UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE ENGENHARIA. GEOGRÁFICA, GEOFÍSICA

E ENERGIA

JOINT INVERSION OF MAGNETOTELLURIC AND SEISMIC DATA FOR CRUSTAL CHARACTERIZATION

Milenko Markovic

DOUTORAMENTO EM CIÊNCIAS GEOFÍSICAS E DA

GEOINFORMAÇÃO

(Especialidade em Geofísica)

2013

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE ENGENHARIA. GEOGRÁFICA, GEOFÍSICA

E ENERGIA

JOINT INVERSION OF MAGNETOTELLURIC AND SEISMIC DATA FOR CRUSTAL CHARACTERIZATION

Milenko Markovic

Tese orientada pelo Prof. Doutor Fernando Monteiro Santos,
especialmente elaborada para a obtenção do grau de
doutor em Geofisica

2013

Abstract

The research that has been performed in this thesis contributes to integration two different geophysical methods (magnetotellurics and seismic refraction). Pattern recognition has not been used for the cooperative inversion of these two methods before. This study demonstrates potential of soft clustering methods for integration of two data sets. Gathering information from both inversions, a multi-parameter model was made. This multi-parameter model contains regions, each of which is characterized by a consistent relationship between the model parameters. Both synthetic and field data results were used to discuss an assumption that mutual exchange of information can lead to better results of inversions in both cases. The method developed in this thesis was successfully applied at experimental data collected in the Iberian Pyrite Belt.

Resumo

A investigação desenvolvida nesta tese é uma contribuição para a integração de dois métodos geofísicos (magneto-telúrico e sísmica de refracção). A técnica de reconhecimento de padrões não tinha sido ainda utilizada na inversão cooperativa de dados destes dois métodos. Este estudo mostra as potencialidades dos métodos de análise de "clusters" na integração dos dois conjuntos de dados. Um modelo "multi-parâmetros" foi desenvolvido para extrair informação de ambos os métodos geofísicos. Este modelo contém regiões caracterizadas por relações consistentes entre os parâmetros. Resultados obtidos com dados sintéticos e de campo foram usados para mostrar que o uso da informação contida nos dois conjuntos de dados conduz a melhores resultados. O método desenvolvido nesta tese foi aplicado, com resultados satisfatórios, a um conjunto de dados reais obtidos na Faixa Piritosa.

Table of contents

List of tables

List of figures

Acknowledgement

I wish to thank my advisor, Fernando Monteiro Santos, with whom I constantly collaborated during these four-years work on my PhD thesis . His support, advice, expertise and helpful suggestions were always available at just the right moments.

I would like to thank all colleagues from IDL- Universidade de Lisboa, especially Alfonso Loureiro for help with processing seismic data and Alexandra Afilhado for helping me with seismic software.
Joao Carvalho(LNEG) provided me with seismic data from Algarve, would like to thank him for his patience and support.
Josef Pek (Geophysical Institute, Academy of Sciences of the Czech Republic) kindly hosted me in the Geophysical Institute and also provided software for my research.
Kerry Key (UCSD) was available for online help all the time, , provided me a lot of usefull comments regarding the OCCAM2D code.
I would like also to thank Colin Zelt(Rice University) and Steven Constable(UCSD), Janos Abony (University of Pannonia) for providing code for my research.
Artem Kashubin (Schlumberger,Upssala) made a lot of comments regarding seismic inversion, he has also written scripts that could help a lot. I would like to thank him also.
Carlos Inverno helped us also, allowing the use of the seismic data acquired by Prospectiuni and LNEG under the scope of the FP7 Promine Project.

This research was funded by FCT(Fundação para a Ciência e a Tecnologia), I am thankfull to them for giving me financial support. (Referência da bolsa: SFRH / BD / 66455 / 2009)

Abbreviations

FCM - fuzzy c menas

MT- magnetotellurics

HCA- hierarchical cluster analysis

GK - Gustafson-Kessel

GG - Gath-Geva

UCS- Unconfined Compressive Strength

RMS - Root mean square

1. Introduction

Here, I have presented my work, related to the integration of two geophysical methods: seismic refraction and magnetotellurics. Both of the methods are widely used in academy and industry. As a result of an inversion process, MT enables imaging electrical conductivity. Inversion of travel time data sets from seismic experiment provides us with seismic velocity distribution. Motivation for this research lies in the fact that electrical conductivity is varying by many orders of magnitude, unlike seismic velocity that varies by factor of 2 . Also the physics of the MT method involves the diffusion equation ,while the seismic method uses wave equation, with much better resolution. In order to get information from both methods we have made inroad into pattern recognition in geophysics, by using fuzzy logic algorithms.

Chapter 1 deals with an inverse problem generally and points out what other scientists have done so far (on integration of geophysical methods), discusses also problems encountered.
Chapter 2 explains the theory behind the travel time forward modeling and inversion, which was used later to obtain seismic P wave velocity model.
In chapter 3, the magnetotellurics method is presented.
In chapter 4, pattern recognition is explained, with latter emphasis on clustering and three algorithms (FCM,Gustafson-Kessel and Gath-Geva) that have been employed for this thesis.
Chapter 5 presents geological setting and data processing
Chapter 6 explains parametrization and methodology
In chapter 7 are given results with synthetic data
In chapter 8 are given results from the Pyrite belt
In chapter 9, the summary is presented and new ideas which could lead to improvements, are discussed
Chapter 10 contains bibliography

Inverse problems arise whenever one searches for causes of observed effects . Inverse problems arise in wide variety of applications such as medical imaging, geophysical exploration, and nondestructive evaluation where measurements made on the exterior of a body are used to determine properties of the inaccessible interior .

