (10^{-1} to 10 s, black ellipse) appears in the phase which is at least for 1% noise not reflected in the apparent resistivities. This shows that in some cases ρ_a can be correct even if the phase is disturbed. A more significant feature is the downshift of the data with a distinct offset between 1 and 10 s in the xy component (Hy with 1% noise). For 10% spikes in Hy most of the ρ_a curve gets shifted by approximately one decade still resulting in a smooth curve. If the original curve was not known, only periods between 10^{-3} and 10^{-1} s would be considered as severely affected by noise (in ρ_a), though they are closest to the true curve. Therefore, a close inspection of the data is essential to avoid false or misleading interpretations: The time series where spikes and steps can easily be seen and the spectra which reveal signals with a dominant frequency are the basic means to get a comprehensive overview of the data and its quality.

Another important parameter for analysing MT data are the so-called induction vectors (eq. A.13). They relate the horizontal magnetic field components (Hx and Hy) to the vertical component (Hz) and are indicators for lateral conductivity variations. Since the effects of *additive noise*, *regular* and *arbitrary spikes* on the data are comparable, only the processing results with *arbitrary spikes* in one or more of the magnetic channels are displayed here. Fig. 2.19 summarizes them, showing that both direction and length of the vectors are changing if synthetic noise is added to the time series. If the noise is not removed this can have a severe impact on the inversion of the data since location, conductivity and the dimensions of conductive bodies in the model would be altered by the noise.
If noise is present in either Hx or Hy the vectors have a distinct direction over the affected period range. Therefore, observed induction vectors which are parallel over a wide period range can be an indication for a strong noise source in a certain direction. In general, the induction vectors with arbitrary spikes show only small variations in length and direction from one period to the next, making it more difficult to clearly identify the periods affected by noise. Hz with 10% spikes is the only exception. Here, the noise is rather obvious since starting from $\sim$ 1 s large induction vectors can be observed with dramatic variations in their orientation.

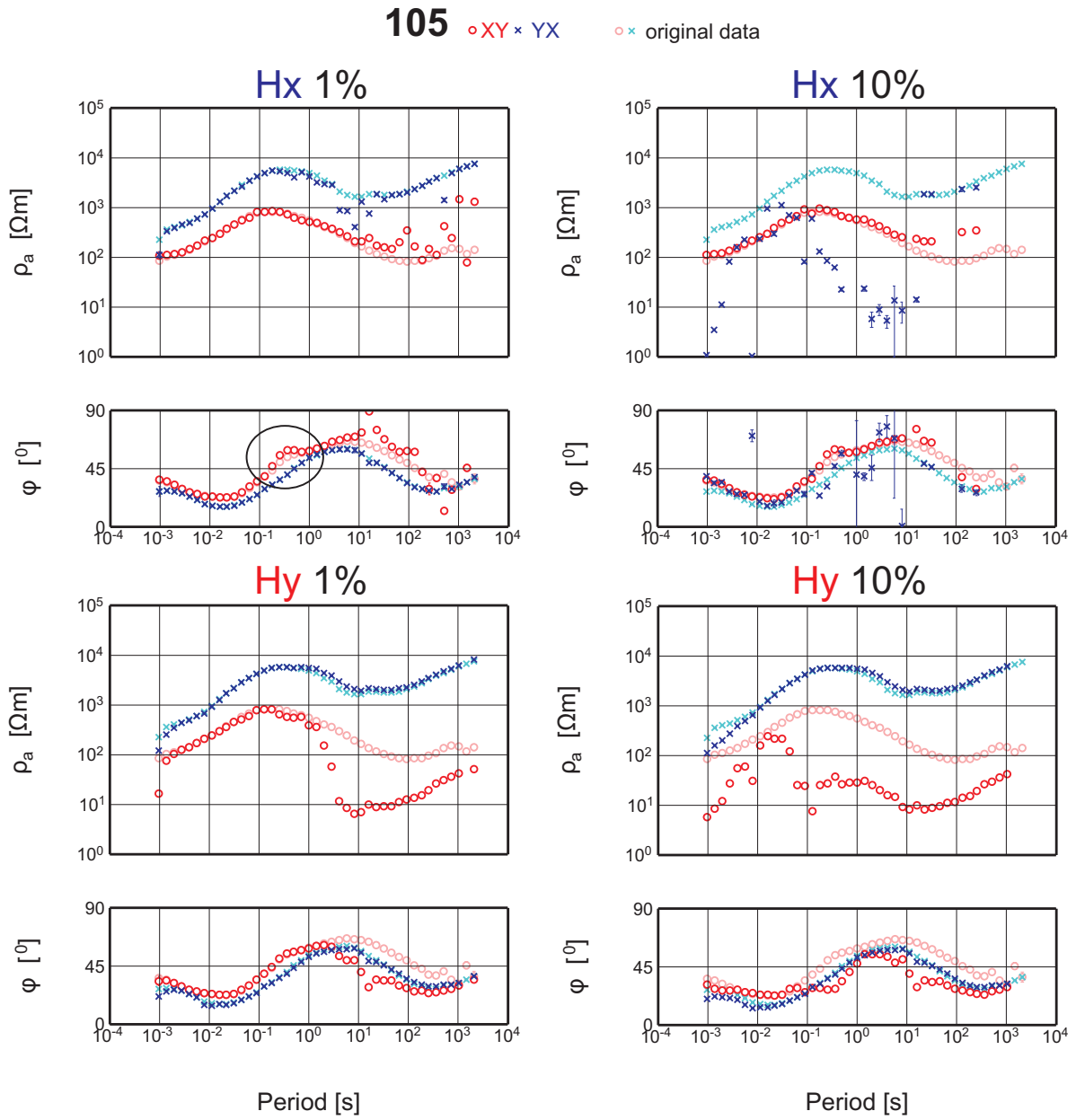


Figure 2.18: Apparent resistivity ρ_a and phase φ for rotated data (10.5^o anticlockwise) with 1 and 10% arbitrary spikes in either channel Hx or channel Hy ($n_{mult} \in [0, 10)$). Original (rotated) curves are shown in the background in lighter colours.

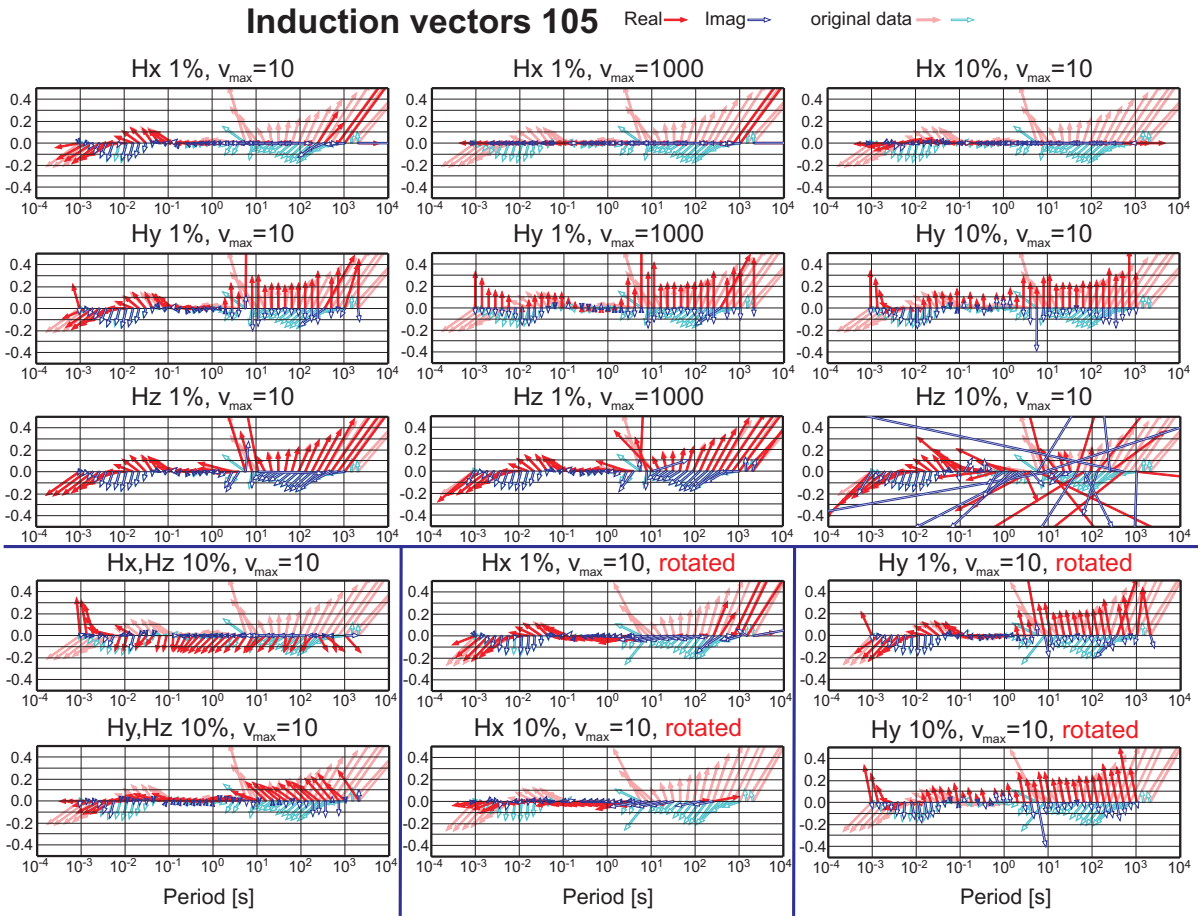


Figure 2.19: Induction vectors for varying values of v_{max} and percentages of arbitrary spikes in channel Hx, Hy and Hz. Data is not rotated unless otherwise stated (declination angle of 10.5^o anticlockwise). Original vectors are always shown in the background in lighter colours.

2.2 Classic processing methods: Remote reference and delay filter

Out of the large number of approaches to remove noise, basically two methods have evolved as standard processing tools for Magnetotellurics: Remote reference (RR) and delay filtering. They were implemented into the *EMERALD* package (Ritter et al., 1998; Weckmann et al., 2005; Krings, 2007; Chen, 2008; cf. sec. A.4) as additional processing tools prior to this thesis. The two methods serve as a first step to improve the data quality of the higly EM noise contaminated MT data set from the Barberton area (sec. 1.2). To further illustrate the performance of the RR and delay filtering method, the techniques are applied to the synthetically distorted data described in section 2.1, as well.

The different types of noise observed in data need to be addressed appropriately in order to effectively remove them. The RR method uses the close relationship of the horizontal magnetic field components at a local and a reference MT site evoked by the natural EM fields to cancel out uncorrelated noise in the data. The delay filter attenuates (noise) signals with a specific base frequency and its multiples. The more figurative term comb filter is used for it as well. In contrast, the also widely used notch filter only attenuates signals at one frequency. The usage of either delay or notch filter depends on the noise observed in the data. If multiples are visible within the frequency spectrum, the application of the delay filter might be more appropriate. In any case, both filters should be tested. For the Barberton data set, the

usage of the delay filter yielded a higher data improvement. Therefore, only the delay filtered processing results will be shown in this work.

The remote reference technique was first suggested and applied by Goubau et al. (1978) and Gamble et al. (1979), respectively. The general idea is to use the magnetic data of a remote site to improve the data quality of a local site. A basis for the method is the close relationship (e.g. in the sense of their coherency) of local and remote horizontal magnetic field components even over long distances (cf. Fig. 2.20 and 2.21, bottom). Naturally, only simultaneously recorded data can be used for this method.
The basic condition for the success of the method is that the noise in the local site is not correlated with the noise in the reference site (Fig. 2.20). If this is fulfilled, the uncorrelated noise is canceled out in both the local and the reference site. An additional benefit of the method is that it circumvents the systematic bias (e.g. Swift, 1967; Sims et al., 1971) in the MT transfer functions (TF) caused by the usage of autopowers of data channels which contain noise within the TF calculations.
Finding and maintaining a reference site throughout a MT field campaign can be challenging. This is due to the higher distance and remoteness usual for such a site and for providing the necessary power for running the measurement equipment. The minimal separation of the reference from the local site depends on the distance where the noise in both of the sites is no longer correlated. Depending on the conductivity of the subsurface and the power of the noise source, this distance may vary between a few kilometres and several hundreds of kilometres. Both, higher subsurface conductivities and a low signal power lead to a faster attenuation of noise. In these cases, the reference site can be closer to the local site.
In the times when the RR method evolved, the maximal distance limitations were due to limitations in telemetry (4.8 km in Gamble et al. (1979)). Today, via GPS exact scheduled measurements are possible to record data at different sites simultaneously making remote sites with a distance up to several hundred kilometres feasible (350 km in Oettinger et al., 2001; 215 km in Shalivahan and Bhattacharya, 2002; 800 km in Schill et al., 2011).
Due to the required similarity of the magnetic fields at the local and remote site, the RR method is usually targeted at longer period data ($> 10^{-2}$ s) especially for far distanced remote sites. However, Schill et al. (2011) even reported data improvements for short periods ($< 10^{-2}$ s) with a reference site 800 km away.
In general, only magnetic reference data is used, which is mainly due to the mostly higher noise level of the electric channels. Furthermore, the electric fields are stronger influenced by the local geology at the sites (e.g. Gamble et al., 1979; Bahr, 1988).

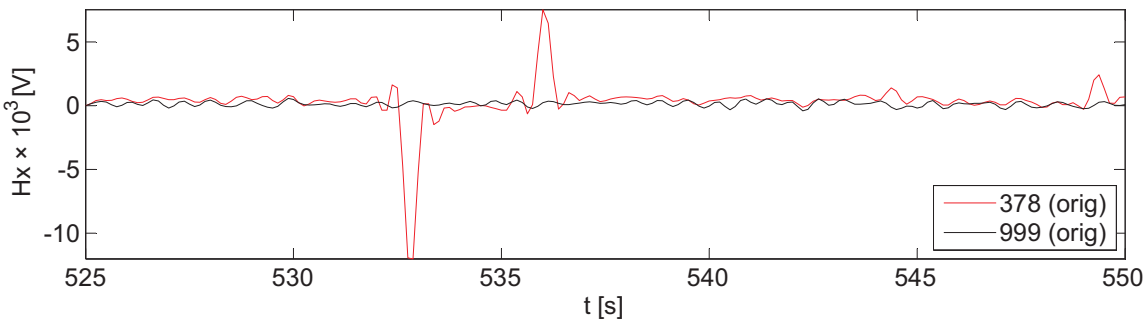


Figure 2.20: Time series of sites 378 and 999 (position of sites cf. Fig. 2.26). Clearly visible is the uncorrelated noise in site 378 but also the similarity of site 378 and 999.

Cultural electromagnetic signals in the form of harmonic oscillations can often be found within both electric and magnetic MT time series (Fig. 2.21, top). They can be caused by any electric device which uses an alternating current and powerlines providing the electricity for the devices. The closer such a noise source is to the MT site, the more the time series are dominated by the cultural signal masking the desired signal caused by the natural electromagnetic fields. The Fourier transform of such a contaminated MT time series shows a peak in the frequency spectrum at the base frequency of the cultural signal (Fig. 2.22, top). Often the integer multiples (harmonics) of this base frequency are dominant and expressed as peaks in the spectrum as well. In order to remove this type of noise from the MT data a delay filter can be used. A digital version of the filter is implemented in the *EMERALD* package as described in Chen (2008). If only a base frequency is observed in the spectrum (and no multiples) a notch filter can be applied to remove the noise.
The general idea of the delay filter is to add the time shifted signal to itself. This causes interferences erasing but also enhancing different parts of the signal, which can be seen in the frequency spectrum as well (Fig. 2.22, bottom). A mathematic form to express the filter is given by

$$\mathrm{ts}_{del}(n) = \mathrm{ts}_{orig}(n) + \alpha \cdot \mathrm{ts}_{orig}(n - D). \tag{2.6}$$

$\mathrm{ts}_{del}(n)$ denotes the delay filtered form of the original time series $\mathrm{ts}_{orig}(n)$ at the sampling points n, α is a scaling factor and D is the value the time series gets shifted by. This shifting value is called delay as well, giving the filter its name.
In most cases, α is either set to 1 or -1. For the described type of noise $\alpha = -1$ is chosen. Hence, signals at D/N, where $N = 1, 2, \cdots$, get filtered while the signals in between get enhanced. Important to note is that if $\alpha = -1$, the delay filter attenuates long period data as well which is a clear disadvantage of the method.
Finally, the delay D needs to be chosen in such a way that it corresponds to the base frequency of the cultural signal. That means that it must be an integer multiple of the period of the noise.
If the base frequency of the noise and its multiples are not equally spaced over frequency the delay filter will not work properly. In this case, using the notch filter (even multiple times) might be more appropriate.

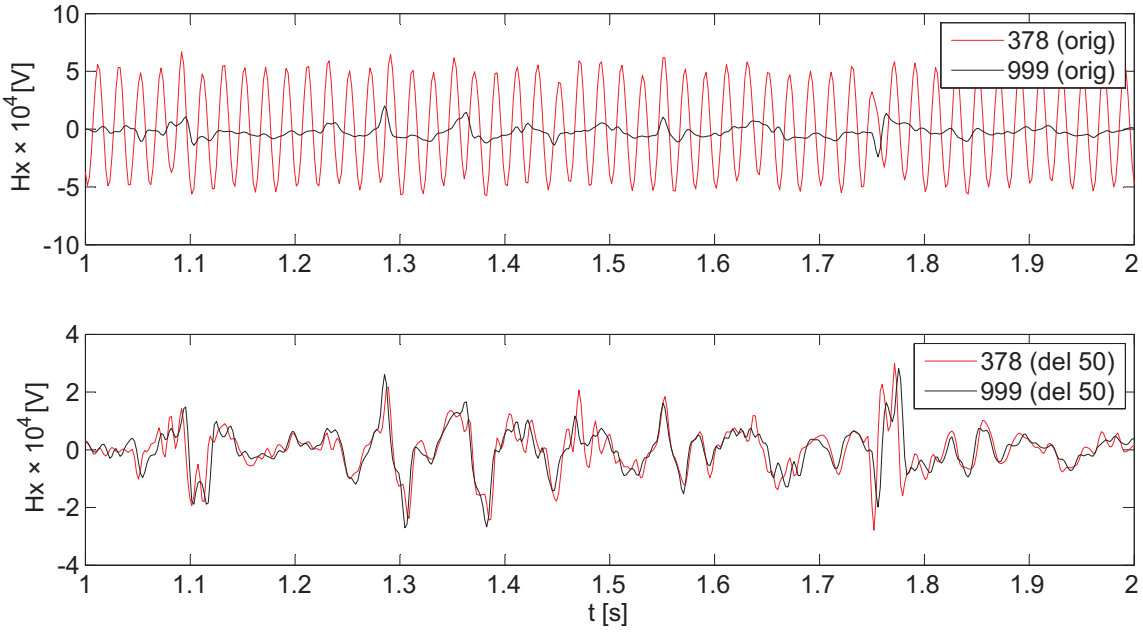


Figure 2.21: Time series of sites 378 and 999 (position of sites cf. Fig. 2.26): Original time series (top) and after applying a delay filter at 50 Hz (bottom).

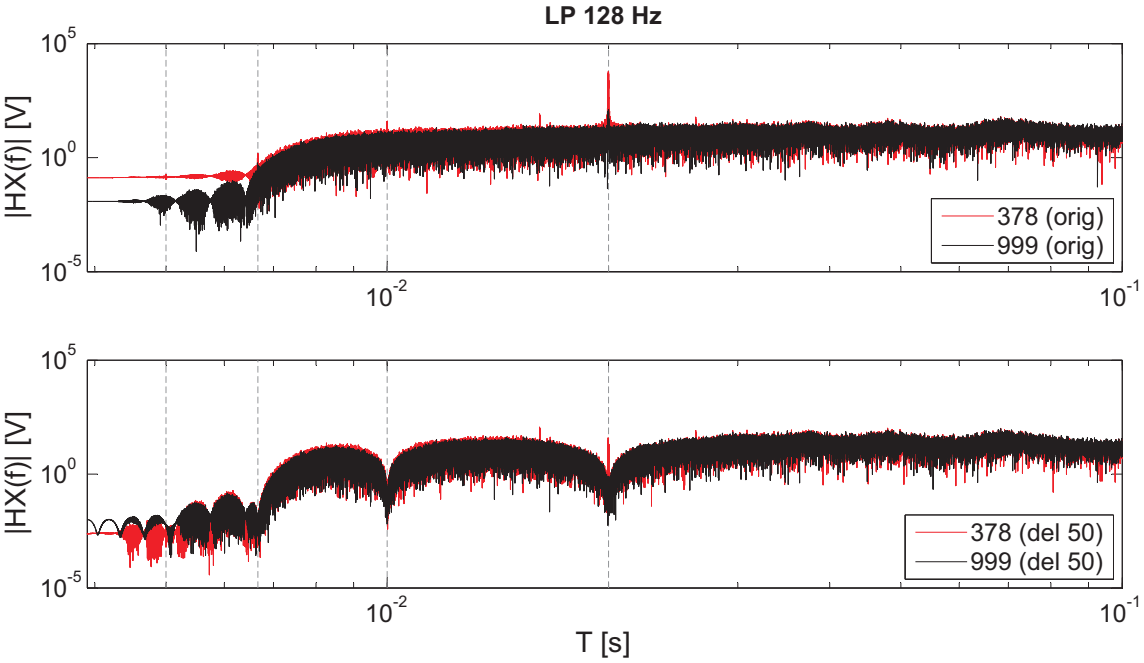


Figure 2.22: Frequency spectra of sites 378 and 999 (position of sites cf. Fig. 2.26): Original spectra (top) and after applying a delay filter at 50 Hz (bottom). Marked with grey dashed lines are 50, 100, 150 and 200 Hz (0.02, 0.01, $0.00\overline{6}$ and 0.005 s, respectively). No. of samples for spectra calculations: LP 128 Hz: 262,144 (for filtering details see Tab. A.1).

Application of remote reference and delay filter

The performance of the two classic methods for removing synthetic noise will exemplarily be shown for *arbitrary spikes* and *steps* (cf. sec. 2.1.2) since they are the most complex type of synthetic noise presented in this work. Moreover, this noise type is closest to the observed 'real' noise in the BGB

data set (e.g. Fig. 1.6, 2.12 and 2.13), as well. Data distorted with either 10% noise in Hy (hereafter abbreviated as ***ex. 1***), in Ex (***ex. 2***) or in all of the channels (***ex. 3***) are chosen as examples. For the RR method site 103 is used as reference site (Fig. 2.2). From the data set of the BGB, the processing results of sites 061, 173 and 378 are exemplarily shown in this section.

Synthetic noise

The remote reference method is applied to data starting from around 10^{-2} to over 10^{3} s (bands LP 128 Hz to LP 128 s, see Fig. A.2). Shorter periods are not included in the processing because the EM fields in this period range have mostly a more local character (Gamble et al., 1979).
Overall, the performance of the remote reference method to remove the influence of the synthetic noise is rather poor for all three examples. For ***ex. 1*** (Fig. 2.23) improvements can only be seen for periods up to $\sim$ 1 s, especially in the phase and the real part of the induction vectors. For the yx component even a worsening of the data quality can be observed between 1 and 10 s which seems to be related to the general lower signal to noise ratio within the 'dead band': Even a slight disturbance of the data in this period range might cause a significant drop in the data quality. The amount of data rejected due to the RR method reduces the amount of data used for the robust statistics as well.
The further data degredation after applying the RR method in ***ex. 2*** (Fig. 2.24) is due to the presence of noise only in the x-component of the electric field. No data improvement can be achieved since only the magnetic channels are treated by the RR method. This example shows that any unnecessary or additional data filtering might alter the data in an unwanted way.
For ***ex. 3*** (Fig. 2.25), the data improvement is limited to an even smaller period range than for ***ex. 1*** ($\lesssim 10^{-1}$ s). Any data points within the apparent resistivity curves that match their original level after the remote referencing, do not show the same improvement for their corresponding phase values. Even if both phase and apparent resistivity would have their original values, identifying the few correct data points is difficult without the comparison to the true (undistorted) curves.

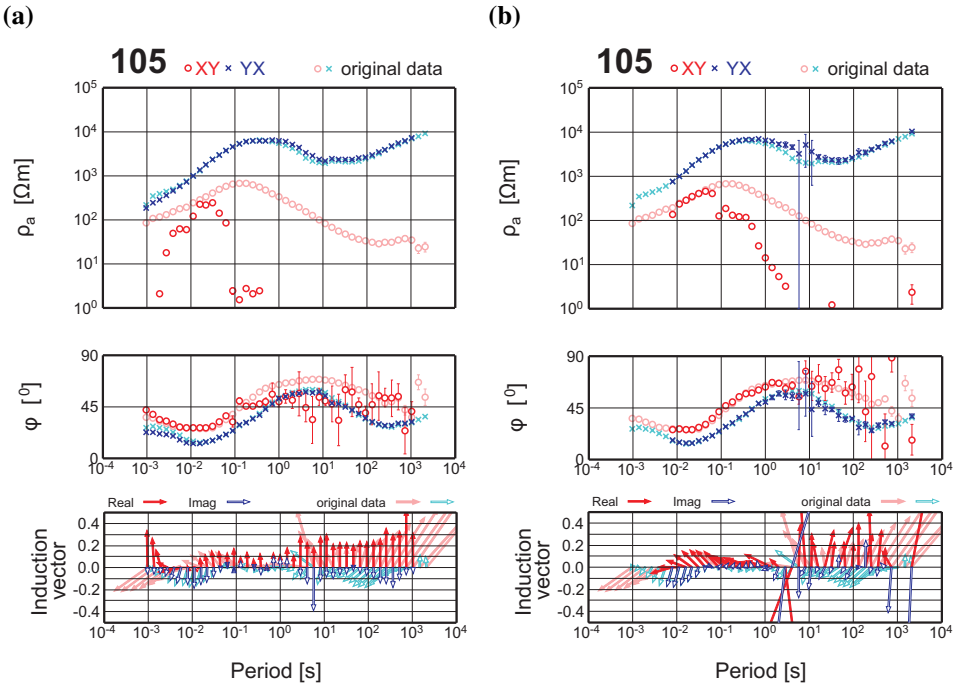


Figure 2.23: ***Ex. 1***: Apparent resistivity ρ_a, phase φ and induction vectors with 10% arbitrary spikes in Hy **(a)** ($n_{mult} \in [0, 10)$) and after applying remote referencing **(b)**. Original curves are shown in the background in lighter colours. Some data points for short and long periods in **(b)** got rejected because they were either missing in the Ex or Ey component when they were assembled to the final transfer functions.

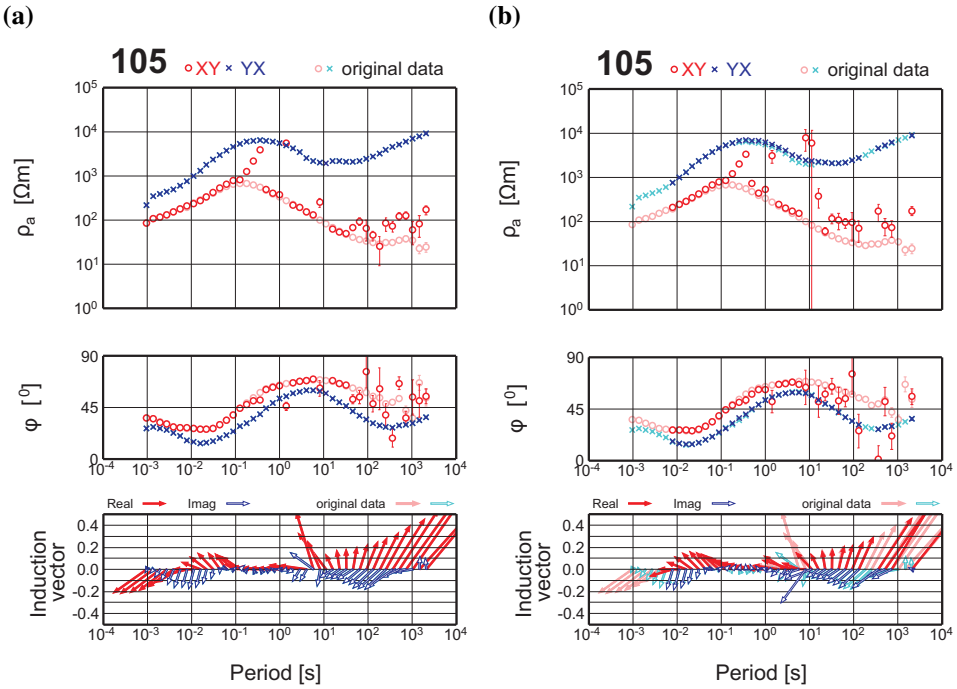


Figure 2.24: ***Ex. 2***: Apparent resistivity ρ_a and phase φ with 10% arbitrary steps in Ex **(a)** ($n_{mult} \in [0, 10)$, $o_{user} = 0.001$) and after applying remote referencing **(b)**. Original curves are shown in the background in lighter colours.

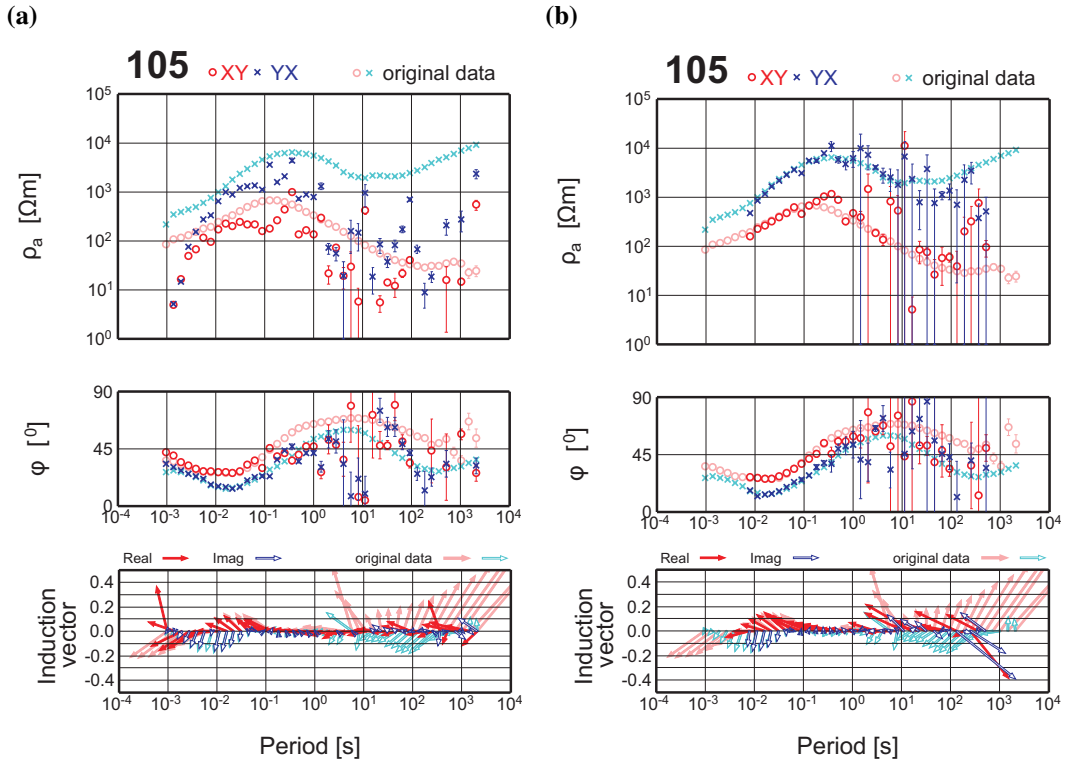


Figure 2.25: ***Ex. 3***: Apparent resistivity ρ_a and phase φ with 10% arbitrary spikes and steps in all magnetic and electric channels, respectively **(a)** ($n_{mult} \in [0, 10)$, $o_{user} = 0.001$) and after applying remote referencing **(b)**. Original curves are shown in the background in lighter colours.

For the synthetic noise examples, testing the performance of the delay filter is impossible. The synthetic noise introduced in section 2.1 does not have one specific frequency but instead exhibits a random distribution within the time series. This can also be seen in the various amplitude spectra (e.g. Fig. 2.7 and 2.11): Instead of a distinct peak of amplitude at one frequency, wide parts or the whole spectrum gets shifted. Selecting a specific frequency necessary as input parameter for the delay filter is therefore unfeasible.

MT data from the Barberton Greenstone Belt

The MT data from the southern part of the Barberton Greenstone Belt (BGB) was collected during two field campaigns. During the first field experiment in 2009, finding a remote site in the vicinity proved to be extremely difficult due to the high level of cultural EM noise. Locations up to 350 km away were tested. Reference sites located further away from the measurement area would have been difficult to be maintained. Two sites (201 and 202, Fig. 2.26) were finally set up, both exhibiting a rather poor data quality (Fig. 2.27, right). Since site 202 has a slightly better quality, it is used as reference for the 2009 data (blue crosses in Fig. 2.26).

For the second field experiment in 2010 (blue stars in Fig. 2.26), a site was set up in an area which was known for its good data quality from a previous MT field campaign. Unfortunately, during the presented measurements it exhibited poor data quality in the dead band and for long periods due to the low solar activity at this time (Fig. 2.27, left), making the application of the RR method challenging. Due to the large distance of about 1,200 km, the data quality of the site could not be checked during the 2010 field campaign and the data was retrieved subsequent to the actual measurements.

If the data of a local site is filtered (e.g. with a delay filter) prior to the application of the RR method, the reference site needs to be filtered as well. Otherwise, the data of the local site is distorted by the RR processing because uncorrelated noise is artificially added (e.g. 50 Hz signal is present in the reference site but not in the local site). For the delay filter, spectra are calculated (Fig. 2.22) to select frequencies where peaks in the amplitude occur. A clear peak can be identified at 50 Hz for nearly all sites. Multiples at 100 and 150 Hz are usually marked as well by an increase of the amplitude. Smaller peaks, e.g. at ~ 0.0161 s (~ 62 Hz, used by electric chipping hammers and other mining equipment), can sometimes be observed as well.

Having a closer look, these peaks are not as narrow as they appear at first. Hence, I tested the delay filter not only with 50 Hz base frequency but also with slightly offset values (ranging between 48.5 and 51.5 Hz). Nevertheless, the results varied only marginally and 50 Hz was selected as delay filter frequency.

For several sites, further tests with other frequencies (e.g. 62 Hz) were carried out, using both delay and notch filter. Even if an improvement of the data quality was observed, data improvement for 50 Hz delay filtering was always more pronounced.

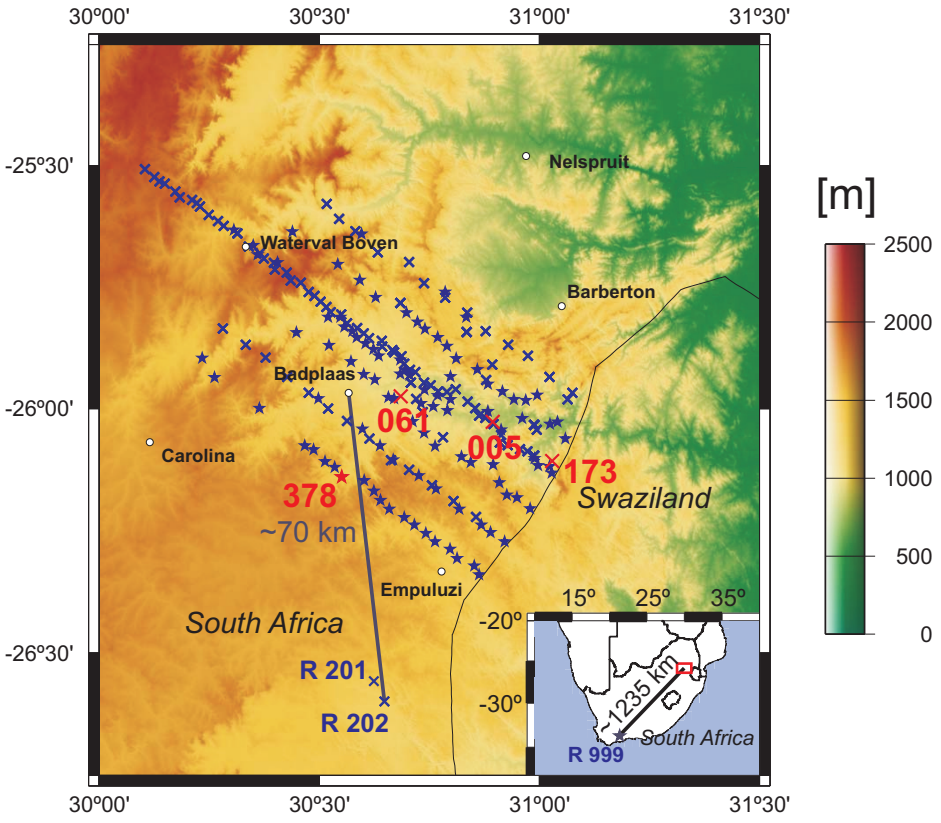


Figure 2.26: Location of reference sites: R 201 and R 202 for sites recorded in 2009 (marked by blue crosses) and R 999 for sites recorded in 2010 (marked by blue stars, inset map).

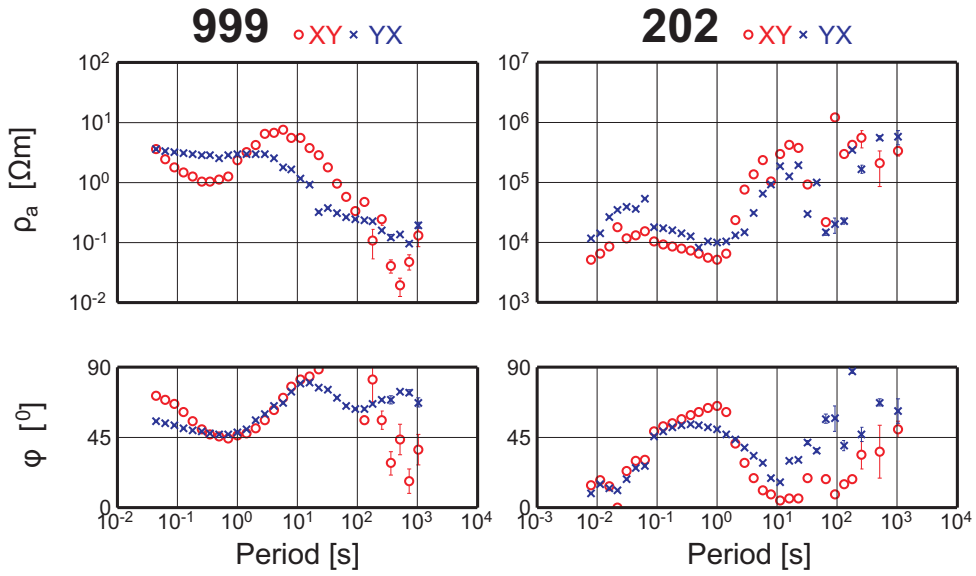


Figure 2.27: Reference sites 202 and 999 (single site processing, position of sites cf. Fig. 2.26).

The first two data examples from the BGB are sites 061 and 173 from the campaign in 2009. During the data recording at site 061 (site position marked in Fig. 2.26), reference site 202 was not yet installed. However, the simultaneously running site 005 could succesfully be used as remote site (single site processing results for site 005 cf. Fig. 1.4).

To clearly see the filtering effect, delay filter and remote reference are applied seperately to the data. For site 061 (Fig. 2.28), the delay filter could only improve the data at a few selected points (between 10^{-2} and 10^{-1} s). However, the RR method helped to recover at least the xy component and induction vectors up to periods over 10^{-1} s.

For site 173 (Fig. 2.29), the delay filter improved the data in the same range as for site 061. The RR method with site 202 as reference helped to restore the data to over 1 s.

For site 378 (Fig. 2.30), which was recorded in 2010, improvements can be seen for the delay filter comparable to the previous examples. For the RR method site 999 is used as reference, restoring the data of site 378 over a wide period range. The better data quality of the yx component of the reference site for periods even around 10^2 s can be viewed as the cause for the higher improvements in the yx component of the local site. Unfortunately, the induction vectors still exhibit a chaotic behaviour which can be attributed to the poor data quality of the induction vectors of the reference site.

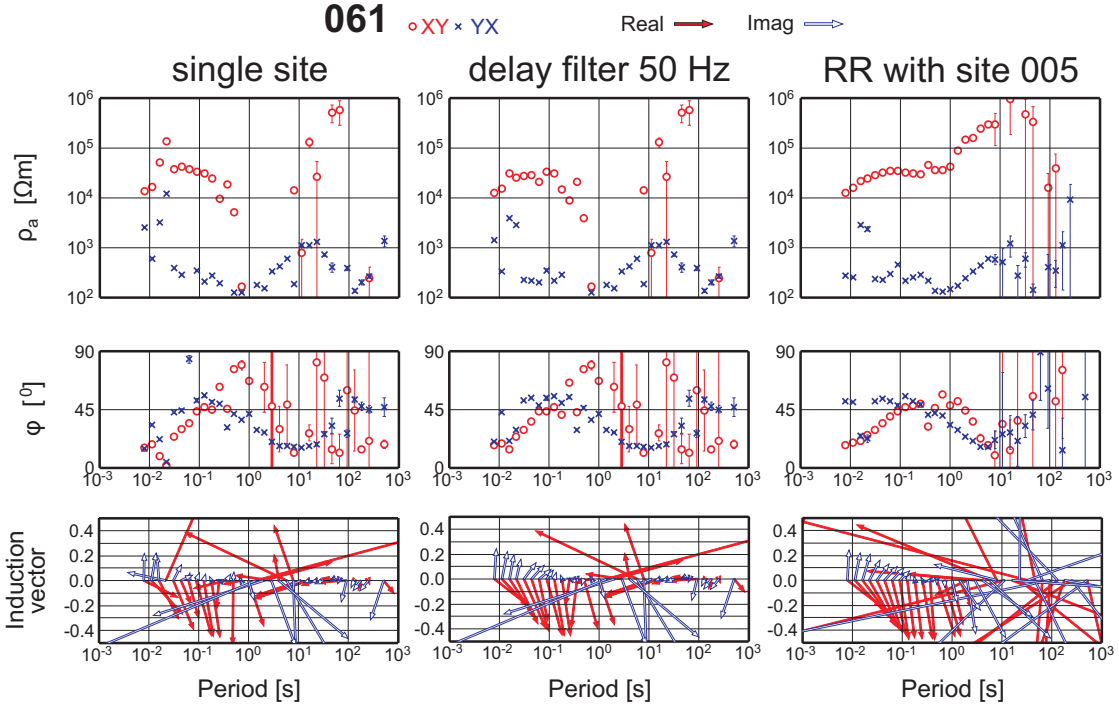


Figure 2.28: Site 061: Results from single site processing (left), delay filtering with 50 Hz (centre) and remote reference processing with site 005 (right, position of sites cf. Fig. 2.26). Some improvements of the data quality can be seen between 10^{-2} and 10^{-1} s for the delay filter. The improvements for the RR processing are more pronounced.

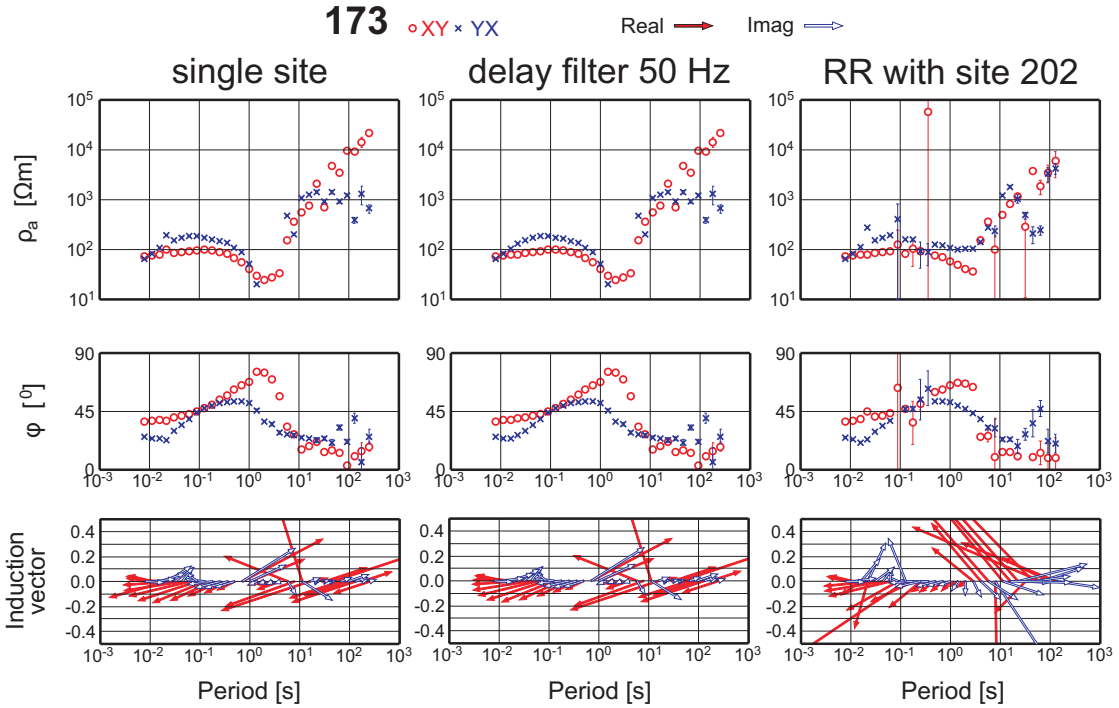


Figure 2.29: Site 173: Results from single site processing (left), delay filtering with 50 Hz (centre) and remote reference processing with site 202 (right, position of sites cf. Fig. 2.26). Some improvements of the data quality can be seen between 10^{-2} and 10^{-1} s for the delay filter. Due to the poor data quality of site 202, the RR processing led to a data distortion of site 173.

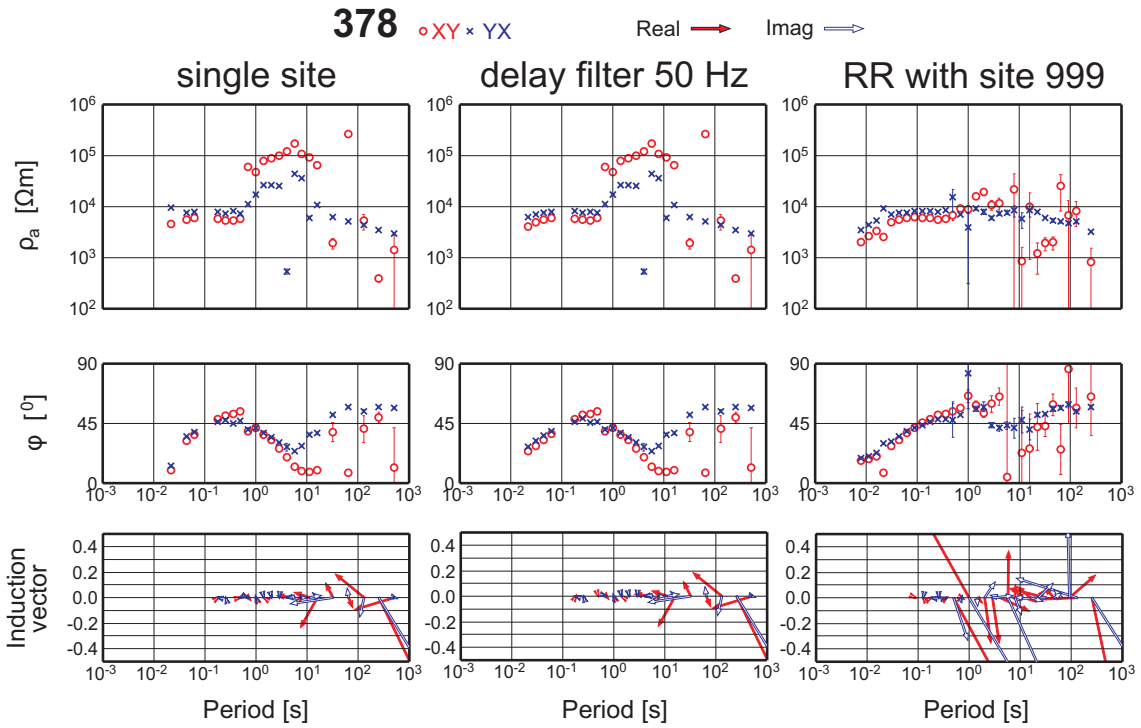


Figure 2.30: Site 378: Results from single site processing (left), delay filtering with 50 Hz (centre) and remote reference processing with site 999 (right, position of sites cf. Fig. 2.26). Both the delay filter and the RR method improved the data quality (of ρ_a and φ) significantly.

Discussion of the filtering results

The delay filter and the remote reference technique could successfully improve the data in a period range between 10^{-2} and 1 s for a large number of the BGB sites. For MT sites from 2009, correlation of noise in the local and reference site 202 for the longest recorded periods cannot be ruled out in this highly resistive geological setting (skin depth of 150 km for 10^2 s and 1,000 Ωm, cf. app. A.2). Including the poor data quality of the reference sites for long periods, this might lead to the generally bad results for periods over 1 s.

Regarding the tests with synthetic noise, the RR method failed over a wide period range to improve the data. This is even more severe considering the excellent data quality of reference site 103 from Namibia (Fig. 2.2). This shows that the RR method is not well adapted to removing the synthetic random spike-type noise even if it only affects one of the magnetic channels. Since similar noise can be found in the BGB data as well, even a high-quality reference site might have not helped to sufficiently improve the data for long periods.

2.3 Treatment of long period MT data (T > 1 s)

To resolve deep reaching geological structures with MT, long period (> 1 s) data are essential. Unfortunately, at most sites the quality of the Barberton MT data is only sufficient for further analysis (e.g. through inversion) up to 1 s despite the extensive application of delay filter and remote reference technique (see sec. 2.2).

One of the main problems for the calculation of transfer functions at longer periods is the considerable

amount of spikes in the magnetic channels corresponding with steps in the electric field (cf. chapter 1). As the synthetic noise tests (sec. 2.1) have shown, this type of random noise impairs in particular the data quality at long periods. Additionally, this is enhanced by the lower amount of available data, due to longer recording times necessary for long periods. The data processing methods applied so far are not capable to remove such noise (cf. sec. 2.2). This is confirmed by applying them to synthetically distorted data, as well.

Spikes and steps are easy to detect visually, however, their manual detection and removal from the large amount of data (193 sites × 5 channels × 4 days of data recorded in my case) is not feasible. Reliable automated techniques are therefore necessary to accomplish such a task. In the long list of filtering and processing techniques developed so far, the IARWR (**I**ntersite **A**ctivity **R**atio **W**iener-filter **R**eplacement) filter introduced by Kappler (2012) appears to be best adapted to tackle random noise identified within the BGB data set. The IARWR algorithm automatically detects spikes and replaces them with predicted data using at least one additional site which was recorded simultaneously. In this section, its implementation as realized in my work is presented. Since the original IARWR algorithm has problems identifying steps, I especially concentrated on the improvement of their detection. Different from Kappler (2012), data of the reference site is only used for the calculation of the predicted data for the local site and is not altered by the filtering process. Hence, the approach is called *reference site filter* throughout this work.

Because of various reasons, a remote MT site might either not be available (as for parts of the 2009 data set) or its usage is not desirable due to, e.g. correlated noise (cf. sec. 2.2). Based on the STA/LTA (**S**hort **T**ime **A**verage/**L**ong **T**ime **A**verage) trigger developed to detect earthquakes in seismological recordings (Trnkoczy, 2002) and the knowledge gained from the implementation of the *reference site filter*, I introduce a filter method which detects spikes and steps in MT time series and removes them by a simple interpolation approach without using any additional data from a reference site. This filter solely works in time domain and is called *STA/LTA filter* throughout this work.

For a better understanding of the filter process and to test the strengths and limitations of the *reference site* and the *STA/LTA filter*, they are not only applied to the BGB MT sites but also to the synthetically distorted data.

2.3.1 Reference Site filter

The original concept of the *reference site filter* was introduced by Kappler (2012) and further developed by Kütter and Weckmann (2012). Its concept consists of the following steps.

1. The time series of a local site and one or more reference site(s) are parsed into windows.
2. By comparing the data variance in a window of the local site with the reference site variances, windows containing spikes or steps are identified and flagged.
3. Using the data of the reference site, a Wiener filter algorithm is applied to calculate new data to replace the flagged windows in the local site.

A detailed description of this approach and the further developments is given in the following.

The *reference site filter* can be called a hybrid method since it uses both time and frequency domain. Whereas the detection of noise and its replacement by predicted data work in the time domain, the replacement data itself is originally calculated in the frequency domain.

1. Data parsing and calculation of window variances

The electric (E) and magnetic (B) time series of a local and one or more reference sites are parsed into windows (Fig. 2.31 a). The window length (WL) has to be a power of two, e.g. $2^8 = 256$ samples, for the Fast Fourier Transform calculations carried out in step 3. Subsequent windows may have a user-defined overlap (OV) of e.g. 16 samples (Fig. 2.31 b). This increases the length of the time series artificially and might help to detect spikes or steps that are directly at the border between two neighbouring windows. Together with the length of the time series, the parameters WL and OV determine the overall number of windows N.

For each channel in a window i the variance value a_i is calculated using

$$a_i \;\widehat{=}\; VAR(\mathbf{Y}) = \frac{1}{WL}\sum_{j=1}^{WL}(y_j - \mu)^2, \tag{2.7}$$

$$\text{with the mean value } \mu = \frac{1}{WL}\sum_{j=1}^{WL} y_j. \tag{2.8}$$

$\mathbf{Y}$ is the data from one channel in window i, consisting of WL data points y_i. The variance describes the range the data points in one specific window deviate from the expected value μ.