In science, most problems are prediction of external characteristics of a known internal cause. However, if we determine physical laws through indirect observations or measurements, the situation changes. From everyday examples, it is easy to recognize inverse problems. They have been of great significance in development of natural sciences and all branches of engineering . Or let's try to put this in historical perspective. Charles W. Groetsch argues that it is the oldest problem in mathematics is linear interpolation problem. Herodotus traced the origins of geometry to a group of linear interpolators: the rope strechers of ancient Egypt. Today we would call the linear interpolation problem an inverse problem (inverse problem may have unique solution,no solution or infinitely many solutions depending on number and nature of the points). The direct problem is to calculate the values of linear function; in the inverse problem we must determine linear function from a couple of points on the graph. This is exactly what the rope-strechers did: they solved in a direct way an inverse problem. (Groetsch,Inverse problems).

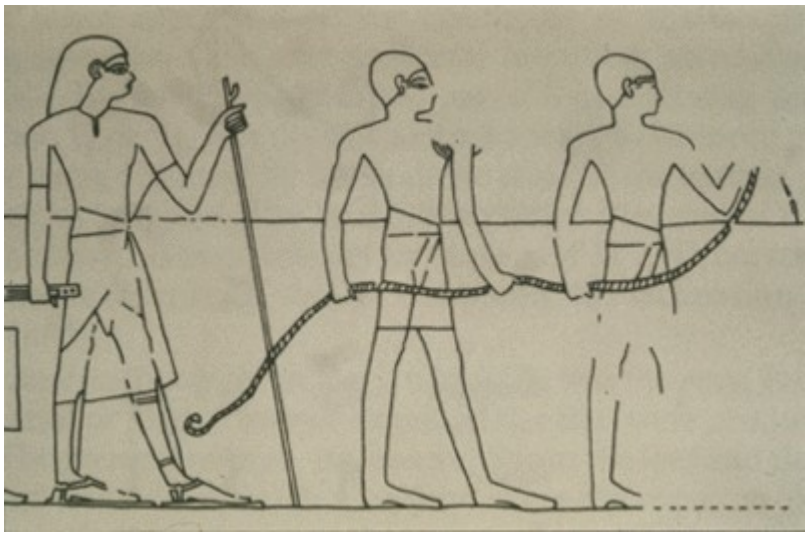


Figure 1.1 shows rope strechers in ancient Egypt

Inverse problems are connected to direct (forward) problems and the choice which one is direct or inverse could be arbitrary (but not in geophysics, inverse problems always consists of using the actual result of some measurements to infer the values of the parameters that characterize the Earth). Usually the direct problem is the more classical one. For example, when dealing with partial differential equations, the direct problem could be to predict the evolution of the described system from knowledge of its present state and the governing physical laws including information on all physically relevant parameters while a possible inverse problem is to estimate (some of) these parameters from observations of the evolution of the system.

This can be graphically described by

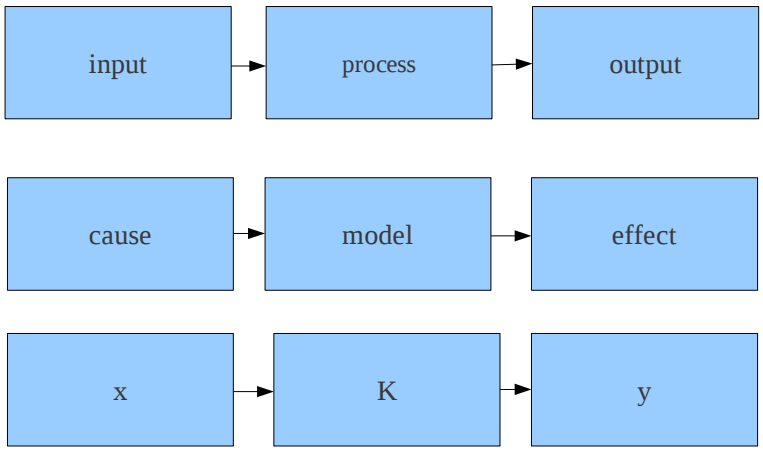


Figure 1.2 shows paradigm of direct problem (Groetch, 1999)

If x is the input and K is the process, then the direct problem is to find Kx. It is quite visible that now two inverse problems are suggested by every direct problem.

One is the causation problem: given a model K and effect y, find the cause of the effect. This inverse problem for K can be reformulated as direct problem for K^{-1}, assuming that that the model is invertible (in geophysics this is usually not the case).

The other inverse problem suggested by the direct problem is the model identification problem:given the cause-effect information, identify model. The direct problem has a unique solution if the process K is a function(y=Kx), but there is no guarantee that the inverse causation and model identifications

problems have unique solutions. (Groetsch, Inverse problems).

In general, the objective of a geophysical inverse problem is to find an earth model, described by a set of physical parameters, that is consistent with the observational data (Barhen *et al., 2000*). But we can find a very broad spectrum in geophysics: the determination of velocity structure (seismic tomography), the location of the earthquakes and determination of their source mechanism (seismology) or to find out about conductivity of electrically conductive half-space (MT).

The solutions of inverse problems are non-unique. With practical data, consisting of a finite set of imprecise observations, infinitely many solutions exist if one does exist at all. There is no scientific theory that could perfectly resolve problem in a sense that eliminates all but one model from the real world.

There are two special issues that need to be addressed with almost every geophysical inverse problem:

- the first problem is the observational data are usually incomplete, they do not contain enough features to resolve the model. The model obtained from the inversion is not necessarily the one we seek, although it complies with the requirements of the inversion process.
- the second problem is that the data contains a noise component. This noise comes from two primary sources, a random component in the observational data and approximations or errors contained in the theory that connects the data and with the model.

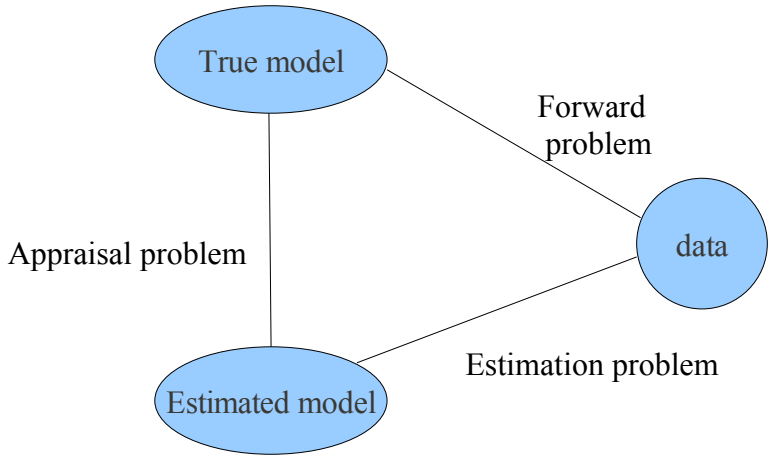


Figure 1.3 The inverse problem view as a combination of an estimation problem plus an appraisal problem (Sneider and Trampert, 1998)

Figure 1.3 shows how we deal with real problems. It is usually assumed that the forward problem, that of calculating simulated data for an earth model, is well enough understood so that reasonably accurate synthetic data can be generated for an arbitrary model. From the data, estimated model is reconstructed (estimation problem). Next step is to check out what relation the estimated model bears to true model. In the appraisal problem we determine which properties of the true model are recovered by the estimated model and what errors are attached it (Trampert, 1998).

INVERSION = ESTIMATION+ APPRAISAL

Assessing reliability of an inversion process is practically the most essential part of complete solution. Although the appraisal stage is of huge importance, there is no general theory to deal with the appraisal problem for a truly nonlinear inverse problem with infinitely many degrees of freedom. Some examples of practices that gave good results:

One approach is to generate not a single model that fit the data within a certain tolerance but to obtain ensemble of models that fit the data within a certain tolerance. Lomax and Sneider have chosen a genetic algorithm (GA) to invert the group- dispersion data because it allows a fully non-linear search of a large solution space and because it produces a set of acceptable solutions. This method gave a useful and direct presentation of the uncertainty and trade-offs in the results (Lomax, 1995) .

Or for example in EM inversion, David Alumbaugh has estimated of how errors in the inversion process, such as data noise and incorrect a priori assumptions, map into parameter error and thus provides valuable information about the uniqueness of the resulting image. Also he derived methods for image appraisal when the iterative conjugate gradient technique is applied (rather than a direct inversion technique) to solve the inverse problem (Alumbaugh, 1998).

In MT inverse problem, problems arise with a non-unique and ill-posed problem of reconstructing the inhomogeneous conductivity distribution of rock formations from the measured electric and magnetic fields at the surface. Usually, practice is to introduce stabilizing functional to stabilize the inversion process (Tikhnov regulrization). But there is no guarantee that the solution will match geological structure properly. Two or even more different conductivity distributions can match the data set. Portniaguine and Zhdanov have developed a new technique called focusing of inversion images, based

on specially selected stabilizing functionals which minimize the area where strong model parameter variations occur. (Portniaguine and Zhdanov,1998).

To improve resolution and sensitivity studies, Pek and Monteiro Santos have developped an algorithm for jointly evaluating the magnetotelluric impedances and their partial derivatives with respect to the parameters of a 1-D generally anisotropic layered medium (Pek and Monteiro Santos, 2004)

In recent years Bayesian inversion has gained a strong popularity in its application to geophysical inverse problems . The popularity is to a large extent a result of the freedom that is taken in controlling the desired model properties through the specification of the a priori model statistics (probabilistic formulation) . We know something about the model and we try to put that in a mathematical form. That is called *a priori* information. Then after we use data set with its statistical properties. Bayesian inversion provides a framework for combining the *a priori* model information with the information contained in the data to arrive at a more refined statistical distribution; that is, the *a posteriori* model distribution. Then the *a posteriori* distribution is what we know after we have assimilated the data and our prior information. The point of using the data is that the a posteriori model information hopefully constrains the model more tightly than the a priori model distribution (Scales,1996). Examples of this are in Bayesian seismic waveform inversion (Gouveia and Scales,1998), or when uncertainty aspect of AVO is assesed using Bayesian approach (Buland, 2000).