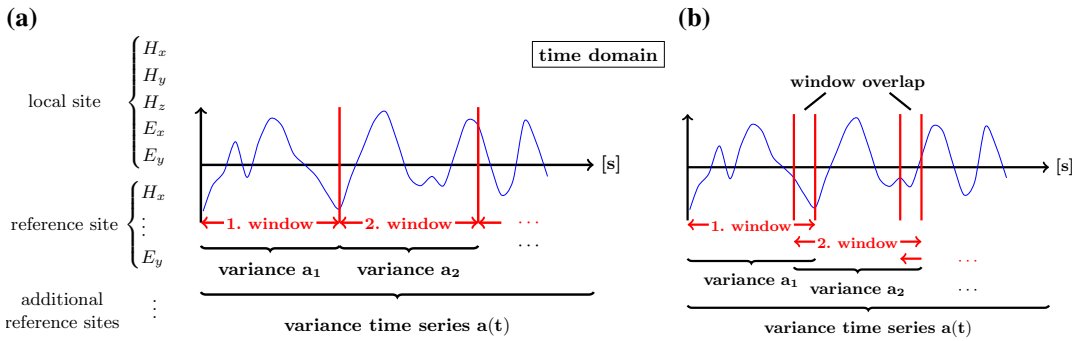


Figure 2.31: Time series of electric and magnetic channels of a local and one or more reference sites are parsed into windows without **(a)** and with overlap **(b)**.

The variance calculations for every window finally lead to a set of variance time series for all channels of the local (Fig. 2.32) and the reference site(s).

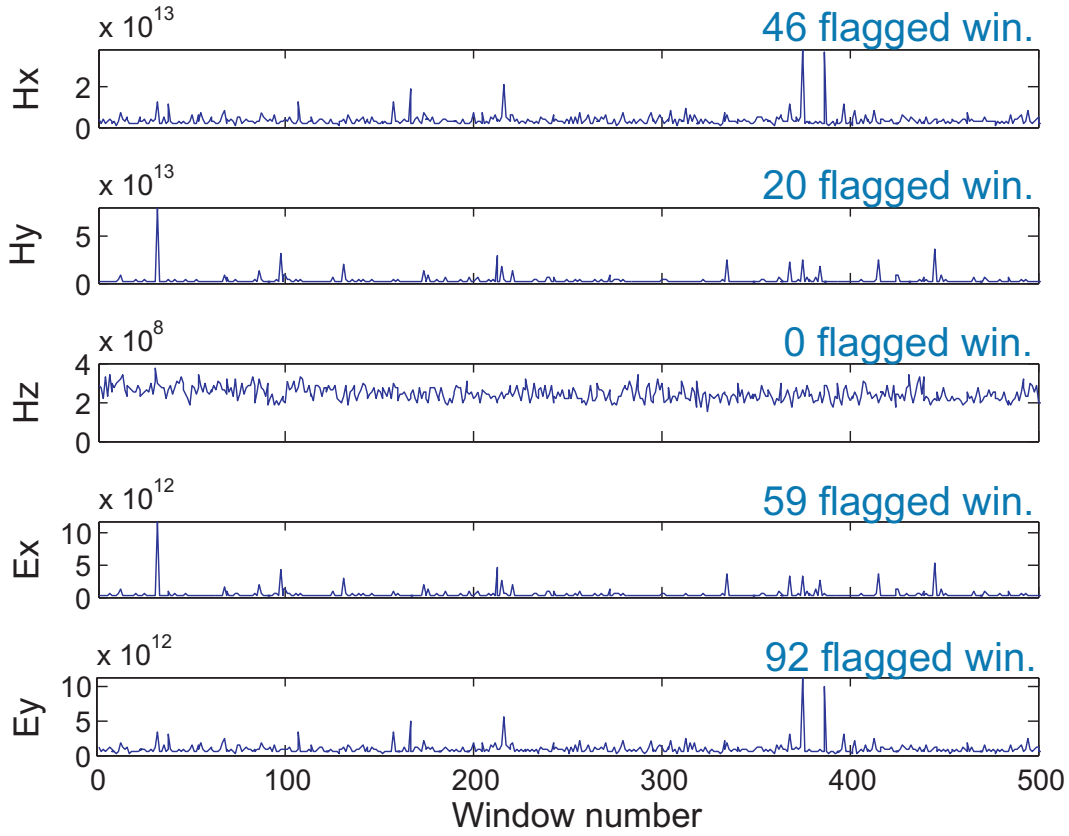


Figure 2.32: Variance time series of five data channels of the local site. Specified are the number of flagged windows for each channel. Note the scale difference between channel Hz, where no windows are flagged, and the other channels.

2. Calculation of variance ratios and window flagging

Outliers in the time series such as spikes or steps result in high variance values a_i. However, elevated variances can also be caused by an episodically higher activity of the electromagnetic source fields. To distinguish between these two cases, variance ratios r of the local and the reference site are used (eq. 2.9).

$$r_{i,j} = log_{10}\left(\frac{a_i(t)}{a_j(t)}\right) \tag{2.9}$$

In equation 2.9 the indices i and j refer to a local (e.g. site 300) and a reference site (e.g. site 999), respectively. If there is more than one reference site, the ratios will be calculated for all combinations of the local site with the reference sites (Fig. 2.33).

If there is a higher EM source activitiy, both the local and the reference site are affected and have higher variances, resulting in low ratios $r_{i,j}$. In the presence of noise only the local site has an elevated variance and the resulting ratio $r_{i,j}$ shows a maximum. Ideally, the reference site should be noise-free to detect spikes in the local site. A noisy reference site decreases the noise detectability of the filter. If the noise in the local site is correlated to the noise in the reference site, the resulting variances $a_{i,loc}$ and $a_{i,ref}$ are similar. This leads to variance ratios close to zero and the noisy window is not detected by the algorithm. In this case, an additional reference site, located further away from the local site to avoid correlated noise, might be helpful.

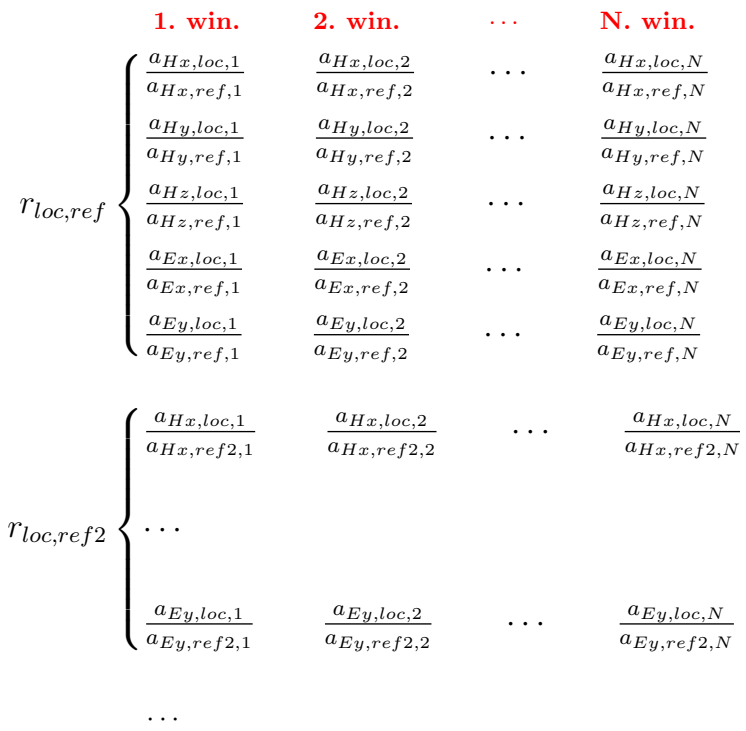


Figure 2.33: Scheme for the calculation of the variance ratios r of a local site (loc) and one or more reference site(s) (ref, ref2, . . .).

To identify the windows which have high variances and thus contain noise, threshold values need to be defined. This comprises the computation of the α-trimmed standard deviation σ_α (eq. 2.10) of all variance ratios given by

$$\sigma_\alpha = \sqrt{\frac{\sum_{i=1}^{N}(r_i - \overline{r})^2}{N-1}}. \tag{2.10}$$

$\overline{r}$ is the mean value of the variance ratios. α-trimming means that a user-defined percentage of the extremal values of the variance ratios will be omitted in the calculations. σ_α is calculated for every channel.

To control the level of the threshold, a user-defined multiplier t_{user} is determined. The multipliers for magnetic and electric channels can be set independently to account for different noise levels in the two channel types.

Since the variance ratios can be positive or negative values, two thresholds are needed. The lower boundary is defined as the negative of $t_{user} \cdot \sigma_\alpha$ (Fig. 2.34).

Variance ratios $r_{i,j}$ contain information about the variance of both local and reference site. Of course, it is possible that the reference site contains noise whereas the local site in the particular window is free of it. To now identify the site which needs to be flagged, the variance values of the local and reference site are compared and the site with the higher variance is chosen. For the subsequent calculation of the replacement data, it is important for the algorithm to memorize the windows of the reference site which contain noise as well. These flagged reference site windows are not used for any further calculations.

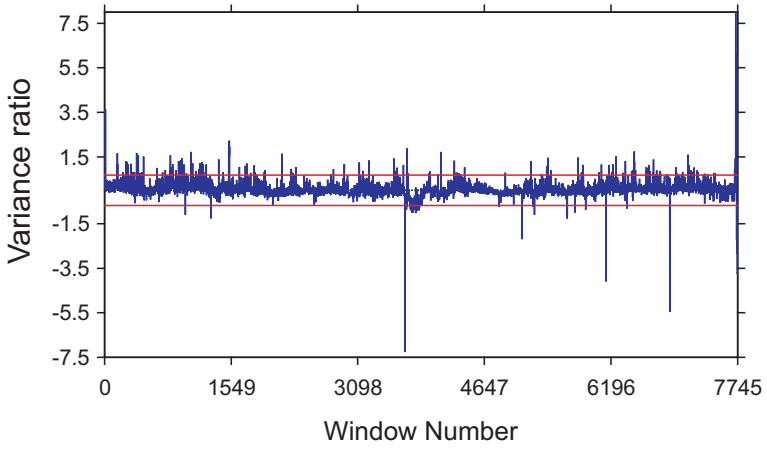


Figure 2.34: Variance ratios $r_{loc,ref}$ between a local and a reference site for all windows of a time series (channel Hx) with thresholds marked in red. For all windows where the variance ratio $r_{i,j}$ exceeds the threshold, the variance of the local and the reference site are compared to finally decide whether a specific window of the local site gets flagged for filtering.

For multiple reference sites, the thresholds are calculated for each local and reference site combination individually. If one of variance ratios exceeds the threshold, the window will be flagged. Hence, if the comparison of the local site variance with one of the reference sites fails to detect a noisy window in the local site, the comparison with another reference site might still be able to detect the noisy window. However, this might also lead to false detections. Therefore, choosing appropriate reference sites with a generally good data quality is certainly fundamental for the success of this filter.
To prevent that too many windows are flagged, e.g. if t_{user} is chosen too low, a minimum value for the threshold can be set. It should be chosen according to the average value of the time series.

3. Calculation of the replacement data

To prevent data gaps by simply cutting out noisy windows, predicted data is used to replace the original data. The calculation of the replacement data is based on the Wiener filter (Wiener, 1949). Originally, the filter is used in signal processing to reduce noise by comparing the noise-free signal spectrum with the noisy signal spectrum.
The input signal $g(t)$ (before filtering) is assumed to be of the form

$$g(t) \quad = \quad s(t) + n(t), \qquad \boxed{\text{time domain}} \tag{2.11}$$

where $s(t)$ is the unknown noise-free signal and $n(t)$ is the additive noise. To simplify the calculation of the noise-free signal, the input signal is transformed via a Fast Fourier Transform (FFT) into the frequency domain: the convolution required for the calculations in time domain (eq. 2.12) becomes a simple multiplication in frequency domain (eq. 2.14).

$$s(t) = \psi(t) * g(t) = \int_{-\infty}^{\infty} \psi(\tau) g(t-\tau) d\tau \qquad \boxed{\text{time domain}} \tag{2.12}$$

Here, $\psi(t)$ is the Wiener filter function in time domain which needs to be applied to the noisy signal $g(t)$ in order to obtain the noise-free signal $s(t)$.
In frequency domain, the Wiener filter coefficients $\Psi(f)$ are defined by the comparison of the spectra of

the noise-free signal $S(f)$ with that of the noisy signal $G(f) = S(f) + N(f)$.

$$\Psi(f) = \frac{|S(f)|^2}{|S(f)|^2 + |N(f)|^2} \qquad \boxed{\text{frequency domain}} \tag{2.13}$$

Then, the calculated noise-free signal spectrum $S_{calc}(f)$ is given by

$$S_{calc}(f) = \Psi(f) \cdot G(f). \qquad \boxed{\text{frequency domain}} \tag{2.14}$$

The basic assumptions underlying this filter are that (i) the signal and the noise follow stochastic processes and (ii) that their spectral properties (i.e. $|S(f)|^2$ and $|N(f)|^2$) are known. Nevertheless, even rather crude assumptions for the spectra $|S(f)|^2$ and $|N(f)|^2$, e.g. estimating them from the noisy signal spectrum $|G(f)|^2$ as sketched in Fig. 2.35, leads to robust filtering results.

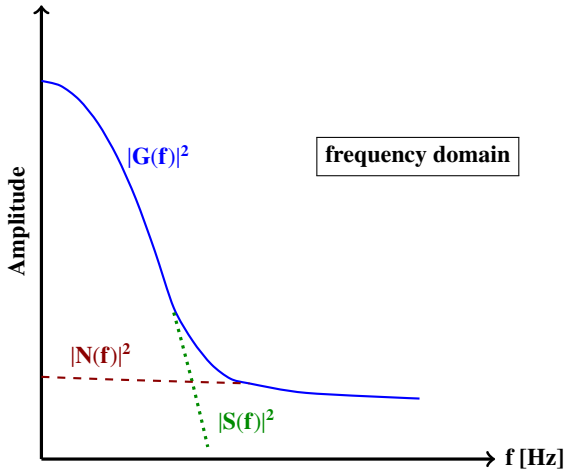


Figure 2.35: Conceptual way of estimating the spectral properties of the noise $|N(f)|^2$ and the noise-free signal $|S(f)|^2$ from the input signal spectrum $|G(f)|^2$ for the Wiener filter.

To calculate replacement data, the Wiener filter algorithm was modified by Kappler (2012). One of the major differences is the determination of the Wiener filter coefficients. Instead of estimating the spectrum of the noise-free signal, the noise-free data from the reference site is used. Therefor, an algorithm searches through the windows of the time series to find unflagged, i.e. noise-free, data. This so-called 'training data' is transformed to the frequency domain by applying a FFT and is then used to assemble a training data vector d_{tr} and the convolution matrix $\mathbf{M}$ (Fig. 2.36 a) and eq. 2.17). The Wiener filter coefficients Ψ can then be calculated by solving the equation system given by

$$d_{tr}(f) = \mathbf{M}\Psi. \tag{2.15}$$

For the solution of this system the Singular Value Decomposition (SVD) is used.
After obtaining the coefficients $\Psi(f)$, the unflagged data in the original window with the flagged channel will be transformed to the frequency domain and again the convolution matrix $\mathbf{M}$ is assembled (Fig. 2.36 b). A simple matrix vector multiplication with a subsequent inverse FFT (FFT^{-1}) leads to the aspired replacement data d_{calc}(t) for the flagged channel(s).

$$d_{calc}(t) = FFT^{-1}(\mathbf{M}\Psi) \tag{2.16}$$

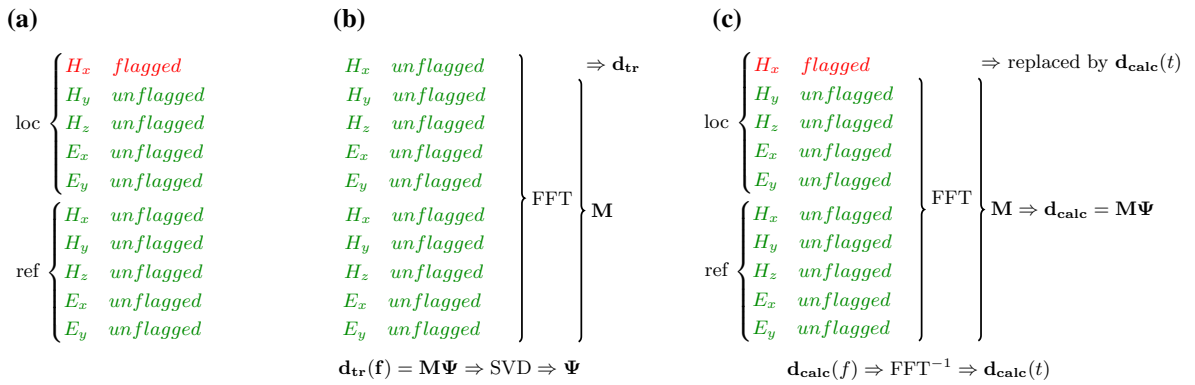


Figure 2.36: Scheme of the calculation of the replacement data: **(a)** Example where window Hx of the local site (loc) is flagged; **(b)** Search for training data where Hx is unflagged, assemble the equation system (convolution matrix **M**, training data vector $d_{tr}(f)$) and solve for the Wiener filter coefficients Ψ (singular value decomposition (SVD)); **(c)** Assemble a new equation system using the unflagged data of the original window and solve for the replacement data $d_{calc}(t)$.

Figure 2.37 shows an example of a detected spike in channel Hx and the data calculated to replace the flagged data in comparison. Prior to the spike, the replacement data has a higher similarity to the original data. After the spike, the replacement data differs considerably from the original data which is possibly due to a slow decay of the spike in the original time series. The spectra corresponding to the flagged data channel Hx, the replacement data d_{calc} and the unflagged data channel Hy are depicted in Fig. 2.38. The spectrum of the replacement data is much closer in its behaviour to the spectrum of the unflagged data.

To guarantee a smooth transition, the replacement data need to be fitted to the original data in the adjacent windows. In a transition zone of a user-defined length on either side of the flagged window, the calculated data is down-weigthed whereas the original data is up-weighted (Fig. 2.39).

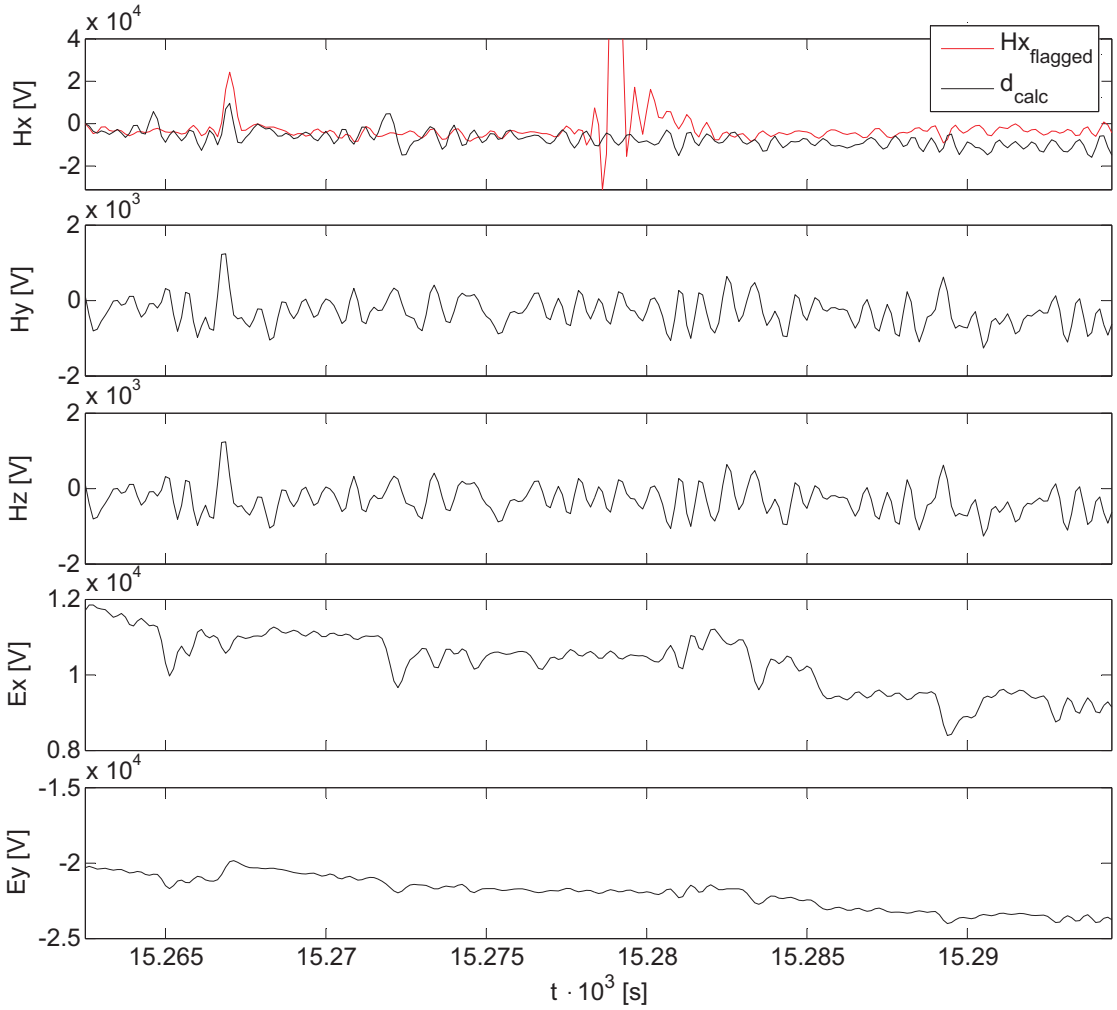


Figure 2.37: Time series window with a detected spike in channel Hx (red curve, $Hx_{flagged}$) and the calculated replacement data $d_{calc}(t)$ (black curve). The shown time series consist of 256 samples (≙32 s).

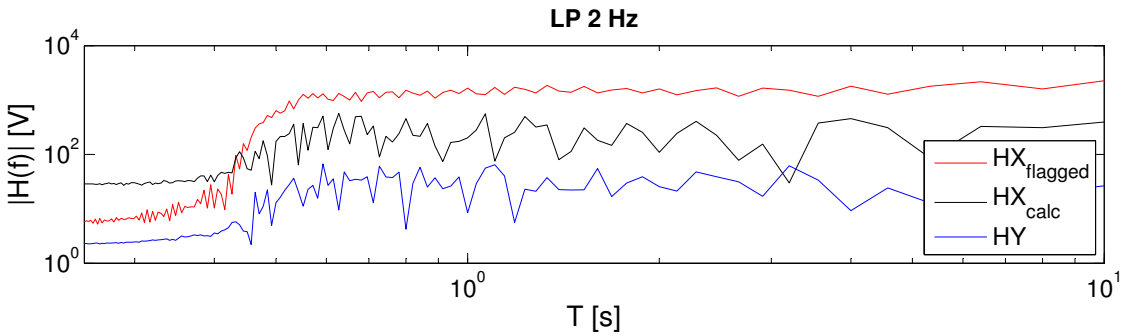


Figure 2.38: Spectra of time series window shown in Fig. 2.37 (256 samples used for spectra calculations). In contrast to the spectrum of the flagged data ($HX_{flagged}$, red curve) the spectrum of the calculated replacement data (HX_{calc}, black curve) resembles that of channel Hy (HY, blue curve).

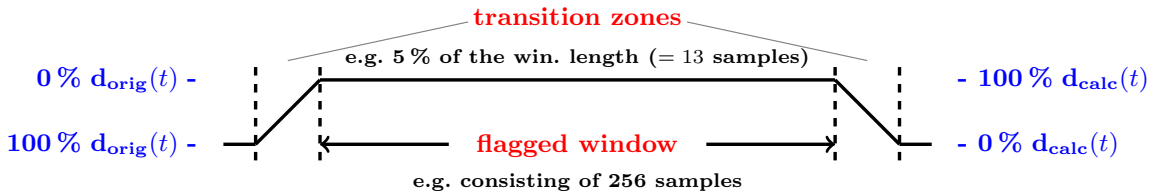


Figure 2.39: The calculated replacement data $d_{calc}(t)$ is fitted to the original time series $d_{orig}(t)$ by weighing the replacement and the original data in transition zones to both sides of the flagged window.

Filter parameters

Important for a successful application of the *reference site filter* is the adjustment of its parameters to the specific data set. The parameter setting depends strongly on the data and the occuring noise and, hence, there are no universally optimal settings. Different parameter settings might be necessary for the xy- and the yx-component since they can be affected differently by noise sources, e.g. depending on the relative position of the noise source in reference to the site and the noise source geometry. Therefore, the following subsections are intended to serve as a guideline for choosing appropriate parameters.
For the following tests, a time series with 1,983,320 samples and a sampling rate of 1/8 Hz is used. Prior to the tests, the time series was bandpass filtered to band LP 2 s (2 s to DC, cf. Tab. A.1, p. 149).

Window length and overlap

The window length (WL) of the time series determines the ability of the *reference site filter* to detect spikes and steps depending on their characteristics. In the presence of spikes, short windows result in higher mean values μ (cf. eq. 2.7) and consequently lower variance values. In this case, the window will not be flagged even though it contains noise. If the window is too wide, the amount of data which is flagged becomes larger as well, consisting not only of the targeted spike but also of time series segments with possibly good quality. Hence, the window length should be chosen to detect noisy time series segments while preserving as much of the original data as possible. Tests show that window lengths varying between 128 and 1024 samples lead to the best results.
In many cases, filtering with and without an overlap (OV) between adjacent windows yields the same results. However, for short time series higher overlaps increase the total number of windows artificially and make the threshold calculations statistically more robust, although windows become dependent on each other and the degrees of freedom won't increase. Good choices for the overlap range from 16 samples to half of the window length, which is the possible maximum of the overlap.

Threshold settings

The threshold for flagging windows with an increased variance ratio r is the most crucial factor in the filtering process. The threshold is defined by the α-trimmed standard deviation σ_α (eq. 2.10) of all variance ratios of a time series multiplied by the user-defined value t_{user}. The multiplier t_{user} can be chosen independently for electric and magnetic channels to account for different noise levels in the two channel types.
Determining the appropriate trimming value α and the t_{user} values should not be done independently as they both equally affect the threshold level.

- **Trimming value** α

  Using the α-trimmed standard deviation σ_α instead of the full standard deviation makes the threshold calculations more robust against outliers. $\frac{\alpha}{2}$ % of the lowest and the highest variance ratios will not be included in the standard deviation calculations:

$$\frac{\alpha}{2}\% < \text{chosen values} < \left(100 - \frac{\alpha}{2}\right)\%.$$

  Hence, the α value can be considered a threshold.

  In the example, a WL of 256 samples and an OV of 0 samples is used to determine α, amounting to 7747 windows. For all windows of the five data channels, variance ratios r are calculated as described in the previous section (see eq. 2.9).

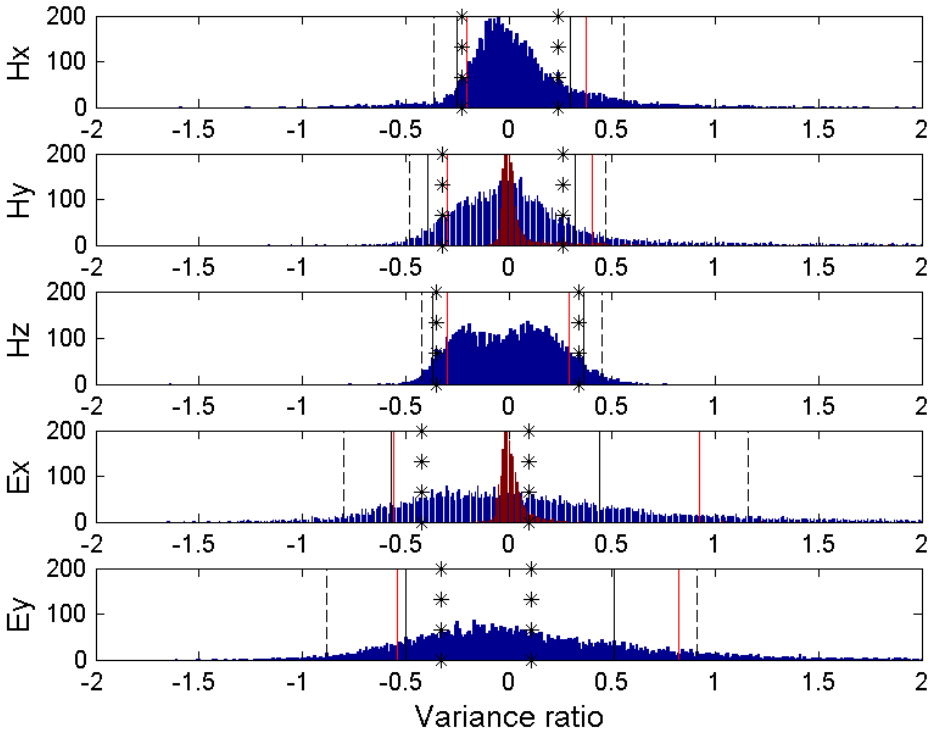


Figure 2.40: Distribution of the variance ratios r (7747 windows, 1601 bins, bin width: 10^{-2}) of all five channels (blue). Picture only shows the section between -2 and 2 for better visulization; the complete distribution ranges from -8 to 8. Marked are the first bins where there are more than 15 (black dashed lines), 40 (black solid lines) and 60 (black asterisk lines) windows in one bin. Bounds for discarding 10% of the minimum and 10% of the maximum values are marked as red solid lines. For channel Hy and Ex distributions for noise-free data are shown (red, cf. site 105 in Fig. 2.2).

Figure 2.40 shows the distribution of the variance ratios for all five data channels (in blue). Regardless of the channel type, the distributions all center around zero as the median values (Tab. 2.3) show.

As Fig. 2.40 illustrates, the distribution of the variance ratios differs for the two channel types and shows a higher asymmetry for the electric than for the magnetic channels. This is not only the case for the chosen example but is a general observation for the BGB MT data set. Whereas for the magnetic channels, there is a sharper boundary between the values centering around zero and the outliers, the variance ratio distribution for the electric channels spreads over a wider range. This

makes it more difficult to set an α threshold for the electric channels and gives a first hint that the detection of steps is more challenging than the detection of spikes.

A comparison of the variance ratio distribution of noisy data with a distribution of almost noise-free data (band LP 2 s of site 105 in Fig. 2.2) is shown in Fig. 2.40 for channels Hy and Ex (noise-free distributions in red). The noise-free distributions show a very low variation of variance ratio values and only few outliers. On the one hand, this is due to the absence of severe noise but on the other hand this is also caused by the spatial proximity and similarity of site 105 and its reference site 103 (cf. map 2.1). Hence, it can be concluded that the higher variation of the variance ratio distribution of the noisy data cannot be attributed to the occurance of noise alone. The comparison also shows that the differences in the distributions between magnetic and electric channels are related to the way noise affects different channels.

Choosing α could be done by analyzing the number of events per bin, considering bins with only few variance ratio values as outliers. This approach can give a hint for the most suitable α but since the distributions are neither Gaussian nor normal distributions, the percentages differ considerably between the minima and maxima values that would be discarded (cf. *min* and *max* in Tab. 2.3). Moreover, choosing different values for each channel or even for each time series, taking into account the large number of sites is not feasible but also not necessary in most cases. If the largest outliers are discarded for the σ_α calculation, the finer adjustment of the threshold can be done with the multiplier t_{user}. Therefore, α is chosen as a fixed value for both, the magnetic and electric channels, for all time series. Cases where a more advanced threshold calculation is necessary are discussed in subsection *Adapted threshold calculation* (p. 51).

In general, α-values of 20 %, i.e. discarding 10 % of the lowest and 10 % of the highest variance ratios, work well. This leaves a much smaller variation of the variance ratios for the magnetic than for the electric channels (cf. red lines in Fig. 2.40).

Table 2.3: Statistical quantities of the variance distributions shown in Fig. 2.40; std: standard deviation. The *min* and *max* values represent the percentage of variance ratios which would be discarded applying the thresholds at the first bins which contain more than 15, 40 or 60 values.

	mean	median	std	*min* 15	*min* 40	*min* 60	*max* 15	*max* 40	*max* 60
Hx	0.042	0	0.382	3.498	6.454	7.913	4.686	13.153	16.832
Hy	0.053	0	0.422	1.278	4.002	8.313	8.197	13.450	17.065
Hz	-0.004	0	0.267	1.071	3.266	4.737	1.756	4.892	6.338
Ex	0.107	0	0.644	2.711	9.229	17.439	6.596	23.338	42.817
Ey	0.084	0	0.586	1.794	11.966	22.409	8.261	19.517	41.306

- **Threshold multiplier t_{user}**

  After deciding on a suitable trimming value α, the second parameter that determines the threshold is the user-defined multiplier t_{user}. For this test a WL of 256 samples and an OV of 0 samples is used amounting again to 7747 windows for the chosen time series example. The multiplier t_{user} for both magnetic and electric channels varies between 1 and 7 for the tests.

  Figure 2.41 shows the resulting apparent resistivity and phase curves with varying t_{user} values. For the two components different threshold settings are necessary to obtain smooth curves. Whereas for the xy component all settings $t_{user} \geq 4$ yield similar results, the yx component requires a lower

threshold level (t_{user} = 2). This reflects the overall better data quality of the xy component which can be seen in the lower amount of flagged windows for channel Hy than for Hx (cf. Tab. 2.4). Appropriate threshold values may vary for different period ranges and even between sites within the same period range. For neighbouring sites which are affected similarly by noise often the same settings can be applied. In general, a set of values should be tested for period ranges that need improvement.

Table 2.4: Percentage of flagged windows for varying values of t_{user} (α = 20%). Coloured are the settings that yield the best results for the xy (red) and the yx component (blue) as shown in Fig. 2.41.

channel \ t_{user}	1	2	3	4	5	6	7
Hx	26.2	14.5	8.8	4.9	3.1	1.9	1.2
Hy	8.1	8.1	7.4	5.1	3.7	2.9	2.1
Hz	54.0	13.2	1.1	0.1	0.1	0.1	0.1
Ex	48.1	17.4	7.9	3.9	2.0	1.2	0.5
Ey	48.5	18.2	7.1	3.2	1.6	0.9	0.2

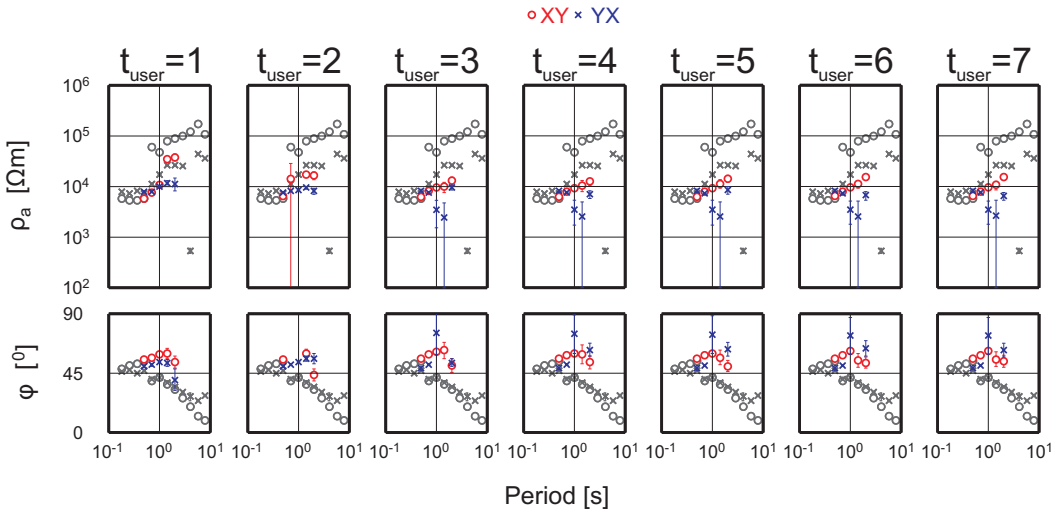


Figure 2.41: Apparent resistivity ρ_a and phase φ for t_{user} values ranging from 1 to 7 for magnetic and electric channels (α = 20%). Note that for the xy component (red markers) t_{user} = 4 and for the yx component (blue markers) t_{user} = 2 yields the best results. For comparison single site processing results are shown as well (in dark grey).

- **Usage of absolute variance ratios for threshold calculation**
  In the distribution of the variance ratios r (Fig. 2.40 and Tab. 2.3) a clear deviation from a normal distribution can be observed. Therefore, setting the thresholds simply to $\pm t_{user} \cdot \sigma_\alpha$ might not always be suitable in order to detect noisy windows. Fig. 2.42 shows the time series of the actual variance ratios (a) and their absolute values (b). Using only absolute values of r for the threshold calculations and applying the trimming value α for the σ_α calculation only to the maxima of $|r|$ can help to detect noise easier in strongly deviated distributions. Hence, I implemented the additional option of using absolute variance ratios within the filter algorithm. Tests for the chosen time series

example show similar results when either the variance ratios or their absolute values are used. However, for the *Adapted threshold calculation* (p. 51) using the absolute values of the variance ratios becomes essential.

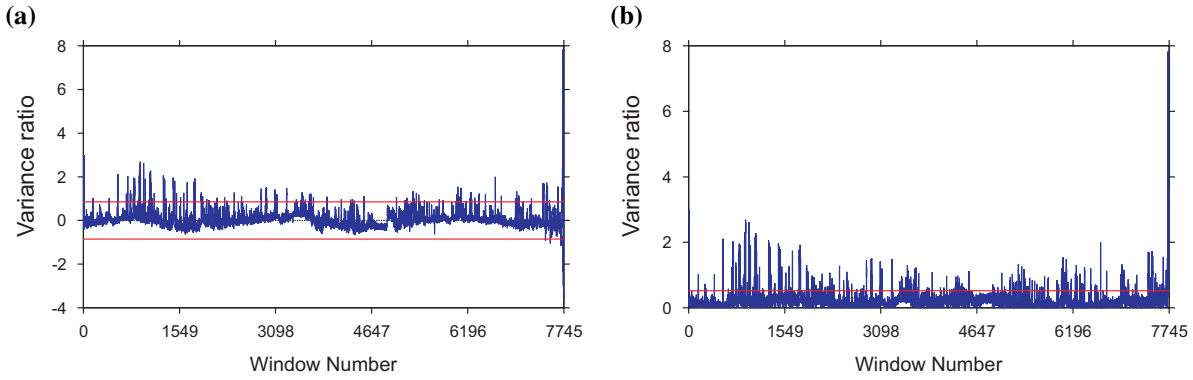


Figure 2.42: Variance ratio time series of channel Hy: **(a)** variance ratios with lower and upper threshold (in red, $t_{user} = 5$) for the window flagging; **(b)** absolute values of the variance ratios are used to calculate only one threshold.

- **Adapted threshold calculation**

  So far, the threshold ($t_{user} \cdot \sigma_\alpha$) for the window flagging process is defined as a constant value for all windows of a time series (see Fig. 2.42). Since outliers are discarded for the threshold calculations (α-trimming of σ_α), this is a robust and reliable method to determine a threshold in an automated way.

  However, the distribution or trend of the variance ratios might make more advanced threshold calculations necessary, especially in the following two cases. (1) Asymmetric variance ratio distributions with a shift to either positive or negative values (e.g. Fig. 2.40, 2.42) might make the lower or the upper threshold, respectively, inappropriate for the flagging since it is either too low or too high to detect extrema of the variance ratios. (2) Independent of the just described asymmetry, variance ratio curves can exhibit fluctuations over time which hinder the determination of appropriate thresholds. In this case, for a higher constant threshold too few noisy windows are detected or if the threshold is set too low numerous windows with good data quality are falsely flagged. The described trends and fluctuations are likely caused by the different geological structures underlying the local and the reference site.

  For better fitting the shape of the variance ratio curves, I propose the usage of an adapted threshold. In the following, two different types of adapted thresholds developed within this work are described.

  (a) Threshold calculation in fixed sections

  For the threshold calculation in sections, the time series of variance ratios is divided into non-overlapping segments each containing an equal number of windows. For each section the threshold is constant and is calculated using $t_{user} \cdot \sigma_\alpha$: t_{user} is fixed while σ_α is calculated separately for every section. The number of sections can be chosen by the user. Note that if the sections become too narrow, only few variance ratios are used for the threshold calculations making them less robust. As the example in Fig. 2.43 shows, it is strongly recommended to use absolute variance ratio values for this type of adapted threshold calculations.

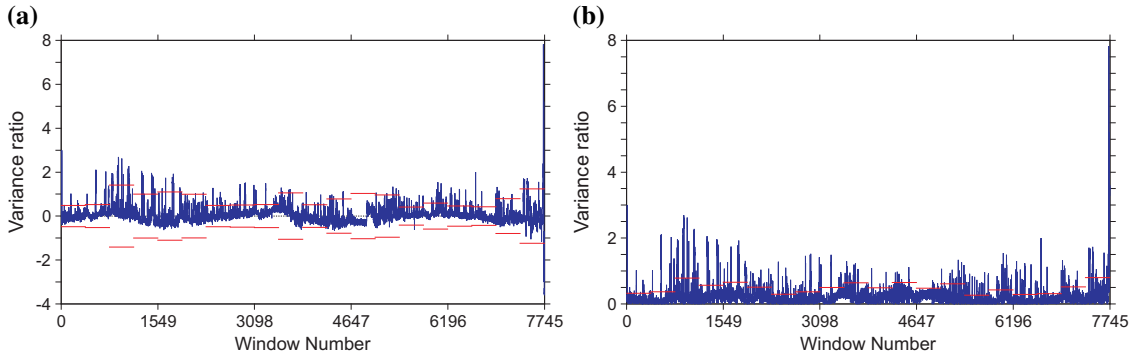


Figure 2.43: Variance ratio time series of channel Hy: Red lines mark the thresholds calculated for the 20 sections with **(a)** lower and upper bounds and **(b)** absolute variance ratio values. As the comparison shows, the section-wise threshold calculation is only adequate if absolute variance ratios are used.

(b) Moving window threshold calculation

For strongly fluctuating variance ratio curves, a moving window threshold might be the best matched method. For this approach, the time series of variance ratios is divided by a user-defined value to determine the length WL_{mov} of the moving window w. In case the obtained window length is even, the algorithm increases WL_{mov} by one to facilitate the later assignment of the calculated threshold values to a specific variance ratio value or the corresponding window.

Fig. 2.44 illustrates the principle of this threshold calculation approach. For the first window w_1, the threshold is calculated applying $t_{user} \cdot \sigma_\alpha$ using the first WL_{mov} variance ratio values. The obtained threshold is used for the flagging of the first $\lceil WL_{mov}/2 \rceil$ windows. $\lceil \cdot \rceil$ denotes the ceiling function. In the next step, the window is moved by one variance ratio value. The threshold for this second window position is only used at $\lceil WL_{mov}/2 \rceil + 1$ in the variance ratio series. The just described procedure is continued until the last window w_n where the calculated threshold value is used for the last $\lceil WL_{mov}/2 \rceil$ windows. As shown by comparing Fig. 2.45 (a) and (b), again, this method works best if absolute variance ratios are used for the threshold calculations.

Since the threshold is calculated $n = N - WL_{mov} + 1$ times where N is the number of windows, this approach is computationally expensive. Therefore, it should only be applied if it is really necessary. The processing example given in Fig. 2.45 (c) shows improvements around 1 s when comparing the *reference site filter* with threshold envelope to the filtering with constant threshold.

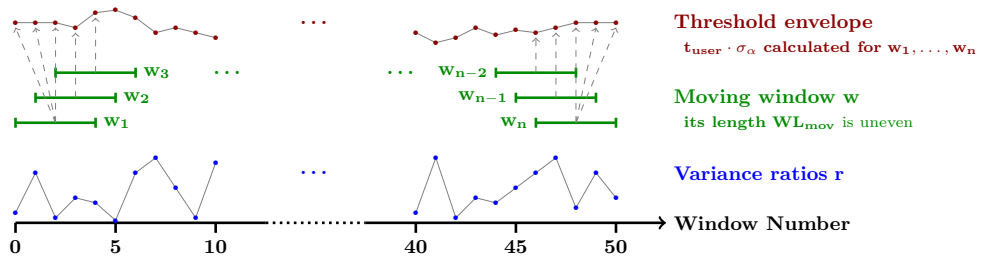


Figure 2.44: Scheme of the moving window threshold calculation: The threshold is calculated for every position of w in the variance ratio series. At the first and last window position (w_1 and w_n), the calculated threshold is used for the first and last $\lceil WL_{mov}/2 \rceil$ windows, respectively. For all windows in between ($w_2, \ldots, w_{n-1}$), the calculated threshold is assigned to the central position of the window w (grey dashed arrows).

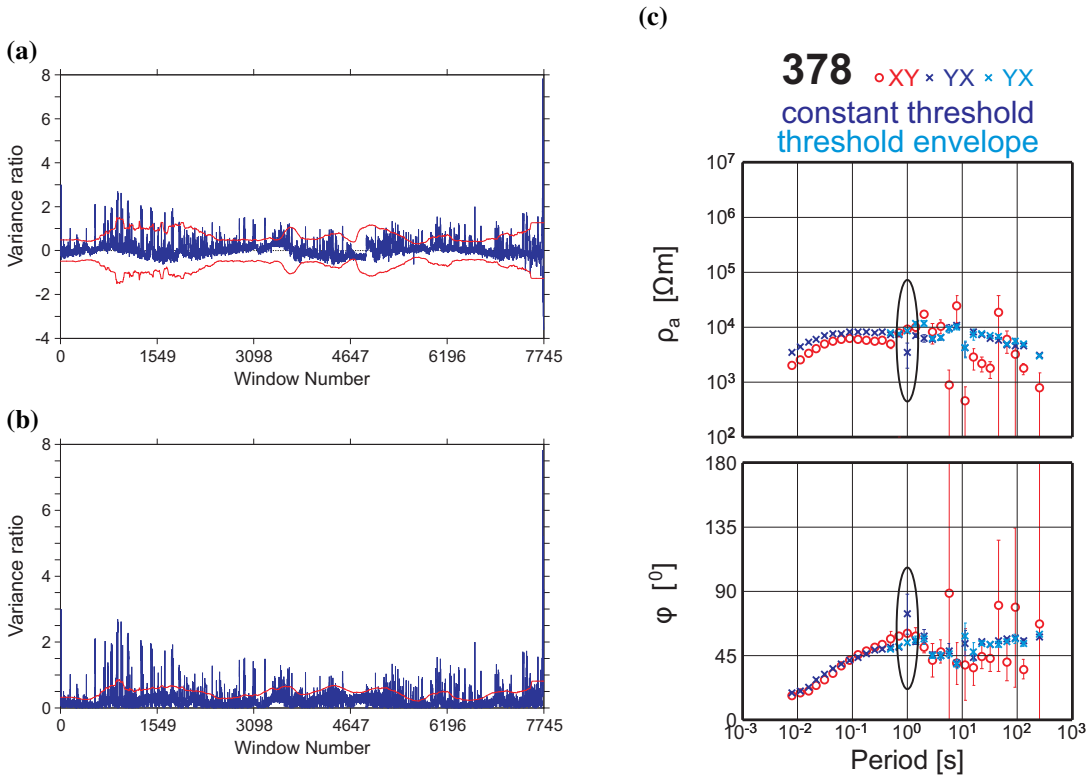


Figure 2.45: Variance ratio series of channel Hy and processing result example: **(a)** variance ratios with lower and upper threshold (in red) for the window flagging; **(b)** absolute values of the variance ratios are used to calculate only one threshold; **(c)** comparison of the yx component after filtering with different threshold settings: with a constant threshold an outlier occurs at 1 s in the TF (dark blue markers in the black ellipses) whereas with a threshold envelope the TF curves are smooth at 1 s (light blue markers).

Training data and Wiener filter length

Writing equation 2.15 in matrix form gives a clearer image of the assembly of convolution matrix **M**, the data vector **d** and the Wiener filter coefficients **Ψ**:

$$\begin{bmatrix} d_{1+\frac{M-1}{2}} \\ d_{2+\frac{M-1}{2}} \\ \vdots \\ d_{T-\frac{M-1}{2}} \end{bmatrix} = \begin{bmatrix} m_{1,1} & \cdots & m_{1,M} & \cdots & m_{K,1} & \cdots & m_{K,M} \\ m_{1,2} & \cdots & m_{1,M+1} & \cdots & m_{K,2} & \cdots & m_{K,M+1} \\ & \vdots & & \vdots & & \vdots & \\ m_{1,T-M+1} & \cdots & m_{1,T} & \cdots & m_{K,T-M+1} & \cdots & m_{K,T} \end{bmatrix} \begin{bmatrix} \Psi_{1,1} \\ \vdots \\ \Psi_{1,M} \\ \\ \vdots \\ \\ \Psi_{K,1} \\ \vdots \\ \Psi_{K,M} \end{bmatrix} \tag{2.17}$$

In this equation, M is the Wiener filter length and T is the training data length. The number of data channels K depends on the number of unflagged data channels in a specific window and the number of reference sites. For the example in Fig. 2.36, K equals 9 (4 channels from the local plus five channels from the reference site).

Parameters T and M determine the size of equation system 2.17:

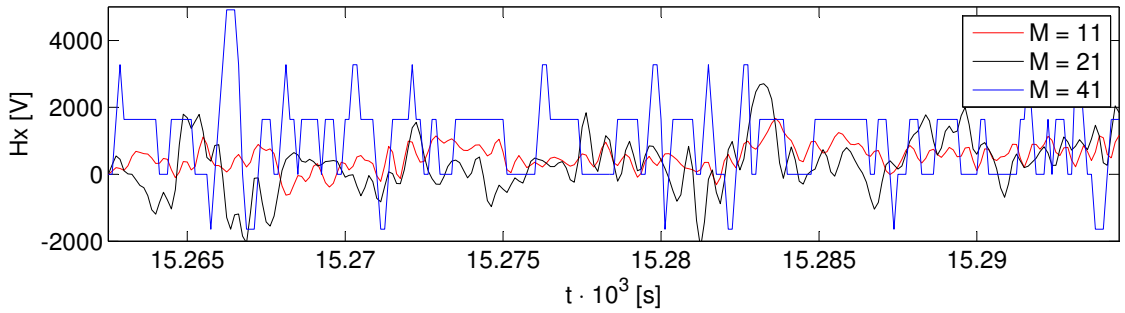
- **d**: vector of size $(T - M + 1)$,
- **M**: matrix of size $((T - M + 1) \times (K \cdot M))$ and
- **Ψ**: vector of size $(K \cdot M)$.

The training data used to assemble $\mathbf{d_{tr}}$ and **M** as input for the Wiener filter are chosen from noise-free data (i.e. unflagged data) from the local and simultaneously recorded reference site. Its length T needs to be at least as long as the window length WL plus the length of the transition zones. Since FFTs are necessary for the calculations $T = 2 \cdot WL$ is the minimum length.

Choosing the Wiener filter length M, a comparison of the calculated replacement data using different M values can be helpful. For the example in Fig. 2.46 (a), calculations with $M = 11$ and $M = 21$ yield reasonable time series. For $M = 41$, the Wiener filter calculation becomes unstable. Both the time series and the spectrum (Fig. 2.46 (b)) for $M = 41$ show a disturbed behaviour.

The size of equation system 2.17 and, thus, the computation time of the algorithm, increase considerably with increasing Wiener filter length M. Filtering of a LP 2 Hz data file requires 3 min 20 sec ($M = 11$), 9 min 40 sec ($M = 21$) and 21 min 2 sec ($M = 41$); in the example time series 341 windows were flagged. For the majority of data, $M = 21$ yields the best results and is used as value for the Wiener filter length throughout this work.

(a)



(b)

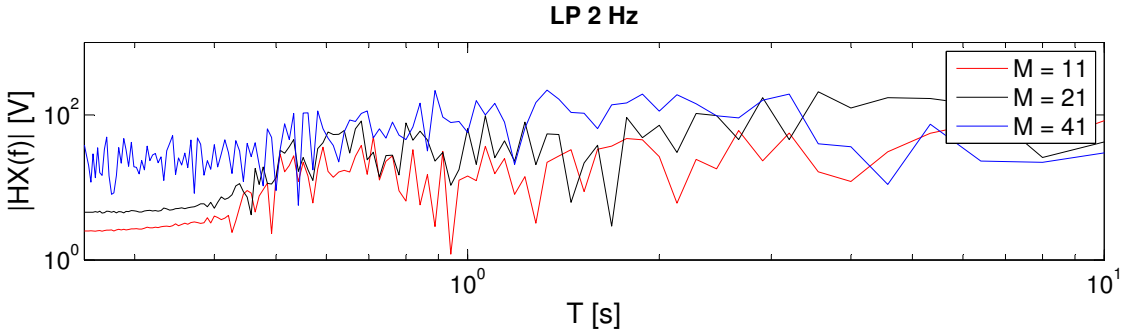


Figure 2.46: Comparison of the **(a)** time series and **(b)** spectra of replacement data for different Wiener filter lengths M: Using either $M = 11$ or $M = 21$ yields stable results, wheras the Wiener filter becomes unstable with $M = 41$. Time series consist of 256 samples (≙32 s) each which are used for the spectra calculations as well.

Differentiation and integration of time series

The detection rate of spikes in magnetic channels is high for appropriately chosen threshold settings. Steps in the electric channels, however, are much harder to detect. Their detection depends mostly on the offset between the mean level of the time series prior compared to the mean level after the step. In general, larger offsets make it easier for the *reference site filter* to identify steps.