The starting point in the analysis of inverse problems is the forward calculation of observables for an earth model . The usual situation is that the observational data *d* represent the solution of the theoretical problem, while the model *m* represents parameters of the equations.

$$d = f[m] \tag{1.01}$$

where

m = model

d = predicted data from model

F = functional operating on model

The general framework is that only the forward problem can be solved in an analytical or semi-analytical sense. The solution of the inverse problem then proceeds by solving the forward problem employing a candidate model *m* in order to obtain simulated data A comparison between the simulated data and the observed data can then be used to make improvements to the candidate model.

$$N(d,m)=||d-f(m)|| \tag{1.02}$$

The requirement that norm *N(d, m)* be a minimum represents a constraint upon the model that incorporates both a theoretical model of the forward problem and observational data.

A variety of other types of constraints can be placed upon the model *m* by specifying certain required or desirable properties that it should have, such as positivity, nearness to a particular model or some measure of smoothness.

These constraints can usually be expressed in the form of equality constraints

c(m) =0,

inequality constraints

c(m)≥0

or a regularization condition

S(m) = minimum.

Two issues need to be addressed here regarding geophysical inverse problems:

1. the fact that the observational data are usually incomplete in the sense that they do not contain enough information to resolve all features of the model.
2. the fact that the data generally contains a noise component

An important consequence of the incompleteness and inaccuracy found in geophysical inverse problems is that there are many possible solutions to the problem. The basic task, in its most complete form, is to describe all of these possible solutions. There are a number of methods that attempt this task by performing a general search of the model space, including grid searches, random searches, and pseudo-random searches.

The most common method of addressing the fundamental non-uniqueness of geophysical inverse problems is to impose additional constraints on the solution and in this way reduce the number of acceptable solutions (Parker, 1994; Oldenburg *et al.,*, 1998). This general process of introducing constraints that restrict the size of the solution space, in many cases reducing it to a single solution, is known as regularization. With regularization non-unique problem becomes unique.

$$\Phi(d,m)=||d-f(m)||+\alpha S(m) \qquad (1.03)$$

Additional term is added to get a well behaved solution (with physical reality). From above leads the formulation of inverse problem as an optimization that involves minimization of an objective function that combines a misfit measure with a regularization measure.

$$min\ \ (\ ||d-f(m)||\ +\ \alpha S(m)\ ) \qquad (1.04)$$

Function is subjected to appropriate equality and inequality constraints.

With this, the geophysical inverse problem reduces to a problem in numerical optimization. The mathematical formulation of the optimization problem, however, is not unique. The formulation can have significant impact on the optimization techniques applicable and on the efficiency of the solution approach .

When the inversion process is performed the obtained geophysical models should map the distribution some of physical property. It is not sufficient just to calculate numerically parameters, further interpretation is required to go from geophysical models to geological models. Generally, a consistent structural model not only fits observation data, but also maintains correct relationships between geological interfaces .

We should bear in mind that models are representation of reality but not the reality itself. When providing a geological model for any kind of physical simulation, one should be able to also provide associated uncertainties. There are two kinds of uncertainties:

-The uncertainties associated with the modeling process, which include simplifications and assumptions
-The uncertainties associated with the data and measurements themselves

Limitations of seismic inversion are numerous. It is necessary to have good quality data,then with good initial model,this leads to good output. Also seismic data has limited frequency content. Therefore it will not be possible to produce a meaningful inversion if the layers are thin. The earth acts as a filter that attenuates the high-frequency components of the seismic data.

Increasing the quantity of geophysical data should not necessarily reduce the uncertainty of geophysical models. This is the consequence of the fact that we are dealing with non-unique inverse problems. Example for this can be seismic refraction, with its inevitable blind zone problem. If blind zone layer is present,it will not appear in refraction profile. (Dobrin, 1976).

Or in geoelectrical methods, equivalence problems (different layered models yield the same result) and suppression (thin layers of small resistivity contrast will be missed) also can not be improved by data redundancy.

Other important issue is that due to varying site and environmental conditions not all techniques are able to be implemented at a specific location .

In the previous section some limitations were explained for geophysical methods, now we come to the point of integration of different geophysical models.
Numerous approaches were developed and here we present some of them.

Sequential inversion

In this method the geophysical data considered to have the best resolution capability are processed first and some of the features (e.g. depths of layers, physical properties of contained bodies, etc.) of the geophysical model generated are selected and used to constrain the processing of the second set of geophysical data (e.g. Lines *et al*., 1988; Nath *et al.*, 2000). Nath *et al.* have developed an algorithm wherein the seismic refraction and DC resistivity inversion routines are amalgamated. Anderson *et al.* (2004) have combined gravity, magnetic and seismic data in The San Bernardino basin research. Scott and Peackok have combined also seismic and geoelectrical data.(Scott *et al.*, 2000). In hydrology Beaujean (Beaujean *et al*., 2010) have used multiple geophysical models in a multi step process to predict seawater intrusion. Apart from the subjectivity in the selection of the features that can be regarded as common among the models, the main disadvantage of this approach is that it seems to disregard the data set deemed to have less resolution and limit its contribution to the final geological model. Yet even with these shortcomings some these methods have found application in industry in recent years.

Joint inversion

By joint inversion, we refer to coupled models that are obtained by simultaneously minimizing a misfit function that includes the data misfit of each data type. Joint inversion concept has been recurrently used in the literature (e.g. Vozzof and Jupp, 1975; Lines *et al*., 1988; Zhang and Morgan, 1996; Haber and Oldenburg, 1997; Berge *et al.*, 2000; Gallardo *et al.*, 2003) and it has been equally applied to the inversion of correlated and uncorrelated data assuming a range of common features between the geophysical models. Two approaches are distinctive in joint inversion: petrophysical and structural approach.

The joint inversion methodologies that follow the petrophysical approach are based on the fact that for some specific geological environments, multiple geophysical parameters can be correlated via physical

or empirical relationships. Examples of such inversion involving direct relationships between parameters are the joint inversion of arrival times of shear and compressional seismic waves using the value of the Vp/Vs ratio (e.g. de Natale *et al.*, 2004) and joint inversion of seismic and gravity data (e.g. Roecker *et al.*, 2004) based on empirical relationships between seismic velocity and density like those derived by Christensen and Mooney (1995).

In some cases it is almost impossible to make a direct relationship between geophysical parameters, valid petrophysical relationships that involve different geophysical parameters can be combined using one or more petrophysical attributes as the common factor . Marquis (Marquis *et al.*,1992) has used porosity to relate electrical resistivities and seismic velocities. Kozlovskaya has performed joint interpretation of seismic and gravity data using the established dependence between seismic velocity and rock density (Kozlovskaya, 2001). These methods require previous knowledge of geological environment to guarantee the validity of each petrophysical relationship.

The joint inversion methodologies that follow the structural approach take the subsurface distribution of the physical properties as a correlating factor. Haber and Oldenburg(1997) have developed a generic approach to invert two data sets when the underlying models are linked by having structural similarity. A joint interpretation is then carried out looking for similarities or differences between the recovered models. They have used Laplacian measure as mathematical indicators of the structural boundaries . Although it has proved successful for some applications, some problems occur regarding this structural approach. First, the directions of the variation are not used and second, it becomes inappropriate when only one model must change to satisfy the data. Gallardo (Gallardo, 2004) has proposed his own algorithm with structural operator being a cross-gradient function. A cross-gradients function quantitatively evaluates the structural similarity between two multidimensional models . The cross-gradients function is incorporated as a constraint in a nonlinear least squares problem formulation, thus enabling the similarity between two models in the form of full collinearity of collocated property changes .They have obtained good results with DC resistivity and seismic refraction experiments .

Still for joint inversion methods, that minimize the one functional, all information is used to constrain the models, which consequently leads to diminishing the inherent ambiguities associated with geophysical model derivation .

Cooperative inversion

This concept is pioneered by Lawrence R. Lines (Lines *et al.*, 1988). He used it for seismic and gravity data sets inversion. The way he defined it “as the estimation of a subsurface model which is consistent with various independent geophysical data sets ”, cooperative inversion can may include either "joint inversion" or "sequential inversion" . Later this concept is used by Hendrik Paache (Paache and Tronicke, 2007), for his work in zonal cooperative inversion of disparate geophysical data sets using fuzzy c-means cluster analysis. Although with the same aim of model integration, this approach is fundamentally different because it involves the application of multivariate statistics.

Motivation for integration of different geophysical models comes from the empirical fact that exists a large class of geological objects that cause anomalies in several geophysical fields.

In our research, inverse problems and pattern recognition are used in a multi-step process to estimate parameters of the subsurface models. Three different clustering algorithms are employed and results compared, striving to exchange the information between two disparate data sets. Subsurface distribution of physical properties should be revealed by iterative procedure.

The goals of our approach can be defined :

1. To create mathematical models that capture key aspects of pattern recognition. We need to prove that observations are closely correlated, so that change in one model changes the other and therefore indicate a systematic relationship. The aim is to provide as much detail as possible on data elements exchanged between geophysical models.
2. Then we come to the similarity issues, what features are representative for model similarity? It is somehow challenge to define what similarities really mean.
3. To discuss the question of how much data we really need to be confident that our results really mean something (not mentioning quality of data which is also important but separate issue)
4. To prove guarantees for algorithms (when will they succeed, and to develop algorithms that provably meet desired criteria and discuss their validity). Are contrasts between structures significant enough to be clearly detected in all the geophysical models?
5. To analyze problems encountered during the whole process. Arbitrary or inappropriate restrictions could be placed on the data thus leading to their variations, will this diminish our capability to gain insight into our problem?
6. Have we exhausted what can be extracted from these two data sets?
7. To suggest ideas for new research.

Travel time tomography

Starting from the meaning of the word tomography which finds its origins in Greek, tomo in Greek means 'slice', and graph means 'describe'. Seismic tomography is method for sectional viewing of the subsurface of the earth.

The starting point of modern seismic tomography is probably the 1974 AGU presentation by MIT's Keiti Aki (Aki *et al.* 1977) in which arrival times of P-waves were for the first time formally interpreted in terms of an 'image' as opposed to a simple one-dimensional graph of seismic velocity versus depth .The name that was later given to the new imaging technique is more than an accidental reference to medical tomography, because the earliest radiologic tomograms also attempted to get a scan of the body that focuses on a plane of interest, although using X-rays rather than seismic waves (Nolet , 2008).