An approach to overcome this difficulty is to differentiate the electric time series. This procedure transforms steps into spikes and allows to use the above described *reference site filter* algorithm nearly unchanged: The additional time series differentiation is carried out prior to the actual flagging and filtering procedure.

For the differentiation, the forward difference quotient is used:

$$y_i'(t) = \frac{y_{i+1}(t + \Delta t) - y_i(t)}{\Delta t}; \tag{2.18}$$

$y(t)$ is the time series and Δt is the time difference between two adjacent data points determined by the sampling rate.

The differentiation changes the time layout of the time series (Fig. 2.47): The original time series is given at points $t_1, t_2, t_3, \ldots$, whereas the differentiated one is given at points $t_{1,2}, t_{2,3}, t_{3,4}, \ldots$ centered between $t_1, t_2, t_3, \ldots$. Due to the differentiation, the time series is shortened by one data sample, whereas Δt is the same for both the original and the differentiated time series.

The differentiated (!) time series are filtered using the *reference site filter*. Afterwards, the filtered time

series need to be integrated again. This is done by a simple summation of the differentiated data. The first sample of the integrated time series is set to zero ($y_1 = 0$); all subsequent points ($y_i, i = 2, 3, \ldots$) are obtained by

$$y_i(t) = \sum_{k=2}^{i} y'_k. \tag{2.19}$$

The integration procedure removes the offset of the time series (since it begins at zero) and transfers the time layout back to its original. The resulting integrated time series has the same number of samples as it had before differentiation.

If only data of the electric channels are differentiated, the ratio between the electric and magnetic field components (i.e. the impedance **Z**) will be disturbed. Hence, apparent resistivity ρ_a and phase ϕ curves become nonsensical. This is enhanced by interpolation of the magnetic data from time points $t_1, t_2, t_3, \ldots$ to $t_{1,2}, t_{2,3}, t_{3,4}, \ldots$ to obtain magnetic data at exactly the same time as the differentiated electric data. This is necessary because both electric and magnetic data combined are used to calculate the replacement data. A second interpolation of the magnetic data back to the original time points is necessary when the electric time series are integrated (Fig. 2.47). To avoid rounding errors due to interpolation and to maintain the same ratio between the electric and magnetic field, the magnetic channels are differentiated and integrated in the same way as the electric channels. In general, this amplifies the spikes and further helps to detect them (see original and differentiated magnetic data in Fig. 2.48).

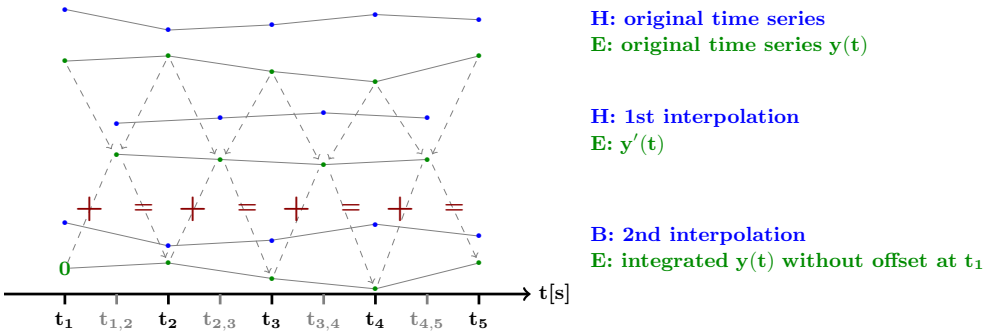


Figure 2.47: Differentiation and integration scheme of time series (from top to bottom): For enhancing the detection rate of steps in electric time series (green), they are differentiated which changes the time layout from $t_1, t_2, t_3, \ldots$ to $t_{1,2}, t_{2,3}, t_{3,4}, \ldots$. For application of the *reference site filter*, magnetic data (blue) needs to be available at the same time and is therefore interpolated. Piecewise summation (symbolized by + and =) leads to the integrated electric time series. The magnetic data needs to be interpolated a second time.

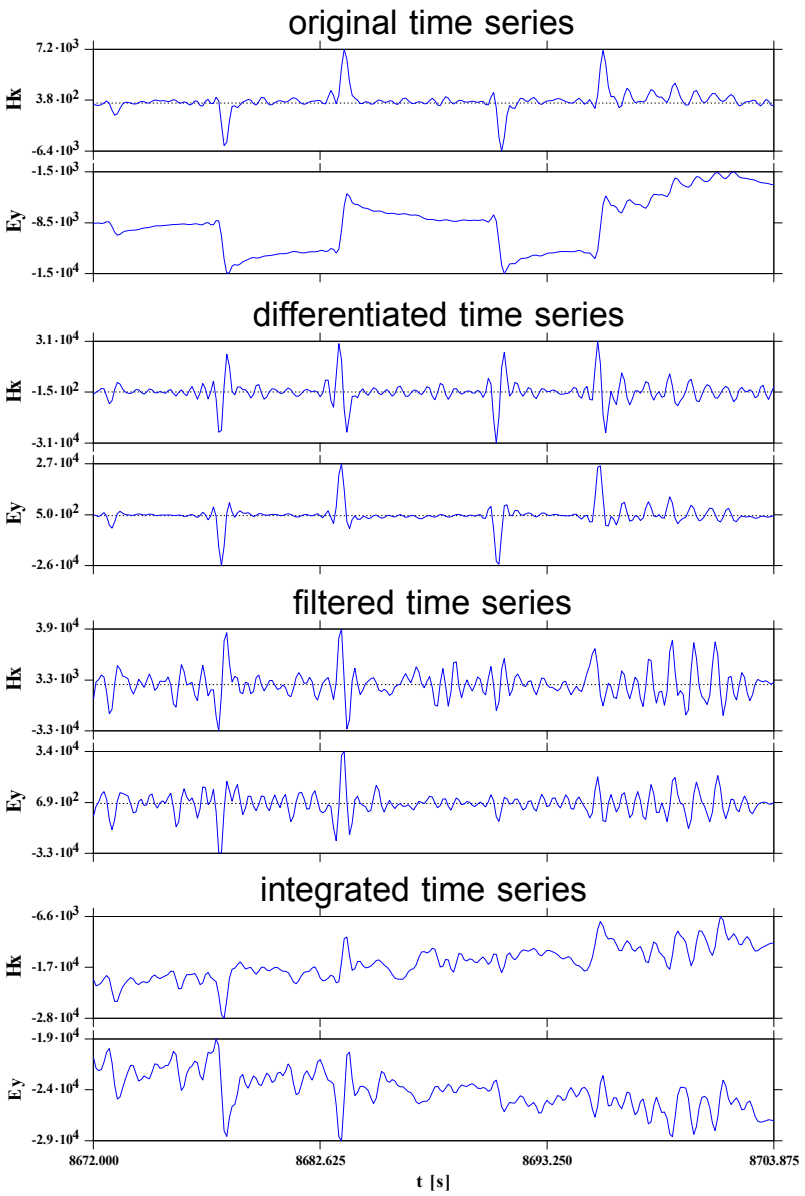


Figure 2.48: *Reference site filtering* progression of time series with differentiation and integration: Original time series window containing spikes in the magnetic channel (Hx) and corresponding steps in the electric channel (Ey). After differentiation, spikes in Hx are enhanced (cf. scales) and steps in Ey are spikes. Both windows are replaced by calculated data during the filtering and are finally integrated.

In almost every case, the calculated data used to replace the flagged windows of the differentiated time series has an offset to the adjacent windows. This offset is substracted from the calculated data to prevent integrated time series with a constant slope.

A comparison of processing results for a site where the *reference site filter* without and with differentiation of the time series is applied is shown in Fig. 2.49. For the filtering with differentiation, a clear improvement of the data quality can be observed for periods between 1 and 100 s. For periods above 100 s, filtering without differentiation in the given example results in smoother curves.

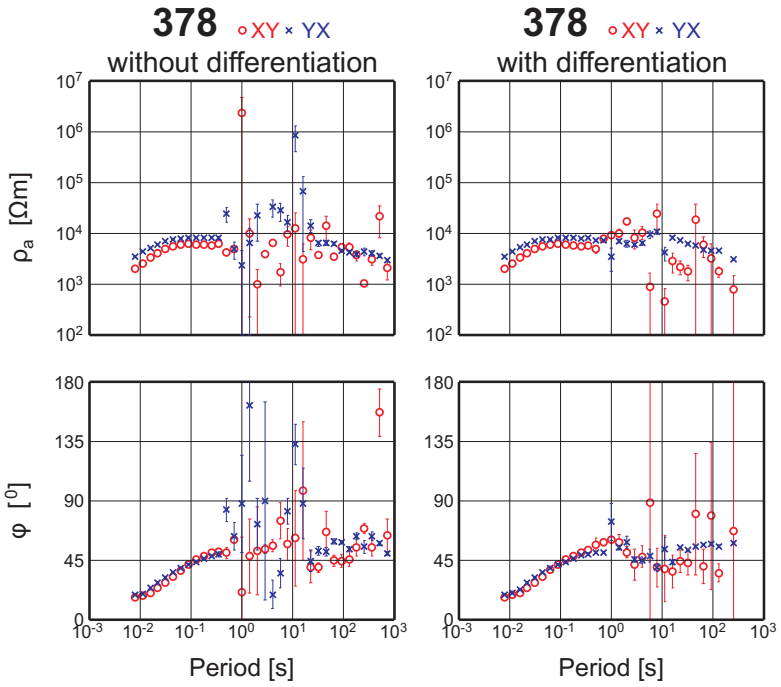


Figure 2.49: Comparison of processing results applying the *reference site filter* without and with differentiation and integration of the time series. Clear improvement for filtering with differentiation can be observed for periods 1 to 100 s. For periods above 100 s, the filter without differentiation works slightly better in this example.

Additional filter settings

The filter implemented in this work offers various additional options. Using these options depends on the quality of the local site and the reference site(s) and can help to increase the detection rate of noisy data:

- **Flagging of corresponding electric channels**
  Frequently, orthogonal data components (either Hx|Ey or Hy|Ex) exhibit noise simultaneously. If the noise is clearly visible in the magnetic channel but not as pronounced in the electric component, only the H channel is flagged. However, using the distorted but unflagged E channel for the replacement data calculations will reintroduce noise to the H channel again. Therefore, if the x- or y-component of the magnetic field is flagged in a particular window and if this option is set, the corresponding orthogonal component of the electric field will be flagged, as well.

- **Flagging of both electric components**
  If noise is observed in both electric channels, but is only detected in one of them by the filter, the usage of the other electric channel for the calculation of the replacement data would bias these calculations and reintroduce the step again into the time series. Hence, this option flags both electric channels, even if the algorithm flagged only one of them.

- **Use only reference site data for replacement data calculations**
  If the data quality of the local site is low containing not only spikes and steps but also noise types which are not detectable by the *reference site filter*, using the local site data for the replacement

data calculations might bias these calculations. Hence, this option allows to only use the reference site data for the calculations even if not all of the local site channels are flagged.
The data of the reference site should be closely related (i.e. coherent) to the local site data when using this option. Therefore, it needs to be a relatively near reference site, especially for the calculation of the replacement data for the electric channels to work.

Exceptional cases during the filter process

There is a number of situations in which the *reference site filter* fails to calculate suitable replacement data. These situations are addressed as follows:

- **No convergence of the SVD**
  If the SVD does not converge within 70 iterations replacement data cannot be calculated. In this case, flagged data is replaced by a constant value, i.e. the mean value of the adjacent windows.
  Reasons for the SVD not to converge are e.g. undetected noise in the data or poor correlation between the local and the reference site(s).

- **Wiener filter calculates corrupted data**
  The quality of the replacement data can be poor, e.g. because of very small and, thus, unstable Wiener filter coefficients Ψ). In this case, the replacement data is not suitable for further use. As a quality control, the spectrum of the calculated data is determined and divided in four sections. The power in these sections is compared to each other. If it exhibits a distinct high in one of the sections, the calculated data is not used and the flagged windows are replaced by a constant value. Overall, this approach works better than comparing the variance of the original data with that of the calculated data as suggested by Kappler (2012).

- **Insufficient amount of training data**
  The length of the training data needs to be a power of two and at least two times the window length WL to avoid gaps between the calculated data and the original data. If there are not enough continuous unflagged windows or if the matching combination of unflagged windows cannot be found in the entire time series no Wiener filter coefficients Ψ and, hence, no replacement data can be calculated. In this case, the flagged data is replaced by a constant value.

It would be possible to just keep the original data even for the flagged windows in the above cases. However, if the filter fails for a large number of windows, the overall data quality stays the same as without applying the filter.

2.3.2 STA/LTA filter

In seismology, the short time average (STA) long time average (LTA) trigger is used for triggering the recording of natural seismic events (i.e. earthquakes) of the incoming seismic signal at single seismometers or seismometer arrays (Trnkoczy, 2002). It relies on the fast change of the STA compared to the inert behaviour of the LTA at the beginning of an earthquake and triggers if their ratio exceeds a chosen threshold value.
In areas where irregular, high-amplitude spikes occur in seismic recordings, this approach is disadvantageous since it is also sensitive to such short term events and records them as well (Trnkoczy, 2002).

Therefore, in order to find spikes within a MT time series this algorithm appears to be promissing. Using the STA/LTA trigger, this approach only needs the data of the local site to detect spikes. Therefore, its application is possible even if no reference site data is available.

Detection of spikes and steps

I implemented two different approaches to detect spikes and steps within the *STA/LTA filter*: the STA/LTA trigger and a threshold algorithm based on standard deviation calculation similar to the *reference site filter* (see sec. 2.3.1, p. 47). For both threshold calculation approaches, the time series of the electric and magnetic channels are first differentiated using the forward difference quotient (eq. 2.18). This removes the offset and trend of the time series and leaves only the pure differences between neighbouring samples. Furthermore, this facillitates the detection of spikes and as in the case of the *reference site filter* allows an easier detection of steps as well (sec. 2.3.1). For the determination of the threshold the absolute values of the differentiated time series are used (cf. adapted threshold calculation in sec. 2.3.1).

STA/LTA trigger

The main parameters for triggering are the length of the STA and LTA windows. In seismology the length of the STA should be in the order of a few multiples of the typically expected period of the seismic signal (Trnkoczy, 2002). Since the duration of spikes is rather short (mostly only a few samples, depending on the frequency range) choosing around 5 to 11 samples for the STA length seems to be sufficient in most cases.

As the LTA is a measure for the general amplitude or level of the time series (baseline value), changing its length also changes the sensitivity to spikes. Short LTA windows only a few times larger than the STA window would only allow for the highest-amplitude spikes to be detected. The wider the LTA window is, the lower the amplitude of a spike needs to be for its identification. Therefore, depending on the desired spike detection level, the LTA window length should be ten to twenty times of that of the STA window.

In order to assign the calculated STA and LTA value to a specific sample of the time series, both window lengths are uneven with their centre located at the sample position the STA and LTA are calculated for (Fig. 2.50). At the beginning and towards the end of a time series both STA and LTA window lengths are shortened to account for their entering in and exiting from the time series. Hence, the windows are not centered at the sample for which the STA and LTA are calculated.

The STA and LTA values are simply defined as the mean of the specific window applying equation 2.8. Subsequently, the ratio r_{STALTA} of the STA and LTA are calculated

$$r_{STALTA} = \frac{STA}{LTA}. \tag{2.20}$$

The next step is to choose an appropriate trigger threshold level and to compare it with the r_{STALTA} values. It is defined similar to the constant threshold for the reference site filter as the mean value of all r_{STALTA} values multiplied by a user-defined value t_{user} (cf. sec. 2.3.1). Again, t_{user} can be chosen independently for electric and magnetic channels. If either magnetic or electric channels exhibit noise, the multiplier t_{user} can be set to higher values ($> 1,000$) for the noise-free channel type. When the r_{STALTA} exceeds the trigger threshold for a specific sample, this sample is marked.

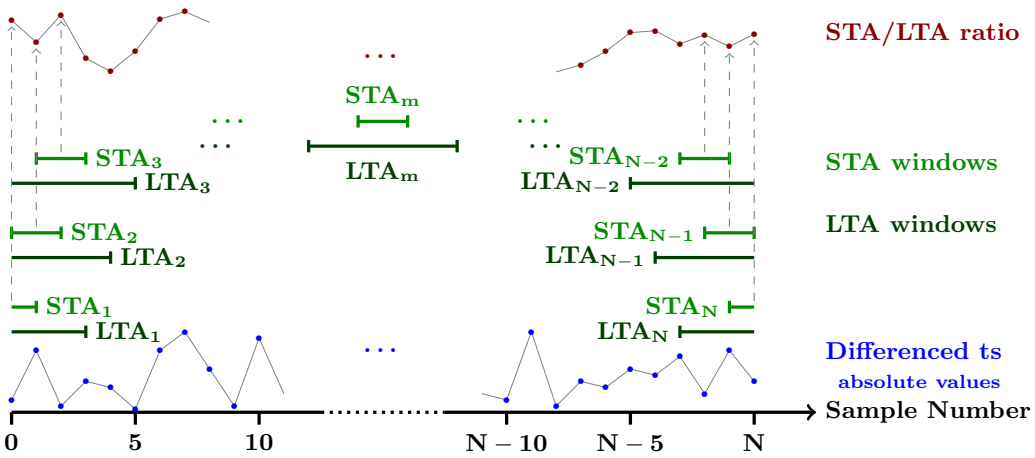


Figure 2.50: Sketch of the STA and LTA windows as they move through the differentiated time series (ts). At the beginning and towards the end of the time series, the STA and LTA length are shortened (e.g. windows STA_1 and LTA_1). Marked by grey dashed arrows are the samples the STA and LTA values are calculated for.

In seismology, a detrigger threshold level needs to be selected as well and the recording stops when the r_{STALTA} value drops below the detrigger level. For the purpose of detecting short time events such as spikes in the data choosing a detrigger value is unnecessary and only r_{STALTA} values above the trigger threshold are singled out. Moreover, since the STA and LTA values are calculated from the differentiated time series where any offset is already removed there is no danger that all samples following a spike will be marked even if the noise level increases.

Standard deviation based threshold algorithm

A simpler way than the STA/LTA trigger in the sense of computational effort is the standard deviation based threshold calculation. The differentiated time series is divided into non-overlapping windows containing a user-defined amount of samples. The last window of the time series can be shorter than the chosen length if the number of samples cannot be divided by the window length without remainder.
For each window, the α-trimmed standard deviation is calculated (cf. sec. 2.3.1, eq. 2.17). The standard deviation multiplied by the user-defined value t_{user} chosen independently for electric and magnetic channels yields the threshold. All samples of the time series that exceed this value get marked.

Spike and step removal

All calculations until now are done for differentiated time series. Therefore, the results after the application of the threshold need to be translated to the original time series: If a differentiated sample is marked, the two samples of the original time series utilized for its calculation get marked as well. The actual filtering described in this paragraph is done in the original time series.
In seismology a pre- and post-event time can be selected for the STA/LTA trigger to record data prior and subsequent to the seismic event to ensure a complete record of the event. Accordingly, a bin width n_{bin} can be chosen which determines how many samples before and after a spike get marked as well, even though they were not marked before by either of the two threshold algorithms.
For the removal of spikes in magnetic channels, the user can choose between two methods. In the first method, the data is linearly interpolated between the sample before the spike and after the spike. For the second method, the marked samples are smoothed applying the Whittaker-Henderson smoothing algo-

rithm (Bohlmann, 1899; Whittaker, 1923) as implemented by Weinert (2007). The method is based on a moving weighted average filtering scheme where the fidelity to the original data is weighted against the smoothness. Lower fidelity increases the smoothness of the resulting data. The intensity of the smoothing is user-defined, ranging between 0 and 1 where lower values correspond to higher smoothing. For time series where larger amounts of consecutive samples are marked smoothing might be more appropriate than linear interpolation.
For removing steps in electric channels, the mean value of the time series before and after the step is calculated. The difference between the mean values is added to or substracted from the time series starting from the first unmarked sample (subsequent to the last marked sample) until the end of the time series. For the calculation of the mean, the user chooses the number n (e.g. 5) of samples prior and after the step that are used for the calculations. The marked samples are linearly interpolated as described above for the spike removal.
An exception occurs if the step is detected within the first n or last n samples of the time series. In this case, these first (or last) samples are set to the values of the first sample after (or the last sample before) the step.
In cases where steps are observed in the magnetic channels and/or spikes in the electric channels, the STA/LTA filter can be set to treat magnetic as electric and electric as magnetic channels removing the noise as described above.

Application of the STA/LTA filter

In this section, the *STA/LTA filter* is tested to remove synthetic noise (as described in sec. 2.1) and the noise found in the MT data from the BGB (see chapter 1). The performance of the filter for synthetic noise will again be exemplarily shown for *arbitrary spikes* and *steps* (cf. sec. 2.2): the data is distorted by either 10% noise in Hy (hereafter abbreviated as ***ex. 1***), in Ex (***ex. 2***) or in all of the channels (***ex. 3***).

Synthetic noise

- **Magnetic and electric multipliers $t_{user,H}$ and $t_{user,E}$**
  The multipliers are the main control parameter of the threshold level for spike and step detection. The trimming value α which also influences the threshold is set to 20% for all calculations. This parameter and the way of choosing it is discussed in section 2.3.1 (p. 48).
  Table 2.5 (for ***ex. 1***) summarizes the marked percentages of samples for varying t_{user} values exemplarily for a time series of band LP 128 s (see Tab. A.1). The marked percentage for ***ex. 1*** of channel Hy appears to converge to a level of about 20%. At first, this seems to be to high considering that only 10% noise was added to the data. However, two factors raise the percentage of detected samples. Firstly, due to the minimum selected bin width $n_{bin} = 1$, samples directly before and after detected spikes get marked as well. Secondly, the differentiated time series is used for the spike detection, making an assignment of the spike exactly to one sample impossible: For a one sample spike two difference quotients are affected, hence, three consecutive samples will be marked at least.
  Regarding the correct identification of the synthetic spikes, $t_{user} = 20$ for ***ex. 1*** yields the best results. For this t_{user} value, all marked samples are related to a synthetic spike and no samples are marked additionaly. 3.1 % of the synthetic spikes ($\hat{=}$ 0.3 % of all samples) are not detected by the

STA/LTA filter since they have much lower amplitudes compared to the amplitudes of the detected spikes (Fig. 2.52 and 2.53).

For lower t_{user} values (< 20), minor amplitude spikes which were already in the data (possibly 'real' noise) get detected as well. Together with them, more data with good quality gets additionally selected, reducing the amount of continuous good quality data. Since detected spikes are simply interpolated, there is a limit to the percentage of marked samples before the time series become useless for any further processing. Therefor, for $t_{user} = 5$ where the percentages of marked samples increase drastically to over 20 % for all channels, the filtering no longer improves the data quality but instead even the yx-component curves get distorted (not shown here).

For higher values of t_{user} (> 30), obviously, less synthetic spikes get marked and removed by the algorithm. This leads to a lower data quality of the processing results of the xy-component curves as well (not shown here).

If either magnetic or electric channels are affected by noise and the respective other field type is clean, filtering only the spiked channels yields the best results (Fig. 2.51). Even though the phase curve could almost be recovered completely, apparent resistivity and induction vectors still show a severe influence of the synthetic noise starting from ~ 1 s. On the one hand, this underlines that the phase is less affected by this type of noise. On the other hand, this shows that for longer periods not enough data windows can be found to calculate correct values. Longer time series windows also might contain more interpolated data after the spike removal. A comparison of the spectra (Fig. 2.54 **(a)**) also reveals that the spectrum of the filtered data is not on the level it had been on before the synthetic noise was added.

As mentioned above, 3.1 % of the synthetic spikes get not detected with $t_{user} = 20$. However, the robust processing can deal with such a low amount of spikes in the data. Nevertheless, if there remain too many spikes in the time series after the filtering, running the filter a second time with the same or a higher value for t_{user} might be better than running the filter only once with a low t_{user} value. Running the filter twice allows for a better control of the threshold level and hence, helps to avoid marking too much of the data.

Table 2.5: ***Ex. 1***: Application of STA/LTA filter: Percentage of marked samples for all channels for varying magnetic and electric multipliers t_{user} (LP 128 s, $n_{WL} = 1,000$, $n_{bin} = 1$, $\alpha = 20$, $n_{step} = 10$). Only the Hy component contains synthetic *arbitrary spikes* (10 %, $v_{max} = 10$). Marked in green is the t_{user} value that yields the best processing results (Fig. 2.51 **(b)**).

$t_{user,H}$ and $t_{user,E}$	5	10	15	20	25	30	40
Samples Hx %	26.5	8.6	4.7	2.3	1.5	0.9	0.2
Samples Hy %	32.0	23.7	22.3	21.4	21.4	21.1	20.9
Samples Hz %	26.2	6.3	2.5	0.8	0.5	0.2	0.1
Samples Ex %	21.8	4.4	1.6	0.7	0.5	0.2	0.1
Samples Ey %	23.4	6.5	2.5	0.9	0.5	0.4	0.0

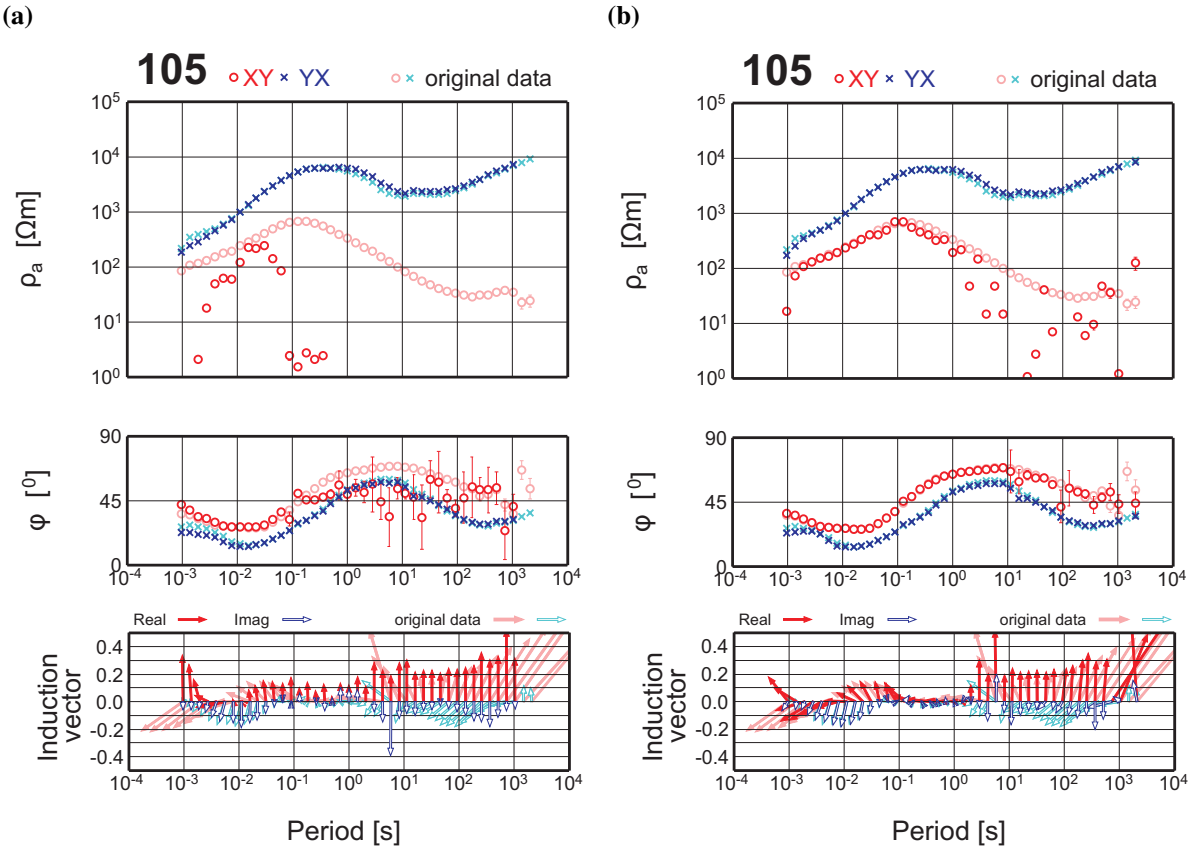


Figure 2.51: ***Ex. 1***: Apparent resistivity ρ_a, phase φ and induction vectors with **(a)** 10% arbitrary spikes in Hy ($n_{mult} \in [0, 10)$) and **(b)** after applying the STA/LTA filter ($t_{user,H} = 20$, $t_{user,E} = 100,000$, $n_{WL} = 1,000$, $n_{bin} = 1$, $\alpha = 20$). Original curves are shown in the background in lighter colours.

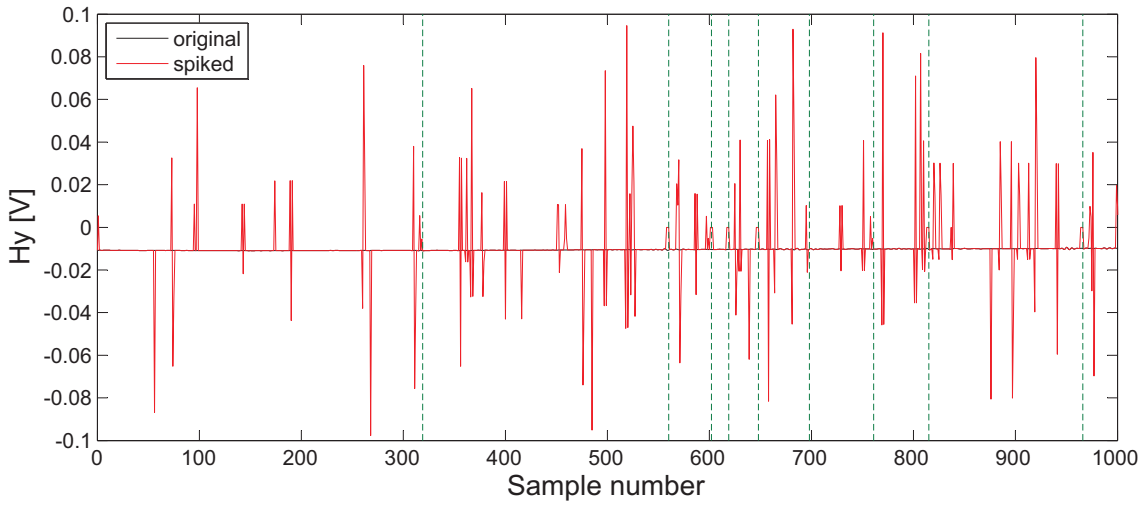


Figure 2.52: ***Ex. 1***: Example of a synthetically corrupted time series (Hy, LP 128 s, 10 % *arbitrary spikes*, $v_{max} = 10$). Synthetic spikes which were not detected by the STA/LTA filter ($t_{user} = 20$, $n_{WL} = 1,000$, $n_{bin} = 1$, $\alpha = 20$, $n_{step} = 10$) are marked by green dashed lines.

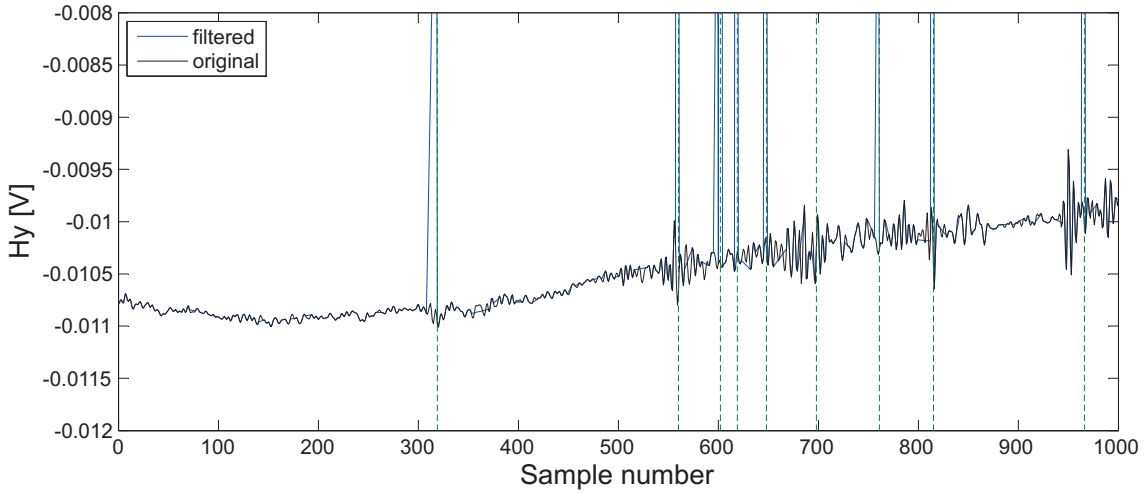


Figure 2.53: ***Ex. 1***: Comparison of the original time series and after STA/LTA filtering is applied (LP 128 s, $t_{user} = 20$, $n_{WL} = 1,000$, $n_{bin} = 1$, $\alpha = 20$, $n_{step} = 10$). Synthetic spikes which were not detected are marked by green dashed lines.

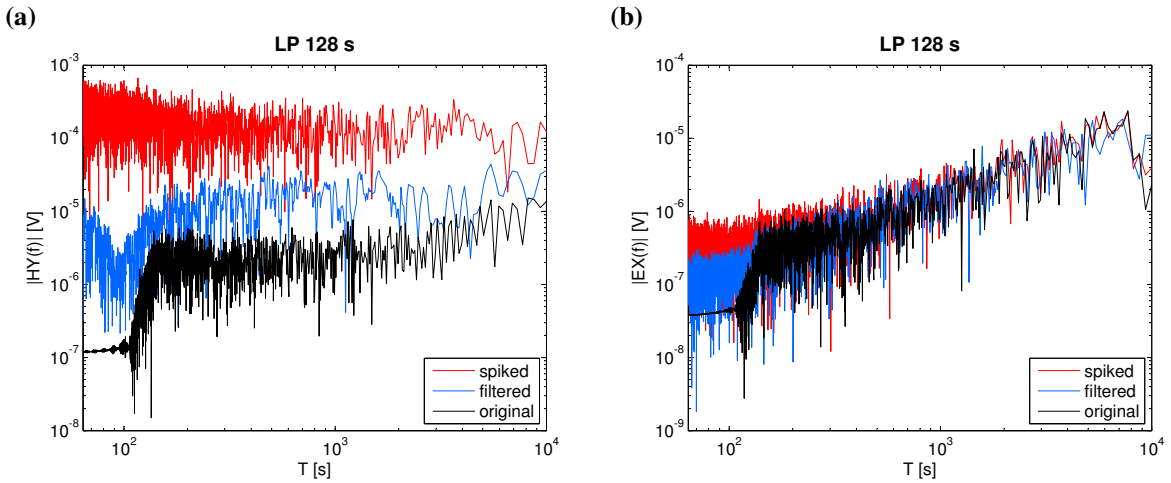


Figure 2.54: ***Ex. 1*** and ***ex. 2***: Amplitude spectrum of channel **(a)** Hy and channel **(b)** Ex, original (black), with 10% *arbitrary noise* added (red, $v_{max} = 10$, $o_{user} = 0.001$)) and after applying the STA/LTA filter ($t_{user,B/E} = 20/15$, $n_{WL} = 1,000$, $n_{bin} = 1$, $\alpha = 20$, $n_{step} = 10$). No. of samples for spectra calculations: LP 128 s: 4,096 (for filtering details see Tab. A.1).

For ***ex. 2***, 10 % steps are added to channel Ex. In this example, the marked percentage of samples decreases continuously with higher $t_{user,E}$ values (Tab. 2.6), not converting to any specific percentage (as for ***ex. 1***). Depending on the step offset, finding the correct threshold for detecting steps is therefor more difficult than for spike detection. This is especially true for removing 'real' noise where the percentage of noise and its distribution in the time series is unknown.

Comparing the original, distorted and filtered time series (Fig. 2.55) shows that the determination of the correct step offset is also a difficult and for 'real' noise even impossible task (especially for rugged steps which affect several consecutive samples). Therefore, the filtered time series might be on a completely different level than the original one. Nevertheless, the general shape of the filtered curve is very similar to the original.

For an easier comparison of the results in the presence of steps, the time series in Fig. 2.56 and 2.57 are differentiated. This form of time series display only shows the difference between neighbouring samples and illustrates the way steps are detected by the STA/LTA filter as well. Marked by green dashed lines are the steps which are detected by the filter showing that some smaller steps were not identified wheras a few minor steps which were already in the data were removed.
For the chosen synthetic noise example, $t_{user,E} = 15$ yields the best results after filtering (Fig. 2.58 **(b)**). For the filtered time series, the spectrum (Fig. 2.54 **(b)** in general has the same level as the original one. Nevertheless, deviations remain explaining the distortion of the ρ_a and ϕ primarily at the transitions from one period band to the next. The slight deviations in the yx-component curves results from the filtering of channel Ey (2.5 % marked samples, see Tab. 2.6).

Table 2.6: ***Ex. 2***: Application of STA/LTA filter: Percentage of marked samples for all channels for varying magnetic and electric multipliers t_{user} (LP 128 s, $n_{WL} = 1000$, $n_{bin} = 1$, $\alpha = 20$, $n_{step} = 10$). Only the Ex component contains synthetic *arbitrary steps* (10 %, $v_{max} = 10$, $o_{user} = 0.001$). Marked in green is the t_{user} value that yields the best processing results (Fig. 2.58 **(b)**).

$t_{user,B}$ and $t_{user,E}$	5	10	15	20	25	30	40
Samples Hx %	26.5	8.6	4.7	2.4	1.4	0.9	0.2
Samples Hy %	27.4	6.9	2.5	0.9	0.2	0.1	0.1
Samples Hz %	26.2	6.3	2.5	0.8	0.5	0.2	0.1
Samples Ex %	26.9	13.5	9.4	6.6	4.6	3.7	1.8
Samples Ey %	23.4	6.5	2.5	0.9	0.5	0.4	0.0

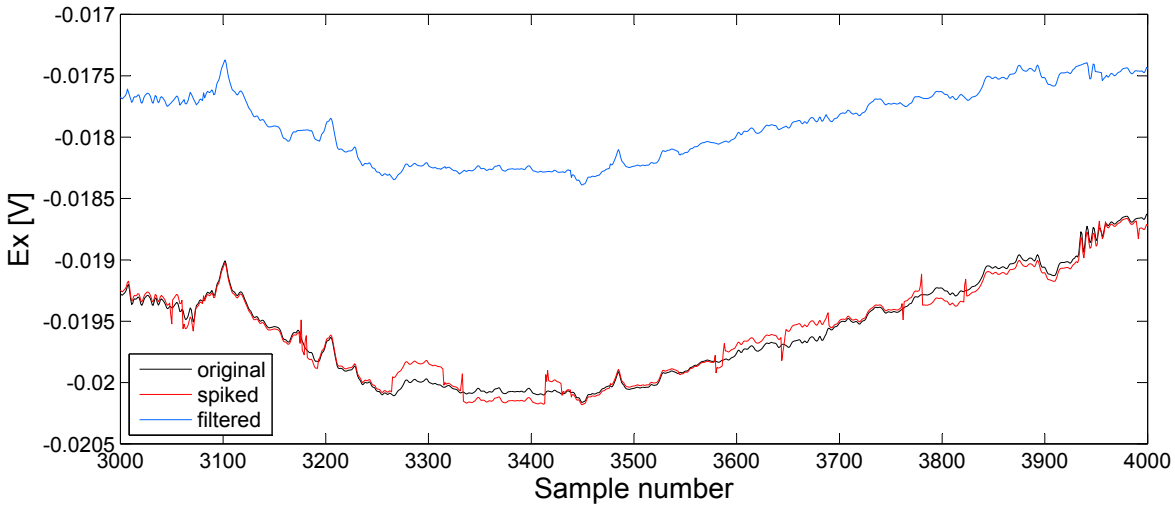


Figure 2.55: ***Ex. 2***: Comparison of original time series (black), time series with 10 % *arbitrary steps* (red, $v_{max} = 10$, $o_{user} = 0.001$) and after STA/LTA filtering is applied (blue, LP 128 s, $t_{user,E} = 15$, $n_{WL} = 1,000$, $n_{bin} = 1$, $\alpha = 20$, $n_{step} = 10$).

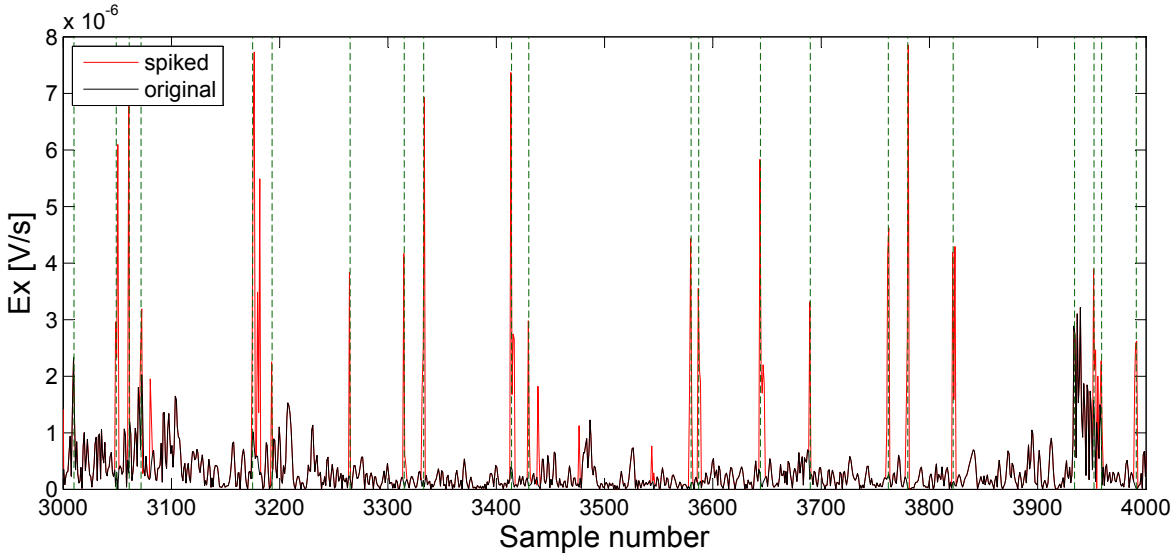


Figure 2.56: ***Ex. 2***: Comparison of differentiated original time series (black) and time series with 10 % *arbitrary steps* (red, $v_{max} = 10$, $o_{user} = 0.001$). Steps detected by the filter are marked by green dashed lines.

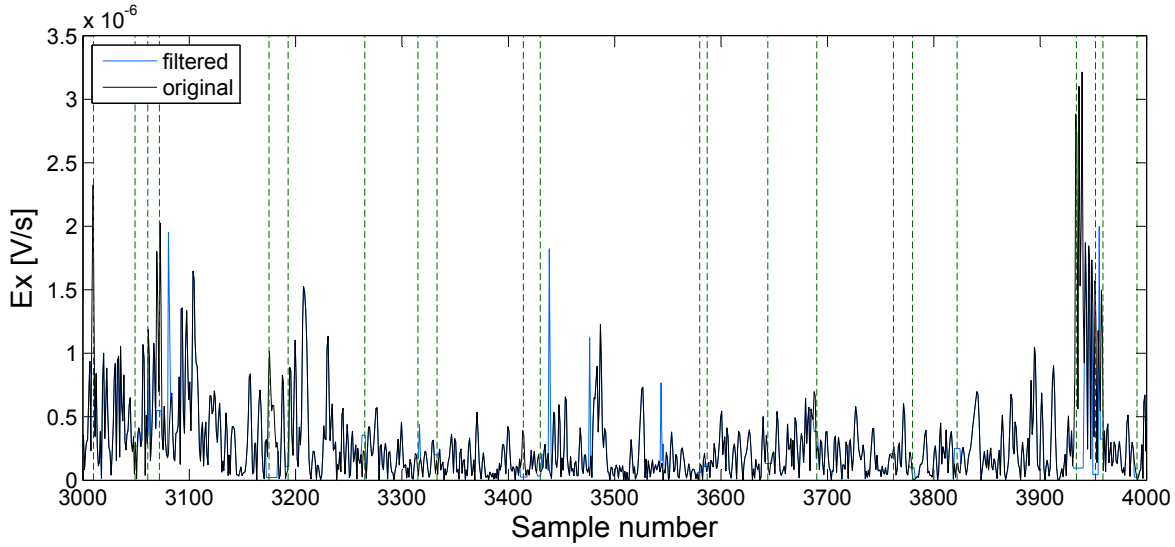


Figure 2.57: ***Ex. 2***: Comparison of differentiated original time series (black) and time series after STA/LTA filtering is applied (blue, LP 128 s, $t_{user,E} = 15$, $n_{WL} = 1,000$, $n_{bin} = 1$, $\alpha = 20$, $n_{step} = 10$). Sample positions where the synthetic steps were detected and removed are marked by green dashed lines.

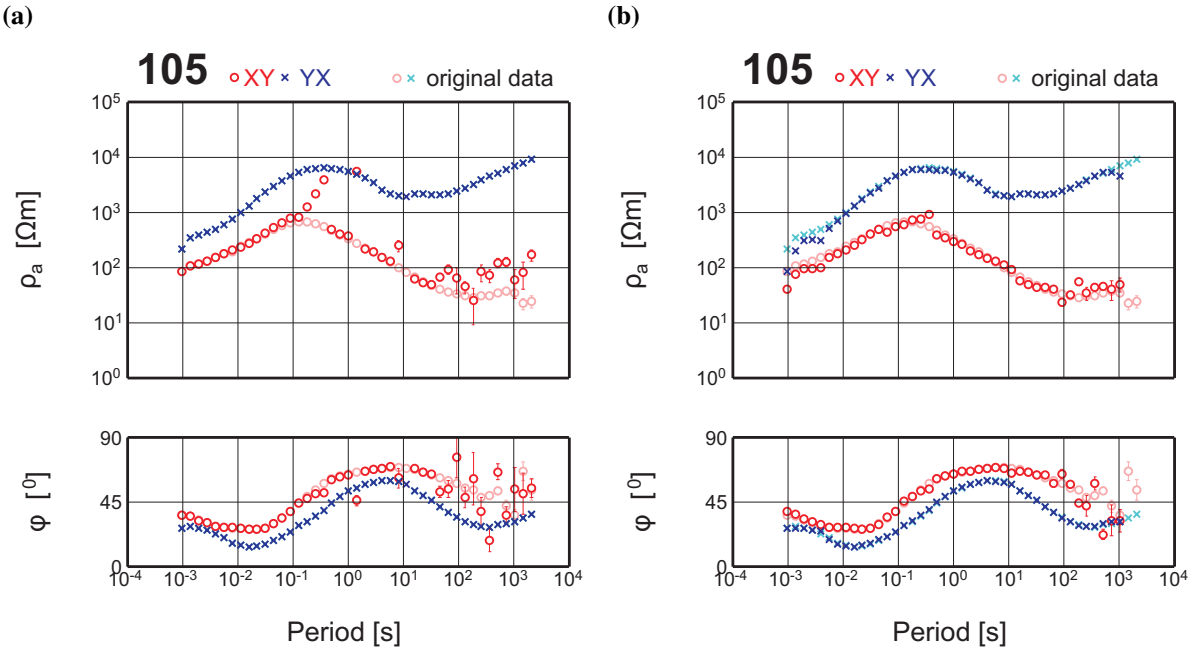


Figure 2.58: ***Ex. 2***: Apparent resistivity ρ_a and phase φ with **(a)** 10% arbitrary steps in Ex ($n_{mult} \in [0, 10)$, $o_{user} = 0.001$) and **(b)** after applying the STA/LTA filter ($t_{user,B} = 100,000$, $t_{user,E} = 15$, $n_{WL} = 1,000$, $n_{bin} = 1$, $\alpha = 20$, $n_{step} = 10$). Original curves are shown in the background in lighter colours.

For ***ex. 3***, 10 % *arbitrary spikes* are added to all magnetic and 10 % *arbitrary steps* to both electric channels. Based on the filtering results of ***ex. 1*** and ***ex. 2***, $t_{user,B}$ and $t_{user,E}$ are set to 20 and 15, respectively. Figure 2.59 **(b)** shows a clear improvement of the data after filtering. For a better understanding of the filtering, the processing results where either the magnetic (Fig. 2.60 **(a)**) or electric channels (Fig. 2.60 **(b)**) are filtered are displayed as well. Solely treating the magnetic channels yields a higher data improvement and for periods shorter than 10^{-1} s the results are even better than for filtering both channel types. However, for periods longer than 10^{-1} s, filtering all channels leads to the best results. This shows that while filtering data, testing different filter settings is essential. Even for data from a single site different filtering approaches might be necessary for different period ranges.

If all data channels are distorted by noise, cleaning the magnetic channels is certainly a main issue. Nevertheless, treating the electric channels is important as well, especially for longer period data.

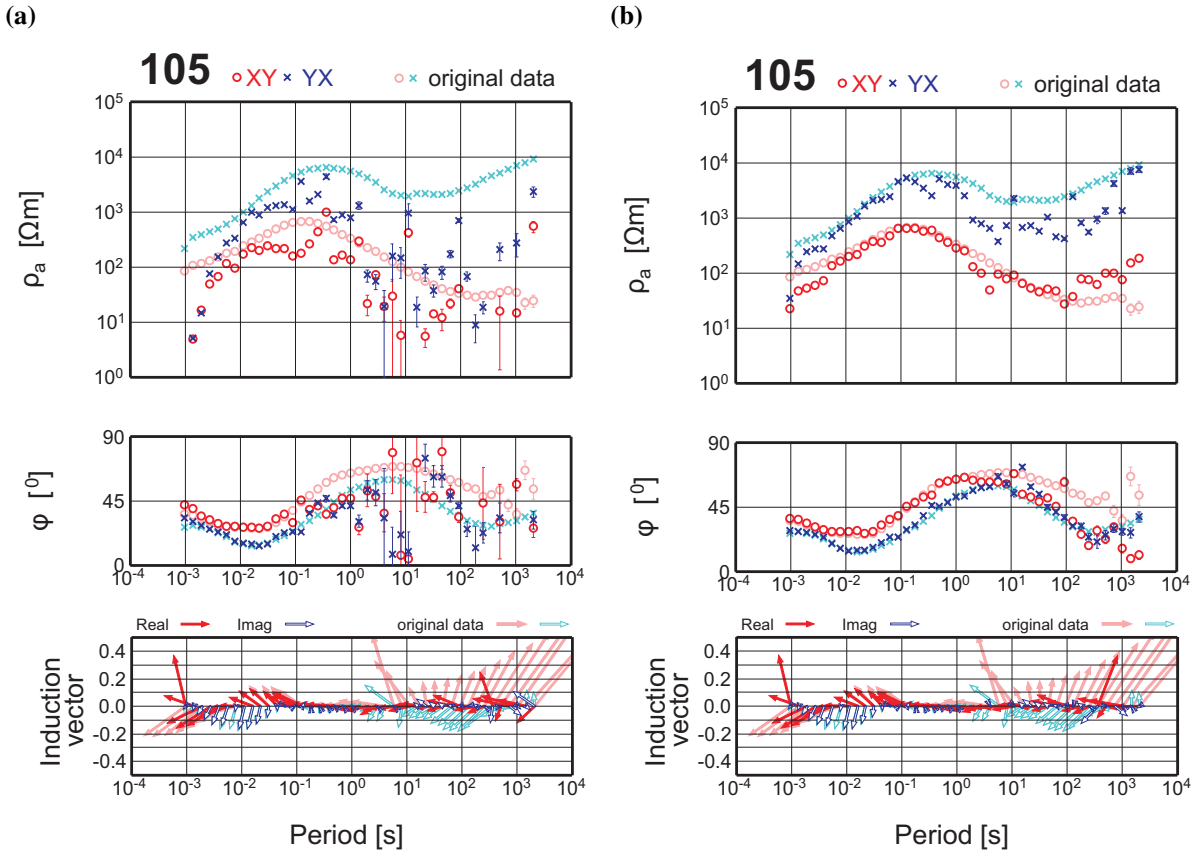


Figure 2.59: ***Ex. 3***: Apparent resistivity ρ_a and phase φ with **(a)** 10% arbitrary spikes and steps in all magnetic and electric channels, respectively ($n_{mult} \in [0, 10)$, $o_{user} = 0.001$) and **(b)** after applying the STA/LTA filter ($t_{user,B} = 20$, $t_{user,E} = 15$, $n_{WL} = 1,000$, $n_{bin} = 1$, $\alpha = 20$, $n_{step} = 10$). Original curves are shown in the background in lighter colours.

- **Additional parameter settings**
  Additional parameters which need to be defined for the STA/LTA filter are

  – the window length n_{WL},

  – the number of pre- and post-event (either spike or step) samples n_{bin} and

  – the number of samples for step removal n_{step}.

  The adjustment of these control values depends on the distribution of the noise within the time series, the width of the noise events and their appearance and characteristics. Choosing an appropriate n_{WL} is similar to defining the window length for the reference site filter (cf. sec. 2.3.1). Values between 100 and 1,000 samples yield the best results.
  The STA/LTA filter is most effective for removing noise which only distorts a few consecutive samples and has high amplitudes since the noisy data is simply interpolated. Hence, n_{bin} and n_{step} should be kept small (< 10 samples) to affect the least possible number of samples.