Seismic tomography aids in the understanding of the planet's internal structure. The basic idea of tomography is to use a large number of travel times for non-coincident sources and receivers located at the Earth's surface to determine velocity distribution. However, due to the different scopes of investigations regarding said structure, most applications of seismic tomography find themselves divided between two different types(global, near-surface):

In global tomography scientists apply tomographic methods to data obtained naturally from earthquakes to understand the structure of the mantle-one of the deeper layers of earth.

Near-surface seismic tomography is concerned with the shallower subsurface, not going further than a few kilometers. Its major application is in exploration geophysics where the subsurface is investigated in a search for specific substances such as oil or mineral resources. As a major difference from global tomography, this method does not rely on naturally occurring earthquakes, but rather artificial seismic-wave sources, such as explosive detonations(our case).

Travel time tomography is a more accurate way of referring to seismic tomography, one that puts the emphasis on the key aspect used to investigate the subsurface-travel time. Different method can be classified by the type of seismic signal used (P or S waves), by the data of modeling (traveltime,full waveform,amplitudes), by the number of dimension in the model (1-D, 2-D, 3-D), by the type of signal paths (refraction,reflection,first arrival) or by the geometry of seismic experiment (normal incidence,wide angle,cross-well).

Research in this thesis is related with 2-D, P wave, refraction, wide angle, travel time tomography.

Seismic first-arrival tomography is a technique currently experiencing growth due to the simplicity of implementation and promising results for delineating a variety of subsurface targets

Travel time tomography work flow:

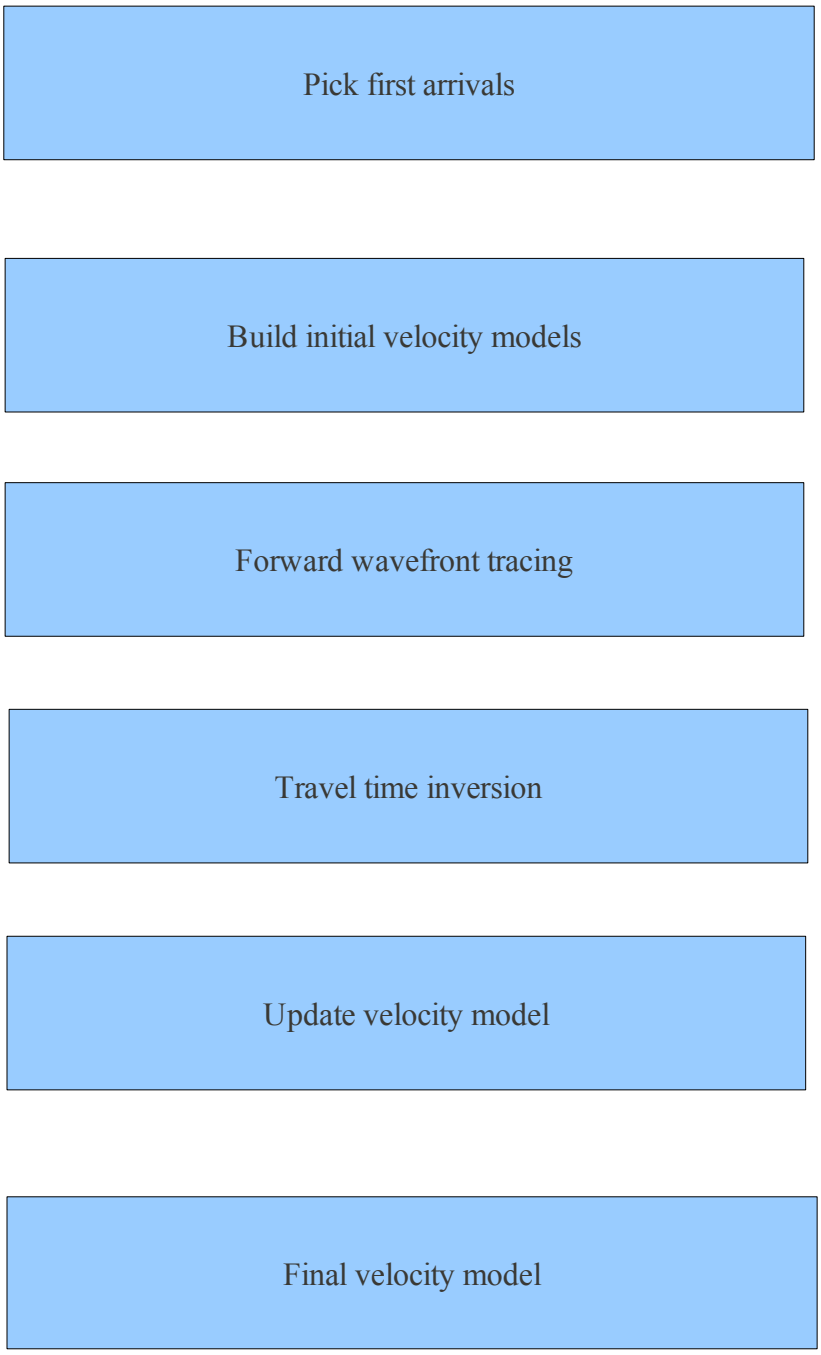


Figure 2.1 shows traveltime tomography workflow

Inversion code-FAST

FAST was written by Dr. Colin Zelt (Rice University). This program package is for 2D and 3D first arrival travel time tomography. Big advantage of first-arrival times is that they are most reliably picked from the seismic records since they are not contaminated by signal-generated noise that can complicate the identification and picking of later-arriving phases. The models are parameterized on a uniform square grid (velocities specified at equal node spacing in the *x, z* directions). In this case geometry was two-dimensional. The inverse grid is cell-based with constant cell size in each direction, but the sizes may be different in the *x, z* directions. The node spacing used for the forward grid must divide equally into the *x, z* lengths of the model, and the cell size of the inverse grid must also divide equally into the *x, z* lengths of the model. The forward calculation of travel times and raypaths uses the Vidale (Vidale, 1990) scheme modified to handle large velocity contrasts according to the method of Hole and Zelt (1995).The regularization is a jumping method in that the constraints are applied to the model perturbation with respect to a background model.

FAST is written in FORTRAN 77, compilation was performed both with g77(g77 - GNU project Fortran 77 compiler) and ifort (Intel Fortran Composer XE 2011) compilers,as well as all input files (some of the files are binary files, like .picks files). New subroutines were written to allow slight changes for initial model and thus make a program more practical for (newbuild instead of velbuild).

Parametrization of the model are of great importance in order to obtain meaningful results . The quality of the forward solution, in particular, strongly depends on parametrization of the velocity field. In this case it was necessary to adapt the mesh,with one that is used for MT inversion. Vector of two dimension will be formed in latter pattern recognition process (every cell has resistivity and velocity) so although inversion is done separately, parametrization must be adequate for both inversions.

The theory of forward modeling

The technique of forward modeling in seismology begins with the numerical solution of the equation of motion for seismic waves, or more specifically, the numerical computation of theoretical or synthetic seismograms, for a given geological model of the subsurface. The idea is then to compare the synthetic seismic traces with real seismic data acquired in the field. If the two agree to within an acceptable level of accuracy, the given geological model can be taken to be a reasonably accurate model of the subsurface. If not, the geological model is altered, and new synthetic traces are computed and compared with the data. This process continues iteratively until a satisfactory match is obtained between the synthetics and the real data. The forward modeling approach is, in a sense, the opposite of the inverse modeling approach in which the parameters of the geological model are computed from the acquired real data. At the end, both methods have the same goal-to determine geological structure and lithology of the subsurface.

The phenomenon of elastic propagation for an isotropic medium can be formulated from the second Newton's law, Hooke's law and the definitions of stress and strain. Search for displacement set up by undirectional point body force acting with time-varing magnitude at a fixed point in homogeneous,unbounded,isotropic,elastic medium. (elastodynamic equation) (Aki and Richards, 1980, p. 64):

$$\rho \frac{\partial^2 u}{\partial t^2} = f + (\lambda + 2\mu)\nabla(\nabla u) + \mu \nabla \times (\nabla \times u) \qquad (2.01)$$

where ρ is the density of material, u is the particle displacement, f is a body force and λ and μ are Lame constants that describe the elastic properties for isotropic media.

In this case the body force f is given by $f_i = X_0(t)\delta(x)\delta_{il}$ (2.02)

and we have initial conditons u(x,0)=0 and $\dot{u}(x,0)=0$ for $x \neq 0$. (2.03)

This displacement field has components

$$u_n(x,t) = X_0 G_{nl} \qquad (2.04)$$

Problems arise with directionality at the source(x_1 at O) and receiver (u at x)

In order to get solution of equation, the scalar problem can be solved, but with directionality of the source,which means that problem must be spherically symmetric

$$\ddot{g} = \delta(x)\,\delta(t) + c^2 \nabla^2 g \qquad (2.05)$$

The solution of the equation is

$$g(x,t) = \frac{1}{(4\pi c^2)} \frac{\delta(t - |x|/c)}{(|x|)} \qquad (2.06)$$

Important properties of solution:

-The solution is product of one factor(the delta function) whose spatial fluctuation is rapid and the other factor whose fluctuation is slow

-Rapidly varying function depends on time relative to arrival time*(|x|/c)*

-Wave shape is same in the time as the time history of the inhomogeneous term in (2.05)

Ray theory gives approximate solutions for the waves that propagate in inhomogeneous media. The basic assumptions is that body waves travel with a local propagation speed along ray paths determined by Snell's law, arriving with an amplitude determined by the geometrical spreading of rays from the source to the receiver. The wavefront is a traveling discontinuity in the displacement of the particles originated by a discontinuity in the seismic source (discontinuity of either the variable or its derivative). Equation for inhomogeneous isotropic medium (full derivation in Aki,Richards)

$$(\nabla T \nabla T - \rho/\lambda + 2\mu)(\nabla T \nabla T - \rho/\mu) = 0 \qquad (2.07)$$

Thus T satisfies the eikonel equation

$$(\nabla T)^2 = 1/c^2 \qquad (2.08)$$

where *c* is velocity

Analytical solutions for eikonel equation are possible only for very simple models,so numerical solution will be employed.