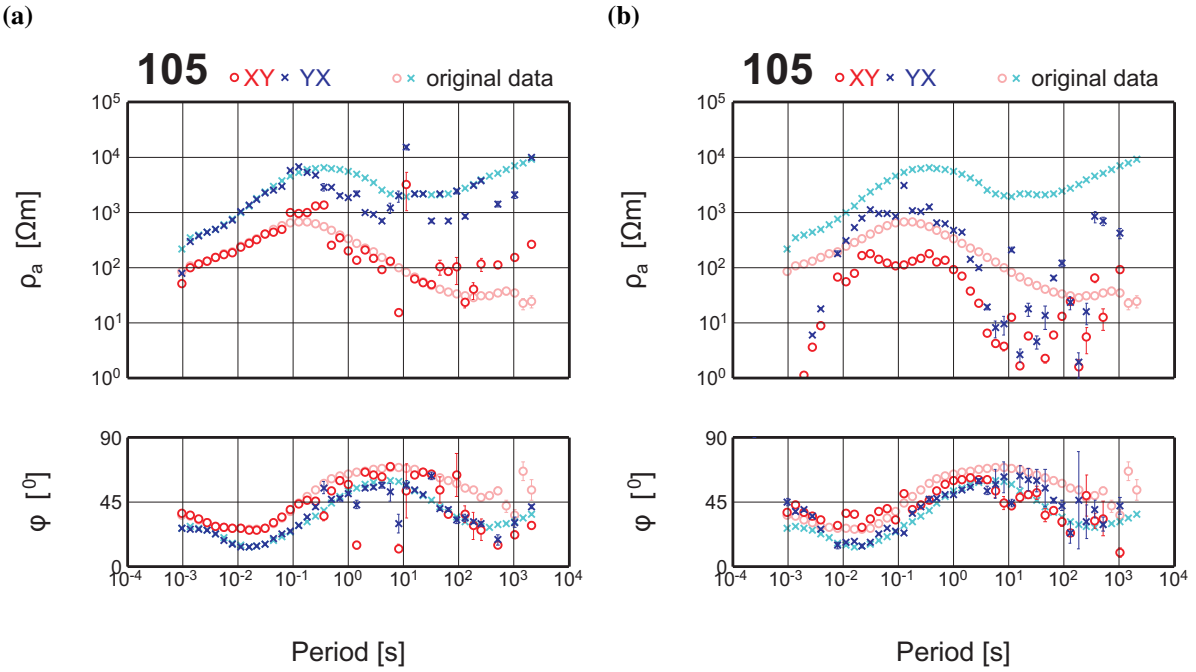


Figure 2.60: ***Ex. 3***: Apparent resistivity ρ_a and phase φ with 10% arbitrary spikes and steps in all magnetic and electric channels, respectively, after applying the STA/LTA filter either on **(a)** the magnetic or **(b)** the electric channels ($t_{user,B} = 20$, $t_{user,E} = 15$, $n_{WL} = 1,000$, $n_{bin} = 1$, $\alpha = 20$, $n_{step} = 10$). Original curves are shown in the background in lighter colours.

MT data from the Barberton Greenstone Belt

As an example for the STA/LTA filtering tests, site *005* is chosen.

For 'real-world-data', the percentage of noise within the time series is not known and can only be roughly estimated. Table 2.7 summarizes the percentages of marked samples for all five channels using different t_{user} values (band LP 128 s, see Tab. A.1). Neither a clear convergence of these percentages to a certain value nor a specifically noisy channel can be identified. The higher percentages for the electric channels can be attributed to the higher value of n_{bin} ($n_{bin,H} = 1$, $n_{bin,E} = 10$) and are not due to a higher noise level of the electric channels. The higher bin value for the electric channels is chosen due to the pronounced effect the steps have on the samples neighbouring them.

Table 2.7: Application of STA/LTA filter to site *005*: Percentage of marked samples for all channels for varying magnetic and electric multipliers t_{user} (LP 128 s, $n_{WL} = 1000$, $n_{bin} = 1$ for magnetic channels, $n_{bin} = 10$ for electric channels, $\alpha = 20$, $n_{step} = 10$).

$t_{user,H}$ and $t_{user,E}$	5	10	20	30	35
Samples Hx %	26.5	7.0	1.8	1.0	0.9
Samples Hy %	27.5	8.1	1.8	1.1	0.9
Samples Hz %	26.0	5.1	0.8	0.3	0.2
Samples Ex %	44.7	11.8	5.6	3.1	2.7
Samples Ey %	48.3	10.1	2.1	0.9	0.6

As for the synthetic examples, a comparison of the time series (Fig. 2.61) and the spectra (Fig. 2.62) of the original time series and the filtered time series can be helpful for determining the best filter settings.

The shape of the time series has clearly improved after filtering with $t_{user} = 10$: Major spikes and steps could be removed from the data. Nevertheless, an unknown amount of noise remains in the data, especially obvious for the highly distorted channel Ex. All spectra of the channels show an decrease in their general amplitude after filtering, especially distinct for the electric channels. Harmonic signals as observed in channel Ey (Fig. 2.62 for $T > 10^3$) could also be eliminated by the filter.
Since it can only be guessed what the time series without noise would look like a simultaneous close analysis of the processing results for the transfer functions is essential (Fig. 2.63). t_{user} values between 10 and 20 yield the best processing results with the highest improvement between 1 and 10 s.

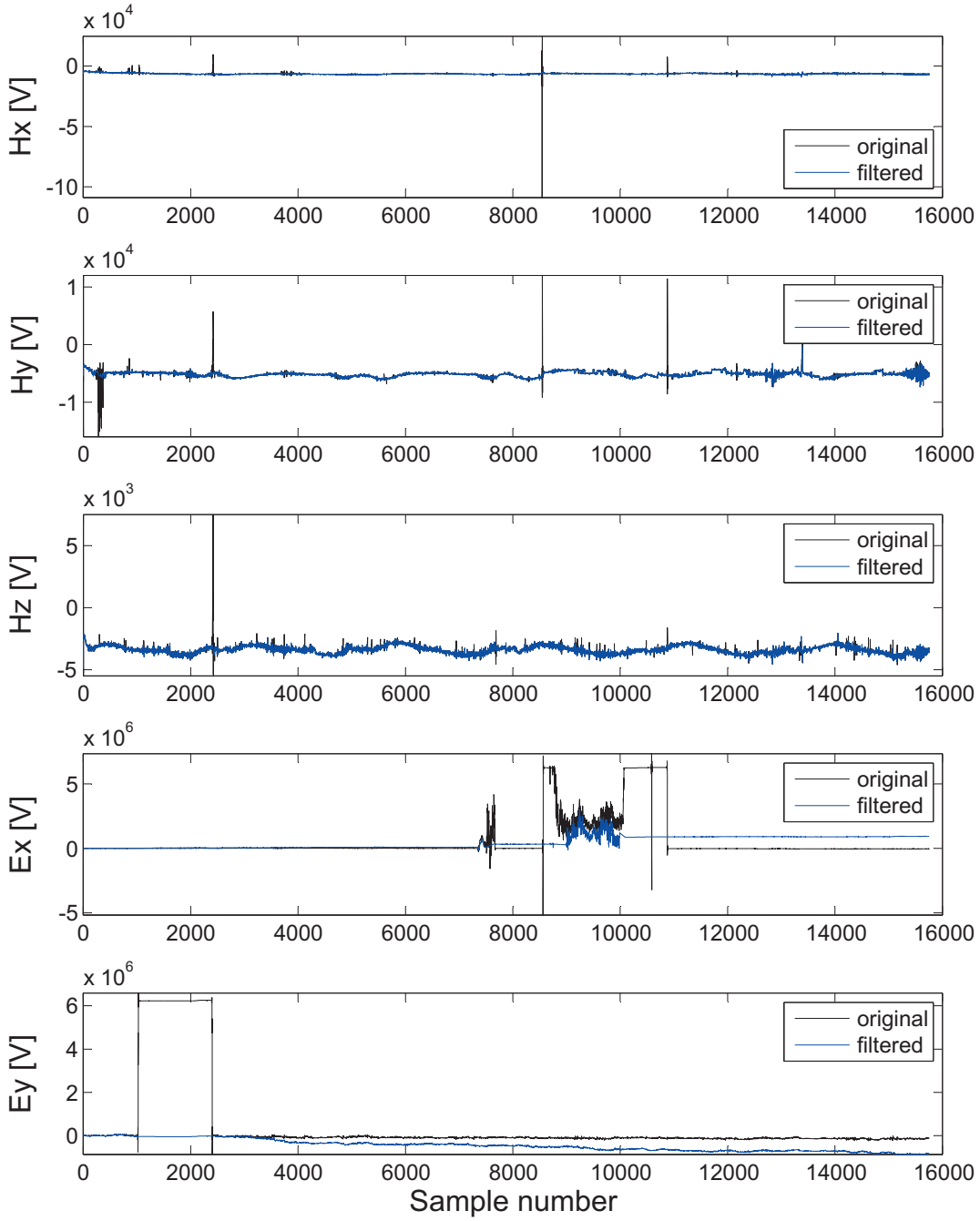


Figure 2.61: Comparison of the original time series and after STA/LTA filtering is applied (site 005, LP 128 s, ~ 6 days, $t_{user} = 10$, $n_{WL} = 1000$, $n_{bin} = 1$ for magnetic channels, $n_{bin} = 10$ for electric channels, $\alpha = 20$, $n_{step} = 10$).

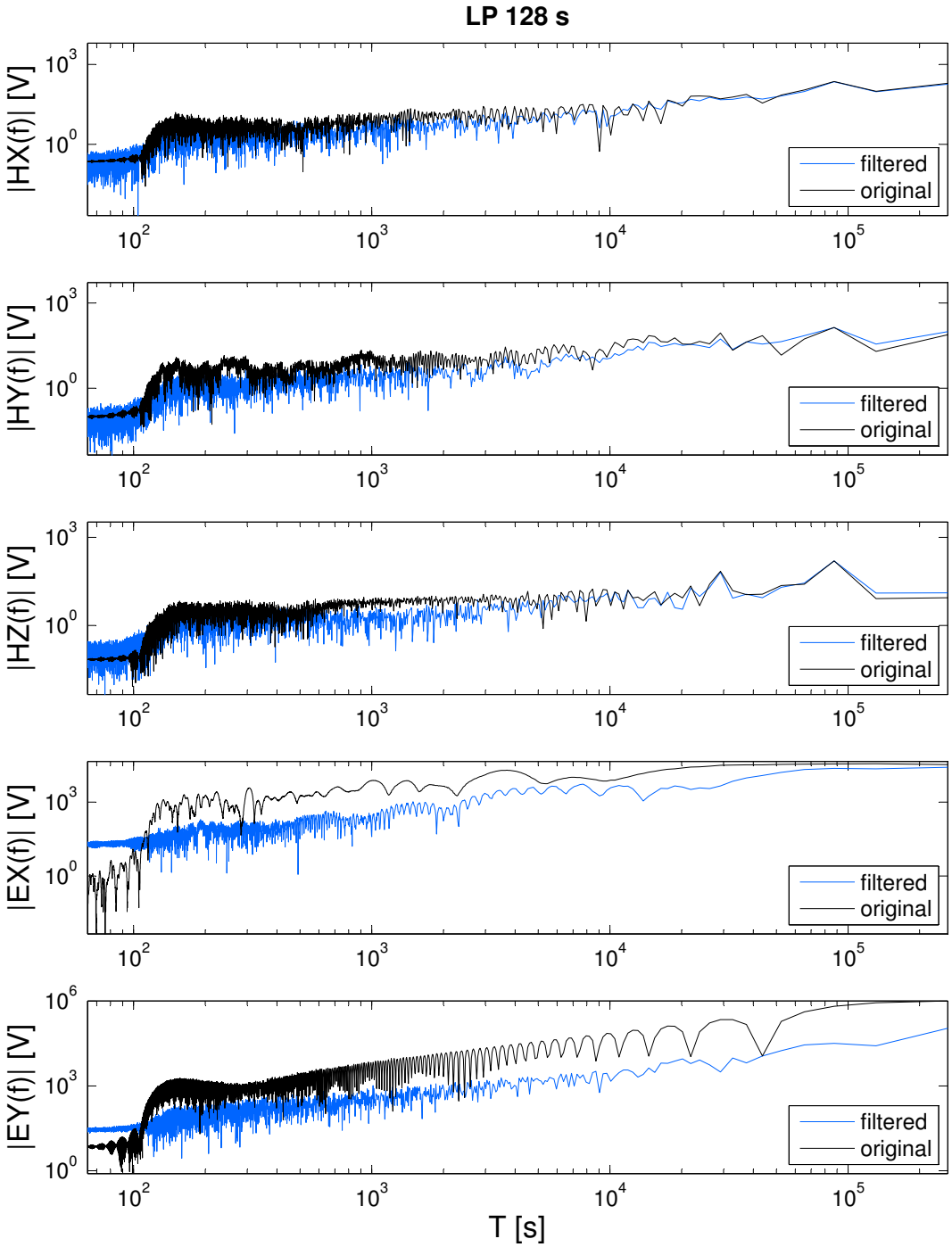


Figure 2.62: Amplitude spectra of site 005 for the original and STA/LTA filtered data (t_{user} = 10, n_{WL} = 1000, n_{bin} = 1 for magnetic channels, n_{bin} = 10 for electric channels, α = 20, n_{step} = 10). No. of samples for spectra calculations: LP 128 s: 8,192.

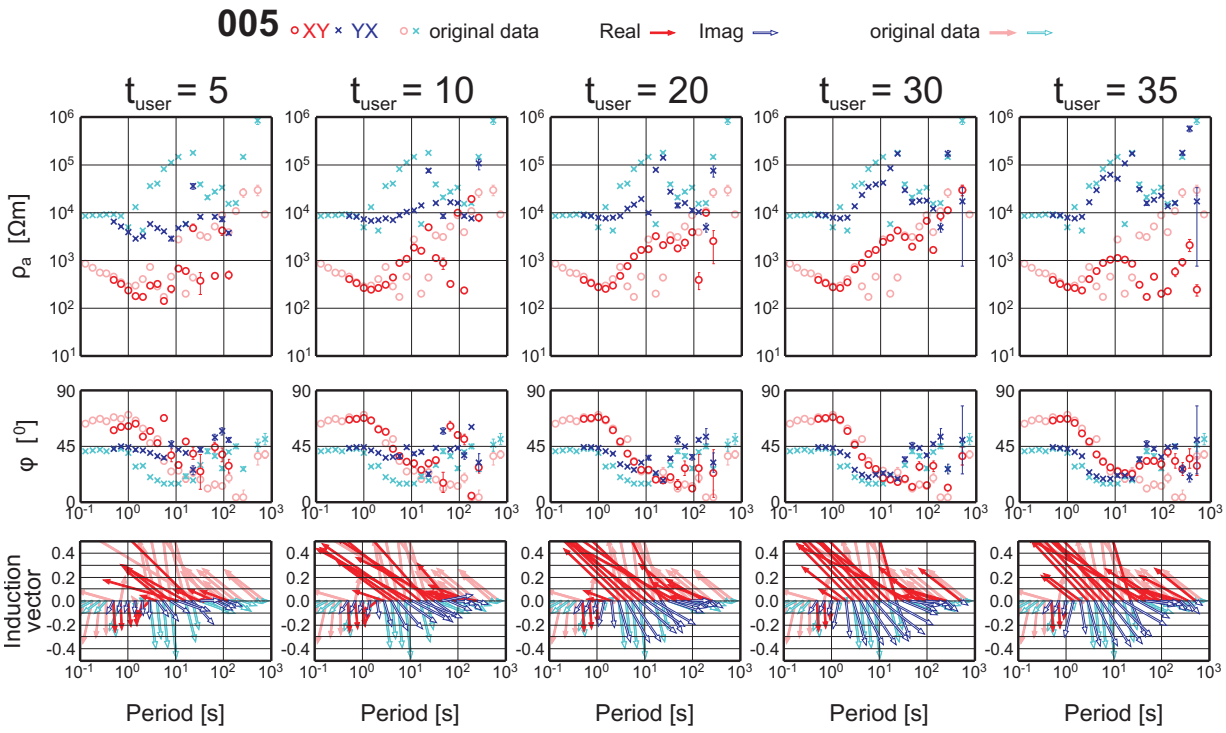


Figure 2.63: Long period data of site 005. Comparison of STA/LTA filter results for varying t_{user} values. t_{user} = 10 or 20 yields the best results. Original curves and induction vectors are shown in the background in lighter colours.

2.4 Processing strategy and post-processing

Filtering MT data is no straightforward procedure. Depending on the observed EM noise (e.g. power line and arbitrary noise), different period ranges might require a different treatment and several filtering methods need to be applied. Even for a specific noise type, the best adapted filtering method might not perform equally well for every MT site of the data set. This is mainly due to the location of MT sites in regard to the EM noise source which are different for every site resulting in a different manifestation of the EM noise in the data. Presence of other EM noise sources which superimpose each other in the MT recordings further hinder the successful application of filtering and processing methods. Hence, processing and filtering settings need to be adapted to the MT sites which can be very time consuming for large data sets.

For the Barberton MT data set, notch and delay filtering was carried out first. Delay filtering with a frequency of 50 Hz produced for most of the MT sites the best results. After applying the delay filter to the reference site data (to avoid bias, see sec. 2.2), all BGB MT sites were processed using the remote reference method. Improvement in the data quality of the MT transfer functions could be achieved up to a period of 1 s and sometimes even up to 10 s with this method.

The two methods for data quality improvement developed in this thesis - *reference site filter* and STA/LTA filter - were only applied to periods longer than 1 s (i.e. bands LP 2 s, LP 16 s and LP 128 s, see Fig. A.2 and Tab. A.1). This was done since (i) arbitrary spikes and steps primarily affect long periods (see sec. 2.1), (ii) the data below 1 s could be improved sufficiently with the delay filter and the RR method (see sec. 2.2) and (iii) since processing of the entire period range requires a lot of time. In general, the *reference site filter* and STA/LTA filter were applied several times to each MT site using different filter

settings (e.g. different window lengths for the filtering and varying threshold levels, see sec. 2.3.1 and 2.3.2): 32 different processing results with the *reference site filter* and 20 different processing results with the STA/LTA filter. The two filter methods were applied to the time series either before or after the bandpass filterering of the data (see sec. A.4). Application of the filters after bandpass filtering yielded normally the best processing results.

After using the *reference site filter* some MT sites still contained spikes or steps in their time series (cf. integrated time series in Fig. 2.48). This is either caused by spikes or steps in the other channels of the local site (which were not detected by the *reference site filter*) or by noise in the reference site data. In these cases, the STA/LTA filter was applied additionaly to remove the spikes and steps.

Together with the processing results after the application of the delay filter and the RR method, almost 60 different transfer functions needed to be compared visually (!) for each of the 193 MT sites of the BGB data set. After discarding the filtering results which led to no data quality improvements, the transfer functions were assembled from the best processing results for the different period ranges (programme for assembly of transfer functions written by Dr. Paul Sass).

Due to severe (!) noise, the application of the mentioned filtering and processing methods did not lead to a (required) major improvement in the Hz component of the MT data. Hence, the magnetic transfer functions (or induction vectors) still exhibit severe distortions even after intense filtering and processing. Therefore, they are not used for any further analysis and are not included in the 2D or 3D inversions.

Figures B.5 and B.6 (in app. B) show pseudosections of the filtered data which is used in the next chapter for the 2D and 3D inversion studies.

Chapter 3

Electrical conductivity structure of the Barberton Greenstone Belt

The Barberton Greenstone Belt (BGB) directly located at the border between South Africa and Swaziland (Fig. 3.1) is one of the few remnants of Archean rocks that have been well-preserved throughout time. The BGB has been extensively studied for over a century mainly in regard to its structure, mineralogy and traces of early life forms (e.g. Hall, 1918; Anhaeusser, 1969; Knoll and Barghoorn, 1977; de Ronde and de Wit, 1994). Studying the BGB allows scientists to gain knowledge of the evolution and the geological mechanisms which played a role in the development of the Earth's crust during its early stages.
Especially over the last several decades very controversial theories regarding the formation of the BGB have been established (e.g. Heubeck and Lowe, 1994; Moyen et al., 2007). These theories focus mainly on the question whether plate tectonics as observed today took place already during Archean times or if other mass transportation processes were active during the generation of the early crust. To determine the plausibility of the theories a detailed understanding of the internal structure of the BGB is necessary. Unfortunatelly, geoscientific studies carried out so far have mainly focused on the surface structure of the BGB and information about greater depths is sparse (de Beer and Stettler, 1988; de Beer et al., 1988). The magnetotelluric (MT) data and its analysis presented in this work is a first step to improve the understanding of the inner structures of the BGB and their origin.
The MT data set collected at the southern part of the BGB consists of ~ 200 sites which are aligned along six profiles. The extensive filtering and cleaning process of the highly noisy data (see chap. 1) described in chapter 2 led to the data set shown in Fig. B.5 and B.6 (in app. B). This data set is used as basis for further analysis applying 2D (RLM2DI algorithm by Rodi and Mackie, 2001, implemented in WinGlink) and 3D (ModEM by Egbert and Kelbert, 2012) inversion methods. The conductivity models obtained by 2D and 3D inversion are interpreted in a geological context which is the main topic of this chapter.

3.1 Geology of the Barberton Greenstone Belt

The geological investigation of the BGB started centuries ago and was mainly fueled by the interest in the mining of natural resources (Hall, 1918; Anhaeusser, 1969; de Ronde and de Wit, 1994; Anhaeusser, 2012; Dirks et al., 2013). Especially, the discovery of gold deposits in the second half of the 19th century expedited the geoscientific exploration of the BGB and led to the foundation of the town Barberton (An-

haeusser, 2012). Additional mineral discoveries of chrysotile asbestos, iron ore, magnesite, talc, barite and the more recent finding of a nickel-sulphide deposit have ensured the more or less uninterrupted mining of the BGB from the 19th century unil today (Anhaeusser, 2012).
In part, the geoscientific studies address the BGB as a whole but often they focus on smaller regions or specific rock types as well (e.g. Heubeck and Lowe, 1994; Parman et al., 1997). This led to a very detailed knowledge of the geologic and tectonic structures at the surface (Fig. 3.1 and 3.2). Due to its origin in the Archean and its good preservation, the BGB is also an ideal place to investigate the formation and evolution of the crust during the early stages of the Earth's development.
Only few geophysical studies have so far addressed the deeper structure of the BGB (de Beer and Stettler, 1988a; de Beer et al., 1988b; sec. 3.2). Hence, the now existing BGB evolution theories are based mainly on geological and chronological data and are highly controversial with regard to the main tectonic forces and processes which were dominant in the Archean. Tectonics on the Early Earth might have been similar to the plate movement and their driving forces (i.e. seafloor spreading and subduction) that we observe today. However, regarding that some fundamental conditions like the thermal framework of the Earth (i.e. heat flow), magmatic production rate and the first crust (and its stability) were considerably different at this time (Poujol et al., 2003) other processes like vertical mass transport might have been a governing factor in early tectonics.

The BGB as part of the Kaapvaal Craton

The BGB is located in the eastern part of the Kaapvaal craton (Fig. 3.1, inset map) and is one of a number of greenstone belts which can be found in this region (Giyani, Pietersburg and Murchison greenstone belt). The accretion of this ~ 1,200,000 km^2 large area occured in several stages and spans a time from 3.5 to 2.5 Ga ago (Poujol et al., 2003). It is stated that the growth of the craton which consists of periods of magmatic arc formation and accretion and tectonic amalgamation of terranes and blocks marks the onset of processes which are similar to modern-day plate tectonics (de Wit et al., 1992). In the north, the Kaapvaal craton is borderd by the Zimbabwean craton. The southern and western boundaries are Proterozoic mobile belts (Namaqua-Natal Mobile Belt). In the east, the edge of the Kaapvaal craton is marked by the Lebombo monocline which is associated with the Gondwana break-up (de Wit et al., 1992). Seismological and petrological studies give a general depth to the Moho of 35 to 37 km and a depth to the base of the craton of about 350 km (de Wit et al., 1992 and references therein; de Beer and Stettler, 1988a).
Since the evolution of the BGB and its southern neighbouring geological units marks the early stages of the Kaapvaal craton formation, understanding the geological history of the BGB is essential for the understanding of the craton development, as well. Stabilization of the eastern Kaapvaal craton took place between 3.1 to 2.7 Ga which coincides with the emplacement of large granite bodies and the last magmatic activities in this area (Schoene et al., 2009). Geoscientists benefit from this stabiliztion since it prevented any further alteration of the BGB and led to its good present-day preservation.
Large parts of the craton are covered by volcanic and sedimentary sequences deposited in Neoarchean to Palaeoproterozoic times (Poujol et al., 2003). This is also the case for the region NW of the BGB where a sedimentary cover belonging to the Transvaal Supergroup can be found (Fig. 3.1, inset map). All MT profiles extend to the sedimentary cover, as well. Unfortunately, no geoscientific information is available for the Transvaal Group sediments in this region which makes the interpretation of the MT data in this

part difficult. However, studies in the Griqualand West sub-basin which is also a part of the Transvaal Supergroup investigate the sedimentary sequences found there (Schröder et al., 2006). Underneath up to 1 000 m thick chert and Banded Iron Formations (BIF), siliciclastic mudstone, limestone, sandstone and conglomerate layers are reported by the authors and the thickness of the sedimentary cover is between 1 150 and 1 900 m in total.

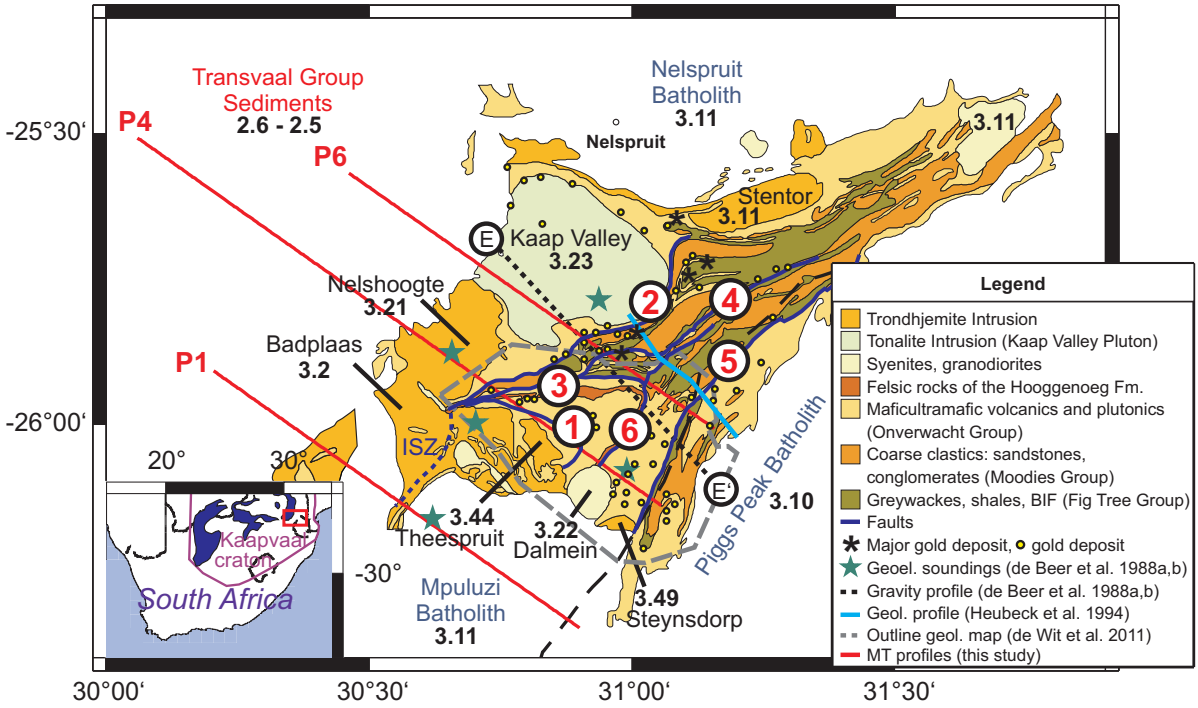


Figure 3.1: Geological map of the Barberton Greenstone Belt and location of selected MT profiles (red lines). Faults (in blue): ① Komati, ② Moodies, ③ Inyoka, ④ Saddleback, ⑤ Maanhaar, ⑥ Kromberg (after de Ronde and de Wit, 1994 and Lowe, 1994). Black numbers in the map denote the ages (in Ga) of geological units taken from de Ronde and de Wit (1994). Location of gold deposits from Dirks et al. (2013). Inset map shows the location of the measurement area within South Africa (red rectangle), the Kaapvaal craton (purple outline) and the Transvaal Supergroup (blue areas, map after Schoene and Bowring, 2007; de Ronde and de Wit, 1994) and Schröder et al..

General geologic setting of the BGB

The BGB is an Archean fold and thrust belt with a complex assembly of geological structures and tectonic features which have formed over a time of about 490 Ma (de Ronde and de Wit, 1994). The depth to the boundary between the Earth's crust and the mantle (Mohorovičić discontinuity or short Moho) is unknown in the Barberton area. South of the BGB in the Kaapvaal craton it is reported to be around 35 km (de Beer and Stettler, 1988a).

Regarding the subsurface structure of the BGB, highly controversial models have been established over time. This is best seen in the strongly varying range of the depth extent of the BGB: Thickness are stated to be between ~4 km (Darracott, 1975) up to 20 km (Anhaeusser, 1986). Theories and observations made so far of the BGB's inner architecture are the subject of the following sections.

At the surface, three different geological units can be identified: the Onverwacht, the Moodies and the Fig Tree Group (Fig. 3.1 and 3.2). These units are in part bordered by shear zones, faults and sutures which can be traced throughout the BGB. The Saddleback-Inyoka Shear Zone, e.g. divides the BGB

in a northern and a southern terrane which is bounded in the SE by the Manhaar fault (de Ronde and de Wit, 1994; MacLennan, 2012). Since the surrounding plutons and batholiths played an important role in the formation of the BGB, they are discussed in this section as well (Fig. 3.2). Regarding their general properties, two major rock units can be distinguished: (i) The low metamorphic grade supracrustal rocks of the BGB itself and (ii) the surrounding higher grade granite-gneiss units (Schoene et al., 2008).

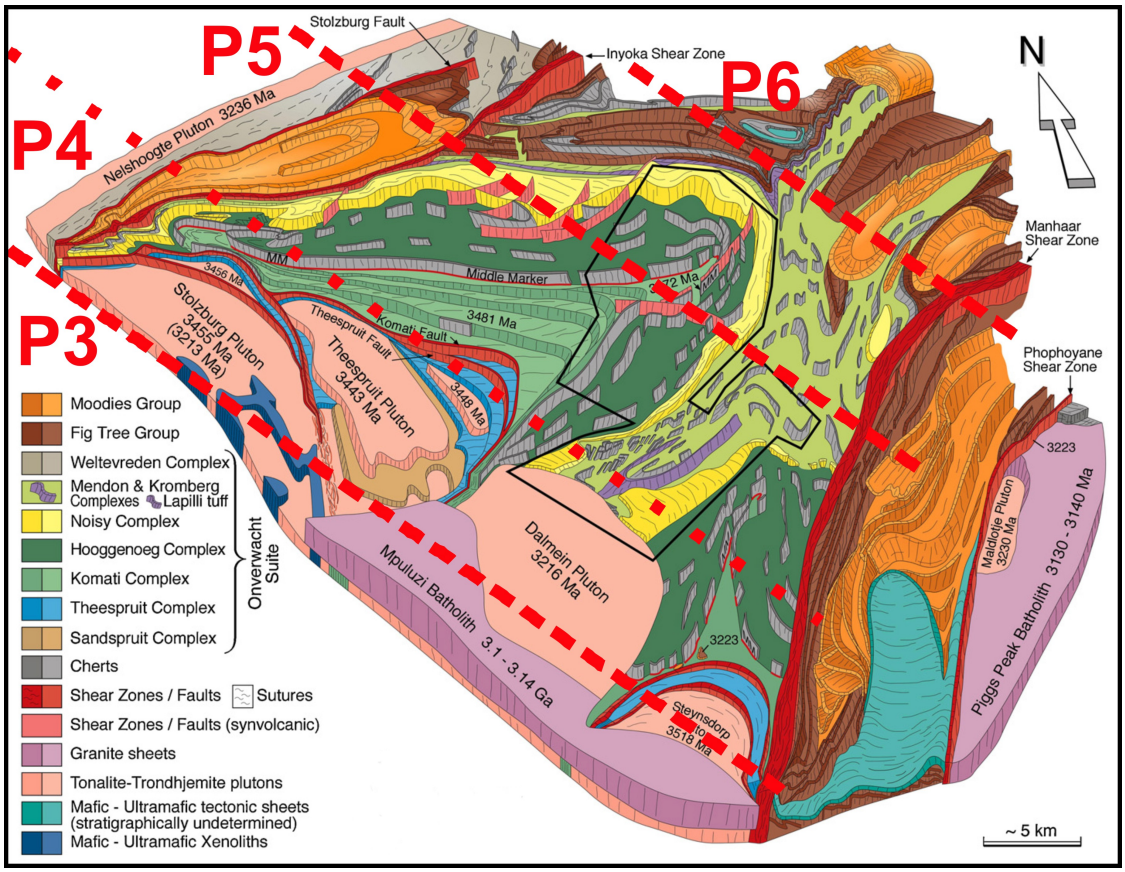


Figure 3.2: Detailed geological map of the southern part of the BGB. MT profiles lines are marked by dashed/dotted red lines. Map from de Wit et al. (2011).

Onverwacht Group

The Onverwacht Group is the lowest and oldest unit in the BGB with ages between 3.53/3.49 and 3.45 Ga (de Ronde and de Wit, 1994; van Kranendonk et al., 2009). First thickness estimates were around 13 km (Viljoen and Viljoen, 1969), later revised to be around 3.5 - 4 km (de Wit, 1982; de Wit et al., 1983). The overestimation was explained by a tectonic repetition.

The Onverwacht Group mainly consists of mafic (silicate mineral rich in magnesium and iron) and ultramafic (igneous rocks with usually more than 90% mafic minerals) intrusive and extrusive rocks (Neuendorf et al., 2005). These rocks make up the different complexes which form the Onverwacht Group (in some newer publications it is called Onverwacht Suite): Weltevreden, Mendon, Kromberg, Noisy, Hooggenoeg, Komati, Theespruit and Sandspruit complex. Chronology data showed that the contact between most of the complexes is of a tectonic and/or unconformable nature (de Wit et al., 2011).

The Theespruit complex which is located SW of the Komati fault is considered to be the oldest part of the BGB with rocks reaching ages of about 3.53 Ga (van Kranendonk et al., 2009). In this complex, interlayered pillow basalts and orthogneisses can be found and one of the few occurences of sedimentary

(diamictite) rocks within the Onverwacht Group van Kranendonk et al., 2009).
Directly north of the Komati fault, the Komati complex is located which consists of (serpentinised) komatiites, (tholeiitic) basalts and felsic tuff. The age of formation of the complex determined around 3.48 Ga (Armstrong et al., 1990) and the age of the all-over metamorphic overprinting around 3.45 to 3.49 Ga suggest that both processes occured simultaneously in a MOR setting with hydrothermal alteration (de Wit et al., 1987).
The rest of the complexes show a similar composition of komatiites and basalts with the exception of the Weltevreden Complex. Here outcrops of BIFs and cherts can be found towards the NE (Ronde and Kamo, 2000; MacLennan, 2012).

Fig Tree Group

The Fig Tree Group was deposited conformably above the Onverwacht Group between 3.26 - 3.22 Ga and is of sedimentary origin (Heubeck and Lowe, 1994b, and references therein). The same authors assumed a thickness between 1 - 3 km for this unit. The siliciclastic sequences are fine grained and contain immature sandstone, tuffaceous and ferruginous sedimentary rocks, cherts and BIFs (Heubeck and Lowe, 1994a).

Moodies Group

The Moodies Group is of the same origin as the Fig Tree Group although it is slightly younger (3.22 - 3.10 Ga, Heubeck and Lowe, 1994a). It was deposited in the final stages of internal thrusting of the BGB (Heubeck and Lowe, 1994b, and references therein). The siliciclastic rocks are coarser-grained than the Fig Tree Group sediments with a assumed thickness of up to 3.7 km of quartzose sandstone, subordinate siltstone and minor polymict (clasts of many different lithologies) conglomerate, shale, and volcanic rock (Heubeck and Lowe, 1994a, and references therein). It is also reported that it contains iron formations and shales (de Beer et al., 1988b).

Sutures, shear zones and faults

The BGB is intertwined throughout with numerous major and minor shear zones and faults (Fig. 3.1). The Saddleback-Inyoka fault system has even been identified by de Ronde and de Wit (1994) as a major suture and interpreted as the collision zone between two arc and trench blocks formed during stage two of the BGB formation (around 3.2 Ga). A southward continuation of the suture in the Inyoni Shear Zone is suggested (Moyen et al. (2006)). Other faults like the Moodies fault are of a similar age as the Inyoka system and supposedly formed during a major collision phase during the BGB's formation (see Fig. 3.3). The fault zones in the central part of the BGB are at the surface marked by zones of strained and foliated rocks up to a width of 40 m (Heubeck and Lowe, 1994b). Most boundaries between the BGB and its surrounding plutons are faults which are vertically to subvertically dipping (Heubeck and Lowe, 1994b). In the final stages of the BGB evolution, the extensive fault sytem most likely facilitated the hydrothermal overprint with ore-rich fluids. These alterations include serpentinization, steatitization, and carbonatization processes (Heubeck and Lowe, 1994b).

Plutonic bodies/granite-gneiss terranes

Along the borders of the BGB numerous intrusive bodies can be found which vary in size from the relatively small Steynsdorp Pluton (age: 3.52 Ga) to the vast structure of the Mpuluzi batholith (age: 3.1 Ga, see Fig. 3.2). These bodies document the magmatic activities which occured in the Barberton area.
Regarding their origin and composition, the granitoid bodies can be divided into two types: (i) the tonalite-trondhjemite-granodiorite (TTG) suite (plutons emplaced between 3.5 - 3.2 Ga), and (ii) the granite-monzogranite-syenite-granite (GMS) suite (batholiths emplaced between 3.2 3.1 Ga).

For the TTG bodies, Clemens et al. (2006) suggest a garnet-rich amphibolitic or eclogitic depleted upper mantle source in a depth of at least 40 km which is also supported by experimental studies. Over time the magma source changed. For the GMS batholiths a plagioclase-rich, garnet-poor source where at least parts of the already existing greenstone crust got mixed in the magma is proposed (Yearron et al., 2003). The Mpuluzi batholith which dominates the southern part of the measurement area is assumed to be a mostly thin (around 500 m) shallowly dipping granite sheets (Kamo and Davis, 1994, and references therein). Underneath, a network of steeply dipping, variably deformed dikes and sheets can be found (Westraat et al., 2005). The exact origin of the batholiths is still questioned, however, their emplacement occured at the same time as a transition from a largely convergent (at ca. 3.23 Ga) to a transtensional tectonic regime in the BGB at around 3.08 Ga (de Ronde and de Wit, 1994).

Evolution theories of the BGB

The evolution theories centre on the type of tectonic processes responsible for the formation of geological units that make up the BGB. Some theories favour solely vertically driven tectonics, early plate tectonic settings or a mixture of both (Fig. 3.3).
The onset of present-day style plate tectonics is heavily discussed. Ages range from 4.3 Ga (Harrison et al., 2008) to 3.3 - 3.5 Ga (Moyen et al., 2006; Smithies et al., 2007; Van Kranendonk et al., 2007a). According to Hamilton (1998), preliminary plate tectonic processes such as rifting and convergence only started around 2.0 Ga and was fully established around 0.8 Ga.
The formation of the BGB spans over ~ 490 Ma (3.57 - 3.08 Ga, de Ronde and de Wit, 1994). Various authors (e.g. de Ronde and de Wit, 1994; Ward, 1995) propose a tectonogenesis of the BGB similar to present-day style plate tectonics. Based on geo-chronological data, the tectonic evolution took place in disjointed intervals and can be divided into different stages. Between 3.49 and 3.45 Ga submarine eruptions in a mid-ocean ridge-like (MOR) setting led to the formation of the ultramafic and mafic volcanic rocks of the Onverwacht Group (cf. left sketch in Fig. 3.3). A thrusting and folding event of existing supracrustal sequences (called D1 deformation) occured around 3.4 Ga. The second stage can be subdivided in a tectonic environment comparable to an intra-oceanic suprasubduction system (3.45 - 3.42 Ga) and an intra-arc and inter-arc setting. This led in the third stage to the amalgamation of the arc (3.23 - 3.08) and the creation of major shear zones. The final stage (around 3.08 Ga) is described as a transition from a convergent to a transtensional regime and is marked by hydrothermal activities as well. The tensional strain led to NW-SE and NE-SE striking fracture network and dyke swarm. The mentioned hydrothermal activities manifested in mesothermal gold mineralization especially pronounced along the BGB fault system and occurred around 3.1 - 3.0 Ga (Kakegawa and Ohmoto, 1999; Dirks et al., 2013). Furthermore, the supracrustal rocks of the BGB have been heavily altered by metasomatism (de Wit et al., 1987; de Ronde and de Wit, 1994). Ages of igneous and metasomatic rocks have been found to be very similar. This indicates that the metasomatic alterations occurred directly after solidification of the igneous rocks (de Wit et al., 1987). Fluid inclusions in sample rocks and studies on C, O and Sr isotope compositions have identified the fluid responsible for most of the alterations as modified seawater (de Ronde and de Wit, 1994). A conceptual model of a MOR where convection of seawater through the newly generated ultramafic and mafic rocks is the most likely explenation for the made observations (de Ronde and de Wit, 1994, and references therein).
The formation of the Fig Tree Group mostly falls before the beginning of the third stage (3.26 - 3.23 Ga)

although some Fig Tree Group rocks even exhibit an age of 3.45 Ga. The siliciclastic rocks of the Moodies Group were deposited during the collisional period of the third stage (3.23 - 3.11 Ga, D2 collision and collapse in Fig. 3.3). The possibility of an oceanic setting during early Archean times with a later onset of shallow subduction and collision of small proto-continents and/or island arcs is also suggested by Lowe (1994), Foley et al. (2003) and Moyen et al. (2006). These authors even compare the evolution of the BGB to an early small scale Wilson cycle with a periodic opening and closing of ocean basins.
A major suture zone identified by de Ronde and de Wit (1994) - the Saddleback-Inyoka fault system (faults ③ and ④ in Fig. 3.1) - marks the juncture between two arc and trench blocks formed during the D2 collision stage (~3.2 Ga) and is regarded as a further prove for early plate tectonics. In the parts where the Saddleback and the Inyoka fault are separated this suture zone has a width of several hundred metres to a couple of kilometres.
Theories on a crustal evolution driven by vertical mass movement are based on the higher temperature gradient, higher heat flow and resulting fast convection and crustal recycling during the Archean. Hence, the presence of extended areas with stabilized crust which is a necessary prerequisite for plate tectonics is questioned (Hamilton, 1998). Hamilton (1998) describes the formation of Archean crust as follows: Periodic (discontinuous) intervals of submarine volcanism produced thick layers of volcanic rocks while during times of volcanic quiet sedimentary rocks were deposited. The author further states that the deformation of the greenstone belt expressed by synclines, anticlines, subvertical fault zones and steep bedding planes was caused during the final stages of the formation by the emplacement of diapiric batholiths which resulted in vertical or horizontal displacement of the geological units. This displacement could also have caused buckle folding and subsequent subhorizontal detachment and delamination in the older parts of the BGB (e.g. Onverwacht and Fig Tree Group rocks, right sketch in Fig. 3.3). These processes would have been facilitated by the difference in mechanic strength between the basement and the supracrustal rocks (mechanical weakness of the Onverwacht Group due to serpentinized ultramafic rocks, Heubeck and Lowe, 1994). Due to the lack of outcrops, the type of contact between the BGB and the basement is an area for speculations. The deposition of the Moodies Group is suggested to have occured during an extensional period above or adjacent to thrust-faulted rocks of the Onverwacht and Fig Tree Group by the authors.
Kranendonk (2011) state that the central parts of the BGB are the result of vertical mass movement: During the ascend of granitic melts around 3.2 Ga denser greenstone material sank into partially molten crust along extensional shear zones.
Various authors (e.g. Moyen et al., 2006; Moyen et al., 2007; Van Kranendonk et al., 2007a) describe the formation of Archean greenstone belts as a mixture of the two tectonic mechanisms. During the initial stages of the greenstone formation vertical tectonics are considered to have been dominant with a subsequent transition to plate tectonics with subduction and accretion processes around 3.2 - 3.1 Ga (left and centre sketch in Fig. 3.3).
The origin of TTG intrusive bodies at the southern and western border of the BGB is described as either the result of extensive vertical mass movement (Hamilton, 1998) or of primitive plate tectonic processes as described by Foley et al. (2003). Magmatic activity and melting processes within a subduction zone are considered to be another source for at least some of the TTGs, probably during a stage when subduction was halted, perhaps by the accretion of micro-cratons to the BGB (Wikipedia, 2014).

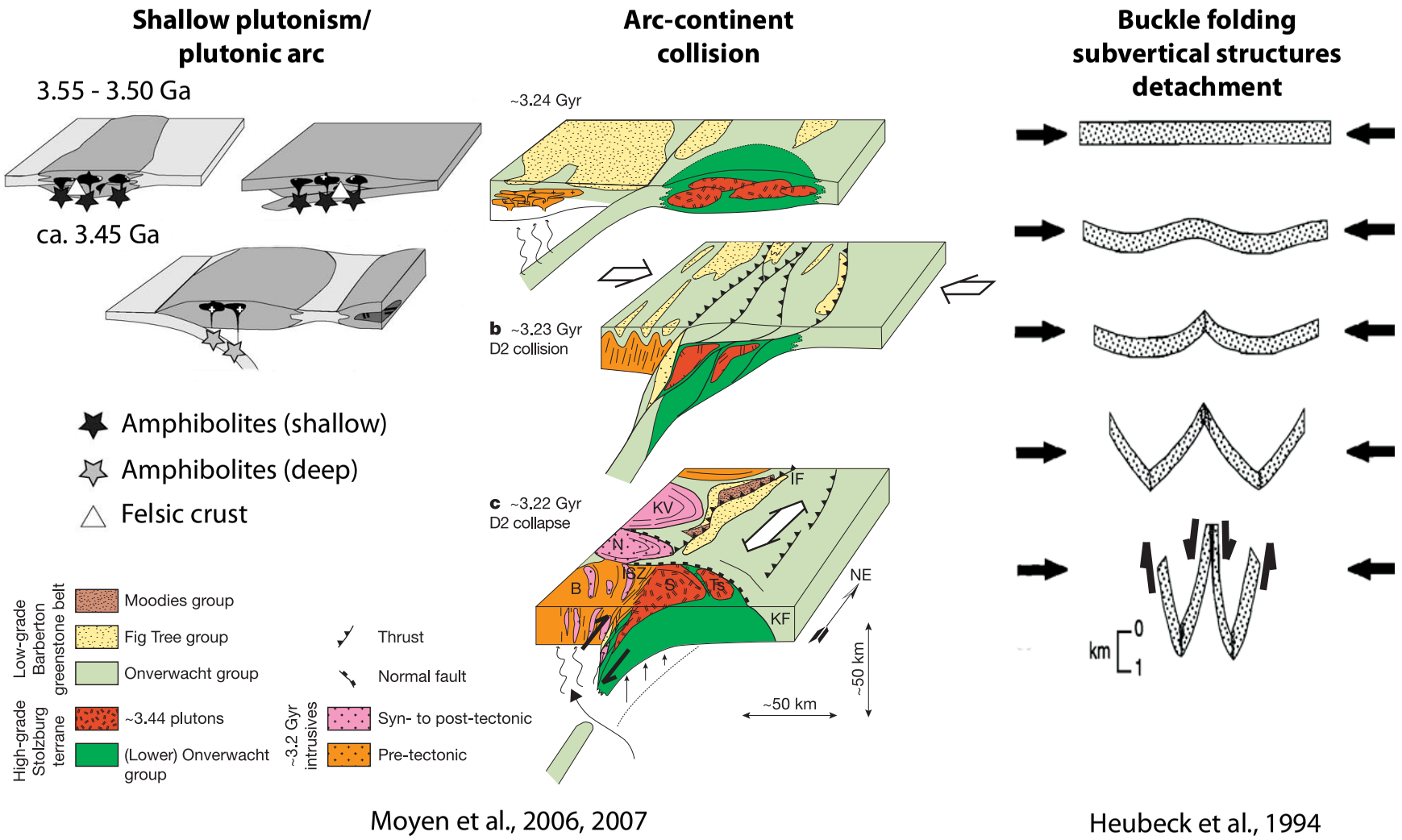


Figure 3.3: Evolution theories explaining the structures that we presently observe in the BGB: Vertical tectonics (left sketch), plate tectonics (centre sketch) and buckle folding (right sketch) schemes were developed over time based on geo-chronological and structural data. ISZ Inyoni Shear Zone, IF Inyoka fault, KF Komati fault, Ts Theespruit pluton, KV Kaap Valley pluton, N Nelshoogte pluton, B Badplaas pluton, S Stolzburg pluton.

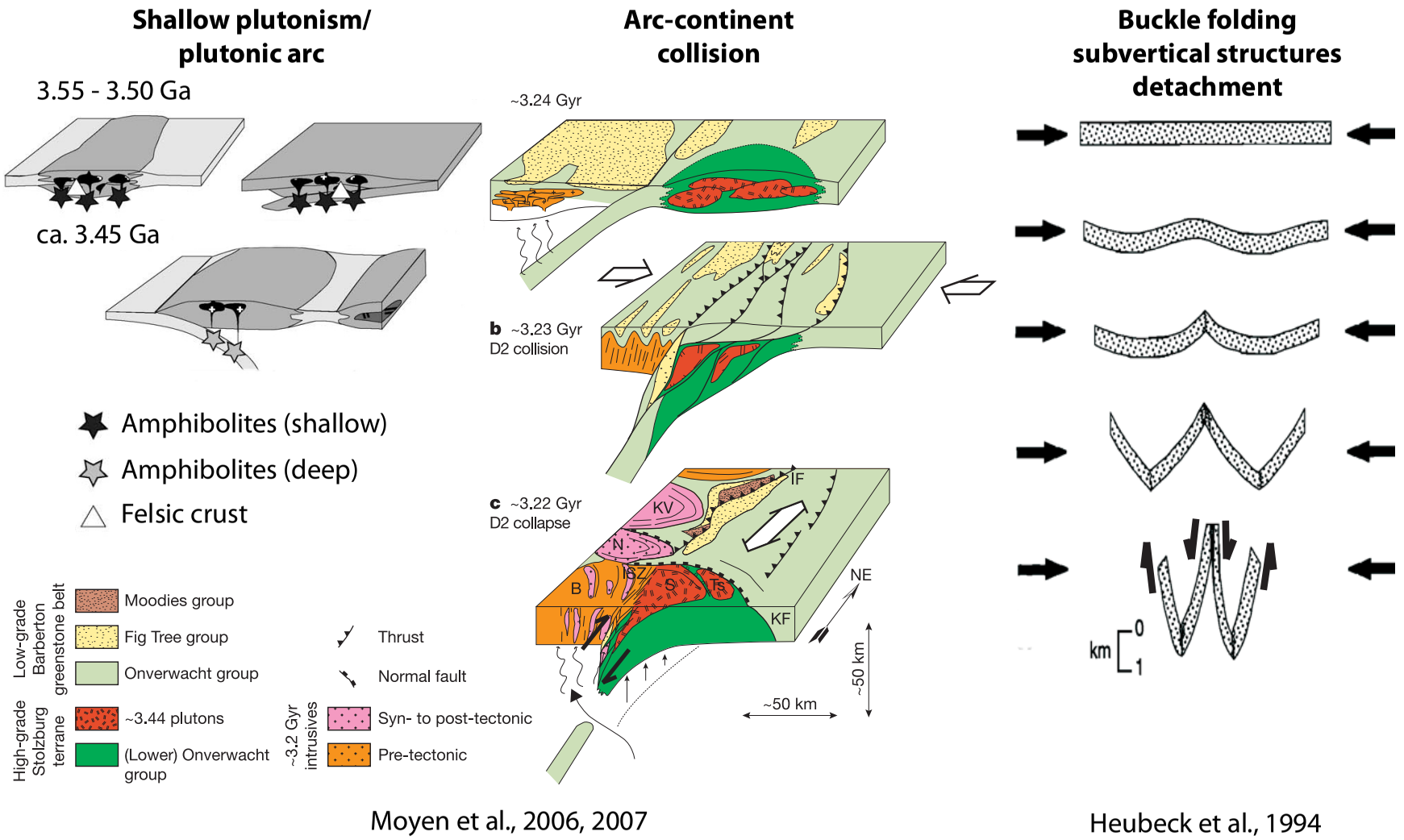


Figure 3.3: Evolution theories explaining the structures that we presently observe in the BGB: Vertical tectonics (left sketch), plate tectonics (centre sketch) and buckle folding (right sketch) schemes were developed over time based on geo-chronological and structural data. ISZ Inyoni Shear Zone, IF Inyoka fault, KF Komati fault, Ts Theespruit pluton, KV Kaap Valley pluton, N Nelshoogte pluton, B Badplaas pluton, S Stolzburg pluton.

The TTGs show a rough inverse relation between their volume and age: They increase in size with their decreasing age (de Ronde and de Wit, 1994). Due to a higher heat flow and hence, a higher melting rate at the MORs, the oceanic crust was much thicker in the Archean than today (25 km at about 3.5 Ga) and could not be subducted as a whole (Foley et al., 2003). During this time, melts necessary for the generation of TTG complexes could only be formed in small amounts by melting processes in the upper oceanic crust after the delamination of the lowermost ultramafic layer and subsequent exposure of the upper oceanic crust to the higher temperatures of the primitive mantle of the Early Earth (Foley et al., 2003). Later on (3.3 to 3.0 Ga), due to the gradual cooling of the Earth, the oceanic crust became thinner and could be subducted as a whole. The subduction of the in parts hydrothermally altered topmost layer allowed a larger generation of TTG melts. This is considered to be another prove for an onset of early plate tectonics around again 3.2 Ga.

3.2 Previous geophysical studies at the BGB

In contrast to the geological surveys, geophysical studies in the Barberton area were so far only carried out in the 1970's and 1980's (de Beer and Stettler, 1988a; de Beer et al., 1988b) with a focus on the central and northern part of the BGB and only minor interest in the southern part.
In de Beer and Stettler (1988a) the used geophysical methods comprise deep geoelectric soundings (Schlumberger sounding), gravity and magnetics measured at greenstone terrains located at the north-eastern margin of the Kaapvaal Craton. For the paper of de Beer et al. (1988b), the same gravity and geoelectric data was again analyzed this time with a focus on the BGB. The following statements and conclusions are taken from these two papers unless otherwise indicated. Locations of the geoelectric soundings and the gravity profile which are relevant for the presented study are shown in Fig. 3.1.
The geoelectric soundings carried out on outcrops of the Onverwacht Group identify an uppermost weathered zone with low resistivities ($\sim 10^1$ Ωm). Resistivities between 10^2 and 10^3 Ωm are characteristic for the deeper unweathered mafic to ultramafic rocks of the Onverwacht Group. This unweathered zone is followed by the granitoid rocks of the basement ($>10^3$ Ωm).
Geoelectric measurements at the TTG bodies and plutons indicate that the upper 6 to 10 km have resistivities in the order of 10^4 to 10^5 Ωm. These high values reflect the solid, mostly unfractured structure of TTG complexes. Based on the low temperatures of the topmost 10 km of cratonic rocks Van Zijl (1978) estimated a water resistivity of 1 Ωm and calculated a water content of about 0.5 % for granitoids. This further strengthens the assumption on their compact and solid structure. Down to the crust-mantle boundary (Moho at about 35 km) lower resistivities in the order of 10^3 Ωm can be found.
The gravity data shows positive anomalies over the BGB reflecting their higher density. One of the major results of the data was that the transition between the rocks of the BGB differs along the southern and northern margins. A sharp vertical to subvertical boundary exists between the BGB and the Piggs Peak Batholith. Only in narrow bands the BGB slightly extends locally underneath the batholith. In contrast, rocks of the BGB underlie over a large area (at least 8 km) the Nelspruit Batholith north of the Stentor Pluton. A maximum thickness of 600 m of the sheet-like structured Nelspruit Batholith cover was ascertained.
Regarding the inner structure of the BGB, gravity results along profile E - E' (Fig. 3.1) depict a thin (~1 km) sedimentary strata of Fig Tree and Moodies Group type rocks only in parts covering Onverwacht Group material. In this region, the Onverwacht Group has a thickness of up to 7 km and is thinning to-

wards the northwest and southeast. A similar BGB structure was observed at its northeastern end. In the central part of the BGB, generally thicker (around 3 km) sequences of Fig Tree and Moodies Group sedimentary rocks can be found which are mainly a result of major folding processes. Profile E - E' also shows that the Kaap Valley pluton partly extends underneath BGB rocks.
A combination of the resistivity and gravity results limitates the overall thickness of the BGB to less than 8 but usually more than 4 km. Similar results regarding the depth extent and geophysical properties were obtained for other greenstone belts located ~230 km to the northwest of the BGB (Murchison, Pietersburg and Giyani greenstone belts).

3.3 Inversion theory and general settings for the inversion studies

Apparent electrical resistivities (sec. A.1 eq. A.11) are an averaged representation of the true resistivity distribution of the subsurface. Inversion algorithms are necessary to obtain a model relating the apparent resistivity dependant on the period to a resistivity distribution with depth. The resistivity distribution $\rho(x, y, z)$ (for a general 3D case) for a subsurface model can be described using M parameters $\mathbf{m} = (m_1, m_2, \dots m_M)^T$ and the model response (i.e. the measured data) is represented by $\mathbf{d} = (d_1, d_2, \dots d_N)^T$. The forward problem is then given by $\mathbf{d} = F(\mathbf{m})$, where F is the, in general, non-linear forward operator. The equation always has an unique solution $\mathbf{d}$. Linearisation using a Taylor series expansion leads to

$$F(\mathbf{m} + \Delta\mathbf{m}) \approx F(\mathbf{m}) + \frac{\delta F(\mathbf{m})}{\delta \mathbf{m}} \Delta\mathbf{m} = F(\mathbf{m}) + \mathbf{J}\Delta\mathbf{m}. \tag{3.1}$$

The sensitivity or Jacobian matrix $\mathbf{J}$ describes the sensitivity of the model response $\mathbf{d}$ towards small changes $\Delta\mathbf{m}$ of the model parameters $\mathbf{m}$.
For the inverse problem, data vector $\mathbf{d}$ is the known and the model vector $\mathbf{m}$ is the unknown parameter. Since this is an ill-posed, under-determined problem (no. of data parameters $N \ll$ no. of model parameters M) the solution (if it exists) is not unique and can reproduce the data only to a certain degree or within error bounds (data error vector $\mathbf{e}$)

$$\mathbf{d} = F(\mathbf{m}) + \mathbf{e}. \tag{3.2}$$

For the solution of the inverse problem a penalty function Φ is determined

$$\Phi(\mathbf{m}, \mathbf{d}) = (\mathbf{d} - F(\mathbf{m}))^T \mathbf{C}^{-1} (\mathbf{d} - F(\mathbf{m})) \tag{3.3}$$

with $\mathbf{C}$ as the covariance matrix of the errors of the data.
Due to the mentioned ill-posed and under-determined nature of the inverse problem, model regularization terms $\Omega(\mathbf{m})$ are commonly added to the penalty function (3.3). These regularization terms are formulated in a way to stabilize the problem and to limit the number of possible solutions. In most cases, this means that the inversion searches for the smoothest model to explain the measured data. The extended penalty function

$$\Phi(\mathbf{m}, \mathbf{d}) = (\mathbf{d} - F(\mathbf{m}))^T \mathbf{C}^{-1} (\mathbf{d} - F(\mathbf{m})) + \lambda\Omega(\mathbf{m}) \tag{3.4}$$

includes the regularization parameter λ which controls the influence of the regularization term during the inversion. In search for a solution, λ determines the trade-off between data fit and the simplest structured model.

As a quality criterion for inversion results and a rough rating between different results the overall root mean square (RMS) value is commonly used.

$$\mathrm{RMS} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\frac{\left(d_n^{obs} - d_n^{mod}\right)^2}{e_n^2}}. \tag{3.5}$$

Here, n is the index for all data points, d_n^{obs} and d_n^{mod} are the measured and modelled data, respectively. The RMS value in this thesis is defined as a normalized value, hence, it has no unit of measurement.
Since the data fit for all data components (e.g. $\rho_{i,j}, \phi_{i,j}$ or $Z_{i,j}$, $i, j = x, y$) for all sites and periods is included in the RMS calculation, this value can only be considered as a first rough estimate of the quality of an inversion and a closer comparison of the measured and modelled data is absolutely necessary.
An extended description of inversion theory in magnetotellurics can, e.g., be found in Chave and Jones (2012).
For the BGB MT data, both 2D and 3D inversions are conducted for each profile separately and for the whole data set, respectively. Knowledge of 2D and 3D inversion techniques, their strengths and weaknesses and their general properties are essential for verifying the significance of electrical resistivity structures in the inversion results. In the 2D and 3D inversion algorithms applied in this thesis, finite differences are used for the discretization of the models. The model discretization should be adapted to the resolution of the MT data: Finer discretization in areas where data coverage is high and of the near subsurface with an increase of the cell size with depth and towards model boundaries.
Due to the much smaller model size in 2D inversions, fine discretization of the model space and, hence, the recovery of a detailed resistivity model given a sufficient data coverage is possible. If the measurement area is characterized by complex geological structures with resistivity variations in all three spatial directions which is the case for the BGB data, 2D inversions might have problems recovering the true subsurface conductivity structure. In such a case, 2D inversions can still be used to obtain a first approximation of the resistivity distribution. 3D inversion studies provide then an additional image of the resistivity distribution not only along but also in between profiles.
3D inversion has become almost a standard tool for MT data interpretation due to the rapid advances in computer technology and 3D inversion algorithms (e.g. Sasaki, 2001; Siripunvaraporn et al., 2004; Egbert and Kelbert, 2012). Hence, 3D MT inversion problems can be solved within a reasonable amount of time (~ one week). Nevertheless, inversion of large data sets (e.g. BGB MT data set) still faces limitations in the refinement of the model grid in order to keep the computational time for the inversion within reasonable bounds (in the order of a few weeks). Successful 3D inversion is further complicated by EM noise in the data, defining regularization techniques to stabilize the inversion process and to narrow down the possible solutions. Regarding the physical and numerical limitations and challenges, 2D and 3D inversions should, hence, be used complementary to obtain a comprehensive view of the investigated area.

3.3.1 2D inversion

To determine if 2D inversion can be applied to the BGB MT data set, the dimensionality of the data needs to be verified: Either the resistivity of the subsurface changes only with depth (1D case), with depth and along one horizontal direction (2D case) or in all three spatial directions (3D case). The dimensionality of the data can, e.g. be identified by the determination of the regional geo-electric strike direction which

is the direction of the main regional current flow in the subsurface. In this thesis, I used the strike analysis after Becken and Burkhardt (2004). The algorithm is based on the assumption that in a general 2D case, the linearly polarized source field generates elliptically polarized secondary EM fields. The geo-electric strike is then the direction where these ellipticities of the EM fields are minimized (i.e. where the EM fields are linearized). For the impedance tensor, linear polarization is expressed by a common phase for the elements of the first ($Z_{x,x}$, $Z_{y,x}$) and second tensor column ($Z_{x,y}$, $Z_{y,y}$; eq. A.9). This method for strike direction calculations has the advantage that it is unaffected by small-scale inhomogeneities which cause galvanic distortion of the data and can lead to erroneous strike directions.

The geo-electrical strike direction is typically displayed in a rose diagram. Due to its definition as the direction where the EM fields are linearized, the strike direction has a 90^o ambiguity, e.g. at 0 and 90^o. To identify the correct direction, the regional geology can provide useful informations.

Figure 3.4 summarizes the results of the strike direction analysis for all six profiles (P1 to P6) for the whole period range (left column), for short periods (0.001 to 0.1 s, middle column) and for long periods (0.1 to 1 000 s, right column). All profiles show strong variations in the strike direction with in general only upto four sites exhibiting the same strike. Therefore, determining one specific or preferred strike direction for each profile used for the rotation of the data prior to the 2D inversion is difficult. Hence, 2D inversions might have difficulties resolving the conductive structures properly.

For profile P1, a strike direction of -34^o off north (clockwise) was chosen since it corresponds to the average strike direction of the sites and also coincides with the main geological strike direction (NE-SW aligned geological structures, de Ronde and de Wit, 1994). This orientation of the regional geo-electrical strike perpendicular to the profile is conform with the 2D assumption where the resistivity ideally only changes in one horizontal direction along the profile. A rotation of -34^o of the data and subsequent 2D inversions were also conducted for profiles P2 to P6. In all tests for these profiles, the inversion terminated after only few iterations (< 5) or could not be started at all. For profiles P2 to P6, when using a strike direction of 0^o (i.e. the north direction which corresponds to unrotated data, see also Fig. 3.4), 2D inversions of the data were possible. Using unrotated data instead of rotating the data to a direction perpendicular to the profile violates the 2D assumption since then, the resistivity changes not only occur in one horizontal direction. Hence, these 2D inversion results can only be seen as a first approximation of the resistivity distribution of the subsurface. However, the results can be verified by using geological informations and by a comparison with the 3D inversion results (see sec. 3.6). A comparison of 2D inversion results of profile P1 of unrotated data and data rotated by -34^o shows only minor changes in the models. This underlines that the 2D inversion results with unrotated data for profiles P2 to P6 still contain significant information of subsurface structures.

The 2D inversions were carried out with the software WinGLink which uses the RLM2DI algorithm by Rodi and Mackie (2001). For the solution of the inverse problem the Gauss-Newton algorithm is applied which is based on the non-linear conjugate gradient (NLCG) method.

Within WinGLink, there are a number of different parameters and settings which influence the outcome of the inversion. The most important settings are the selection of the starting model and the MT data used in the inversion and the model grid. Since no relevant deep reaching geological or structural information is available for the BGB a homogeneous half-space is used as starting model for all inversions. Different resistivity values ($\rho = 10^2, 10^3, 10^4, 5 \cdot 10^4 \Omega m$) for the half-space were tested leading only to minor structural changes in the inversion results or the overall resistivity. For the preferred models (sec. 3.4), a resistivity of 10^3 Ωm is chosen for the starting model due to the in general high resistivity of the BGB.

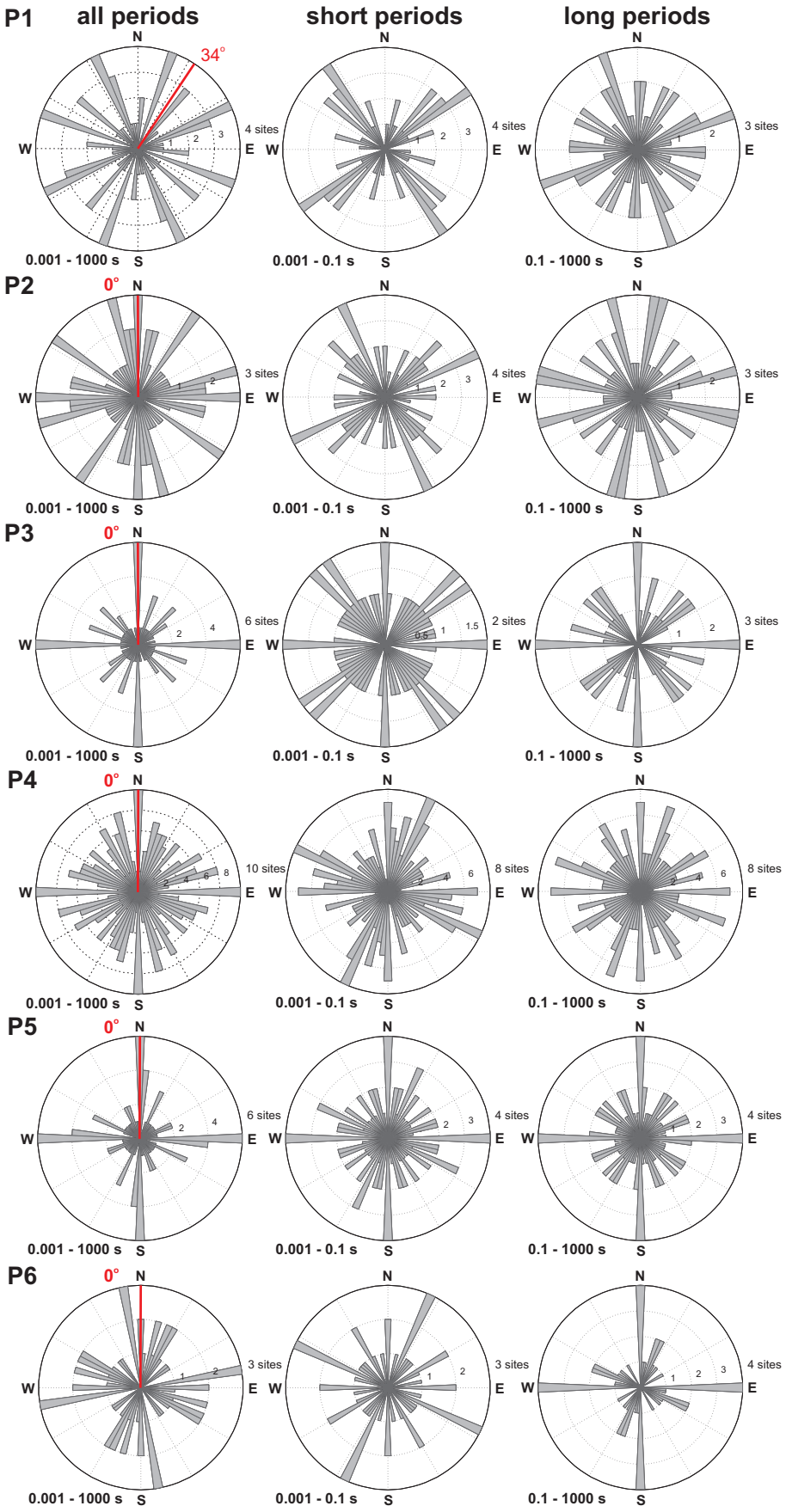


Figure 3.4: Strike analysis after Becken and Burkhardt (2004) for all profiles: Rose diagrams for all periods (left), short periods (middle) and long periods (right). Marked in red is the selected strike direction for 2D inversion.

For each profile, the MT sites used in the inversion were chosen in a way to obtain an (if possible) equally spaced site coverage (green markers in Fig. 3.5 identify chosen sites). This facilitates the generation of a regularly spaced model grid which is important for a successful inversion without numerical artifacts due to a distorted grid. MT data which exhibited a strong 3D character was also excluded from the 2D inversions. Due to the poor data quality of the magnetic transfer functions (eq. A.13 and chap. 1) no induction vectors were included in the inversions. TE and TM mode data was inverted simultaneously and also separate from each other. The period of the inverted data ranges (if available) usually between 10^{-2} and 10^{2} s.

A parameter which influences the model fit and the general structure of the inversion results is the regularization parameter τ which controls the model smoothness. Low τ values correspond to very detailed models with low RMS values whereas larger τ values result in smooth models with a higher RMS. To choose a proper τ, so-called L-curves can be used where the model roughness (depending on τ) is plotted against the RMS value of the inversion. The point of maximum curvature is considered to be the optimum value for τ and the (ideal) trade-off between model smoothness and data fit. For profile P1, $\tau = 10$ was selected based on the L-curve criterion. For the other profiles, conducting inversions with $\tau \geq 10$ was impossible since these inversions usually terminated after only a few iterations. Hence, the highest possible τ values (usually around 7) which resulted in physically meaningful inversion models were chosen.

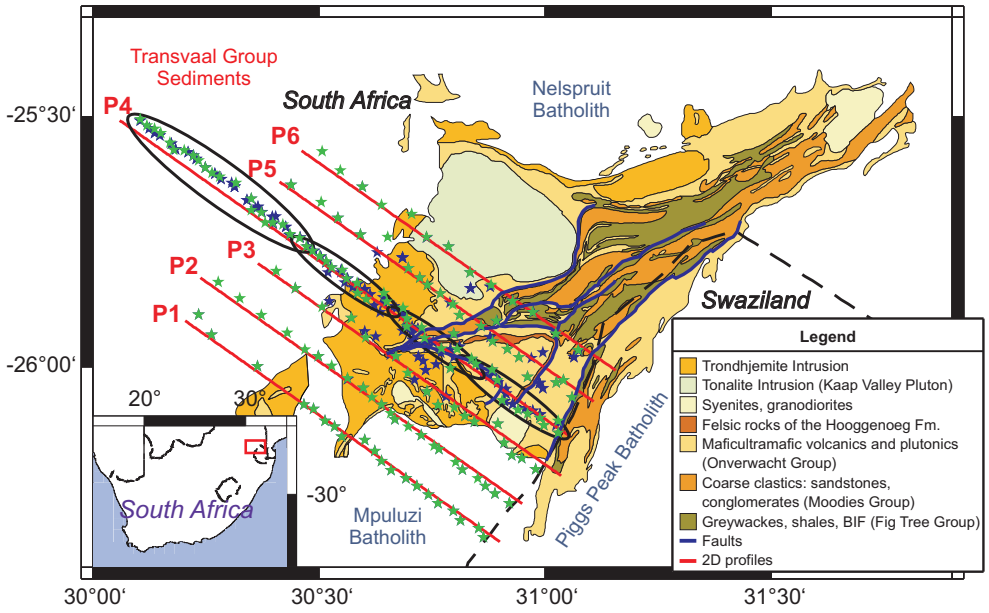


Figure 3.5: Location of MT sites used for 2D inversions (green markers, sites not used have blue markers). Ellipses mark the sections of profile P4 for which additional 2D inversions of data subsets were carried out. Inset map shows the location of the measurement area within South Africa (red rectangle, map after Schoene and Bowring, 2007 and de Ronde and de Wit, 1994).

3.3.2 3D inversion

3D inversions are on the one hand adapted to include complex 3D structures in a model but on the other hand there are limits to the possible refinement of the model due to computational limitations. The latter aspect might hinder the resolution of small scale structures especially for large investigation areas and MT data sets requiring large 3D models. 3D inversion is still computationally and numerically challeng-

ing and is in parts still poorly understood.

For the 3D inversions the modular inversion code ModEM by Egbert and Kelbert (2012) and Meqbel (2009) was used. The modular implementation of the ModEM package allows for an easier extension if new algorithms are incorporated or added. The programme package is divided in three layers: The numerical discretization, the generic inversion and the interface between these two layers. The first layer includes the forward problem, its discretization and solution. The third layer consists of inversion algorithm codes and sensitivity calculations. The interface layer defines the solution space, data functionals and the used solver.

The used parallelization of the 3D MT package of ModEM distributes the calculation of the forward models and the sensitivity to different processors for the different frequencies and the two source field polarizations (Meqbel, 2009). The ideal number of processors (eq. 3.6) used for an inversion is, hence, defined by the number of frequencies N_f. One additional computer node regulates the parallelization process.

$$\text{No. of nodes} = 2 \cdot N_f + 1 \tag{3.6}$$

The regularization term used for the 3D inversion is made up in the following way

$$\Omega(\mathbf{m}) = (\mathbf{m} - \mathbf{m_0})^T \, \mathbf{C_s}^{-1} \, (\mathbf{m} - \mathbf{m_0}) \tag{3.7}$$

The smoothness covariance matrix $\mathbf{C_s}$ contains the parameters α_x, α_y and α_z which regulate the smoothness in x-, y- and z-direction, respectively. The parameters can be arbitrarily chosen from the closed interval [0,1]. Furthermore, the regularization incorporates the prior model $\mathbf{m_0}$. Equation (3.4) in the 3D case favours smooth models which are as close to the prior model as possible.

After linearization (eq. 3.1), the solution of the penalty function (eq. 3.4) is obtained by applying a NLCG algorithm.

Due to the complexity of the geology in the Barberton area and the severe noise in the time series, the BGB MT data set is not easy to work with and to invert using 3D algorithms. Furthermore, the presented array data with 193 MT sites is a fairly large data set for 3D inversion when compared to other studies (e.g. 220 sites in Heise et al., 2010). The chosen subset of 158 MT sites (green markers in Fig. 3.6 identify chosen sites) spans a rectangle with an edge length of 90 by 50 km. The selection of sites used in 3D was adapted to the chosen model grid to avoid more than one site per cell. Sites with better data quality and/or more data points were favoured. Site spacing along the profiles varies between 2 and 13 km while the distance between profiles is about 10 km. All components of the impedance tensor $\mathbf{Z}$ were inverted simultaneously. Magnetic transfer functions were not used in the inversion due to their poor data quality. The period range of the inverted data is between 10^{-2} and 10^2 s (if available). Periods where less than three sites had usable data were skipped in order to facilitate the inversion process. A data error floor of 5% for all data components proved to be best choice for the 3D inversions.

In some 3D inversion studies, the data was rotated according to the dominant strike direction of the investigation area which proved to be essential in the presence of dominant 2D structures (e.g. Tietze and Ritter, 2013). Since the strike direction analysis of the BGB MT data set did not show a specific preference for one direction, the data was left unrotated for the 3D inversion studies. The coordinate system for the 3D models is chosen according to geographic coordinates (x axis pointing north, y axis pointing east, z axis is positive downwards). No a priori information was included in the models and homogeneous half-space models were used as a starting (and prior) model. As for the 2D inversions, a half-space with a resistivity of 10^3 Ωm was chosen. The top layers of the preferred 3D model were set to be small starting

from a thickness of 0.02 km with an increasing factor of 1.15 with depth to accomodate the effects from small-scale near surface structures. Other cell sizes were tested as well: In x- and y-dirtection 1 000, 1 200, 2 000 and 3 000 m and in z-direction 20 and 100 m for the cell size in the centre of the model. The factors for increasing the cell size with depth and towards the model boundaries were set to either 1.15 or 1.2 for all directions for the tests. Five to ten padding cells were used at the model boundaries.
Validation of 3D inversion results is difficult. As suggested, e.g. by Heise et al. (2010) and Patro and Egbert (2008, 2011), the regularization of the model was changed for this purpose. Different grids and smoothing values (covariance parameters were set to $\alpha_{x,y,z} = 0.2, 0.3$ or 0.4) were used. In general, the horizontal smoothing was set to higher values (for $\alpha_{x,y}$) than vertically (α_z) to accomodate for the difference in site to profile spacing (2 to 10 km). The resulting models vary especially in smaller details but the overall structures remain relatively stable. A comparison of the 2D and 3D results as further validation for the inversion models along the profiles is conducted in sec. 3.6.

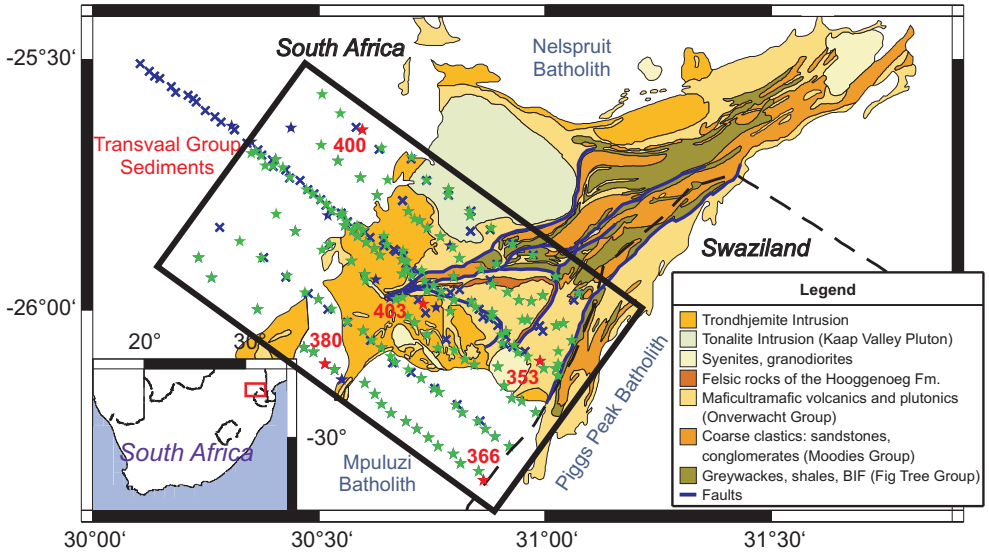


Figure 3.6: Location of MT sites used for 3D inversions (green markers, sites not used have blue markers, red markers for sites where the data fit is shown in Fig. 3.25). Inset map shows the location of the measurement area within South Africa (red rectangle, map after Schoene and Bowring, 2007 and de Ronde and de Wit, 1994).

3.4 2D MT Inversion results

The 2D MT inversions for the six profile lines are carried out using the RLM2DI algorithm by Rodi and Mackie (2001, implemented in the software WinGLink). This section presents the preferred 2D inversion results for the profiles. Inversion settings for each model are stated in the caption for the respective inversion result. The number of cells used to discretize the models for the profile are given in the form rows x columns, e.g. 152 x 141 corresponds to 152 rows and 141 columns.

Profile P1

Profile P1 is the southernmost profile of the Barberton MT data set and crosses mainly the Transvaal Group sediments in the northwest and the Mpuluzi Batholith in the southeast (cf. Fig. 3.5). The profile intersects the southernmost extension of the BGB.
The preferred model for profile P1 is shown in Fig. 3.8 (top). For the inversion, TM and TE mode

data of 20 sites are used which were rotated clockwise by 34°. In Fig. 3.7, the data fit for this model is exemplarily shown for three sites. The positions of these sites are marked in Fig. 3.8 (top) as red triangles. The shape and value of apparent resistivity and phase curves is especially well fitted for TM mode data. TE mode data shows that it is in parts affected by static shift. This is expressed as a fixed offset in the apparent resistivity curves over the entire period range (e.g. site 373).

For a better visibility of the conductive structures especially within the Mpuluzi Batholith, the model is displayed with a constricted resistivity scale between 10^3 and $5 \cdot 10^4\,\Omega$ m (Fig. 3.8, bottom). With this altered colour scale an area with an increase of conductivity can be identified between profile kilometre 50 and 75 km (marked as area A) which appears to reach in parts from the surface to at least 10 km depth. To further investigate this conductive structure and to determine whether it could be linked to a southward continuation of known faults within the BGB, 2D inversions with tear zones are conducted. Tear zones can be included in the inversion model to disrupt regularization between neighbouring cells. However, if such a sharp contrast is not required by the data the inversion will ignore this tear zone.

Since in general inversions favour smooth models with contiguous structures ensured by regularization terms in the inversion algorithm, there ia a minimum threshold in width of conductive features within resistors so that they are included by the inversion. The same regularization also causes that sharp boundaries appear as smooth and larger transition zones. Faults which are important tectonic features for understanding the formation of the BGB are e.g. such structures challenging to be resolved in the inversion. To represent faults within the model I inserted vertical tear zones reaching from the surface to the bottom of the model. For the tear zone inversion studies, I changed the number of tear zones, their position along the profile and their width (i.e. number of columns comprising the tear zone). The starting model in all studies was a homogeneous half-space of $10^3\,\Omega$ m. Tear zone inversion studies for the other profiles were carried out in the same way.

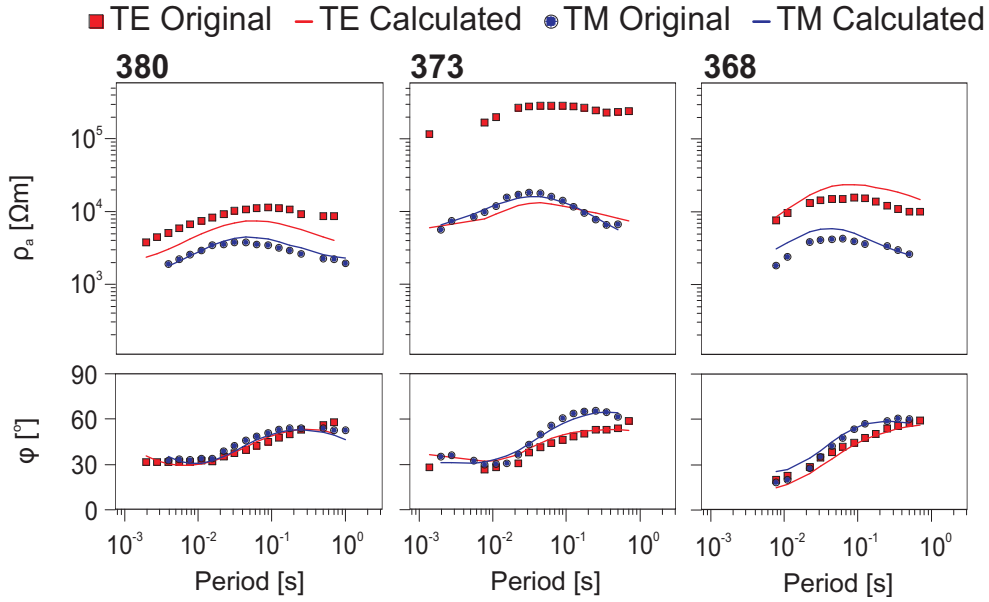


Figure 3.7: 2D data fit profile P1 (model cf. Fig. 3.8): Data of three sites is shown exemplarily. Locations of the sites is marked in Fig. 3.8 by red triangles. The data fit for TM mode is in general better. TE mode data is affected by static shift.

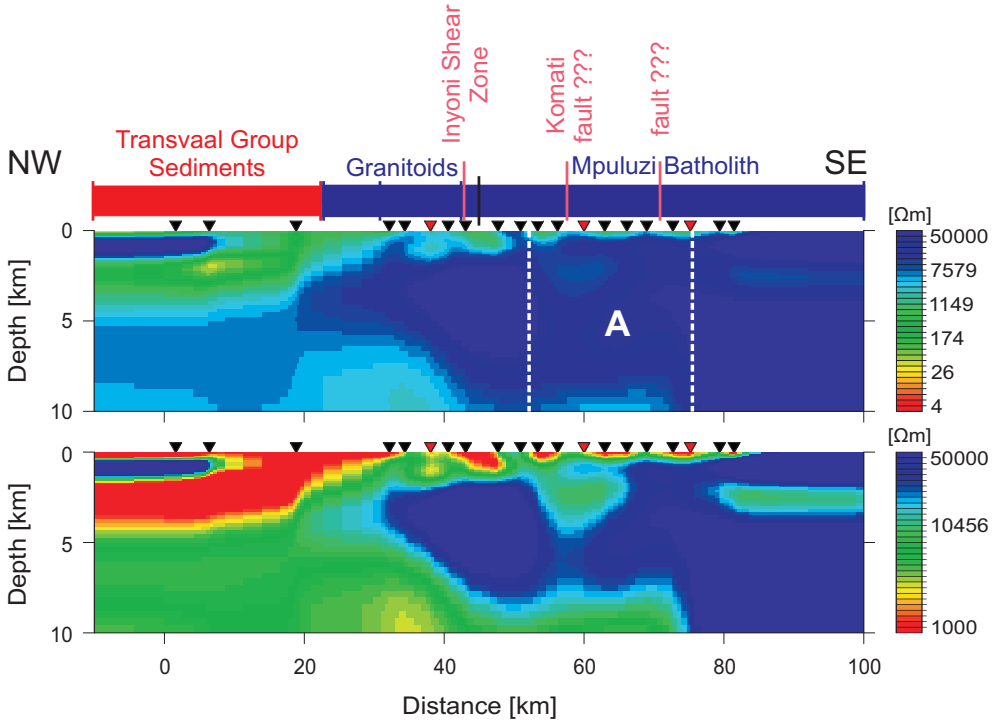


Figure 3.8: 2D inversion result along profile P1: To enhance the visibility of the conductive structures within the Mpuluzi Batholith the inversion result (top) is plotted with a different colour scale for the electrical resistivity (bottom). Site locations are marked by triangles; sites for which the data fit is shown in Fig. 3.7 are marked in red. Settings: 10^3 Ωm half-space; 152 x 141 cells; error floor: **TE**: ρ 100 %, ϕ 1.45 ^{o}, **TM**: ρ 30 %, ϕ 1.45 ^{o}; τ = 10; 300 iterations; RMS 1.46.

The tear zone model with the best fit of the modelled to the measured data and also the lowest RMS value is shown in Fig. 3.9. This model includes two tear zones which are both two columns wide and are close to a presumed southward continuation of faults. The figure shows a comparison of the RMS for each site of the model with and without the tear zones as well. Especially for the sites in the vicinity of the tear zones the RMS is lower for the model with tear zones. To further validate the tear zone structures, I set the resistivity in the tear zones to the value of the cells surrounding the tear zones ($\sim 3 \cdot 10^4\,\Omega$ m), keeping the rest of the model structures unchanged and started the inversion again. This test resulted in an inversion model almost completely identical to the one shown in Fig. 3.9. Considering all the described tests, the tear zone structures appear to be plausible features for this profile.

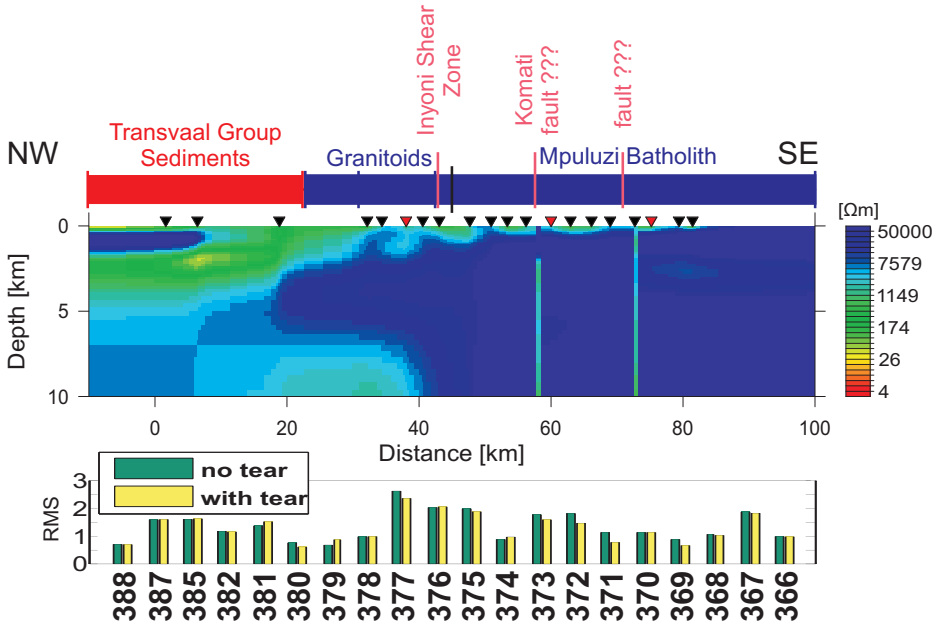


Figure 3.9: 2D inversion result along profile P1: Two tear zones (no smoothing at the tear zone boundaries) were included in the model at the location where enhanced conductivity was observed in the 2D inversion result without tear zones (cf. Fig. 3.8 bottom). The graph at the bottom shows a comparison of the RMS values of the models with and without tear zones. Settings: 10^3 Ωm half-space; 152 x 141 cells; error floor: **TE**: ρ 100 %, ϕ 1.45 ^{o}, **TM**: ρ 30 %, ϕ 1.45 ^{o}; τ = 10; 300 iterations; RMS 1.38.

Profile P2

Profile P2 crosses the same geological units as profile P1 and is hence controlled by the Transvaal Group sediments in the northwest and the Mpuluzi Batholith in the southeast. The joint inversion of TE and TM mode data (unrotated) of the 22 sites along this profile results in the model shown in Fig. 3.10. As expected, the conductivity distribution is similar to profile P1.

The data fit of this profile is shown exemplarily for three sites in Fig. 3.11. Especially for sites close to pronounced conductivity contrasts (e.g. site 310) and at boundaries between geological units (e.g. site 306), the data is not always fitted by the inversion over the entire period range. Since this profile is closer to the more complex geological structures of the BGB (than profile P1), the apparent resistivity and phase curves reflect this higher complexity as well and 2D inversions are not able to fit them entirely causing higher RMS values for the sites when compared to profile P1 (cf. RMS plots in Fig. 3.9 and 3.12).

An area of slightly increased conductivities can be observed at the boundary between the granitoids and the Mpuluzi Batholith (area A, marked in white in Fig. 3.10). Again, tear zone inversion tests were carried out for this profile to investigate the possibility of a southward continuation of the known BGB faults. In the first tests, I inserted three tear zones located at the estimated positions of the faults (marked in the geological section in Fig. 3.10, top). These tear zones were ignored during the inversion process and showed hardly any conductivity contrasts to the neighbouring cells in the final inversion result. In subsequent tests, only one tear zone was inserted and its position was changed within the boundaries of area A. The tear zone inversion shown in Fig. 3.12, where the tear zone is one column wide, results in the best fit of the data and the lowest overall RMS. Due to its position, the tear zone probably does not reflect a remnant of a southward continuation of a fault, but is possibly caused by a conductive mineralization along the contact between the Mpuluzi Batholith and the granitoid complex.

The comparison of the RMS values for each site for the model without and with tear zone (Fig. 3.12, bottom) shows that the tear zone model fits the data better in general. However, the northernmost sites which cover the Transvaal Group sediments have higher RMS values in the tear zone model and their data is (especially for longer periods) not well fitted. This suggests that the higher conductive rocks of the Transvaal Group sediments reach deeper (cf. Fig. 3.10) than suggested in the tear zone model.

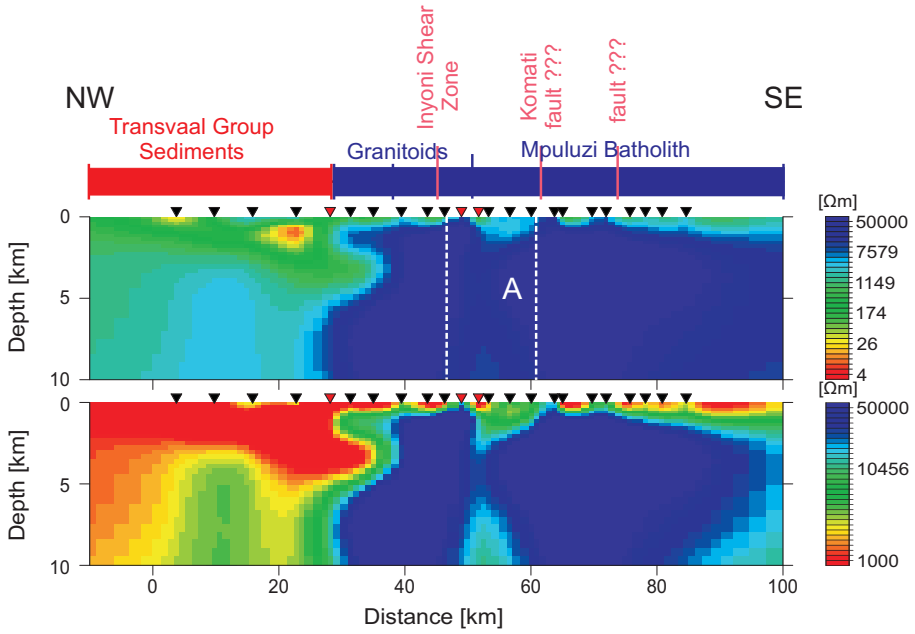


Figure 3.10: 2D inversion result along profile P2: Model shows a faint imprint of increased conductivity (area A, top). To enhance its visibility the inversion result is plotted with a different colour scale for the electrical resistivity (bottom). Site locations are marked by triangles; sites for which the data fit is shown in Fig. 3.11 are marked in red. Settings: 10^3 Ωm half-space; 152 x 143 cells; error floor: **TE**: ρ 100 %, ϕ 1.45 ^{o}, **TM**: ρ 30 %, ϕ 1.45 ^{o}; τ = 7; 300 iterations; RMS 2.28.

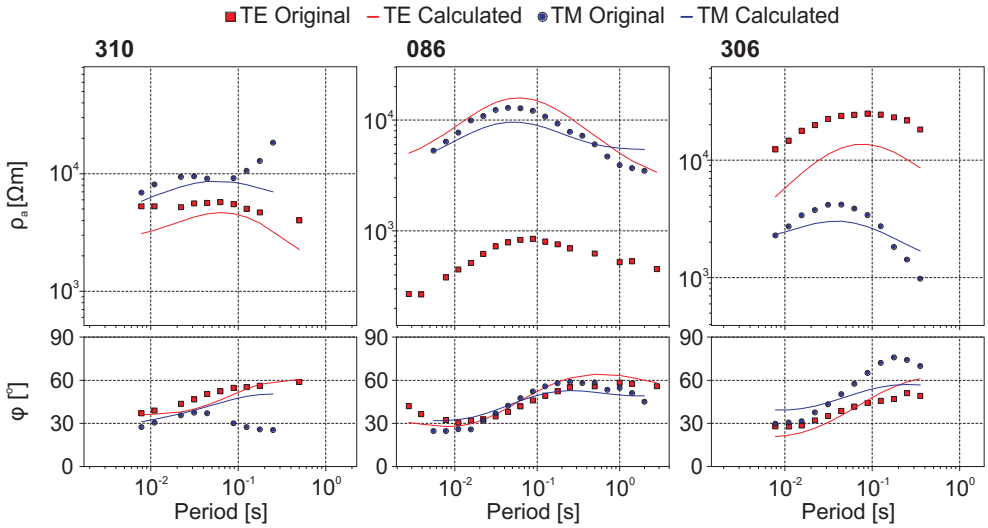


Figure 3.11: 2D data fit profile P2 (model cf. Fig. 3.10): Data of three sites is shown exemplarily. Locations of the sites is marked in Fig. 3.10 by red triangles. The data fit for TM mode is in general better. TE mode data is affected by static shift. Sites close to boundaries between geological units (e.g. site 310) have a bad data fit (see text for further discussion).

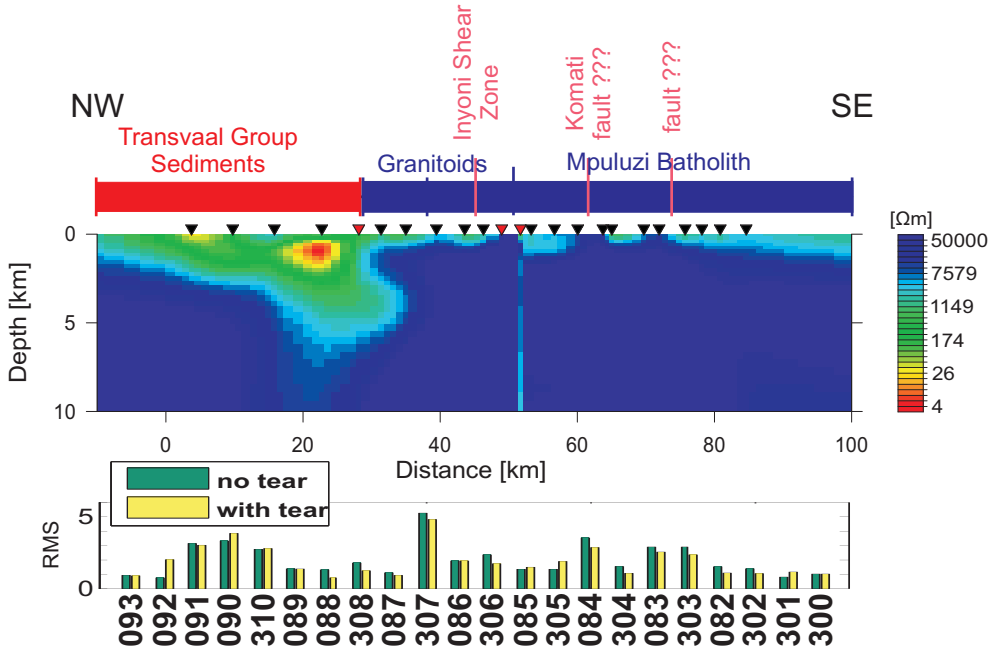


Figure 3.12: 2D inversion result along profile P2: One tear zones (no smoothing at the tear zone boundaries) was included in the model at the location where enhanced conductivities were observed in the 2D inversion result without tear zone (cf. Fig. 3.10). The graph at the bottom shows a comparison of the RMS values of the models with and without tear zones. Settings: 10^3 Ωm half-space; 152 x 143 cells; error floor: **TE**: ρ 100 %, ϕ 1.45 ^{o}, **TM**: ρ 30 %, ϕ 1.45 ^{o}; τ = 7; 300 iterations; RMS 2.11.

Profile P3

For profile P3 it was unfortunately not possible to obtain a plausible model using 2D inversion. Various tests changing the inversion (e.g. smoothing parameters, error floor) and model settings (e.g. starting model, model discretization) were carried out. These tests led to either inversion runs that did not converge, to inversion results where the modelled data did not fit the measured data (RMS >16) or to models with geologically impossible structures and artefacts, as e.g. very high conductivities for model cells in the vicinity of the sites (Fig. 3.13). These difficulties are maybe due to the location of the profile at the direct contact between the Mpuluzi batholith in the south and the rocks of the BGB in the north. Therefore, major parts of the dominant structures are not perpendicular but parallel to the profile. The profile marks also the boundary between the geologically simpler structures in the south and the complex structures of the BGB north of the profile. This causes violations of assumptions for 2D inversions (cf. sec. 3.3.1). Although the 2D inversion model for profile P3 (Fig. 3.13) shows artefacts in the vicinity of the sites and an overall (geologically implausible) vertical elongation of conductive structures, major features could be resolved which correlate to the Transvaal Group sediments, the granitoid complexes and the Dalmein Pluton. Area A and especially area B (i.e. Komati fault) might be traces of faults. Due to the problems during the inversion for this profile described above, further investigation by tear zone inversions was not possible.

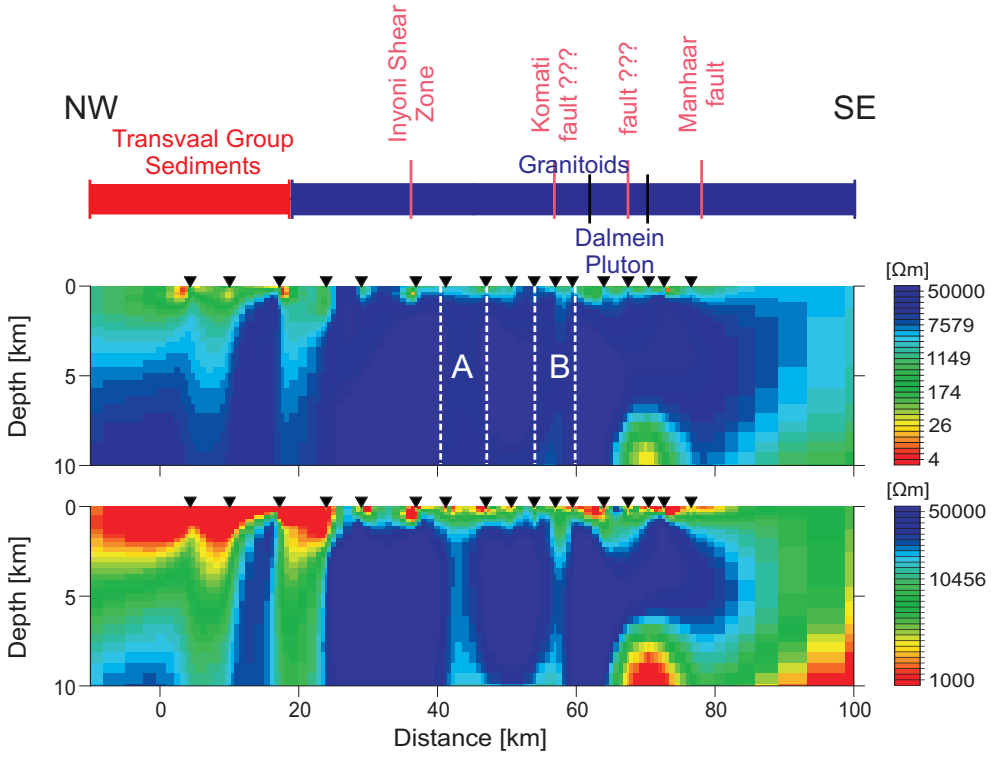


Figure 3.13: 2D inversion result along profile P3: Model shows faint imprints of increased conductivity (areas A and B, top). To enhance their visibility the inversion result is plotted with a different colour scale for the electrical resistivity (bottom). Settings: 10^3 Ωm half-space; 127 x 158 cells; error floor: **TE**: ρ 100 %, ϕ 1.45 ^{o}, **TM**: ρ 30 %, ϕ 1.45 ^{o}; τ = 7; 300 iterations; RMS 16.78. Range for the RMS for the sites: 2.05 ... 54.03 with the majority around 10.

Profile P4

Profile P4 crosses a central part of the BGB and all major faults (e.g. Inyoka and Manhaar fault) and reaches further north into the Transvaal Group sediments than the other profiles. For the inversion result shown in Fig. 3.14 TE and TM mode data (unrotated) of 50 sites was used. Overall, the modelled data fits the measured data well as can be seen for the three sites displayed in Fig. 3.15. As for the previous profiles, the TE mode data is in parts affected by static shift (e.g. sites 030 and 352).

For profile P4, I tried to insert tear zones as well within the section of the profile corresponding to the rocks of the Onverwacht Group. Due to the complexity of the data and hence, the model, tear zone inversions failed to result in a geologically comprehensible model. Especially in the area where the Moodies, Inyoka and Komati fault join, it was impossible to separate the individual faults by using tear zones within the inversion. Tear zone inversions were conducted for profiles P5 and P6 but failed as well due to the same reasons.

To further investigate the extent (mainly considering their depth) of the faults and their structure (e.g. their dipping angle), I subdivided profile P4 into partly overlapping subprofiles: P4N, P4W, P4C and P4E (Fig. 3.14, bottom). This enables me to use a finer discretization and to include more of the in total 83 sites measured along this profile in the inversions.

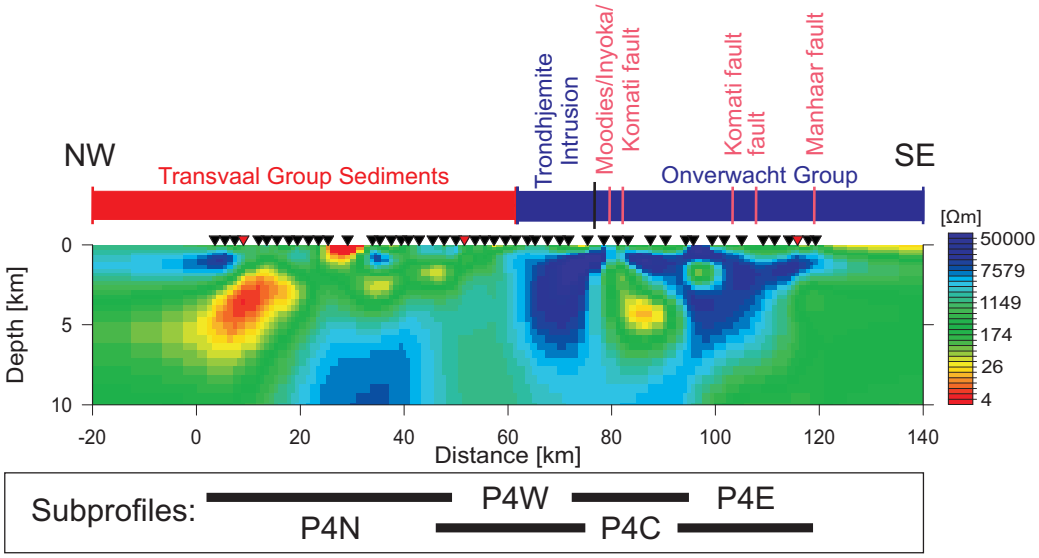


Figure 3.14: 2D inversion result along profile P4: Model shows clearly the transition from the conductive Transvaal Group sediments to the resistive rocks of the BGB. Within the BGB conductive areas can be seen which relate to known faults. Bottom sketch shows the way profile P4 was divided into the subprofiles P4N, P4W, P4C and P4E. Site locations are marked by triangles; sites for which the data fit is shown in Fig. 3.15 are marked in red. Settings: 10^3 Ωm half-space; 127 x 144 cells; error floor: **TE**: ρ 100 %, ϕ 1.45 ^{o}, **TM**: ρ 30 %, ϕ 1.45 ^{o}; τ = 7; 300 iterations; RMS 2.82.

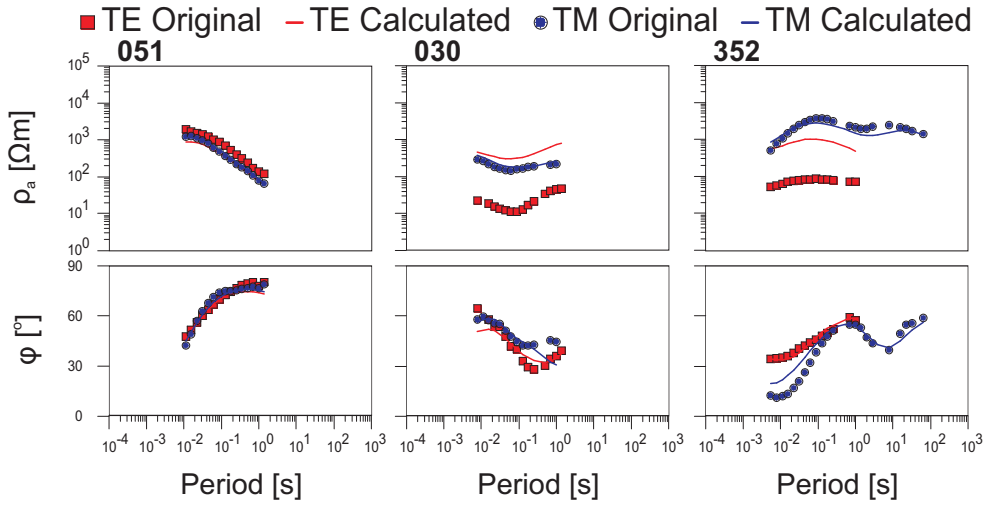


Figure 3.15: 2D data fit profile P4 (model cf. Fig. 3.14): Data of three sites is shown exemplarily. Locations of the sites is marked in Fig. 3.14 by red triangles. The data fit for TM mode is in general better. The TE mode data is affected by static shift.

Figures 3.16 to 3.19 show a section of the inversion result for profile P4 (top) compared to the result for the corresponding subprofile (middle) and a graph of the RMS values for the sites. In cases where the green RMS value bar is missing, the site was not included within the inversions.

In general, the inversion results of the subprofiles are in good agreement with the inversion over the entire profile line. The actual structure of conductive and resistive features within profile P4 could be resolved in more detail as well.

Pronounced differences between profile P4 and the subprofiles appear in the models for P4W, P4C and

P4E where the highly resistive rocks of intrusions and the BGB appear as deeper reaching features than in profile P4. Since the RMS values are in general lower and the modelled data fits the measured data better for the subprofiles, the greater depth extent of the resistive rocks seems to be consistent. This and other discrepancies between profile P4 and its subprofiles can be attributed to the finer discretization of the subprofiles allowing for more complex structures within the model. Furthermore, the additional sites used for the subprofile inversions provide additional information of the subsurface.

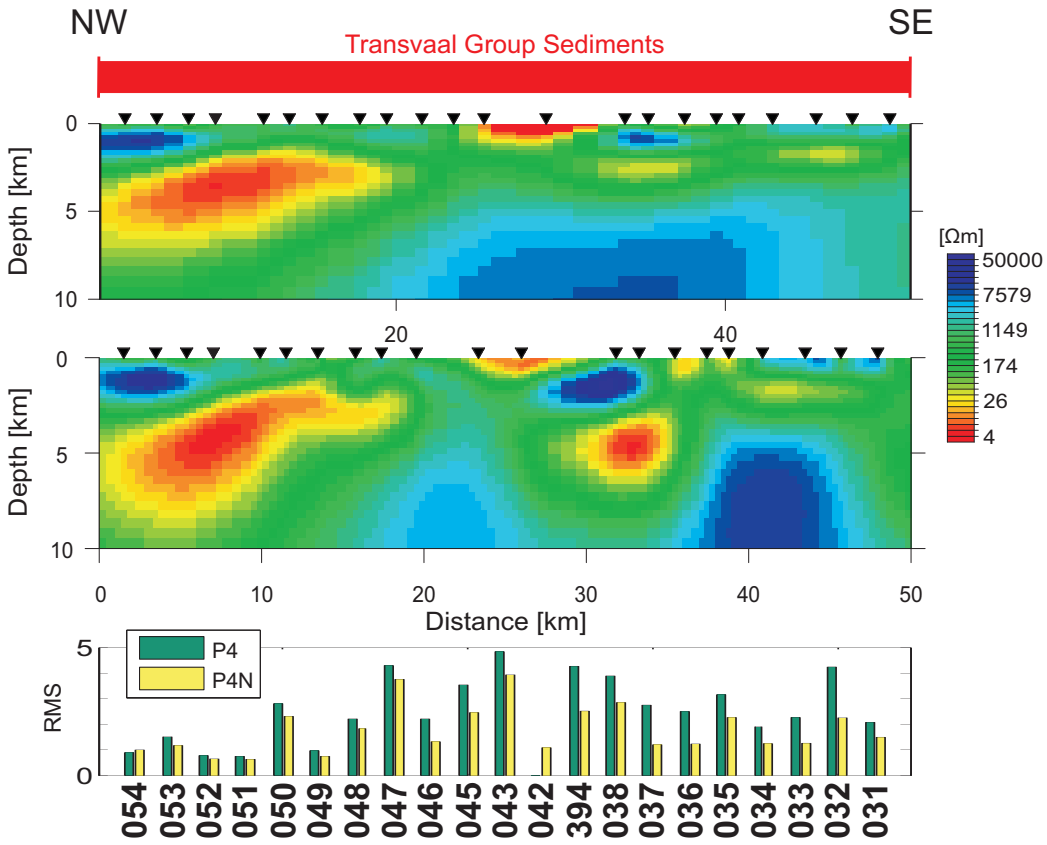


Figure 3.16: 2D inversion result along profile **P4N**: Comparison with the section of profile P4 (top) corresponding to **P4N** (middle). The graph at the bottom shows a comparison of the RMS values of the P4 and P4N models. Settings: 10^3 Ωm half-space; 152 x 167 cells; error floor: **TE**: ρ 100 %, ϕ 1.45 ^{o}, **TM**: ρ 30 %, ϕ 1.45 ^{o}; τ = 5; 300 iterations; RMS 1.75.

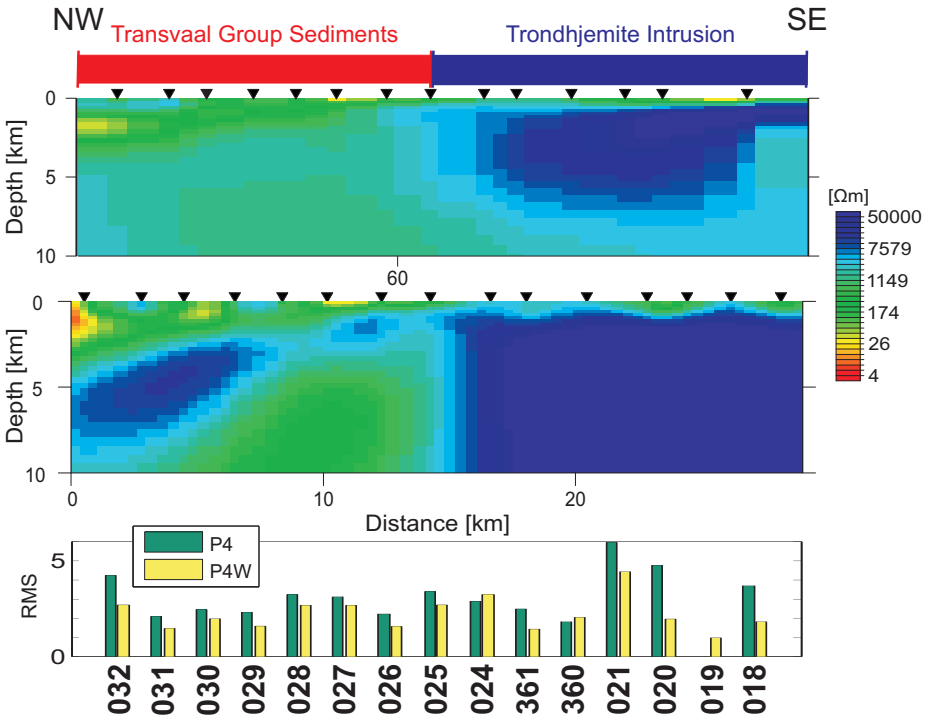


Figure 3.17: 2D inversion result along profile **P4W**: Comparison with the section of profile P4 (top) corresponding to **P4W** (middle). The graph at the bottom shows a comparison of the RMS values of the P4 and P4W models. Settings: 10^3 Ωm half-space; 152 x 151 cells; error floor: **TE**: ρ 100 %, ϕ 1.45 ^{o}, **TM**: ρ 30 %, ϕ 1.45 ^{o}; τ = 5; 300 iterations; RMS 2.31.

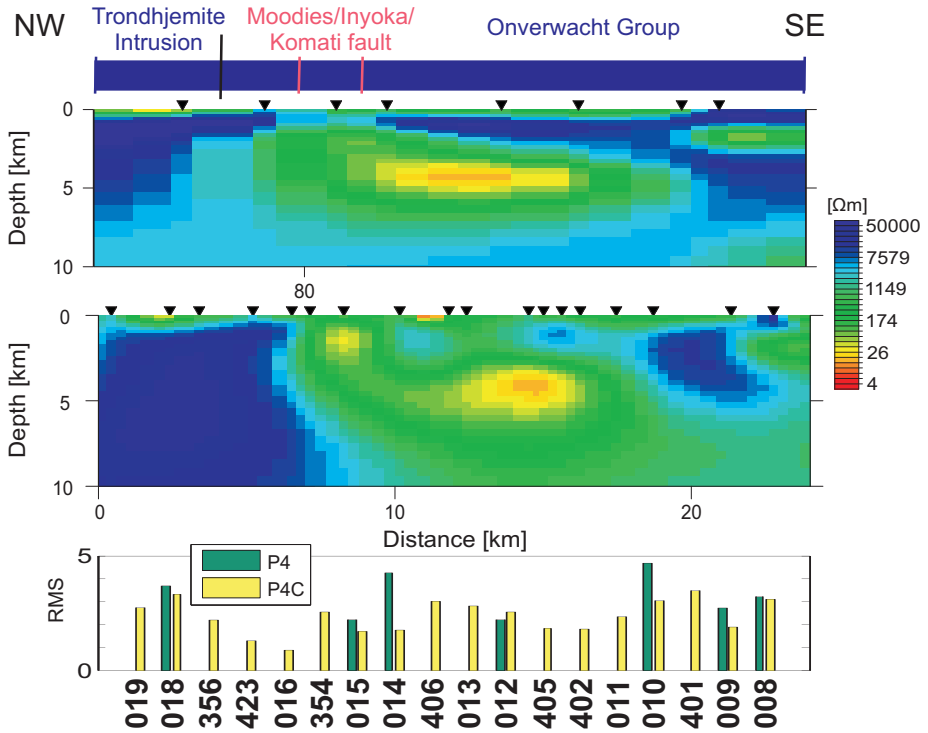


Figure 3.18: 2D inversion result along profile **P4C**: Comparison with the section of profile P4 (top) corresponding to **P4C** (middle). The graph at the bottom shows a comparison of the RMS values of the P4 and P4C models. Settings: 10^3 Ωm half-space; 152 x 178 cells; error floor: **TE**: ρ 100 %, ϕ 1.45 ^{o}, **TM**: ρ 30 %, ϕ 1.45 ^{o}; τ = 5; 300 iterations; RMS 2.33.

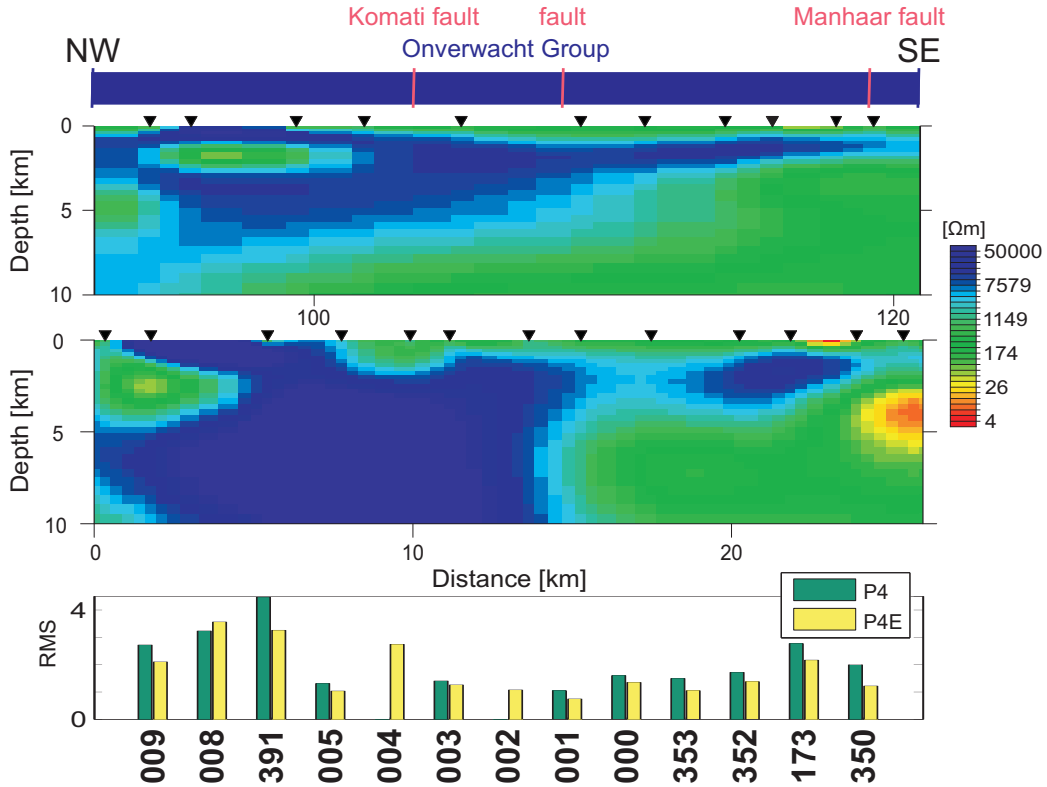


Figure 3.19: 2D inversion result along profile **P4E**: Comparison with the section of profile P4 (top) corresponding to **P4E** (middle). The graph at the bottom shows a comparison of the RMS values of the P4 and P4E models. Settings: 10^3 Ωm half-space; 152 x 167 cells; error floor: **TE**: ρ 100 %, ϕ 1.45 ^{o}, **TM**: ρ 30 %, ϕ 1.45 ^{o}; τ = 5; 300 iterations; RMS 1.90.

Profile P5

Profile P5 crosses from north to south the Transvaal Group sediments, the Trondhjemite intrusion and rocks of the Onverwacht Group which is the oldest part of the BGB. The Onverwacht Group along this profile is partially overlain by units of the Moodies and Fig Tree Group.

TE and TM mode data (unrotated) of 21 sites was inverted to obtain the model in Fig. 3.20. In parts it was difficult to fit the measured rather complex data with a 2D model. This is especially true for the section marked with an A. Combined with a strong static shift in the data (sites 347 and 346), this leads to the observed numerical artefact of high resistivities at the surface. An inversion test without sites 347 and 346 leads to a more homogeneous model with resistivities around 10^3 Ωm for area A.

Apart from the areas around site 346 and 339, the modelled data fits the measured data (Fig. 3.21) reasonably well for such a geologically complex subsurface. The modelled data of the southern sites (from 338 and 329) in particular shows a good agreement with the measured data.

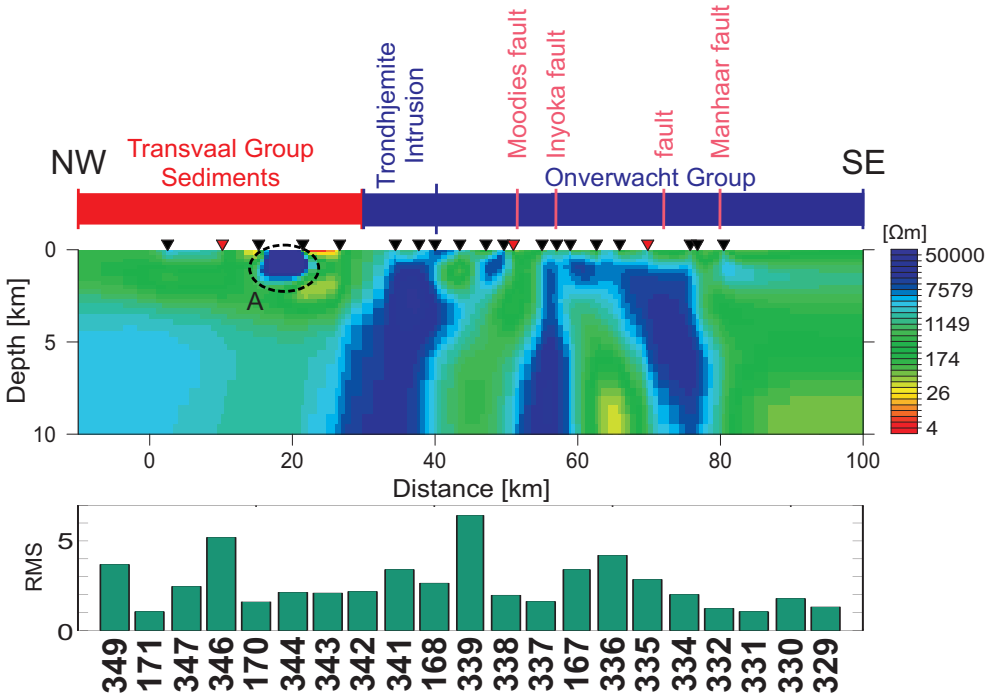


Figure 3.20: 2D inversion result along profile P5: Model shows a numerical artefact in area A due to static shift in the data. The graph at the bottom shows the RMS values of all sites. Site locations are marked by triangles; sites for which the data fit is shown in Fig. 3.21 are marked in red. Settings: 10^3 Ωm half-space; 152 x 150 cells; error floor: **TE**: ρ 100 %, ϕ 1.45 ^{o}, **TM**: ρ 30 %, ϕ 1.45 ^{o}; τ = 7; 300 iterations; RMS 2.62.

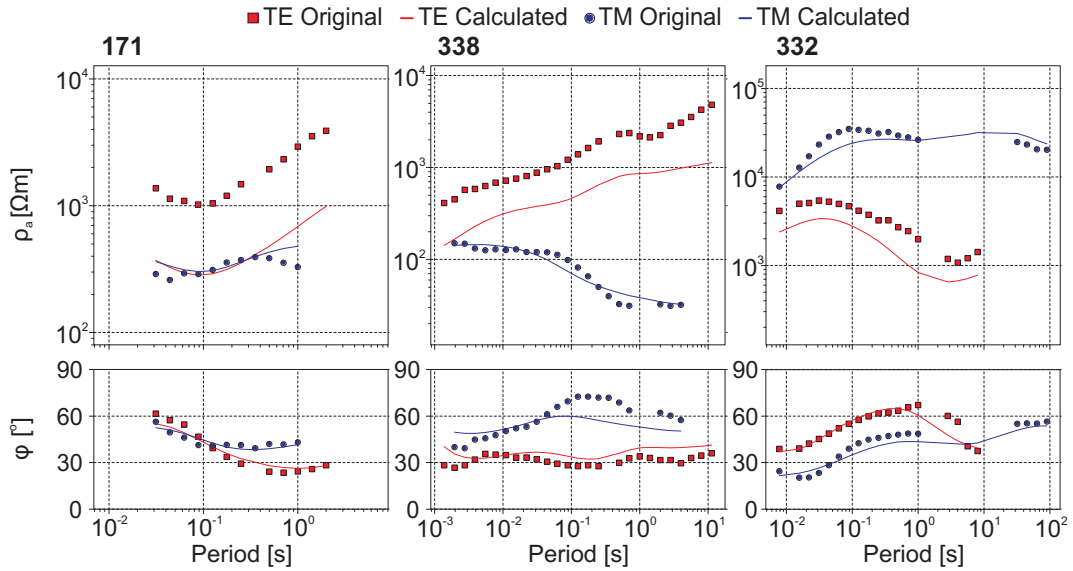


Figure 3.21: 2D data fit profile P5 (model cf. Fig. 3.20): Data of three sites is shown exemplarily. Locations of the sites is marked in Fig. 3.20 by red triangles. The data fit for TM mode is in general better. TE mode data is affected by static shift.

Profile P6

The northernmost profile P6 of the BGB MT data set covers the Transvaal Group sediments, the Trondhjemite intrusion and the Kaap Valley pluton in the north. The southern part of the profile is dominated

by rocks of the Fig Tree Group with numerous outcrops of the Onverwacht Group.
The model for profile P6 shown in Fig. 3.22 was obtained after a combined inversion of TE and TM mode data (unrotated) from 13 sites. MT sites in the northern part of the profile with (at least for the Barberton area) relatively low apparent resistivities in the order of 10^2 to 10^3 Ωm suggest that the Transvaal Group sediments crossed by profile P6 are again characterized by low resistivities. Although, due to static shift these resistivities might be underestimated.
The data fit for profile P6 is exemplarily shown for three sites in Fig. 3.23 and shows in general a good agreement between the measured and the modelled data. However, for sites south of the boundary between the Transvaal Group sediments and the Trondhjemite intrusion the 2D inversion has difficulties to fit the measured data properly.

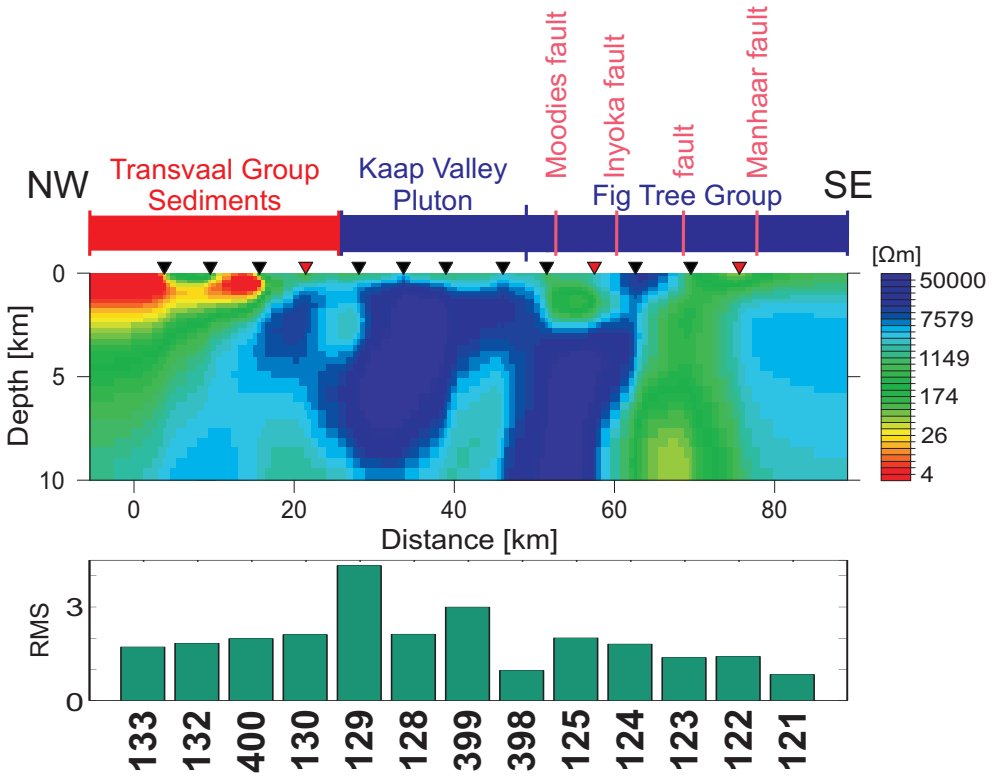


Figure 3.22: 2D inversion result along profile P6. The graph at the bottom shows the RMS values of all sites. Site locations are marked by triangles; sites for which the data fit is shown in Fig. 3.23 are marked in red. Settings: 10^3 Ωm half-space; 152 x 130 cells; error floor: **TE**: ρ 100 %, ϕ 1.45 ^{o}, **TM**: ρ 30 %, ϕ 1.45 ^{o}; τ = 7; 300 iterations; RMS 2.13.

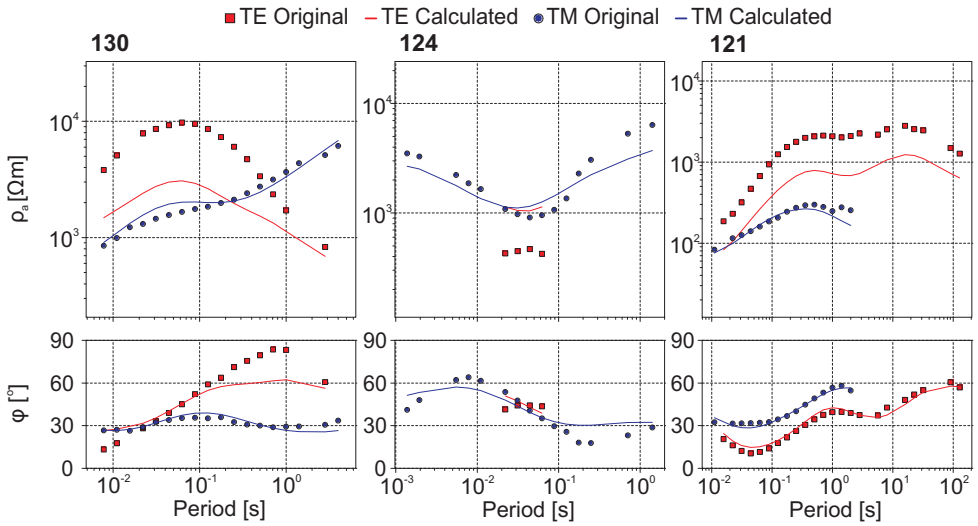


Figure 3.23: 2D data fit profile P6 (model cf. Fig. 3.22): Data of three sites is shown exemplarily. Locations of the sites is marked in Fig. 3.22 by red triangles. The data fit for TM mode is in general better. TE mode data is affected by static shift.

3.5 3D MT Inversion results

The 3D inversions of the Barberton MT data set were carried out using the ModEM inversion system by Egbert and Kelbert (2012) and Meqbel (2009). For the inversions a subset of 158 sites was selected (see Fig. 3.6). To obtain the preferred model presented in this section, all four impedance tensor components of the sites were simultaneously inverted. The MT data was used unrotated. Hence, the data and the model are aligned along the geographic coordinate sytem with the x, y and z axis pointing north, east and positive downwards, respectively.

Four selected depth slices which show the main features of the 3D model are shown in Fig. 3.24 from close to the surface to a depth of ~ 10 km. The inversion changed the starting model down to a depth of ~ 95 km. Starting from a depth of ~ 25 km the subsurface is dominated by high resistivities (around 10^4 Ωm).

The shown depth slices of the 3D model (Fig. 3.24) provide an overview of the areal distribution of conductive and resitive structures. As observed in the 2D inversion results the northwestern part of the model (area A, northwest of the dashed line) is dominated by a more or less interconnected zone of lower resistivities (~ $10^1 \ldots 10^2$ Ωm) starting at the boundary of the BGB and the Transvaal Group sediments. This zone is rather shallow (hundreds of metres) in the southeast reaching greater depths of at least 10 km towards the northwestern part of the model. The discontinuous structure of this high conductivity zone (area A) for the upper hundreds of metres is most likely caused by the in parts poor MT site coverage and also by near surface static shift and distortion effects.

Large resistive deep reaching (> 10 km) structures dominate the area of the BGB (areas B1, B2 and B3) which are in parts interrupted by conductive features. These conductive features (area C) are mainly aligned along faults or are enclosed by them. A southward continuation of the faults within the Mpuluzi Batholith is not discernible in the model. However, the cell size with a minimum of 1.7 km in the centre of the 3D model is unable to resolve the minute structure of the faults. Even in the 2D inversion studies where the horizontal cell size is in the order of hundreds of metres only faint imprints of higher

conductivities were observed (see profile P1 in sec.3.4). This is in part due to the regularization but even the tear zone inversions still show resistivities in the order of 10^3 Ωm. Hence, a southward extension of the faults cannot be ruled out due to the lower resolution power of the 3D model when compared to the 2D models. A finer discretization of the 3D model when using the entire BGB MT data set is not possible due to computational limitations in implementing large 3D models. Inversions focusing only on specific areas (e.g. around the major fault traces) would be necessary for a more in depth analysis concerning the subsurface structure of the BGB fault system. However, further developments of the 3D inversion algorithm and its regularisation are required to address this issue. Currently, artefacts and other side effects usually occur, when cell dimensions differ significantly.

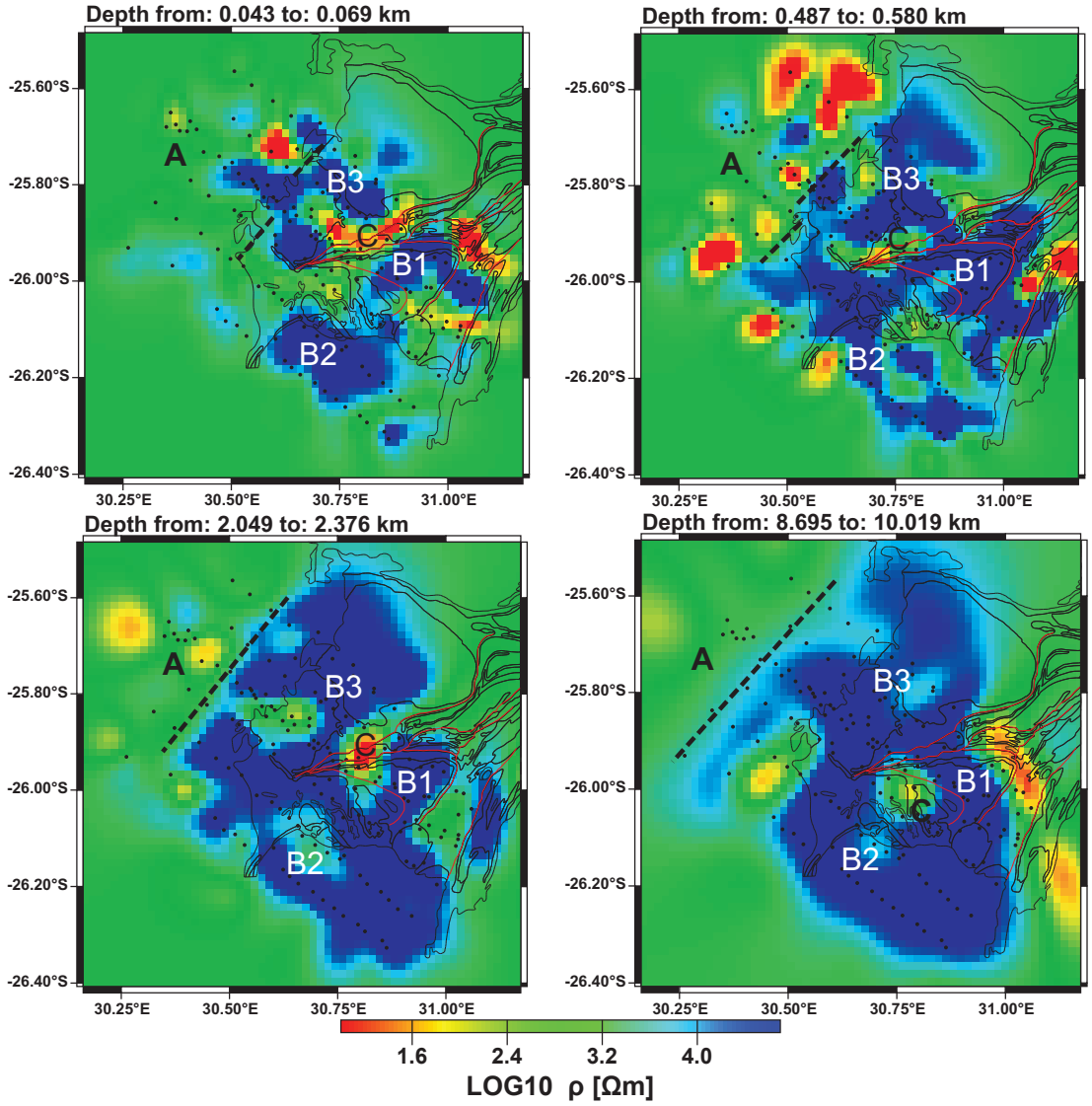


Figure 3.24: Depth slices of the 3D inversion result: Top layers show extended structures of higher conductivities especially towards the northwest of the model area; greater depths within the model are dominated by higher resistivities. Position of geological units as depicted in Fig. 3.6 are outlined in black; faults are marked by red lines and MT site positions are represented by black dots. For areas A, B and C see description in the text. Settings: 10^3 Ωm half-space; $100 \times 100 \times 60$ cells; $1.7 \times 1.7 \times 0.02$ km edge length of cells in the model centre; full impedance tensor inversion; error floor: 5 % for all components; smoothing: 0.3, 0.3, 0.2; 266 iterations; RMS 5.81.

To illustrate the misfit of the 3D inversion result (Fig. 3.24) the modelled and the measured data are shown in comparison in Fig. 3.25 as apparent resistivity and phase curves. Since the components of the impedance tensor **Z** are used in the 3D inversion as input data (instead of ρ_a and ϕ for 2D inversions) and to give an overview over the entire model area, the RMS is shown colour coded in map view as well (Fig. 3.26). No systematic misfit of the data was observed either limited to a particular area, a specific impedance tensor component or a certain period range.

The 3D inversion result has a comparatively high overall RMS value of 5.81 (error floor: 5 % for all impedance components). Looking at the fit of the modelled to the measured data in detail (Fig. 3.25) and the RMS values for all impedance tensor components (Fig. 3.26) reveals the reason for the elevated RMS value. Most sites have a good data fit whereas a few selected sites have an exceptionally bad data fit with corresponding high RMS values (usually only for one of the impedance tensor components) which increases the overall RMS of the model. Larger misfits of the data occur at sites located either close to the boundary between the Mpuluzi Batholith and the BGB rocks, or at sites recorded on top of BGB formations in the southeastern part of the model. These areas are in particular geologically complex. Due to the cell size ($\sim$ 1.7 km), the inversion is not able to include all (minor) structures within the model which are reflected in the MT data resulting in the larger misfits. The large survey area ($\sim 100 \times 100\,\mathrm{km}^2$) made it impossible to select smaller cell dimensions. Inversions with the shown model settings and parallelisation over frequencies require already a computation time of over a week on the cluster available at the GFZ Potsdam. Stronger regularization (higher smoothing) was used laterally (0.3) than vertically (0.2) to accomodate for the spacing of 10 km between the profiles and the limited amount of data available for periods $>$ 10 s (due the the contamination of longer periods by EM noise). This further hinders sharper resistivity contrasts and small scale structures developping during the inversion iterations.

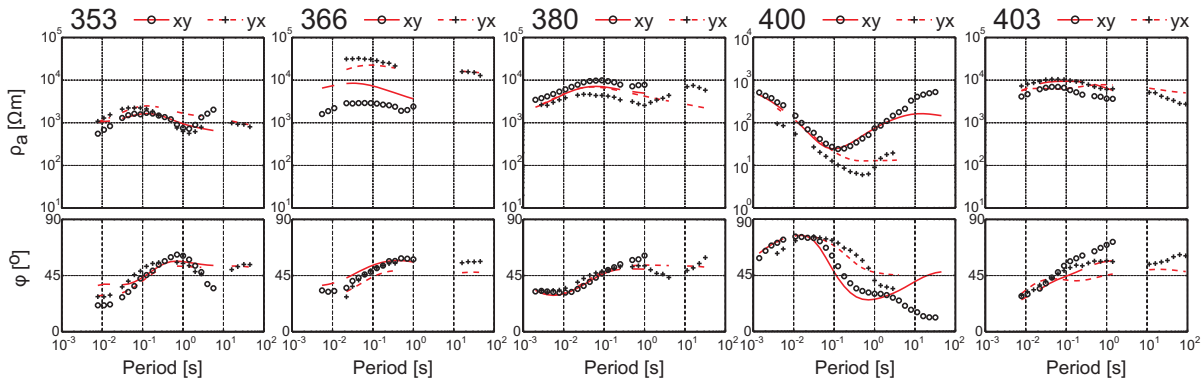


Figure 3.25: 3D data fit (model cf. Fig. 3.24): Apparent resistivity and phase curves calculated from the off diagonal elements Z_{xy} and Z_{yx} of five sites are shown exemplarily. Locations of the sites are marked in Fig. 3.6 by red stars.

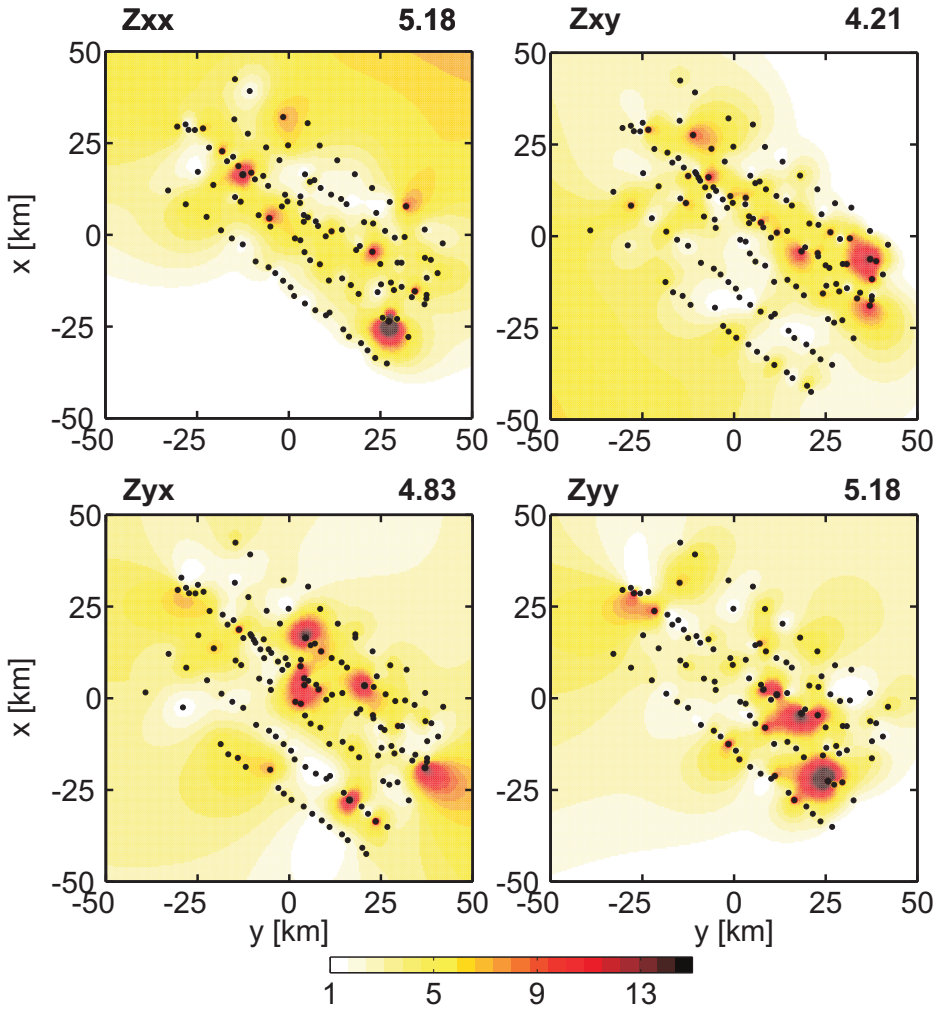


Figure 3.26: Map view of the colour coded RMS values for all MT sites (black dots) for the 3D inversion result shown in Fig. 3.24 for all components of the impedance tensor **Z**. In general RMS values are between 1 and 5, however, a few select sites mainly towards the southeast of the model exhibit larger RMS values (see text for further discussion).

3.6 Joint interpretation and discussion of the 2D and 3D inversion results

A direct comparison of the 2D and 3D inversion results along profiles P1 to P6 is shown in Fig. 3.27 through 3.33 (except Fig. 3.29). Numerous structural similarities occuring in both, 2D and 3D models, can be observed for all profiles and are discussed in detail in this section. Differences between the models can mainly be attributed to the previously discussed difficulties using 2D inversion for the MT data for profiles P2 to P6 (see sec. 3.3.1 and 3.4). Different model grid sizes used for 2D and 3D inversions resulting in different resolution capacities of geological structures lead to further variations between the 2D and 3D results.

In 2D, the most important assumption is that the electrical conductivity solely changes in one horizontal direction along the profile but is constant perpendicular to it. Off-profile conductivity structures violate this assumption and are either projected into the 2D inversion result or hinder a succesful 2D inversion. Especially in the NE part of the measurement area, off-profile structures and the general geological complexity of the area (e.g. the extensive fault system of the BGB) might bias the 2D results. In 3D

inversions, the electrical conductivity is allowed to vary in all spatial directions. Structures located in between profiles are included in the model in the accurate place during the inversion process if they are properly resolved. Since the resolvability characteristics of 3D inversion algorithms are still not fully understood a validation of the observed electrical structures is indispensable. This is done by comparing the 3D with the 2D results and discussing the plausibility of the observed structures using the knowledge from geological studies.

For profile P1 (Fig. 3.27), the 2D (top) and 3D model (bottom) are in good agreement. Both results show the same general division with electrically conductive structures (in the NW) representing the Transvaal Group sediments and a rather homogeneous electrically resistive zone for the granitoids and the Mpuluzi batholith (in the SE). At the surface, the boundary between the Transvaal Group sediments and the granitoids is marked by an escarpment where the sediments unconformably overlie the granitoids (MacLennan, 2012). Both the 2D and 3D models confirm that the Transvaal Group sediments are interspersed with more or less localized resistive and conductive features. This can be observed in all profiles. The resistive body R appears as a shallow around 1 km thick structure in 2D and 3D. Its extension towards northwest is unclear since the northernmost MT site used in the 2D inversion could not be included in 3D (since it would have been in the same cell as its neighbouring site). Whereas in 2D, the northwestern border of the granitoids is marked by conductors C1 and C2, they appear as a continuous deep reaching (>20 km) structure in 3D (conductor C). The geological structures responsible for the last mentioned resistive and conductive features are unclear due to the lack of geoscientific studies of the Transvaal Group sediments in this region. However, in other parts of the Transvaal Supergroup (i.e. Griqualand West sub-basin, ~ 900 km WSW of the BGB) up to 1 000 m thick chert and Banded Iron Formations (BIF) are reported (Schröder et al., 2006). Similar formations could be an explanation for the conductive structures found NW of the BGB. Underneath the BIF layer, siliciclastic mudstone, limestone, sandstone and conglomerate layers can be found (Schröder et al., 2006) which are a likely reason for intermediate and higher resistivities throughout the Transvaal Group sediments covered by the BGB MT sites.

The resistive structure of the granitoids and the Mpuluzi batholith is in part disrupted by localized conductive features. A possible southward continuation of some faults within the Mpuluzi batholith is suggested by the 2D tear zone studies (see sec. 3.4). For profiles P2 and P3 an influence of off-profile conductive features cannot be ruled out since the profiles are close to the conductive fault system of the BGB. However, for profile P1 such an influence on the conductivity structure of the tear zones seems unlikely since the profile is located well within the resistive Mpuluzi batholith. In the 3D model, only a shallow conductive imprint possibly related to the Komati fault (conductor KF) can be observed. Tear zone studies to further investigate KF are in principle possible with the 3D inversion programme ModEM. Practically, they are not feasible in order to investigate narrow fault zones due to the nearly infinite ways to include a tear zone in a 3D model. Furthermore, the minimum possible grid size in 3D is still too large for such narrow structures. In contrast to the 2D results, a conductive structure (ISZ) apparently related to the Inyoni Shear Zone is clearly visible in 3D. The ISZ is tilted towards NW, reaching ~ 4 km depth. The ISZ as well as the KF (both formed ~ 3.22 Ga ago during the D2 collapse phase, Moyen et al., 2006; Fig. 3.3) might have served as the 'path of the least resistence' during the ascend of the granitoids and the Mpuluzi batholith (~ 3.1 - 3.14 Ga ago, de Wit et al., 2011). Hence, the emplacement of these large plutonic bodies might have led in parts to the destruction, to an overprint and/or deformation of pre-existing faults and shear zones (KF and ISZ). This is supported by the inversion results as they show conductive structures that vary significantly from profile P1 to P3. Some of the pre-existing faults (KF

and ISZ) might also have been reactivated after the emplacement of the plutons resulting in the (fault related) higher electrical conductivities in the southern MT profiles.

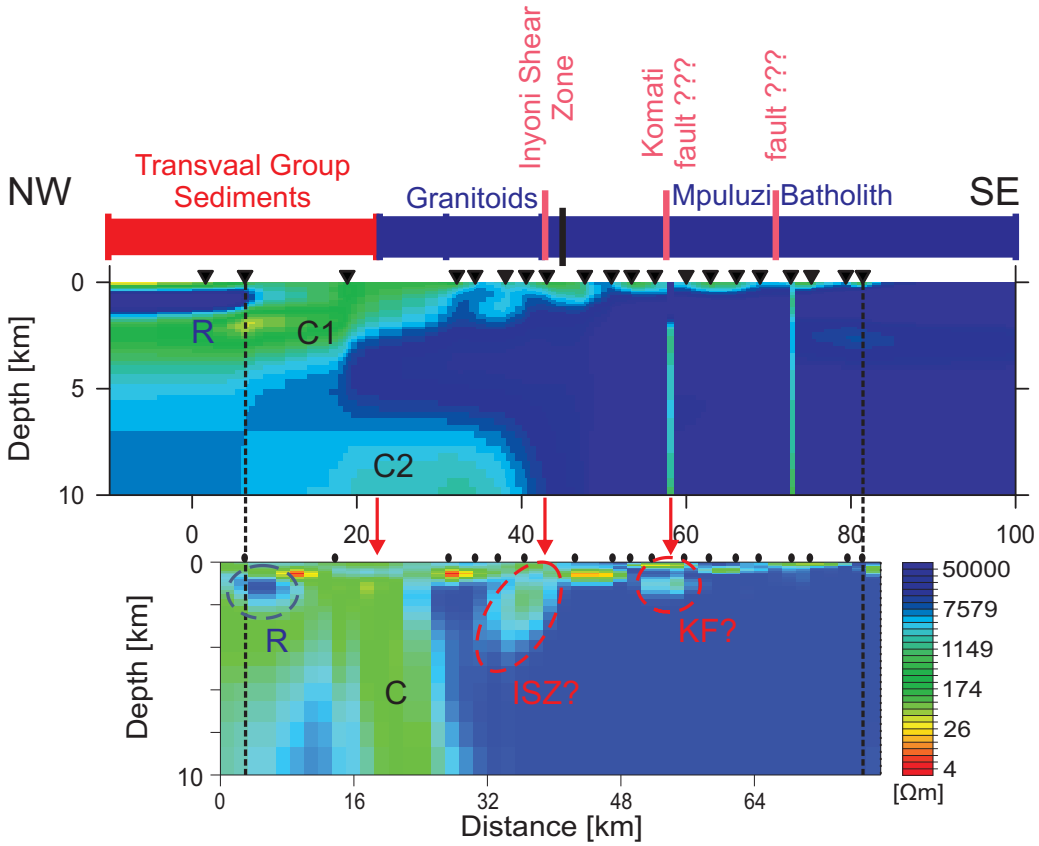


Figure 3.27: 2D (top) and 3D inversion results (bottom) along profile P1: For 2D, the inversion result with tear zones (no smoothing applied between the tear and the neighbouring model cells) is shown. ISZ: Inyoni Shear Zone, KF: Komati fault. The electrically conductive and resistive structures are discussed in the text. Site locations are marked by black triangles (in 2D, top) and circles (in 3D, bottom).

The 2D and 3D inversion results for profile P2 (Fig. 3.28) show a distinct difference in the geo-electrical struture for the Transvaal Group sediments. The conductive structures C1 and C2 in the 3D model are much deeper reaching than conductor C in the 2D result. Furthermore, the resistive structure corresponding to the large granitoid body extends far underneath the Transvaal Group sendiments in the 3D model (resistor R) and separates structures C1 and C2.

Another interesting feature of the 3D result is the conductor IKF which is centred around the position where a possible fault was detected in the 2D tear zone studies. The IKF conductor is bounded by the Inyoni Shear Zone in the NW and a possible continuation of the Komati fault in the SE. It appears to be unconnected to neighbouring conductive bodies (cf. depth slice 2.049 to 2.376 km in Fig. 3.24, north of structure B2) and, hence, it is unclear whether it is a remnant of the fault system or the result of the hydrothermal activities and gold mineralizations which mark the final stages of the BGB evolution (~ 3.126 and ~ 3.084 Ga, de Ronde et al., 1991). In contrast, the deeper conductor IKFd is clearly connected to conductive structures which cross profiles P3 to P5 (Fig. 3.29)

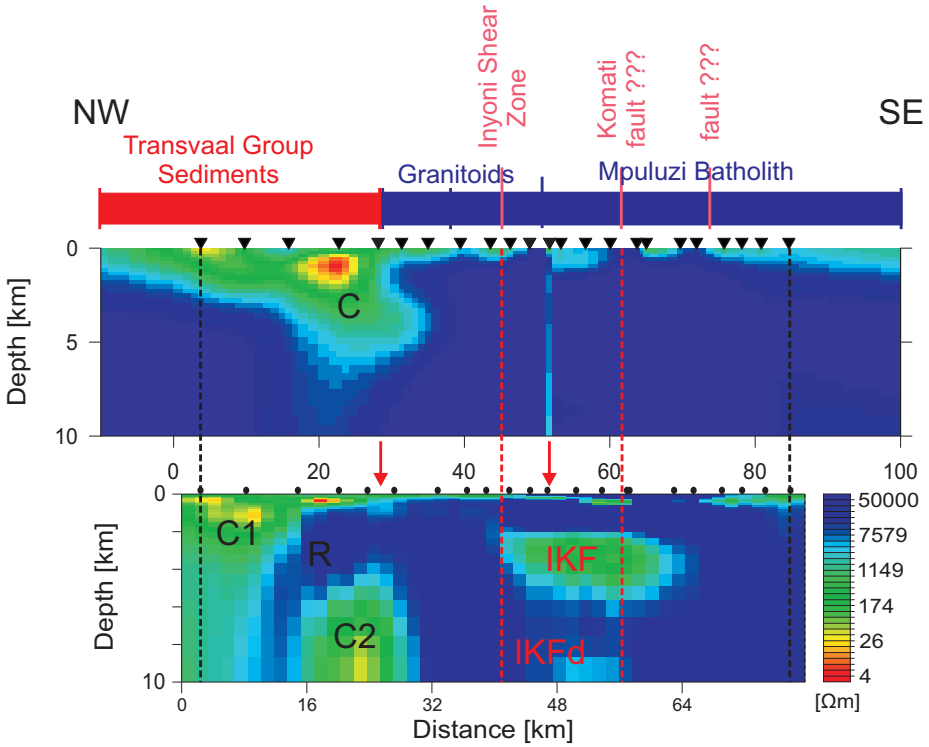


Figure 3.28: 2D (top) and 3D inversion results (bottom) along profile P2: For 2D, the inversion result with tear zone (no smoothing applied between the tear and the neighbouring model cells) is shown. IKF: Inyoni Shear Zone - Komati fault, IKFd: IKF deep. The electrically conductive and resistive structures are discussed in the text. Site locations are marked by black triangles (in 2D, top) and circles (in 3D, bottom).

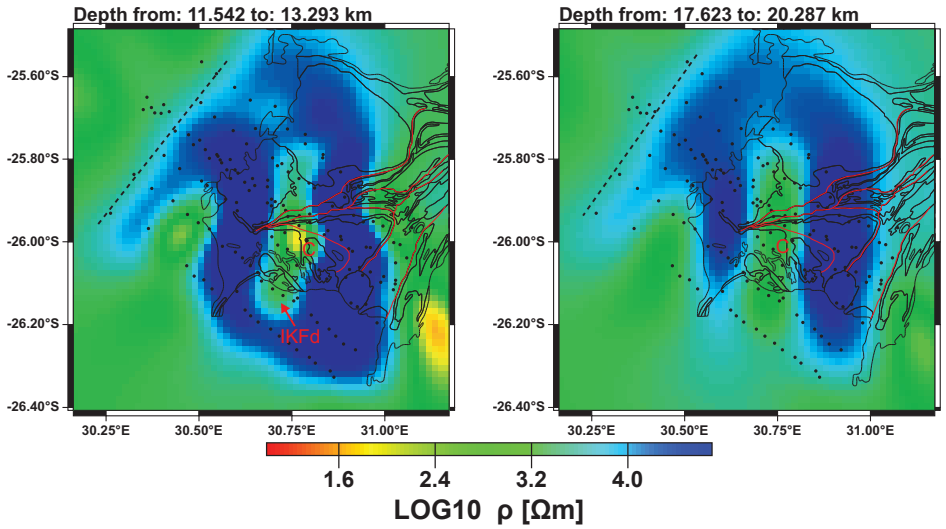


Figure 3.29: Depth slices of the 3D inversion result: Deeper model layers show extended structures of higher conductivities probably connected to hydrothermal activities and gold mineralizations (de Ronde et al., 1991); the dashed line marks the extend of the roots of intrusive bodies and plutons underneath the Transvaal Group sediments. Position of geological units as depicted in Fig. 3.6 are outlined in black, faults are marked by red lines and MT site positions are represented by black dots. IKFd: Inyoni Shear Zone - Komati fault deep conductive structure.

Despite the difficulties during the 2D inversion of profile P3, the results are quite similar to the 3D model (Fig. 3.30). The main problems in the 2D model appear to be the resolution of the upper 2 to 3 km and of the SE border of the profile which is reflected by highly conductive anomalies. Profile P3 (3D) is nearly identical in its geo-electrical structures to profile P2 and can similarly be interpreted. The conductor IKFd is a bit more shallow in profile P3 and a faint trace of its connection to the surface is visible (conductor C in Fig. 3.24 and 3.29). Some of the shallow conductive structures NW of the Inyoni Shear zone might be an imprint of a hydrothermal system related to the Badplaas hot springs.
Instead of the Mpuluzi batholith (in P1 and P2), the Dalmein pluton and neighbouring TTG intrusive bodies dominate the resistive structure of the SE part of the profile. This is only disrupted by a small conductive feature (conductor MF) which is most likely linked to the hydrothermal activity close to the Maanhaar fault supported by the gold deposits found along this shear zone (Fig. 3.1, Dirks et al., 2013). To the NE in profiles P4 to P6 the conductor related to the MF becomes more pronounced (Fig. 3.31 to 3.33).

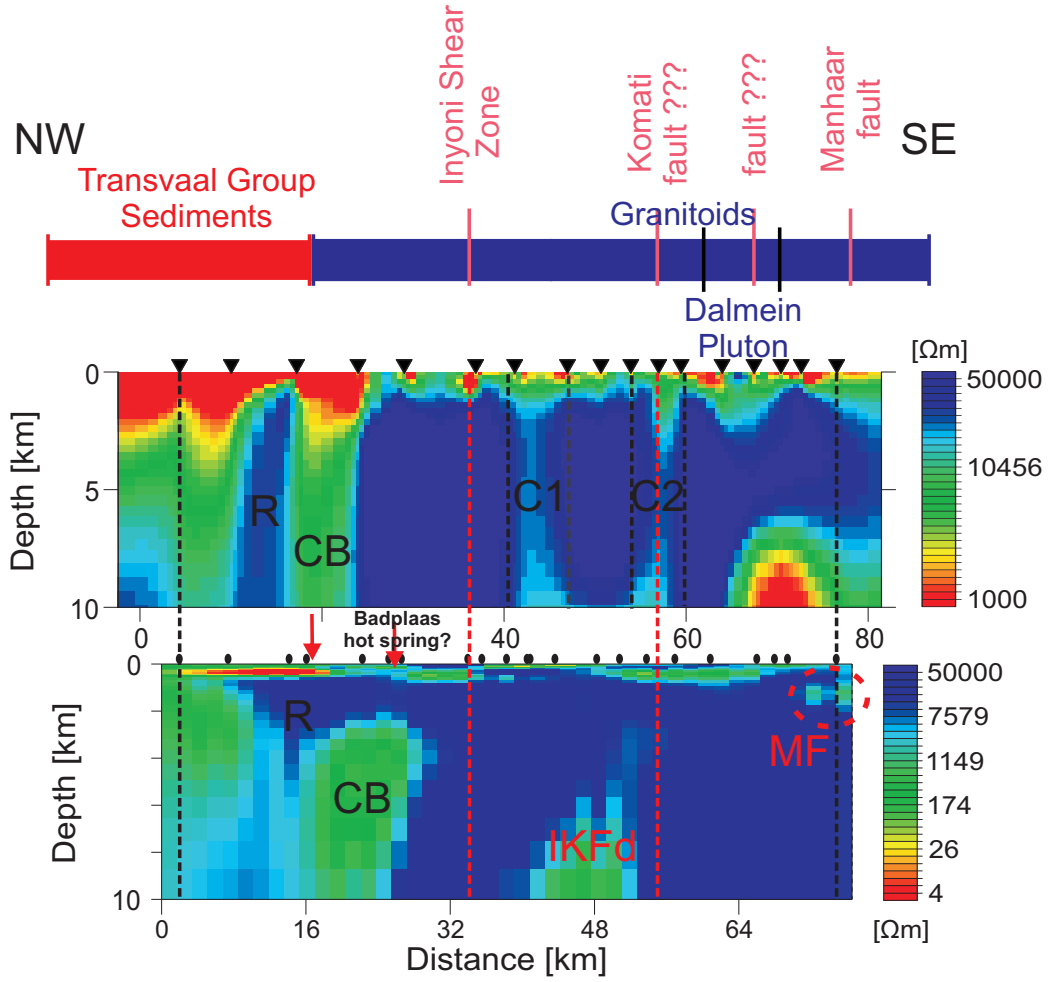


Figure 3.30: 2D (top) and 3D inversion results (bottom) along profile P3: For 2D, the inversion result with a different colour scale for the electrical resistivity is shown for a better visibility of the conductive structures. IKFd: Inyoni Shear Zone - Komati fault deep conductive structure, MF: Manhaar fault. The electrically conductive and resistive structures are discussed in the text. Site locations are marked by black triangles (in 2D, top) and circles (in 3D, bottom).

Starting from profile P4 (Fig. 3.31), the differences between the 2D and 3D results are more pronounced. This is most likely a direct effect of the complex three-dimensional structure of the BGB in this part of the measurement area making it nearly impossible for 2D inversion to properly resolve all significant structures. However, the detailed 2D results for subsections of pofile P4 (Fig. 3.16 through 3.19) confirm the deeper extent of the resistive units of the Trondhjemite intrusion and parts of the Onverwacht Group which can be observed in the 3D model in profile P4 as well as in profile P5 (Fig. 3.32). A major difference between 2D and 3D results is the resolution of the conductive structures related to the BGB fault system. In profiles P4 and P5 (2D results), the conductive structures marked as MIK and MIF, respectively, appear as $\sim$ 10 km wide areas reaching depths of at least 10 km. In contrast, in 3D fault related conductors are wider in the shallow parts of the model (down to $\sim$ 4 km) and narrow with greater depth. Due to the abundance of gold deposits along the faults at the surface (Fig. 3.1, Dirks et al., 2013), the fault system can be considered as a network which allowed the ascend of hydrothermal fluids containing ore mineralizations. In greater depth (>4 km) and in the SW of the measurement area (profile P4), the network is spatially limited to few faults and narrow shear zones. Towards the surface and the NE of the measurement area (profiles P5 and P6), the faults appear to diverge and split-up. Especially in profile P6, the rocks of the Fig Tree and the underlying Onverwacht Group appear to have been strongly hydrothermally altered. This is marked by the generally higher conductivities in the SE section of the profile (conductors MIF, MF and CF, Fig. 3.33). The inversion models are not able to resolve every single one of these faults but instead show extended areas of intermediate and higher electrical conductivities. In profile P4, the conductor MIK (in 3D) resembles a flower structure, which occur in strike-slip fault zones (Neuendorf et al., 2005) and are assumed to be related to a subduction plate boundary (Schoene et al., 2008). In a narrow section of P4, the Komati fault is also parallel and in very close proximity to the MT profile (above the flower structure, see Fig. 3.1). Such a structure cannot properly be resolved in 2D inversions, but appears as a connection of the MIK conductive structure to the surface in 3D (long red arrow in Fig. 3.31). Kisters et al. (2003) located in the area (south of the Komati fault, at the Theespruit Formation, Fig. 3.2) a 1 km wide regional extensional detachment at the margin of the supracrustal rocks of the BGB and the granitoids. The conductive structure CB in profile P4 (3D) appears to be linked near the surface to the fault related conductive structure MIK and might, hence, be another hydrothermally altered zone.

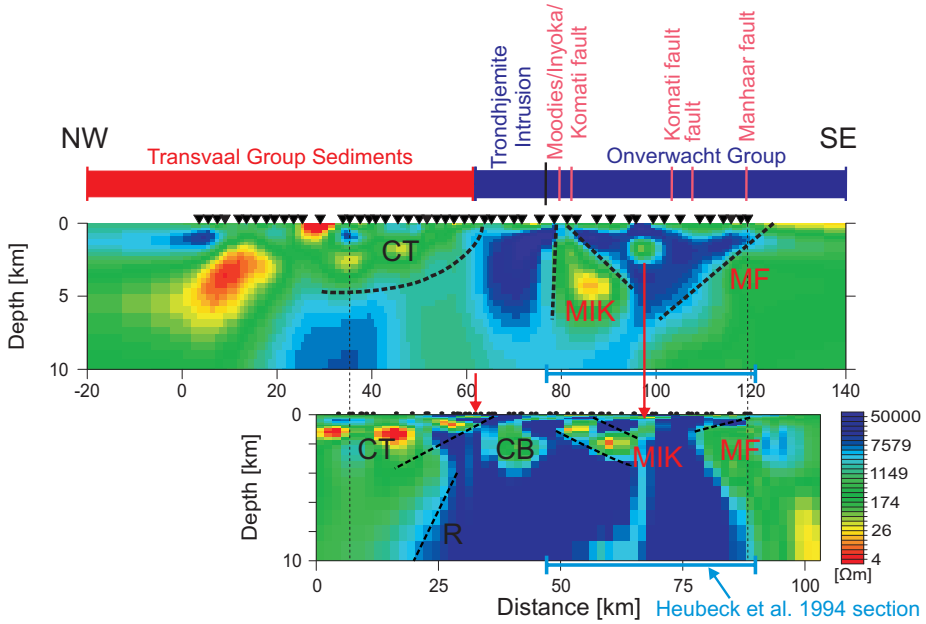


Figure 3.31: 2D (top) and 3D inversion results (bottom) along profile P4. For the geological profile by Heubeck and Lowe (1994) see Fig. 3.33. Light blue lines mark the section for which the geological profile is relevant. MIK: Moodies - Inyoka - Komati fault, MF: Manhaar fault. The electrically conductive and resistive structures are discussed in the text. Site locations are marked by black triangles (in 2D, top) and circles (in 3D, bottom).

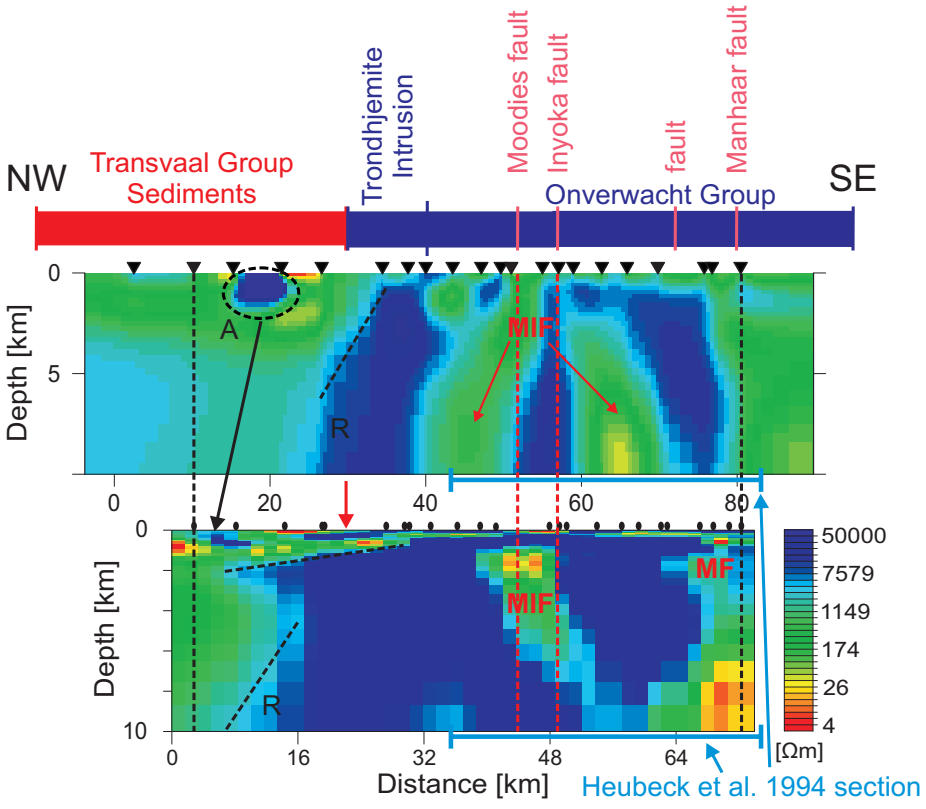


Figure 3.32: 2D (top) and 3D inversion results (bottom) along profile P5. For the geological profile by Heubeck and Lowe (1994) see Fig. 3.33. Light blue lines mark the section for which the geological profile is relevant. MIF: Moodies - Inyoka fault, MF: Manhaar fault. Site locations are marked by black triangles (in 2D, top) and circles (in 3D, bottom).

Profile P6 (Fig. 3.33) crosses large parts of the Fig Tree Group which is rich in Banded Iron Formations and can be considered as the main reason for the shallow conductors (down to ~ 2 km depth) SE of the Moodies fault. This conductive zone is interrupted by a resistive zone which correlates with an outcrop of the Onverwacht Group.

Comparing the obtained 2D and 3D inversion results to the evolution theories described in sec. 3.1 gives indications to which tectonic processes are most likely to have caused the geological structures that we observe today in the BGB.
Sequences of geologic units along profiles P4 to P6 (Fig. 3.31 through 3.33) correlate in part with a geological section through the central part of the BGB postulated by Heubeck and Lowe (1994b). The exact position of the geological section can be found in Fig. 3.1 (light blue line). In the central part of the BGB, Moodies Group rocks are folded in a sequence of overturned synclines forming subvertical structures dipping towards SE which are separated by narrow faults (Heubeck and Lowe, 1994b). The authors suggest that this part went through a period of buckle-folding with a related formation of a detachment in 3 to 6 km depth. Concentrating solely on the shallow structures (down to ~ 4 km), this conceptual model is supported by the (subvertical) fault structure in the 3D MT results in profiles P4 and P5. A finer discretization of the 3D model grid would be necessary in order to resolve the faults in more detail. Although, a complete separation of the faults, especially where they converge between profiles P3 and P4 (conductor MIK in P4) might be impossible even with a very fine discretization. Regarding now the deeper model layers (below 4 km), the inversion results are contrary to the Heubeck and Lowe (1994b) model. Trough profiles P4 to P6 no indication for a detachment can be found. The postulated detachment would appear in a depth range between ~ 2 and 7 km (see geological section in Fig. 3.33). Primarily, it should be conductive due to mineralizations along the detachment shear plane. The most prominent candidate for a conductor in a continental detachment would be graphite (e.g. Ritter et al., 2003; Weckmann et al., 2003; Heinson et al., 2005). The subsequently occuring hydrothermal alteration of the BGB would have further increased the conductivity of a detachment. Hence, this layer could not be missed by MT investigations and the inversion results clearly contradict the existence of a detachment and the buckle-folding scheme.
In contrast, the geological scheme described by Kisters et al. (2003) seems a more likely explanation for the observed structures in the inversion models. Within a shallow 1 km wide regional detachment mapped by the authors south of the Komati fault (see Fig. 3.2) strongly lineated mylonites were found which are usually associated with planar zones of localized ductile deformation (Neuendorf et al., 2005) of, in case of the BGB, mid-crustal rocks with a NE-SW directed stretching. The overlying BGB formations underwent at the same time an episode of shearing in a NE direction (dip-slip faulting). This occured during the D2 orogenic collapse of the greenstone belt and coincides with a period of exhumation (~ 18 km). The (sub-) vertical structures observed in the upper few kilometres of the BGB sequence are attributed by Kisters et al. (2003) to the diapiric emplacement of plutonic bodies towards the end of the BGB formation around 3.2 - 3.1 Ga ago. Similar observations were made by Lana et al. (2010) at the contact between the BGB and the Steynsdorp pluton. This scenario might also be valid for the central parts of the BGB crossed by profiles P5 and P6 where the ascend of the Kaap Valley Pluton and the Piggs Peak Batholith might have caused at least some of the observed deformations of the BGB.

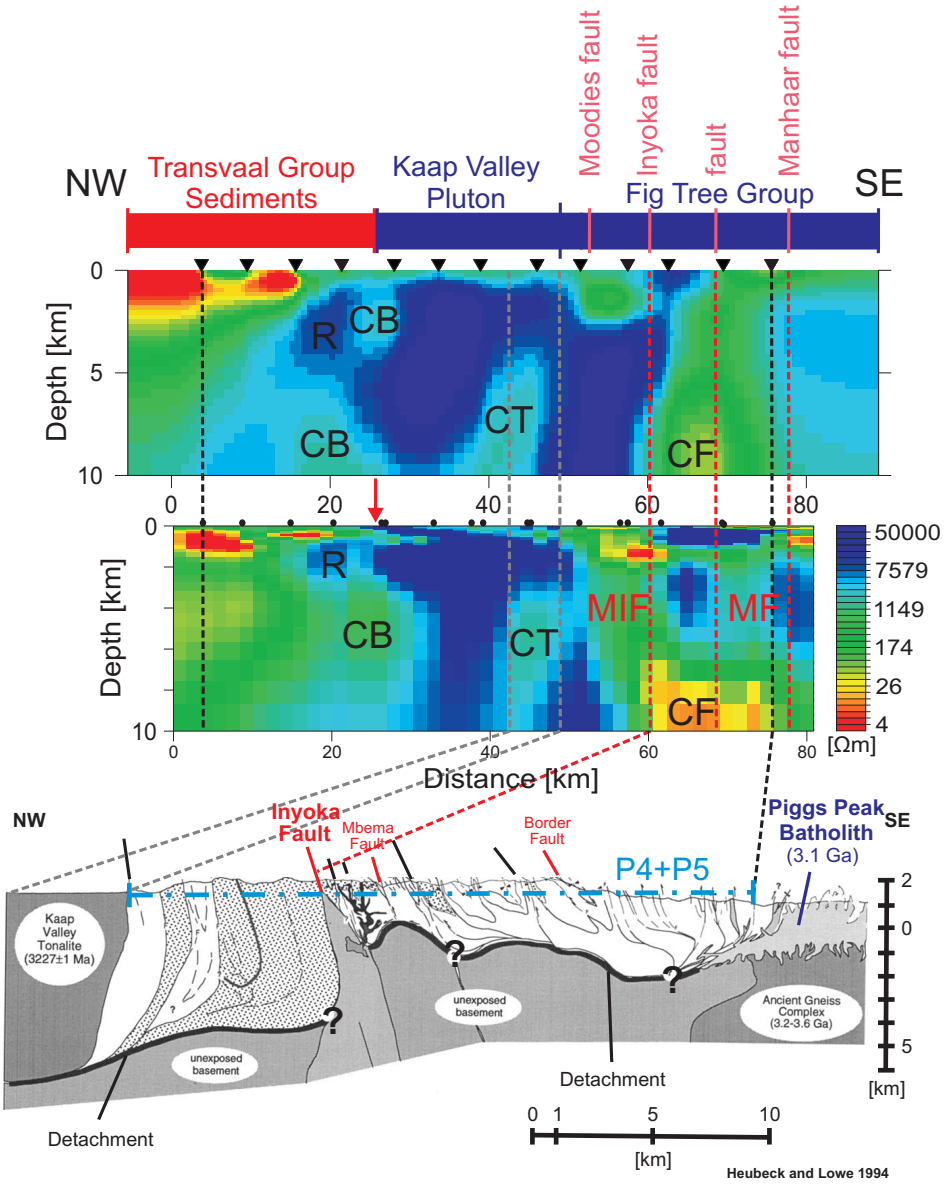


Figure 3.33: 2D (top) and 3D inversion (middle) results along profile P6. For comparison, a geological profile is shown as well (Heubeck and Lowe, 1994; for exact profile location see Fig. 3.1). The section of the geological profile which is relevant for MT profiles P4 and P5 is marked in light blue (dash-dot line). MIF: Moodies - Inyoka fault, MF: Manhaar fault. Site locations are marked by black triangles (in 2D, top) and circles (in 3D, middle).

The TTG plutonic bodies emplaced between 3.5 to 3.2 Ga ago and the younger batholiths formed ~ 3.1 Ga ago (see Fig. 3.2) which flank the southern part of the BGB appear as more or less homogeneous bodies regarding their electrical properties in the inversion results. Their vertical extent which reaches in parts at least a depth of about 20 km in the models supports the theory that they are the result of large vertical mass movement and diapiric emplacement in the upper crust (Kisters et al., 2003, and references therein). Due to the differences in their age and their composition, the emplacement of these bodies was not a continously ongoing process but occured periodically with the magma source changing over time (e.g. Kisters and Anhaeusser, 1995). Although, Lana et al. (2011) state that the emplacement of some of the TTG plutons around 3.2 Ga is a result of subduction zone-related magmatism, the BGB appears to have underwent several periods where vertical tectonic processes were dominant.

Geological units of the BGB as well as the plutonic bodies at the NW border of the BGB appear to

largely extend underneath Transvaal Group sediments (resistor R in profiles P3 to P6). This is similar to the structure observed by de Beer et al. (1988b) where BGB rocks are assumed to extend over a large area underneath the Nelspruit batholith.
The existence of high-pressure, low-temperature terranes was verified in the southern part of the BGB (Diener et al., 2005; Lana et al., 2010). This indicates the presence of cool and rigid continental crust which is essential for the initiation of plate tectonic processes. A reported (for Archean times) low apparent geothermal gradient of about $20\,^{o}$C/km for southern parts of the BGB further suggests an environment where plate tectonics are possible (Diener et al., 2005). Hence, the observed fault related zones of high conductivities (profiles P4 to P6) are interpreted as remnants from an arc-continent collision and subduction system (Moyen et al., 2006; Moyen et al., 2007) during the accretion of the BGB.

3.7 Outlook

The geological knowledge gathered over the last century as well as the MT study in this thesis mark the BGB as a complex geological system where different tectonic mechanisms contributed during the different stages of evolution to the BGB's formation. Studies at other Archean greenstone belts reflect a similar difficulty in finding a conclusive evolution theory. These theories vary likewise from a plate tectonic to a vertical tectonic driven geological setting: For the slightly younger Abitibi greenstone belt in Canada an arc-arc collision and formation of a subduction zone (around 2.7 Ga, Mueller et al., 1996) is suggested whereas for the Pilbara craton and its greenstone belts vertical tectonics are supposed to have been dominant (Van Kranendonk et al., 2007; Hamilton, 1998).
The presented thesis and papers mentioned in this section show that there is still much research needed to answer the question regarding the onset of plate tectonics in a more definite way. For a more detailed resolution of the fault system and the geological units of the BGB, a much finer model discretization in 3D (than the one used in this thesis) using the MT BGB data set would be required. This poses to be difficult with the presently available computer resources due to the huge size of inverse problems for 3D models. However, with the rapid advances in computer technology and the growing understanding of inversion algorithms used for 2D and especially 3D analysis a more detailed investigation of large data sets such as the BGB MT data becomes more and more feasible. Based on the knowledge gained so far by the analysis of the MT data a few more MT sites located at selected positions (e.g. around the junction of the Komati and Inyoka fault) could further help to resolve the subsurface in greater detail. However, the extensive amount of EM noise sources in the Barberton area make data collection difficult. Longer recording times might be necessary, especially for obtaining long period MT TFs. Focusing on areas of interest (e.g. the fault system and borders between plutons and BGB rocks) with 3D inversions of MT data subsets should be a next step in future studies.
In the presented work, no topography was included in the inversion models. In the Barberton Mountain Land the altitude ranges from 600 to 1 800 m above sea level. Hence, some of the topmost layers in the inversion models might be affected by topography. However, since the topography varies more or less smoothly over the measurement area its effects are expected to be very small.
One of the main difficulties while interpreting the obtained electrical conductivity models of the BGB is the lack of recent geophysical studies in the area. Additional deep reaching measurements (e.g. seismics) would be helpful for imaging the inner structure of the BGB and answering the questions regarding its formation.

Acknowledgements

The presented thesis would not have been possible without the help and support of a number of people that I want to mention here. I sincerely apologize to anyone I forgot.

First and foremost, I am deeply grateful to my PhD supervisor Dr. Ute Weckmann. Her patient support, encouragement and our many constructive discussions helped me a great deal in order to go through with the thesis and to finally finish it. I am thankful for the opportunity to work in such an interesting and diverse project, to participate in many MT field campaigns which gave me an innumerable amount of invaluable experiences and to present and discuss my work at conferences which led me to places all over the world. Many thanks to you for your (endless) patience!

I would also like to express my thanks to Prof. Michael Weber for his support of the project, giving me the opportunity to work at the GFZ and his help with many organisational and administrative aspects that come along with a thesis.

Field work is not possible without many, many helpers. Many thanks to Ute Weckmann, Xiaoming Chen, Norbert Ebel, Diane Eydam, Reinhard Klose, Brady Lob, Scott McLennan, Stefan Rettig, Manfred Schüler, Paul Sass, Gerald Schmidt, Euston Witbooi and all the participants of the first field campaign at the BGB. Due to these people my very first field campaign was a very enjoyable experience and taught me a lot.
Special thanks also to Rod Green and Corné Maritz for keeping the reference sites running. Without that data I would have been totally lost!
Many thanks to all the tribal communities, farmers and rangers of the game reserves who gave us access to their land and made lots of adventurous car rides possible.
Abel Sihlongonyane's help was invaluable for getting the permitting for all the sites and to find some of the more remote locations. Many thanks also to the guards who kept people and animals likewise away from the MT sites.
Many thanks to the workers who accompanied us during the field work. Many holes for our equipment would not have been dug without them. Although, you had a strong competition in Manfred Schüler for digging holes.
A very warm thank you to the team of Travelport for the cordial hospitality, for keeping our bellies full and the braais.
Many thanks to the Geophysical Instrument Pool Potsdam (GIPP) for providing the instruments.
I am very grateful to Daniela Oestreich, without her, the equipment would have never been send to South Africa and I would have given up on the custom permits.

A big thank you to Dr. Karl Kappler who allowed me to use the first version of the *reference site filter*. Without his basic ideas for this filtering approach my data would still be noisy. Thanks also for the interesting discussions we had at every conference we met.

Many thanks to Dr. Xiaoming Chen for answering all my questions while we shared an office. I also really enjoyed pulling your car out of the only mud puddle in the Barberton area. The one and a half hour cross-country drive was really worth it.

Many thanks also to Dr. Paul Sass for allowing me to use his data selection programme for MT transfer functions and for the good times we had while we shared an office.

Without Dr. Gerard Muñoz I would have been lost in the world of data processing and I would have never started to write Powershell scripts. They made the processing so much easier and faster. Thank you!

Dr. Kristina Tietze was the unfortunate person who got to read most of the raw version of this thesis. Thanks to her this thesis became reader friendly. Thank you also for broadening my understanding of MT, data inversion and answering all my little questions.

Dr. Naser Meqbel answered all my questions regarding 3D inversion and kept the code and the 3D plotting tools running. Many thanks for that.

Special thanks to Dr. Gerhard Kapinos and the team who did the field measurements in Namibia for collecting and providing the excellent data for my synthetic noise studies. It is so easy to ruin almost perfectly clean data.

During my time as a PhD student, I was fortunate enough to meet Prof. Maarten de Wit several times and to discuss the BGB from a geological point of view. You have my sincere thanks for helping me understand this complex region and for making this project possible.

During his time at the GFZ Potsdam Scott MacLennan deepened my geological knowledge. Thank you also for explaining the geological structures and rock formations in the field.

I would like to express my thanks to all my present and former colleagues in the MT working group at the GFZ and the Free University Berlin, especially Dr. Oliver Ritter and Dr. Heinrich Brasse, for an amazing working atmosphere, the numerous suggestions, discussions when I presented parts of my research and new ways to look at my work.

Finally, my thanks go to the NSF and DST for their support of the MT field experiment and the participating students, to the DFG for funding the experiment, and the GFZ Potsdam for funding my PhD research.

List of Figures

List of Tables

Bibliography

[1] A. Adam, K. Landy, and Z. Nagy. New evidence for the distribution of electric conductivity in the Earth's crust and upper mantle in the Pannonin Basin as a 'hotspot'. *Tectonophysics*, 164: 361–368, 1989.

[2] A. Adam, L. Szarka, and J. Verö. Natural and Man-made EM Variations in the Komlo Coalfield. *Physics of the Earth and Planetary Interior*, 53:207–213, 1989.

[3] C. R. Anhaeusser. *Archaean gold mineralization in the Barberton Mountain Land. In: C. R. Anhaeusser (ed.) Mineral deposits of southern Africa*. Geol. Soc. S. Afr., 1, 113-154.

[4] C. R. Anhaeusser. *The stratigraphy, structure and gold mineralisation of the Jamestown and Sheba Hills areas of the Barberton Mountain Land*. PhD thesis, University Witwatersrand, Johannesburg, South Africa, 1969.

[5] C. R. Anhaeusser. The history of mining in the barberton greenstone belt, south africa, with an emphasis on gold (1868-2012). International Mining History Congress, Johannesburg, South Africa, 2012.

[6] R.A. Armstrong, W. Compston, M.J. de Wit, and I.S. Williams. The stratigraphy of the 3.5-3.2 ga barberton greenstone belt revisited: a single zircon ion microprobe study. *Earth and Planetary Science Letters*, 101:90–106, 1990.

[7] K. Bahr. Geological noise in magnetotelluric data: a classification of distortion types. *Physics of the Earth and Planetary Interior*, 66:24–38, 1991.

[8] Karsten Bahr. Interpretation of the magnetotelluric impedance tensor: regional induction and local telluric distortion. *J. Geophys*, 62:119–127, 1988.

[9] M. Becken and H. Burkhardt. An ellipticity criterion in magnetotelluric tensor analysis. *Geophysical Journal International*, 159:69–82, 2004.

[10] M. Becken, O. Ritter, P. A. Bedrosian, and U. Weckmann. Correlation between deep fluids, tremor and creep along the central San Andreas fault. *Nature*, 480:87–92, 2011.

[11] G. Bohlmann. Ein ausgleichungsproblem. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen*, Mathematisch-Physikalische Klasse:260–271, 1899.

[12] H. Brasse and A. Junge. The influence of geomagnetic variations on pipelines and an application for large-scale magnetotelluric depth sounding. *Journal of Geology*, 55:31–36, 1984.

[13] L. Cagniard. Basic theory of the magneto-telluric method of geophysical prospecting. *Geophysics*, 18:605–635, 1953.

[14] A. D. Chave and A. G. Jones. *The Magnetotelluric Method: Theory and Practice*. Cambridge University Press, 2012.

[15] X. Chen. Filterung von geophysikalischen Zeitreihen mit periodisch auftretenden multifrequenten Störsignalen. Master's thesis, Technical University, Berlin, 2008.

[16] J.D. Clemens, L.M. Yearron, and G. Stevens. Barberton (south africa) ttg magmas: Geochemical and experimental constraints on source-rock petrology, pressure of formation and tectonic setting. *Precambrian Research*, 151:53–78, 2006.

[17] A. S. Orange G. M. Hoversten Constable, S.C. and H. F. Morrison. Marine magnetotellurics for petroleum exploration Part I: A sea-floor equipment system. *Geophysics*, 63:816–825, 1998.

[18] B. W. Darracott. The interpretation of the gravity-anomaly over the barberton mountain land, south africa. *Trans. geol. Soc. South Afr.*, 78:123–128, 1975.

[19] G. J. K. Dawes. Feasibility study for a transputer-based upgrade of the Short-Period Automatic Magnetotelluric (S.P.A.M.) system. *Univ. of Edinburgh, N.E.R.C. report F3/G6/S43*, 1990.

[20] J.H. de Beer and E.H. Stettler. Geophysical characteristics of the southern african continental crust. *Journal of Petrology, Special Lithosphere Issue*, pages 163–184, 1988a.

[21] J.H. de Beer, E.H. Stettler, and J.G. du Plessis. The deep structure of the barberton greenstone belt: a geophysical study. *South Afr. J. Geol.*, 91(2):194–197, 1988b.

[22] C.E.J. de Ronde and M. de Wit. Tectonic history of the barberton greenstone belt, south africa: 490 million years of archean crustal evolution. *Tectonics*, 13(4):983–1005, 1994.

[23] C.E.J. de Ronde, S. Kamo, D.W. Davis, M.J. de Wit, and E.T.C. Spooner. Field, geochemical and u-pb isotopic constraints from hypabyssal felsic intrusions within the barberton greenstone belt, south africa: implications for tectonics and the timing of gold mineralization. *Precambrian Research*, 49:261–280, 1991.

[24] M. de Wit, C. Roering, R.J. Hart, R.A. Armstrong, C.E.J. de Ronde, R.W.E. Green, M. Tredoux, E. Peberdy, and R.A. Hart. Formation of an archean continent. *Nature*, 357:553–562, 1992.

[25] M. J. de Wit. Gliding and overthrust nappe tectonics in the barberton greenstone belt. *J. Struct. Geol.*, 4:117–136, 1982.

[26] M. J. de Wit, R. E. P. Fripp, and I. G. Stanistreet. Tectonic and stratigraphic implications from new field observations along the southern part of the barberton greenstone belt. *Spec. Publ. Geol. Soc. South Afr.*, 9:21–29, 1983.

[27] M. J. de Wit, H. Furnes, and B. Robins. Geology and tectonostratigraphy of the onverwacht suite, barberton greenstone belt, south africa. *Precambrian Research*, 186:1–27, 2011.

[28] M.J. de Wit, R.A. Hart, and R.J. Hart. The jamestown ophiolite complex, barberton mountain belt: a section through 3.5 ga oceanic crust. *Journal of African Earth Sciences*, 6(5):681–730, 1987.

[29] J. Diener, G. Stevens, A. Kisters, and M. Poujol. Metamorphism and exhumation of the basal parts of the barberton greenstone belt, south africa: Constraining the rates of the mesoarchaean tectonism. *Precambrian Reseach*, 143:87–112, 2005.

[30] P. H. G. M. Dirks, E. G. Charlesworth, M. R. Munyaib, and R. Wormalda. Stress analysis, post-orogenic extension and 3.01 Ga gold mineralisation in the Barberton Greenstone Belt, South Africa. *Precambrian Research*, 226:157–184, 2013.

[31] G. Egbert. Robust multiple-station magnetotelluric data processing. *Geophysical Journal International*, 130:475–496, 1997.

[32] G. D. Egbert and J. R. Booker. Very long period magnetotellurics at Tucson observatory: Implications for mantle conductivity. *Journal of Geophysical Research*, 97:15099, 1992.

[33] G. D. Egbert and A. Kelbert. Computational recipes for electromagnetic inverse problems. *GJI*, 189:251–267, 2012.

[34] C. G. Farquharson and J. A. Craven. Three-dimensional inversion of magnetotelluric data for mineral exploration: An example from the McArthur River uranium deposit, Saskatchewan, Canada. *Journal of Applied Geophysics*, 68:450–458, 2009.

[35] G. Fischer, L. Szarka, A. Adam, and J.T. Weaver. The magnetotelluric phase over 2-D structures. *Geophysical Journal International*, 108:778 – 786, 1992.

[36] Stephen F. Foley, Stephan Buhre, and Dorrit E. Jacob. Evolution of the Archaean crust by delamination and shallow subduction. *Nature*, 421:249–252, 2003.

[37] S. Fontes, T. Harinarayana, G. Dawes, and V. Hutton. Processing of Noisy Magnetotelluric Data Using Digital Filters and Additional Data Selection Criteria. *Physics of the Earth and Planetary Interior*, 52:30–40, 1988.

[38] T D Gamble, W M Goubeau, and J Clarke. Magnetotellurics with remote magnetic reference. *Geophysics*, 44(1):53–68, 1979.

[39] T.D. Gamble, W.M. Goubau, and J. Clarke. Error analysis for remote reference magnetotellurics. *Geophysics*, 44:959–968, 1979.

[40] W.M. Goubau, T.D. Gamble, and J. Clarke. Magnetotelluric data analysis: removal of bias. *Geophysics*, 43:1157–1166, 1978.

[41] A. L. Hall. The geology of the barberton gold mining district. *Memoir Geological Survey South Africa*, 9:347, 1918.

[42] W.B. Hamilton. Archean magmatism and deformation were not products of plate tectonics. *Precambrian Research*, 91:143–179, 1998.

[43] T.M. Harrison, A.K. Schmitt, M.T. McCulloch, and O.M. Lovera. Early (≥ 4.5 Ga) formation of terrestrial crust: Lu-Hf, δ^{18}O, and Ti thermometry results for Hadean zircons. *Earth and Planetary Science Letters*, 268:476–486, 2008.

[44] M. Hattingh. The Use of Data-adaptive Filtering for Noise Removal on Magnetotelluric Data. *Physics of the Earth and Planetary Interior*, 53:239–254, 1989.

[45] G. Heinson, A. White, and F. E. M. Lilley. Rifting of a passive margin and development of a lower-crustal detachment zone: Evidence from marine magnetotellurics. *GEOPHYSICAL RESEARCH LETTERS*, 32, 2005.

[46] W. Heise, G.T. Caldwell, and S.L. Bennie. Three-dimensional electrical resistivity image of magma beneath an active continental rift, taupo volcanic zone, new zealand. *Geophys. Res. Lett.*, 37, 2010.

[47] C. Heubeck and D. R. Lowe. Depositional and tectonic setting of the archean moodies group, barberton greenstone belt, south africa. *Precambrian Research*, 68:257–290, 1994a.

[48] C. Heubeck and D. R. Lowe. Late syndepositional deformation and detachment tectonics in the barberton greenstone belt, south africa. *Tectonics*, 13(6):1514–1536, 1994b.

[49] G. J. Hill, T. G. Caldwell, W. Heise, D. G. Chertkoff, H. M. Bibby, M. K. Burgess, J. P. Cull, and R. A. F. Cas. Distribution of melt beneath mount st helens and mount adams inferred from magnetotelluric data. *Nature Geoscience*, 2:785–790, 2009.

[50] M. Ikeda, S. Kato, N. Nishizaka, Y. Ohno, K. Matsuo, and M. Kishimoto. Magnetotelluric imaging of the Median Tectonic Line in western Shikoku, southwest Japan: Implications of the fault-related low-resistivity zone. *Tectonophysics*, 601:78–86, 2013.

[51] V. Iliceto and G. Santarato. On the interference of man-made EM fields in the magnetotelluric 'dead band'. *Geophysical Prospecting*, 1999.

[52] H. Jödicke, J. H. Kruhl, C. Ballhaus, P. Giese, and J. Untiedt. Syngenetic, thin graphite-rich horizons in lower crustal rocks from the Serre San Bruno, Calabria (Italy), and implications for the nature of high-conducting deep crustal layers. *Physics of the Earth and Planetary Interior*, 141:37–58, 2004.

[53] A. Junge. Characterization of and Correction for Cultural Noise. *Surveys in Geophysics*, 17(4): 361–391, 1996.

[54] T. Kakegawa and H. Ohmoto. Sulfur isotope evidence for the origin of 3.4 to 3.1 ga pyrite at the princeton gold mine, barberton greenstone belt, south africa. *Precambrian Research*, 96:209–224, 1999.

[55] S.L. Kamo and D.W. Davis. Reassessment of archean crustal development in the barberton mountain land, south africa, base on u-pb dating. *Tectonics*, 13(1):167–192, 1994.

[56] K.N. Kappler. A data variance technique for automated despiking of magnetotelluric data with remote reference. *Geophysical Prospecting*, 60:179–191, 2012.

[57] A.F.M. Kisters and C.R. Anhaeusser. Emplacement features of archaean ttg plutons along the southern margin of the barberton greenstone belt, south africa. *Precambrian Research*, 75:1–15, 1995.

[58] A.F.M. Kisters, G. Stevens, A. Dziggel, and R.A. Armstrong. Extensional detachment faultig and core-complex form in the south barberton granite-greenstone terrain, south africa: evidence for a 3.2 ga orogenic collapse. *Precambrian Research*, 127:355–378, 2003.

[59] A. H. Knoll and E. S. Barghoorn. Ancient microfossils showing cell division from the swaziland system of south africa. *Science*, 198:396–398, 1977.

[60] M. J. Van Kranendonk. Cool greenstone drips and the role of partial convective overturn in barberton greenstone belt evolution. *Journal of African Earth Sciences*, 60:346–352, 2011.

[61] Thomas Krings. The Influence of Robust Statistics, Remote Reference, and Horizontal Magnetic Transfer Functions on Data Processing in Magnetotellurics. Master's thesis, Inst. fr Geophysik, Westflische Wilhelms-Universitt Mnster, Germany, 2007.

[62] S. Kütter. Three-dimensional finite element simulation of magnetotelluric fields incorporating digital elevation models. Master's thesis, Institute of Geophysics, TU Bergakademie Freiberg, 2009.

[63] S. Kütter and U. Weckmann. Magnetotelluric measurements across the southern barberton greenstone belt, south africa: data improving strategies. In R.-U. Börner and K. Schwalenberg, editors, *Proceedings of the 24th Schmucker-Weidelt-Colloquium for Electromagnetic Depth Research*, pages 146–152, 2012.

[64] C. Lana, A. Kisters, and G. Stevens. Exhumation of mesoarchean ttg gneisses from the middle crust: Insights from the steynsdorp core complex, barberton granitoid-greenstone terrain, south africa. *Geological Society of America Bulletin*, 122(1-2):183–197, 2010.

[65] C. Lana, I. Buick, G. Stevens, R. Rossouw, and W. De Wet. 3230e3200 ma post-orogenic extension and mid-crustal magmatism along the southeastern margin of the barberton greenstone belt, south africa. *Journal of Structural Geology*, 33:844–858, 2011.

[66] J. Larsen, R. Mackie, A. Fiordelisi, A. Manzella, and S.A. Rieven. Robust processing for removing train signals from magnetotelluric data in central Italy. *Proceedings World Geothermal Congress*, 1995.

[67] J.C. Larsen. Transfer functions: smooth robust estimates by least-squares and remote reference methods. *Geophysical Journal International*, 99:645–663, 1989.

[68] J.C. Larsen, R.L. Mackie, A. Manzella, A. Fiordelisi, and S. Rieven. Robust smooth magnetotelluric transfer functions. *Geophysical Journal International*, 124:801–819, 1996.

[69] D. R. Lowe. Accretionary history of the archean barberton greenstone belt (3.55-3.22 ga), southern africa. *Geology*, 22:1099–1102, 1994.

[70] S. MacLennan. Geological and geophysical Characteristics of an Archean suture zone, Barberton Greenstone Belt, South Africa. Master's thesis, Nelson Mandela Metropolitan University, South Africa, 2012.

[71] N. M. Meqbel. *The electrical conductivity structure of the Dead Sea Basin derived from 2D and 3D inversion of magnetotelluric data*. PhD thesis, Free University Berlin, 2009.

[72] N. M. M. Meqbel, O. Ritter, and DESIRE Group. A magnetotelluric transect across the Dead Sea Basin: electrical properties of geological and hydrological units of the upper crust. *Geophysical Journal International*, 193:1415–1431, 2013.

[73] J. Moyen, G. Stevens, and A. Kisters. Record of mid-archean subduction from metamorphism in the barberton terrain, south africa. *Nature*, 443(3):559–562, 2006.

[74] J.-F. Moyen, G. Stevens, A. F. M. Kisters, and R. W. Belcher. *TTG Plutons of the Barberton Granitoid-Greenstone Terrain, South Africa*, volume 15, chapter 5.6. Elsevier, 2007.

[75] G. Muñoz. Exploring for Geothermal Resources with Electromagnetic Methods. *Surveys in Geophysics*, pages 1–22, 2013.

[76] W.U. Mueller, R. Daigneault, J.K. Mortensen, and E.H. Chown. Archean terrane docking: upper crust collision tectonics, abitibi greenstone belt, quebec, canada. *Tectonophysic*, 265:127–150, 1996.

[77] D. Nascimento. *Magnetotelluric Instrument Development and Application*. PhD thesis, 1997.

[78] National Geophysical Data Center. Magnetic declination calculator, last retrieved: 08.02.2015. URL `http://www.ngdc.noaa.gov/geomag-web/`.

[79] K. Neuendorf, J. P. Mehl, Jr., and J. A. Jackson. *Glossary of Geology*. American Geological Institute, 5th edition, 2005.

[80] G. Oettinger, V. Haak, and J.C. Larsen. Noise reduction in magnetotelluric time-series with a new signal-noise separation method and its application to a field experiment in the Saxonian Granulite Massif. *Geophysical Journal International*, 146:659–669, 2001.

[81] S.W. Parman, J.C. Dann, T.L. Grove, and M.J. de Wit. Emplacement conditions of komatiite magmas from the 3.49 ga komati formation, barberton greenstone belt, south africa. *Earth and Planetary Science Letters*, 150:303–323, 1997.

[82] P.K. Patro and G.D. Egbert. Regional conductivity structure of Cascadia: Preliminary results from 3D inversion of USArray transportable array magnetotelluric data. *Geophys. Res. Lett.*, 35(20): 1–5, 2008.

[83] M. Poujol, L.J. Robb, C.R. Anhaeusser, and B. Gericke. A review of the geochronological constraints on the evolution of the kaapvaal craton, south africa. *Precambrian Research*, 127:181–213, 2003.

[84] W. Qian and L. B. Pedersen. Industrial interference magnetotellurics: an example from the Tangshan area, China. *Geophysics*, 56:265–273, 1991.

[85] O. Ritter. *An audiomagnetotelluric investigation of the Southern Upland Fault: novel instrumentation, field procedures and 3D modelling*. PhD thesis, 1995.

[86] O. Ritter and G. Dawes. A transputer-based multi-station multi-technique geophysical data acquisition system – S.P.A.M. MkIII. *Protokoll über das Kolloquium "Elektromagnetische Tiefenforschung"*, Borkheide, 1992.

[87] O. Ritter, A. Junge, and G.J.K. Dawes. New equipment and processing for magnetotelluric remote reference observations. *Geophysical Journal International*, 132:535–548, 1998.

[88] O. Ritter, U. Weckmann, T. Vietor, and V. Haak. A magnetotelluric study of the Damara Belt in Namibia 1. regional scale conductivity anomalies. *Physics of the Earth and Planetary Interior*, 138:71–90, doi:10.1016/S0031–9201(03)00078–5, 2003.

[89] O. Ritter, A. Hoffmann-Rothe, P. A. Bedrosian, U. Weckmann, and V. Haak. Electrical conductivity images of active and fossil fault zones. In *in High-Strain Zones: Structure and Physical Properties, Bruhn, D. and Burlini, L. (eds.)*, volume 245, pages 165–186. Geological Society of London Special Publications, 2005.

[90] W. Rodi and R. L. Mackie. Nonlinear conjugate gradients algorithm for 2D magnetotelluric inversion. *Geophysics*, 66:174–187, 2001.

[91] C.E.J. De Ronde and S.L. Kamo. An archaean arc-arc collisional event: a short-lived (ca 3 myr) episode, weltevreden area, barberton greenstone belt, south africa. *Journal of African Earth Sciences*, 30(2):219–248, 2000.

[92] Y. Sasaki. Full 3-d inversion of electromagnetic data on pc. *J. Appl. Geophys.*, 46(1):45–54, 2001.

[93] E. T. Schill, H. Kohl, Deckert, and et al. EGS Technologie Entwicklung: Risikominimierung bei der Exploration. Technical report, Technische Informationsbibliothek Hannover (TIB), 2011. 4.1, 84-107.

[94] B. Schoene, S.A. Bowring, and M.J. de Wit. The 3.2-3.1 ga tectonothermal history of the barberton greenstone belt.: the importance of crustal extension in cratonisation. *T*, in revision, 2008a.

[95] B. Schoene, M. J. de Wit, and S. A. Bowring. Mesoarchean assembly and stabilization of the eastern kaapvaal craton: A structural-thermochronological perspective. *Tectonics*, 27, 2008b.

[96] B. Schoene, F. O. L. Dudas, S. A. Bowring, and M. de Wit. Smnd isotopic mapping of lithospheric growth and stabilization in the eastern kaapvaal craton. *Terra Nova*, 21(3):219–228, 2009.

[97] Blair Schoene and Samuel A. Bowring. Determining accurate temperature-time paths from u-pb thermochronology: An example from the kaapvaal craton, southern africa. *Geochimica et Cosmochimica Acta*, 71(1):165–185, 2007.

[98] S. Schröder, J. P. Lacassie, and N. J. Beukes. Stratigraphic and geochemical framework of the agouron drill cores, transvaal supergroup (neoarcheanpaleoproterozoic, south africa). *South African Journal of Geology*, 2006.

[99] Shalivahan and B. B. Bhattacharya. How remote can the far remote reference site for magnetotelluric measurements be? *JGR*, 107(B6, 2105), 2002.

[100] W.E. Sims, F.X. Bostick, and H.W. Smith. The estimation of magnetotelluric impedance tensor elements from measured data. *Geophysics*, 36:938–942, 1971.

[101] W. Siripunvaraporn, G. Egbert, Y. Lenbury, and M. Uyeshima. Three-Dimensional Magnetotelluric Inversion: Data Space Method. *Physics of the Earth and Planetary Interior*, 150:3–14, 2004.

[102] R.H. Smithies, M.J. Van Kranendonk, and D.C. Champion. The mesoarchean emergence of modern-style subduction. *Gondwana Research*, 11(1-2):50–68, 2007.

[103] C. Swift. *A magnetotelluric investigation of an electrical conductivity anomaly in the southwestern United States*. Ph.D. Thesis, M.I.T. Cambridge, Mass., 1967.

[104] Laszlo Szarka. Geophysical Aspects of Man-made Electromagnetic Noise in the Earth - A Review. *Surveys in Geophysics*, 9:287–318, 1988.

[105] W.M. Telford, L.P. Geldart, and R.E. Sheriff. *Applied Geophysics*. Cambridge University Press, Cambridge, 1990.

[106] K. Tietze and O. Ritter. 3D magnetotelluric inversion in practice - the electrical conductivity structure of the San Andreas Fault in Central California. *Geophysical Journal International*, 195: 130–147, 2013.

[107] A. N. Tikhonov. On the determination of electric characteristics of deep layers of the Earth's crust. *Dokl. Akad. Nauk SSSR*, 73:295–297, 1950.

[108] A. Trnkoczy. *New Manual of Seismological Observatory Practice (NMSOP)*, chapter Understanding and parameter setting of STA/LTA trigger algorithm. IASPEI, GFZ German Research Centre for Geosciences, Potsdam, 2002.

[109] T. Uchida, Y. Song, T.J. Lee, Y. Mitsuhata, S.-K. Lim, and S.K. Lee. Magnetotelluric Survey in an Extremely Noisy Environment at the Pohang Low-Enthalpy Geothermal Area, Korea. *Proceedings World Geothermal Congress*, 2005.

[110] M. Unsworth, W. Wenbo, A. G. Jones, S. Li, P. Bedrosian, J. Booker, J. Sheng, D. Ming, and T. Handong. Crustal and upper mantle structure of northern tibet imaged with magnetotelluric data. *Journal of Geophysical Research*, 109, 2004.

[111] M. van Kranendonk, A. Kroner, E. Hegner, and J. Connelly. Age, lithology and structural evolution of the c. 3.53 ga theespruit formation in the tjakastad area, southwestern barberton greenstone belt, south africa, with implications for archaean tectonics. *Chemical Geology*, 261:115–139, 2009.

[112] M. J. Van Kranendonk, R. H. Smithies, A. H. Hickman, and D. C. Champion. Review: secular tectonic evolution of Archean continental crust: interplay between horizontal and vertical processes in the formation of the Pilbara Craton, Australia. *Terra Nova*, 19:1–38, 2007a.

[113] J.S.V. Van Zijl. The relationship between the deep electrical resistivity structure and tectonic provinces in southern africa. *Trans. Geol. Soc. S. Afr.*, 81:129–142, 1978.

[114] M.J. Viljoen and R.P. Viljoen. An introduction to the geology of the barberton granite-greenstone terrain. *Spec. Publ. Geol. Soc. S. Afr.*, 2:9–27, 1969.

[115] K. Vozoff. The magnetotelluric method in the exploration of sedimentary basins. *Geophysics*, 37: 98–141, 1972.

[116] J.H.W. Ward. Geology and metallogeny of the barberton greenstone belt: a survey. *Journal of African Earth Sciences*, 21(2):213–240, 1995.

[117] U. Weckmann. Making and breaking of a continent: Following the scent of geodynamic imprints on the african continent using electromagnetics. *SG*, 33:107–134, 2012.

[118] U. Weckmann, A. Magunia, and O. Ritter. Effective noise separation for magnetotelluric single site data processing using a frequency domain selection scheme. *Geophys. J. Int.*, 161:635–652, 2005.

[119] U. Weckmann, A. Magunia, and O. Ritter. Effective noise separation for magnetotelluric single site data processing using a frequency domain selection scheme. *Geophysical Journal International*, 161(3):635–652, doi:10.1111/j.1365–246X.2005.02621.x, 2005.

[120] U. Weckmann, A. Jung, T. Branch, and O. Ritter. Comparison of electrical conductivity structures and 2D magnetic model along two profiles crossing the Beattie Magnetic Anomaly, South Africa. *South Afr. J. Geol.*, 110:449–464, 2007.

[121] Ute Weckmann, Oliver Ritter, and Volker Haak. A magnetotelluric study of the Damara Belt in Namibia 2. MT phases over 90° reveal the internal structure of the Waterberg Fault / Omaruru Lineament. *Physics of the Earth and Planetary Interior*, 138:91–112, 2003.

[122] P. Weidelt and P. Kaikkonen. Local 1-D interpretation of magnetotelluric B- polarization impedances. *Geophysical Journal International*, 117:733–748, 1994.

[123] H. L. Weinert. Efficient computation for whittaker-henderson smoothing. *Computational Statistics & Data Analysis*, 52:959–974, 2007.

[124] K. A. Weitemeyer, S. Constable, and A. M. Trehu. A marine electromagnetic survey to detect gas hydrate at Hydrate Ridge, Oregon. *Geophysical Journal International*, 187:45–62, 2011.

[125] J.D. Westraat, A.F.M. Kisters, M. Poujol, and G. Stevens. Transcurrent shearing, granite sheeting and the incremental construction of the tabular 3.1 Ga Mpuluzi batholith, Barberton granitegreenstone terrane, South Africa. *Journal of the Geological Society*, 162:373–388, 2005.

[126] E. T. Whittaker. On a new method of graduation. *Proceedings of the Edinburgh Mathematical Society*, 41:63–75, 1923.

[127] N. Wiener. Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications. *Journal of the American Statistical Association*, 47(258), 1949.

[128] H Wiese. Die Streichrichtung der Untergrundstrukturen des elektrischen Widerstandes, erschlossen aus geomagnetischen Variationen. *Pure and Applied Geophysics*, 52:83–103, 1962.

[129] Wikipedia. Kaapvaal craton — wikipedia, the free encyclopedia, 2014. URL `http://en.wikipedia.org/w/index.php?title=Kaapvaal_Craton&oldid=637458876`. [Online; accessed 29-March-2015].

[130] L. Yearron, J. Clemens, G. Stevens, and C. Anhaeusser. Geochemistry and petrogenesis of the granitoids of the Barberton Mountain Land, South Africa. EGS - AGU - EUG Joint Assembly, abstract No. 2639, 2003.

Appendix A

Basics of the Magnetotelluric Method

Magnetotellurics (MT) is a geophysical (exploration) method for imaging the electrical conductivity structure of the subsurface. It uses natural electromagnetic (EM) field variations (primary field) as its source which induce electric and magnetic fields (secondary fields) in the earth. The properties of the secondary fields depend on the electrical conductivity distribution of the subsurface.
With decreasing frequency, the EM fields penetrate deeper into the Earth. Therefore, high frequency MT data carry information about shallow structures (small induction space) and low frequencies contain information of the deeper subsurface as well (larger induction space). Since EM fields spread through the entire volume at the investigation area, the measured MT data reflects a bulk conductivity of the subsurface and has a limited spatial resolution capacity of small-scale structures. The resolution capacity strongly depends on the density of the MT site coverage. Hence, comprehension of this bulk conductivity is essential for the interpretation of MT data.
Many factors influence conductivity, e.g. presence of conductive constituents in the rock matrix (granite: $\sim 10^{-6}$ S/m vs. graphite: $\sim 10^{3}$ S/m), porosity, pore content (meteoric water: $\sim 10^{-2}$ S/m, saline water: 10^{2} S/m; values e.g. in Telford et al., 1990) and interconnectivity of pores and conductive materials. Especially, in geothermal or volcanic areas the temperature plays also an important role.
The MT method has been employed in many research areas. Among them are the investigation of geothermal fields (review paper by Muñoz, 2013, and references therein), of natural deposits and reservoirs (e.g. petroleum: Constable and Morrison, 1998; gas hydrates: Weitemeyer et al., 2011; ore deposits: Farquharson and Craven, 2009) and of geological settings and tectonics (e.g. Weckmann, 2012; Hill et al., 2009; Becken et al., 2011; Ikeda et al., 2013) as will be demonstrated in this thesis. Depending on the geological target, investigation depths range from the Earth's crust down to the upper mantle (e.g. Adam et al., 1989; Unsworth et al., 2004; Meqbel et al., 2013).
For all investigations, an understanding of the MT method is necessary. Therefore, this section gives an introduction to the method, how data is collected and processed. More extensive descriptions of the MT method can be found, e.g. in Chave and Jones (2012), which was used for the summary presented here, as well.

The magnetotelluric (MT) method was first introduced and developed by the independent works of Cagniard (1953) and Tikhonov (1950). They describe the usage of the naturally occuring electromagnetic fields for investigating the electrical conductivity distribution of the subsurface. The source fields originate from variations in the geomagnetic field due to its interaction with the solar wind and thunder

activity (especially in the tropical regions) causing random fluctuations. This leads to a wide range of frequencies of the source signals between 10^{-5} and 10^5 Hz. At the transition between the two mechanisms for the source field generation (solar wind - ionosphere interaction and thunderstorm activity) at ~ 1 Hz a source signal minimum occurs (Iliceto and Santarato, 1999). This leads to a low signal to noise ratio and a general worse data quality around this frequency.
The MT method has undergone major developments, considering at first only simple cases of a 1D structured earth (e.g. Vozoff, 1972), e.g. a homogeneous or a layered substratum, but extending over time to 2D and 3D assumptions for the underground (e.g. Patro and Egbert, 2008; cf. 3.3.1 and 3.3.2, respectively).

A.1 Theoretical basics

The basis of electromagnetic processes is summarized in Maxwell's equations.

$$\begin{aligned} \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \quad \text{(a)} & \nabla \times \mathbf{H} &= \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \quad \text{(c)} \\ \nabla \cdot \mathbf{D} &= q \quad \text{(b)} & \nabla \cdot \mathbf{B} &= 0 \quad \text{(d)} \end{aligned} \tag{A.1}$$

where $\mathbf{E}$ is the electric field, $\mathbf{H}$ is the magnetic field, $\mathbf{B}$ is the magnetic flux density, $\mathbf{D}$ is the electric displacement current, $\mathbf{J}$ is the induced current density, q is the charge density and t is the time.
This set of partial differential equations describes the relationship between electric and magnetic fields. They can be divided into two groups. The first comprising the rotational properties of the fields (curl operator $\nabla\times$): any time varying magnetic field induces a variation in the electric field (Faraday's law, eq. (A.1a)) and every electrical current and time varying electrical field generates a magnetic field (Ampre's law, eq. (A.1c)). The second group describes the sources of the fields (divergence operator $\nabla\cdot$): every distribution of electric charge results in an electric field (Gauss's law, eq. (A.1b)) and that there are no magnetic monopoles (Gauss's law for magnetism, eq. (A.1d)). They are not independent equations as the divergence equations ((A.1b) and (A.1d)) can be derived from the curl equations ((A.1a) and (A.1c)).
For homogeneous isotropic media, the constitutive equations and Ohm's law (A.2) are used, introducing the permeability of free space μ_0, the relative magnetic permeability μ_r of a medium, the permittivity of free space ε_0 and the relative dielectric permittivity ε_r. The key parameter of MT included within Ohm's law is the electrical conductivity σ. Its reciprocal value, the electrical resistivity ρ is often used, as well. Equations (A.2) help to relate the material properties, e.g. of the subsurface, to the electromagnetic fields.

$$\mathbf{B} = \mu_0\mu_r\mathbf{H} \qquad \mathbf{D} = \varepsilon_0\varepsilon_r\mathbf{E} \qquad \mathbf{J} = \sigma\mathbf{E} \tag{A.2}$$

For MT, several assumptions can be made which simplify the equations just presented and facilitate the analysis of the data considerably. Since the electromagnetic fields serving as sources for MT are located far from the measurement positions, they can be viewed as uniform and plane-polarized with an angle of incidence at the surface of about 90^o. This holds for most regions except at polar and equatorial latitudes and allows us to formulate the fields as

$$\mathbf{B} = \mathbf{B}_0 e^{i\omega t} \quad \text{and} \quad \mathbf{E} = \mathbf{E}_0 e^{i\omega t}. \tag{A.3}$$

$\mathbf{B}_0$ and $\mathbf{E}_0$ are the amplitudes of the fields at the origin and ω is the angular frequency. These fields are time-dependent and quasi-static.

For frequencies between 10^4 and 10^{-4} Hz used generally in MT, the variations of the electric displacement currents with time are negligible and (A.1c) becomes

$$\nabla \times \mathbf{H} = \mathbf{J}. \tag{A.4}$$

Within the constitutive equations (A.2) the variations of the material properties ε_r and μ_r can be neglected in MT yielding

$$\nabla \cdot \mathbf{E} = \frac{q}{\varepsilon_0} \tag{A.5}$$

for equation (A.1b).
Within a homogeneous half-space or layered subsurface, no free electric charges accumulate (i.e. source-free medium), therefore, $q = 0$, leading to

$$\nabla \cdot \mathbf{J} = 0 \quad \text{and} \quad \nabla \cdot \mathbf{E} = 0. \tag{A.6}$$

However, this assumption is violated if there are lateral inhomogeneities. In this case, charges gather at the lateral interfaces generating additional electromagnetic fields (galvanic distortion and static shift effects, cf. sec. 2).
Starting from Maxwell's equations and considering the mentioned assumptions, leads to the diffusion equations (derived in appendix A.2)

$$\nabla^2 \mathbf{A} = i\omega\mu_0\sigma\mathbf{A}, \qquad \mathbf{A} = \mathbf{H}, \mathbf{E}. \tag{A.7}$$

They describe the propagation of electromagnetic fields and their attenuation with depth in a conductive media.

An important parameter for the interpretation of MT data is the impedance tensor $\mathbf{Z}$ which relates the horizontal components of the magnetic to the electric field.

$$\mathbf{E} = \mathbf{Z}\,\mathbf{H} \tag{A.8}$$

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{pmatrix} \begin{pmatrix} H_x \\ H_y \end{pmatrix} \tag{A.9}$$

The form of the impedance tensor reflects the dimension of the subsurface structure. In 1D, the tensor is reduced to a scalar as the off diagonal elements (Z_{xy}, Z_{yx}) only have opposite signs and the main diagonal elements (Z_{xx}, Z_{yy}) are zero. For ideal 2D structures, Z_{xx} and Z_{yy} are zero as well but Z_{xy} and Z_{yx} are independent of each other. In the most general case of a 3D structured subsurface, $\mathbf{Z}$ is a full tensor where all elements are independent.
Usually the coordinate system in which the EM fields are measured is not aligned with the regional geo-electric strike θ so that even in a 2D case all impedance components exist. However, applying the rotation matrix $\mathbf{R}$

$$\mathbf{R} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}, \tag{A.10}$$

$\mathbf{Z}$ can be rotated in such a way that the main diagonal elements get minimized (ideally becoming zero in 2D). This rotation results in a 90^o ambiguity of the strike direction. Only by considering additional information of the geological setting, a differentiation between the actual strike and the direction perpendicular to it becomes possible.

In 2D, equations (A.1a) and (A.4) can be decoupled into separate systems for the so-called E- and B-polarization or TE and TM mode, respectively. For E-polarization the electric field component $E_{\parallel}$ parallel to the strike direction induces magnetic field components $H_{\perp}$ perpendicular to strike and H_z in vertical direction. The other components of the EM fields are zero. Accordingly, only $H_{\parallel}$, $E_{\perp}$ and E_z exist for B-polarization. If x is the strike direction, the xy component of $\mathbf{Z}$ is consistent with E-polarization and the yx component is consistent with B-polarization.
Since $\mathbf{Z}$ consists of complex numbers they have an absolute value and a phase. They are defined by

$$\rho_{ij} = \frac{1}{\omega\mu} \mid Z_{ij} \mid^2 \tag{A.11}$$

which is also called apparent electrical resistivity and the phase

$$\varphi_{ij} = arg(Z_{ij}) \tag{A.12}$$

with $i, j = x, y$.
Another important parameter for interpreting MT data are induction vectors which relate the horizontal to the vertical magnetic field.

$$H_z = \begin{pmatrix} T_x & T_y \end{pmatrix} \begin{pmatrix} H_x \\ H_y \end{pmatrix} \tag{A.13}$$

T_x and T_y are complex values for which amplitude and an angle can be calculated. In this work, induction vectors are calculated in Wiese-Convention (Wiese, 1962), hence, they point away from conductive bodies. $\mathbf{T} = \begin{pmatrix} T_x & T_y \end{pmatrix}^T$ is also called vertical magnetic transfer function.
Since a vertical magnetic field is only observed in the presence of lateral conductivity contrasts, induction vectors are non-existent in 1D and are indicators for changes in the lateral conductivity in 2D and 3D.
The presented parameters (equations (A.11), (A.12) and (A.13)) are all dependent on the frequency of the EM fields. Thus, they also contain a certain depth information since with decreasing frequency the penetration depth of the fields increases (cf. skin depth appendix A.2). Therefore, they are usually plotted against frequency or its inverse, the period.

A.2 Diffusion equation and skin depth

The diffusion equations describe the propagation and attenuation of the electromagnetic fields within a conductive media. They can be derived from the Maxwell's equations (A.1) and the constitutive equations (A.2) by applying the assumptions made in MT summarized in section A.1. Starting from

$$\nabla \times \mathbf{H} = \sigma \mathbf{E} \quad \text{and} \quad \nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t} \tag{A.14}$$

and taking the curl of these equations leads to

$$\nabla \times (\nabla \times \mathbf{H}) = \sigma (\nabla \times \mathbf{E}) \quad \text{and} \quad \nabla \times (\nabla \times \mathbf{E}) = -\mu_0 \frac{\partial}{\partial t} (\nabla \times \mathbf{H}) . \tag{A.15}$$

Applying the Laplacian operator defined by

$$\nabla^2 \mathbf{A} = \nabla (\nabla \cdot \mathbf{A}) - \nabla \times (\nabla \times \mathbf{A}) \tag{A.16}$$

valid for all vector fields $\mathbf{A}$ yields

$$\nabla^2 \mathbf{H} = -\sigma (\nabla \times \mathbf{E}) + \nabla (\nabla \cdot \mathbf{H}) \quad \text{and} \quad \nabla^2 \mathbf{E} = \mu_0 \frac{\partial}{\partial t} (\nabla \times \mathbf{H}) + \nabla (\nabla \cdot \mathbf{E}) . \tag{A.17}$$

Using equations (A.14), (A.1d) and (A.6) results in

$$\nabla^2\mathbf{H} = \mu_0\sigma\frac{\partial\mathbf{H}}{\partial t} \quad \text{and} \quad \nabla^2\mathbf{E} = \mu_0\sigma\frac{\partial\mathbf{E}}{\partial t}. \tag{A.18}$$

Considering equations (A.3) which represent plane-polarized fields with sinusoidal time variations, derivatives of time become a simple multiplication of the fields with $i\omega$. Thus yielding the diffusion equations

$$\nabla^2\mathbf{H} = i\omega\mu_0\sigma\mathbf{H} \quad \text{and} \quad \nabla^2\mathbf{E} = i\omega\mu_0\sigma\mathbf{E}. \tag{A.19}$$

Equations (A.19) have solutions in the form of

$$\mathbf{H}(z) = \mathbf{H}_0e^{-kz} + \mathbf{H}_1e^{kz} \quad \text{and} \quad \mathbf{E}(z) = \mathbf{E}_0e^{-kz} + \mathbf{E}_1e^{kz} \tag{A.20}$$

for a homogeneous half-space. $\mathbf{H}_0, \mathbf{H}_1$ and $\mathbf{E}_0, \mathbf{E}_1$ are the amplitudes of the fields at the surface, z is the depth and k is the complex wave number.

$$k = \sqrt{i\omega\mu_0\sigma} = (1+i)\sqrt{\frac{\omega\mu_0\sigma}{2}} \tag{A.21}$$

Since the subsurface is assumed to be source-free the fields vanish for $z \to \infty$, hence, $\mathbf{H}_1$ and $\mathbf{E}_1$ must be zero. The real part of e^{-kz} (cf. eq. (A.21)) describes a downward travelling attenuated wave. The depth δ where the fields are reduced to $1/e$ of their initial values is called skin depth and defined as

$$\mathbf{H}(\delta) = \mathbf{H}_0e^{-1} = \mathbf{H}_0e^{-\delta\sqrt{\frac{\omega\mu_0\sigma}{2}}} \quad , \quad \mathbf{E}(\delta) = \mathbf{E}_0e^{-1} = \mathbf{E}_0e^{-\delta\sqrt{\frac{\omega\mu_0\sigma}{2}}} \quad , \tag{A.22}$$

$$\delta \approx 503\sqrt{\frac{\rho}{f}}. \tag{A.23}$$

A.3 MT data measurements

For the MT method, variations of natural magnetic (H) and electric (E) fields are measured all over the area of interest. Site spacing is chosen depending on the geological target and setting and sites are either aligned along a profile perpendicular to the target's strike (cf. sec. A.1 eq. (A.10) and sec. 3.3.1) or in an array.

To measure all three components of the magnetic field either a fluxgate magnetometer or three induction coils are used. In this study, induction coils were used for the measurements in the period range between 10^{-3} and 10^{-3} s (Fig. A.1, left). The induction coils need to be aligned orthogonal to each other with two horizontal (Hx and Hy) and one vertical (Hz) coil. In general, the horizontal coils can be arbitrarily orientated. However, to facilitate the site setup and data analysis all sites are setup with the Hx coil in north-south and the Hy coil in east-west direction. The horizontal coils are buried in a leveled position in 30 to 40 cm deep holes, whereas the hole for a vertical coil should have such a depth that only the topmost 10 to 20 cm of the coil are above ground. Burying the equipment is necessary to avoid any motion of the coils due to larger animals or wind, and therefore, to minimize the noise induced by their motion within Earth's magnetic field. For the same reason, cables connecting each device should be kept as close to the ground as possible and should even be fixed by a stone or soil for the vertical coil. Additionaly, burying the coils minimizes influences on them due to the daily temperature changes.

To measure the electric field non-polarizable electrodes, e.g. silver-silver chloride (Ag/AgCl) electrodes,

are used. To ensure a good contact to the ground the electrode membrane at the bottom is packed in with wet bentonite and the whole electrode is buried as well. Four electrodes aligned north-south (Ex) and east-west (Ey) measure the two horizontal components of the electric field. A fifth electrode in the centre of this cross is used for grounding the instrument. The spacing between the two electrodes is limited by the length of the cable which was in the presented measurements up to $\sim 60\,\mathrm{m}$.
Electrodes and coils are connected to a sensor box (Sbx), which contains electronic filters and is used for amplifying the incoming signals. Another set of cables connects the Sbx to the data logger which converts the continuous analog to a discrete digital signal and records the data according to a set sampling frequency (Fig. A.1, right). Depending on the data logger, data is recorded in different bands using different sampling frequencies. Connected with the data logger is a GPS antenna used for exact time measurements and a battery. To reduce the EM noise due to the electronic equipment, the data logger and the battery should be kept as far away from the measuring devices as possible.
Based on the target depth and the required period range needed for its investigation, the sites record data for at least a few days.
Measurements at different sites do not need to be carried out simultaneously since the Earth's conductivity distribution only changes slowly over a geological time scale. Therefore, a roll along scheme measuring only a few sites at the same time before moving the equipment to the next locations while gradually covering the target area is possible. Hence, for inversion using data from measurements carried out in different years is feasible as well (data inversion, cf. chapter 3).

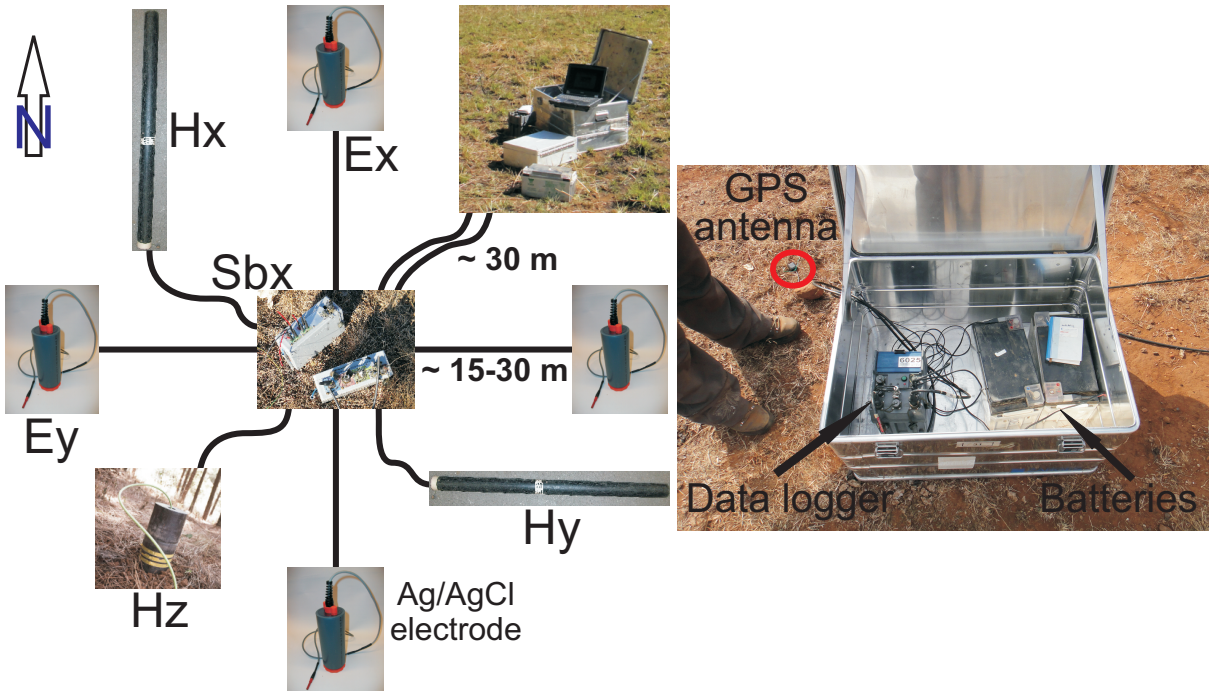


Figure A.1: Schematic setup of a magnetotelluric site (left, view from above, Sbx: Sensor box) with induction coils measuring the magnetic field components (Hx, Hy and Hz) and Ag/AgCl electrodes measuring the electric field components (Ex and Ey, pictures from Gerald Schmidt and Gregor Willkommen). View into the box with data logger, GPS antenna and batteries (right).

A.4 MT data processing

In order to extract the desired information about the electrical conductivity distribution of the subsurface from the recorded magnetic and electric time series, data processing is necessary. The data is converted from time to frequency domain and apparent resistivity ρ_a, phase φ and vertical magnetic transfer functions **T** (cf. sec. A.1) are calculated. Therefore, the linear bivariate equations for the impedance tensor components (A.9) and the magnetic transfer functions (A.13) need to be solved. The transfer functions can already, even before data inversion, give a first idea regarding the conductivity structure underneath a specific site (cf. sec. A.1 and appx. A.2).
The measured time series of the magnetic and electric field components are processed applying the *EMERALD* (**E**lectro-**M**agnetic **E**quipment, **R**aw-data **A**nd **L**ocations **D**atabase) processing routines first implemented by Ritter et al. (1998) and further developed by Weckmann et al. (2005). The processing comprises in general the following steps.

- Bandpass filtering of the time series (preliminary bands, Tab. A.1) to simplify the filtering into sub-bands;
- The data is divided into sub-bands, their centre frequencies are equally distributed over a logarithmic scale;
- Parsing the time series into adjacent or overlapping windows (= data segments/events) with fixed lengths (e.g. 256 samples) depending on the frequency band;
- Windows are (cosine tapered and) Fourier transformed from the time to the frequency domain;
- Auto- and cross-spectra are calculated for every data segment for all frequency bands;
- The spectra are stacked to compute final response functions (A.24), the stacking is accomplished by a linear regression in a least squares sense.

$$Z_{ij} = \frac{-\left\langle H_j^* H_i \right\rangle \left\langle H_i^* E_i \right\rangle + \left\langle H_i^* H_i \right\rangle \left\langle H_j^* E_i \right\rangle}{\left\langle H_i^* H_i \right\rangle \left\langle H_j^* H_j \right\rangle - \left\langle H_i^* H_j \right\rangle \left\langle H_j^* H_i \right\rangle} \tag{A.24}$$

  $\langle \cdots \rangle$ denotes the auto- and cross-spectra and * the conjugate complex of a matrix ($i, j = x, y$, for a derivation of (A.24) see e.g. Krings, 2007).

Additional criteria, included in the stacking process, are used to improve the processing results. These are, e.g. the phase and coherency criterion. For the first one, only phases which lie in the 'correct' quadrant of the coordinate sytem are used. Considering a 1D or 2D structured underground, the phase should be in the first quadrant between 0 and 90^o for xy-polarization and in the third quadrant between 180 and 270^o for yx-polarization (Weidelt and Kaikkonen, 1994).
The magnetic field components can be predicted using the measured electric components. For the coherency criterion, the predicted and the measured signal are compared. Ratios which are lower than a user-defined threshold are rejected. This threshold can be between zero and one, for the Barberton data set it is mostly set to 0.9.

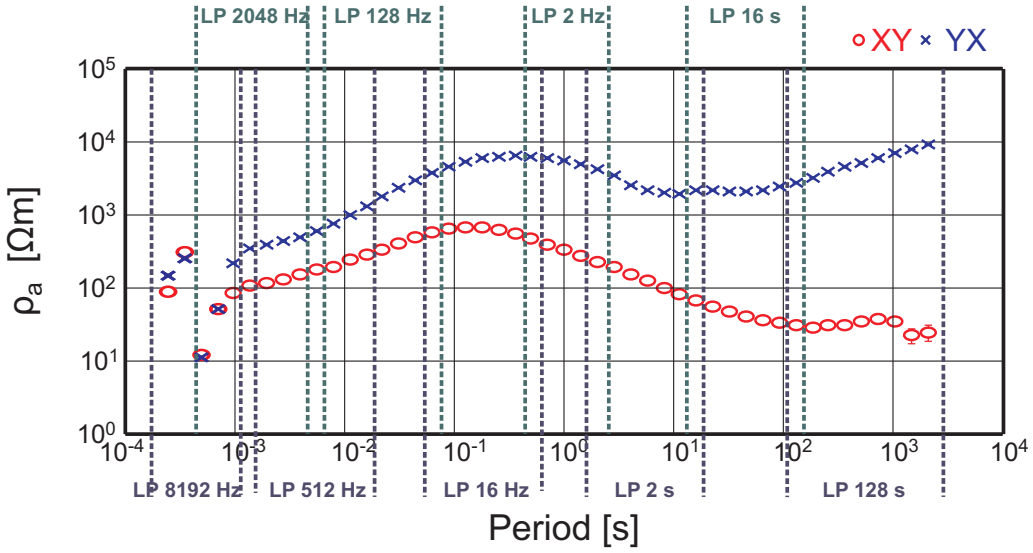


Figure A.2: Period range of the overlapping bands from LP 8192 Hz to LP 128 s used during processing. See also Tab. A.1.

Table A.1: Properties of the preliminary bands: Name of the low pass (LP) filtered band and the corresponding minimum and maximum periods (T_{min}, T_{max}). Note that the period bands are overlapping. See also Fig. A.2.

LP	8192 Hz	2048 Hz	512 Hz	128 Hz	16 Hz	2 Hz	2 s	16 s	128 s
T_{min} [s]	1/4096	1/2048	1/512	1/16	1/16	1/2	2	16	128
T_{max} [s]	1/1024	1/256	1/64	1/128	1/2	2	16	128	2048
f_s [Hz]	32768	8192	2048	512	64	8	2	1/4	1/32
f_{LP} [Hz]	8192	2048	512	128	16	2	1/2	1/16	1/128
f_{HP} [Hz]	0	0	0	0	0	0	0	0	0

For the transfer function estimation (ordinary least squares method) it is in general assumed that the noise which is part of every 'real-world' data (see chap. 2) is normally (= Gaussian) distributed and that it only affects the output channels of the system (Ex, Ey and Hz) (Ritter et al. 1998). These assumptions only hold in few cases and more advanced methods need to be used to deal with noise in the data.
For considering noise which is not normally distributed, robust processing schemes are applied (e.g. Larsen et al., 1996; Ritter et al., 1998). They reduce the influence of outliers and gaps in the data since they are constructed as data-adaptive weighting schemes. Within the *EMERALD* package the robust algorithm in its latest version was implemented by Krings (2007). An important factor for robust methods is the amount of outliers they can handle before failing. Depending on the robust scheme 10 to 50 % (or even more) outliers in the data are possible (Krings 2007 and references therein).
Nevertheless, even robust processing methods can fail to produce smooth transfer functions in the presence of severe noise (cf. sec. 1.2). Hence, to reduce the noise in the time series or in the spectra prior to the stacking and to consider noisy input channels (Hx and Hy) as well, filtering methods (chap. 2) become more and more important.

Appendix B

MT data

This section shows data examples mainly of profiles P2, P3 and P5 which could not be shown in the main part of this thesis due to a shortage of space.

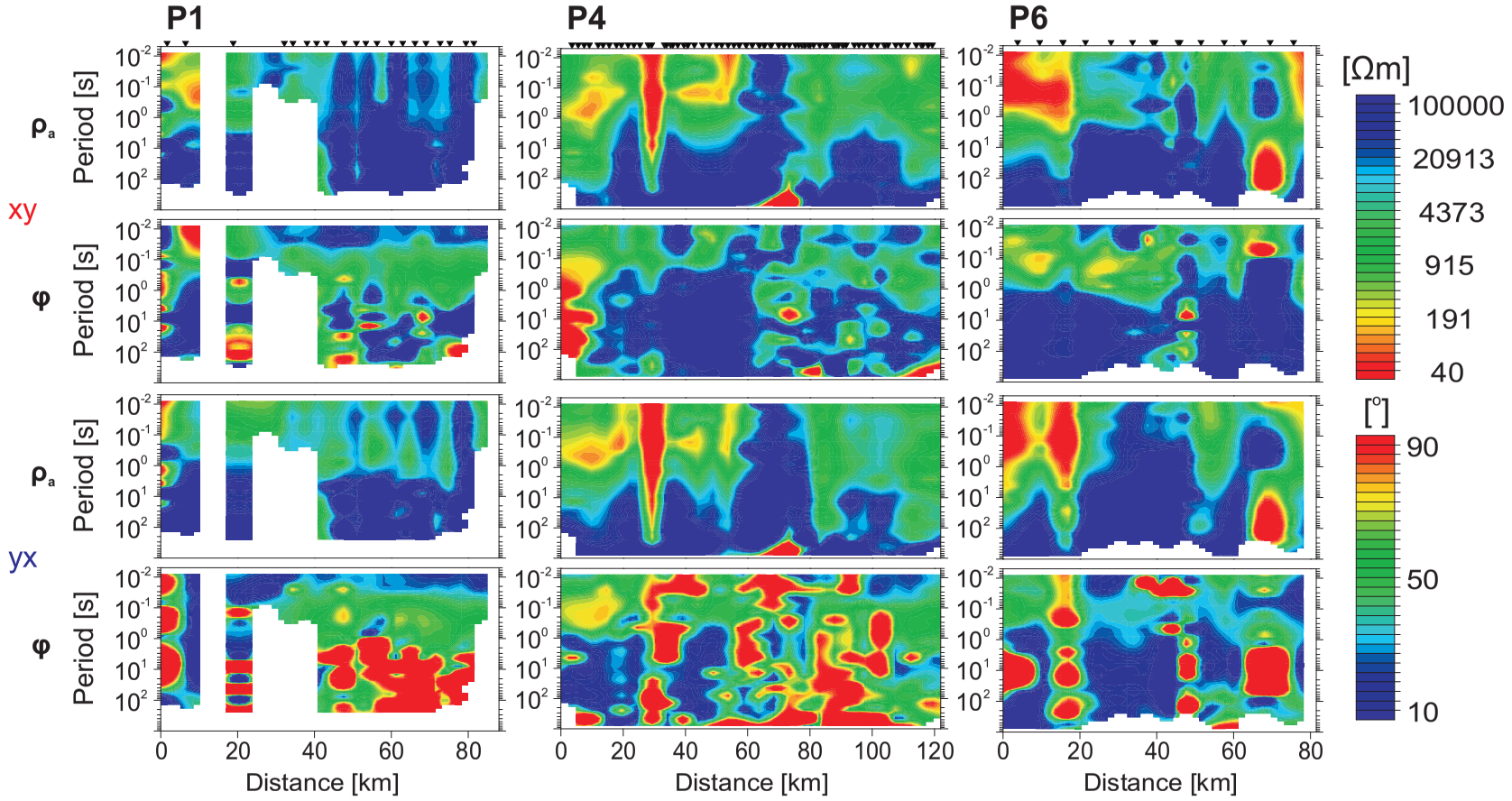


Figure B.1: Pseudosections of apparent resistivity ρ_a and phase φ for components xy and yx of the impedance tensor **Z** after standard processing (see app. A.4) for profiles P1, P4 and P6 (see map 1.2).

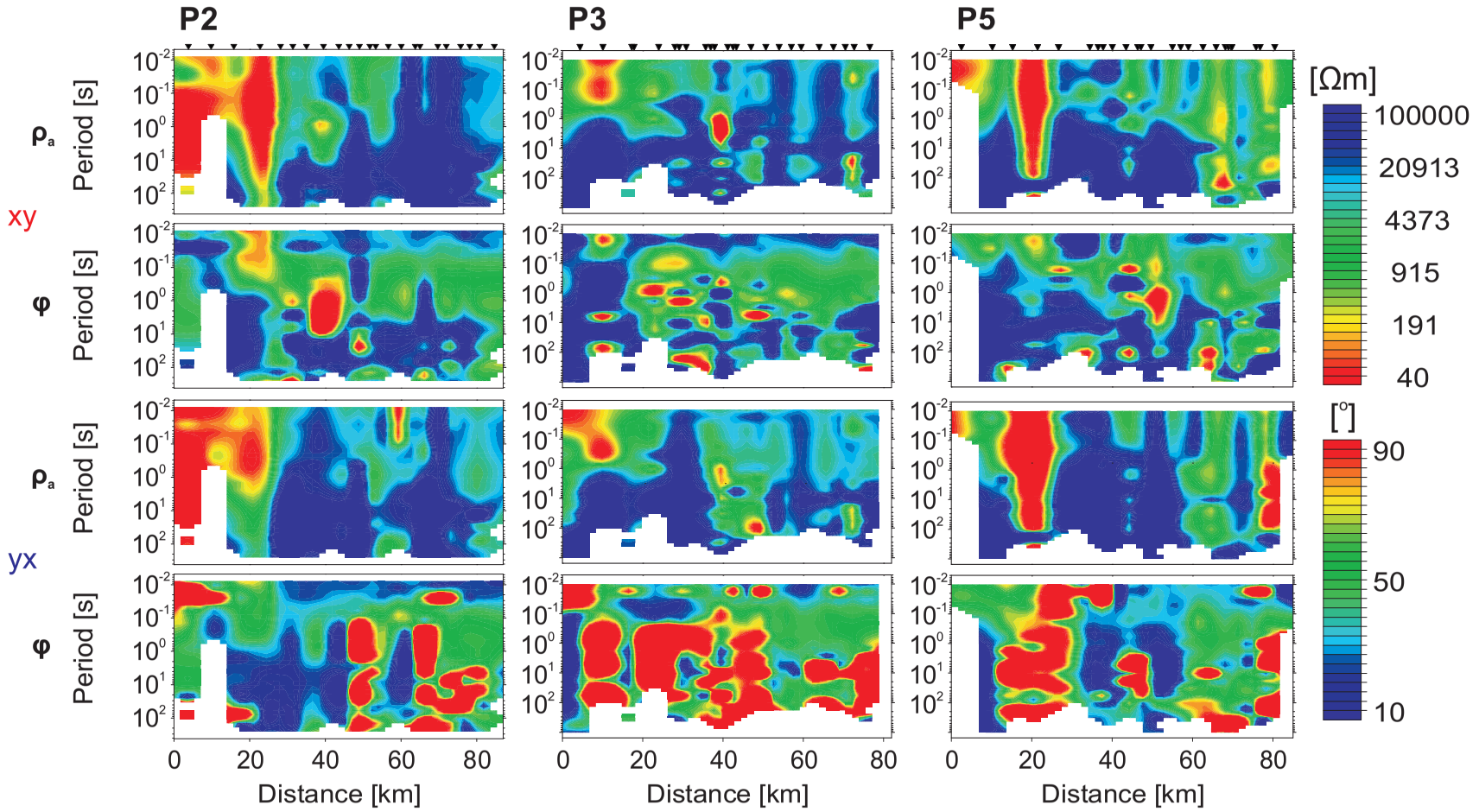


Figure B.2: Pseudosections of apparent resistivity ρ_a and phase φ for components xy and yx of the impedance tensor **Z** after standard processing (see app. A.4) for profiles P2, P3 and P5 (see map 1.2).

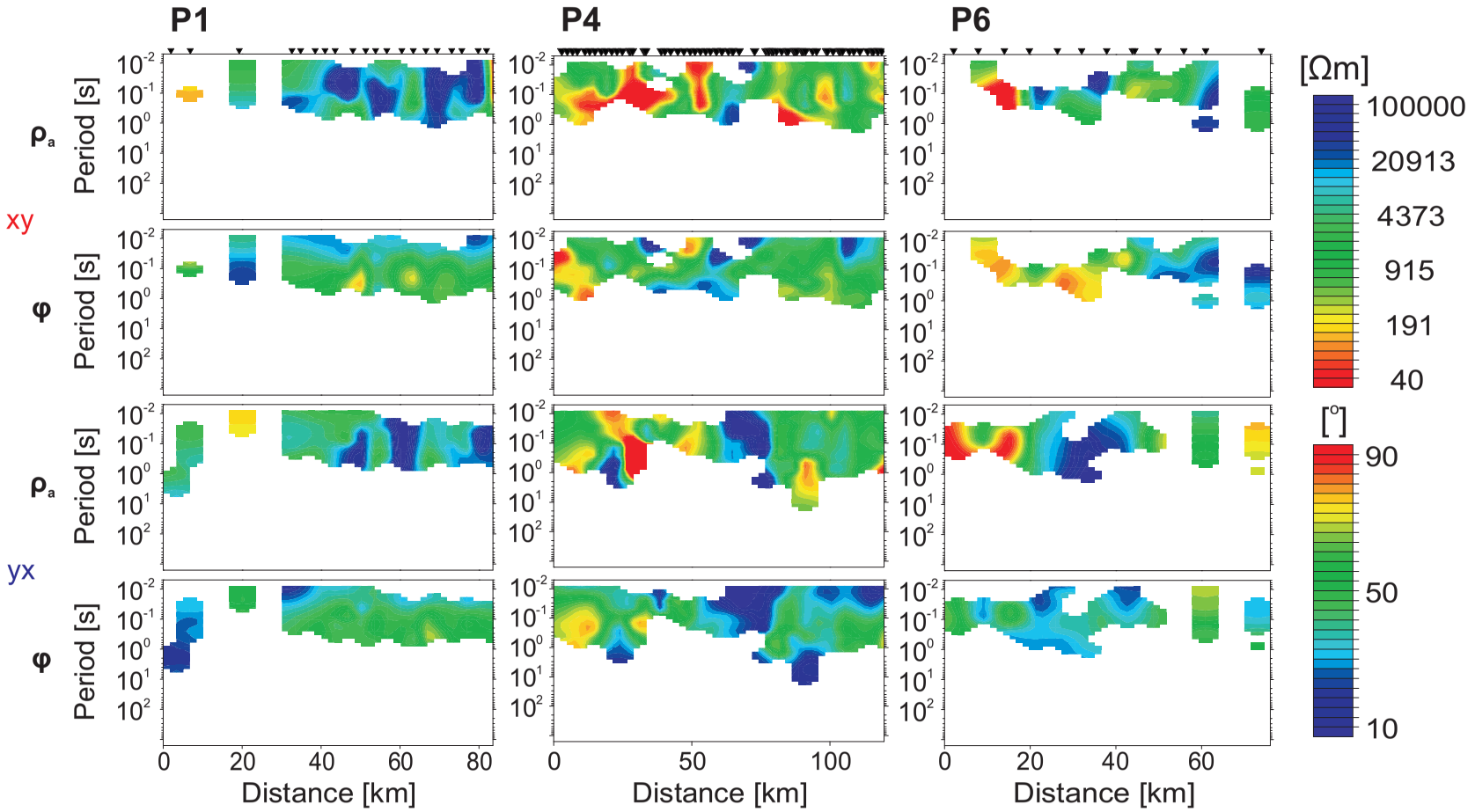


Figure B.3: Pseudosections of apparent resistivity ρ_a and phase φ for components xy and yx of the impedance tensor **Z** after standard processing and masking of noisy data (see app. A.4) for profiles P1, P4 and P6 (see map 1.2).

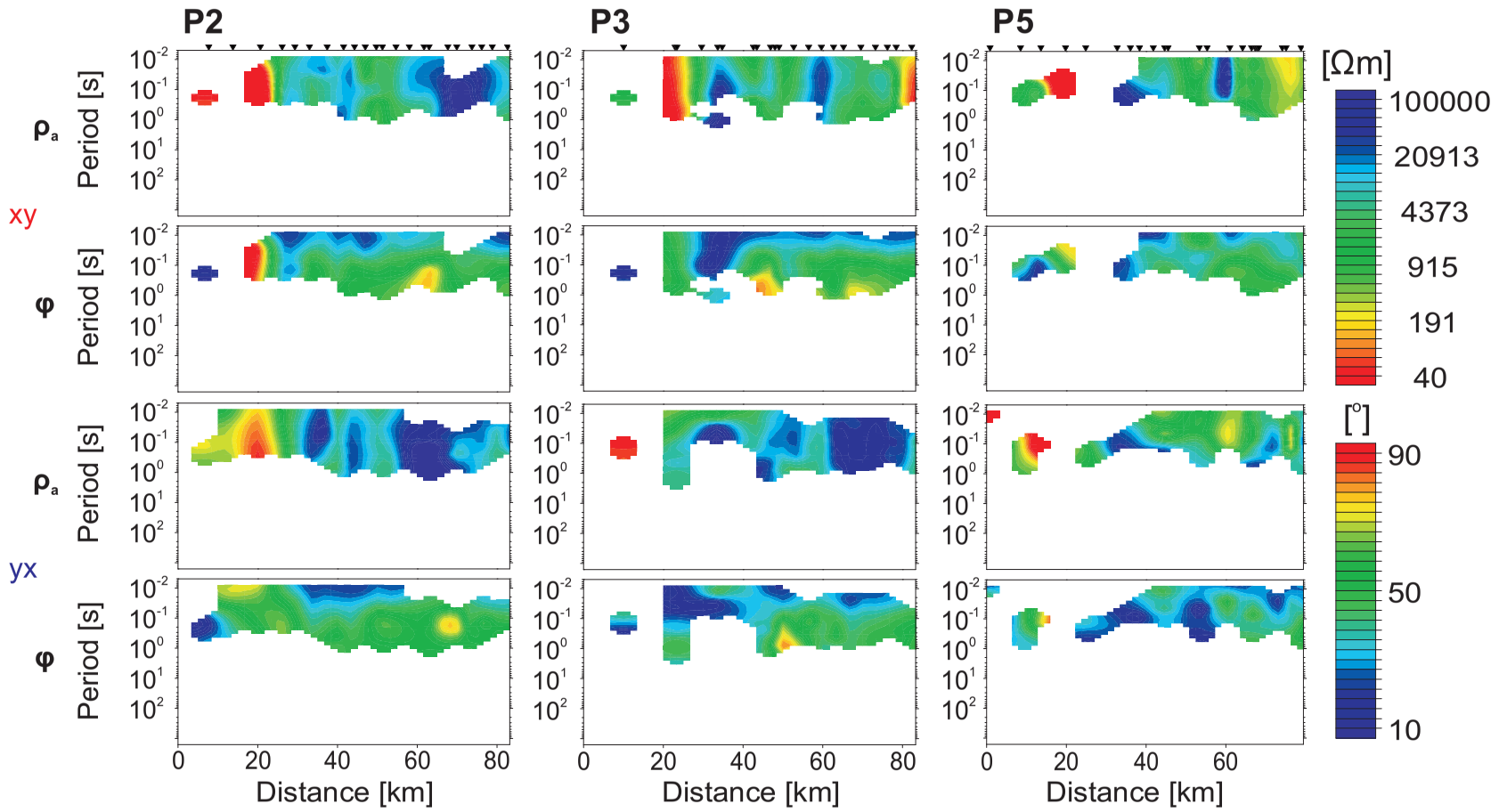


Figure B.4: Pseudosections of apparent resistivity ρ_a and phase φ for components xy and yx of the impedance tensor **Z** after standard processing and masking of noisy data(see app. A.4) for profiles P2, P3 and P5 (see map 1.2).

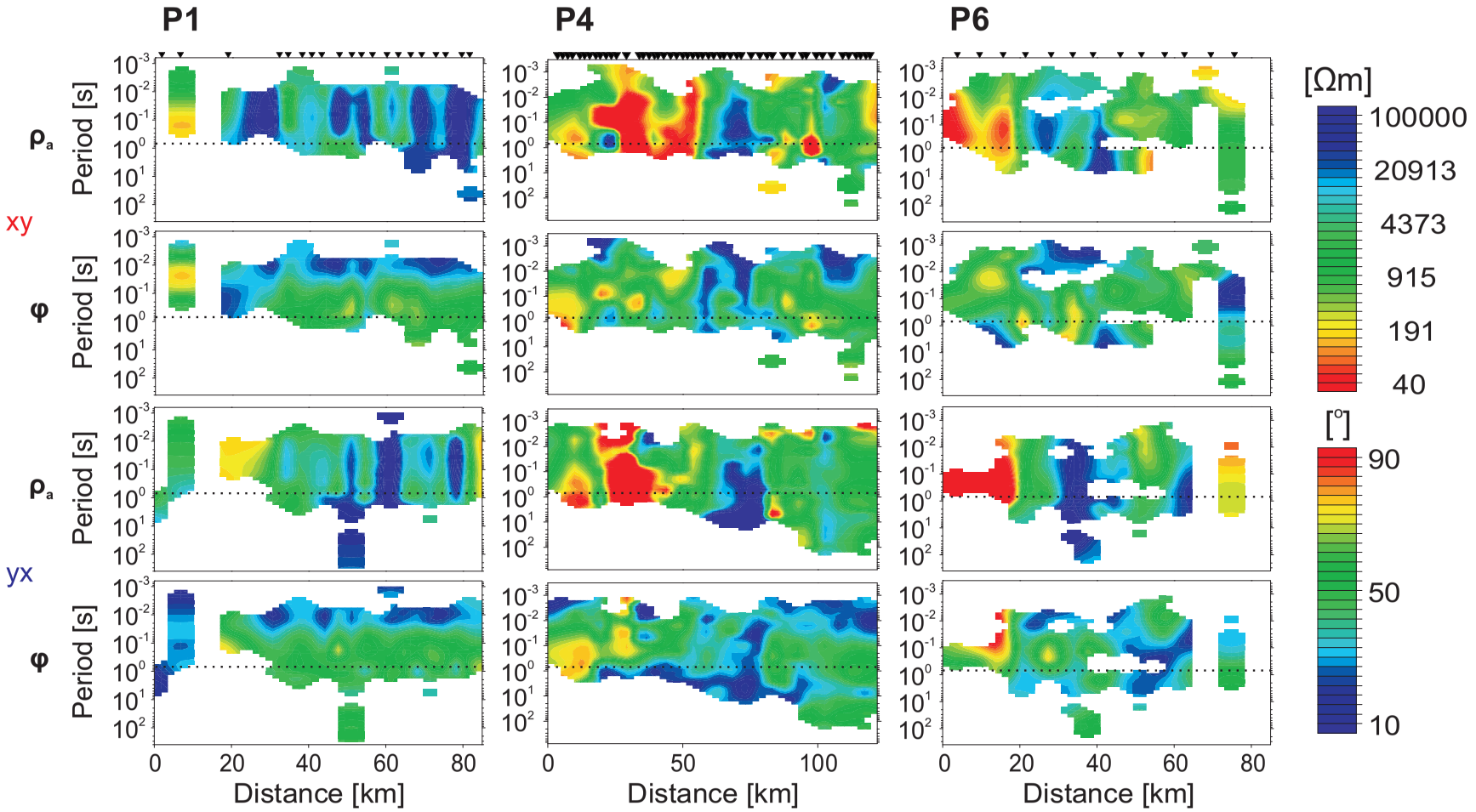


Figure B.5: Apparent resistivity ρ_a and phase φ for the xy and yx component after data filtering and processing as described in section 2.4 for profiles P1, P4 and P6 (see map 1.2). For profiles P2, P3 and P5 see Fig. B.6 in app. B. For periods below the dashed line mainly remote reference and a delay filter (50 Hz) was used for data improvement, for longer periods the reference site and STA/LTA filter were applied.

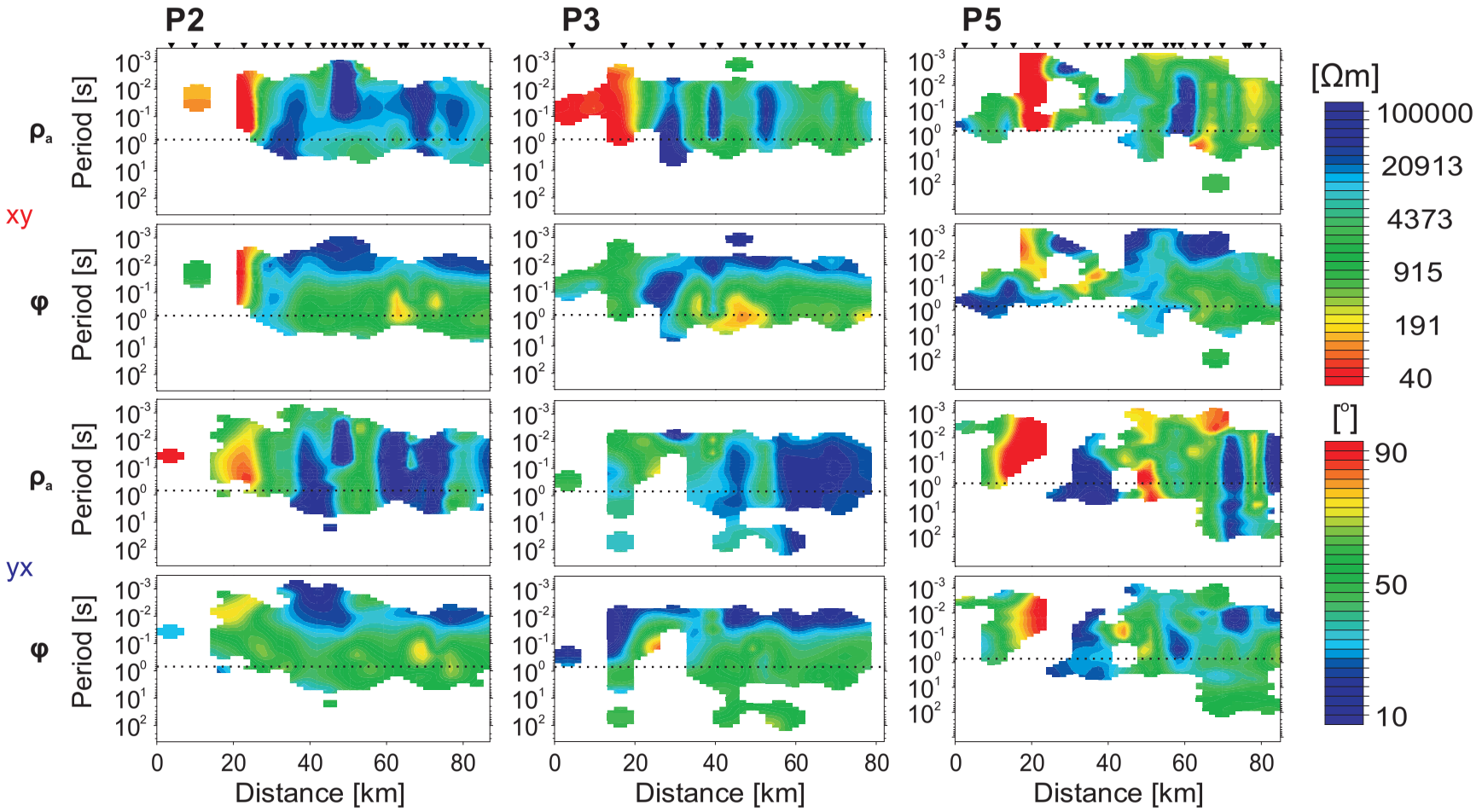


Figure B.6: Pseudosections of apparent resistivity ρ_a and phase φ for components xy and yx of the impedance tensor **Z** after filtering as described in section 2.4 for profiles P2, P3 and P5 (see map 1.2). For periods below the dashed line mainly remote reference and a delay filter (50 Hz) was used for data improvement, for longer periods the reference site and STA/LTA filter were applied.