Techniques based on ray tracing have disadvantages in terms of processing time, along with the usual problems with shadow zones and multiplicity of raypaths (Vidale, 1988; Zhang and Toksöz, 1998). In contrast, alternative methodologies based on wavefronts are attractive because the travel times are swiftly computed and the wavefronts can penetrate into shadow zones, find diffracted trajectories and head waves (Vidale, 1988; Qin *et al.*, 1992; Zhang and Toksöz, 1998). Because of this, tomographic methods based on wavefronts (or combined wavefronts and raytracing) have become increasingly popular (e.g. Zhang and Toksöz, 1998; Zelt and Barton, 1998; Zelt *et al.*, 1999; Korenaga *et al.*, 2000). The method of Vidale (1988) in particular is recognized to be one of the fastest methods to compute travel times (Matarese, 1993 in Zhang and Toksöz, 1998) .

John E . Vidale proposed a method to calculate the traveltimes of first arriving seismic rays through velocity structures on a three-dimensional numerical grid using finite-difference extrapolation . He has also tested this approached and compared that it is superior to ray tracing methods.

The method is formulated for a velocity structure that is sampled at discrete points in a 3-D space, with equal horizontal and vertical spacing. The calculation begins with the identification of the grid point nearest the source location . The traveltime to each grid point is calculated by integrating the slowness from the origin to the point, assuming a straight raypath . The propagation expanding rectangular rings as Figure 2.2 shows (for 2d). Rectangular rings are increased in every iteration. The process continues until to the whole subsurface model is covered, producing this way the field of travel times.

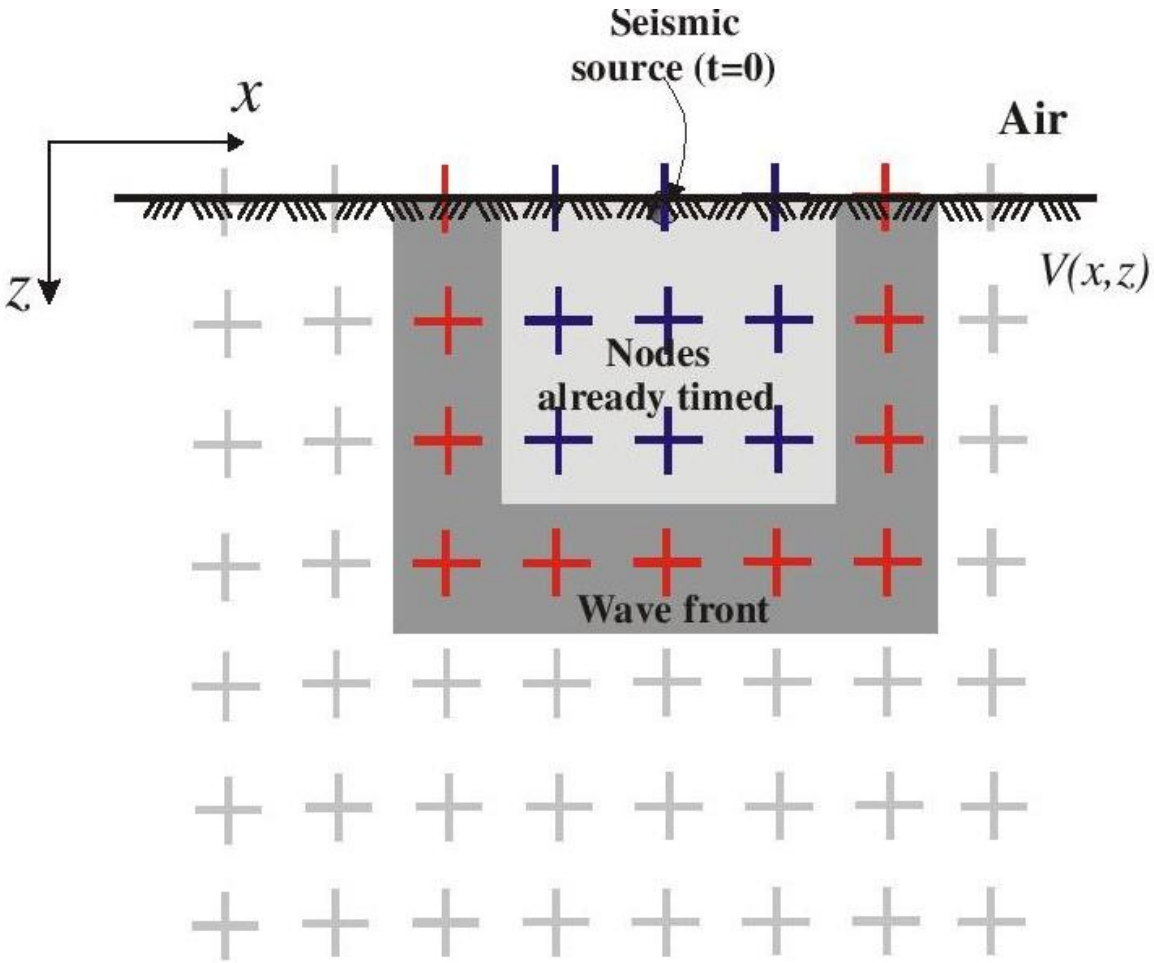


Figure 2.2 shows the propagation of a square wavefront as used in Vidale's (1988) progressive finite difference scheme. The crosses denote the nodes of the seismic grid (timed or to be timed). (Galardo, 2004)

The method to compute traveltimes in this thesis uses three different schemes for three different situations described below.

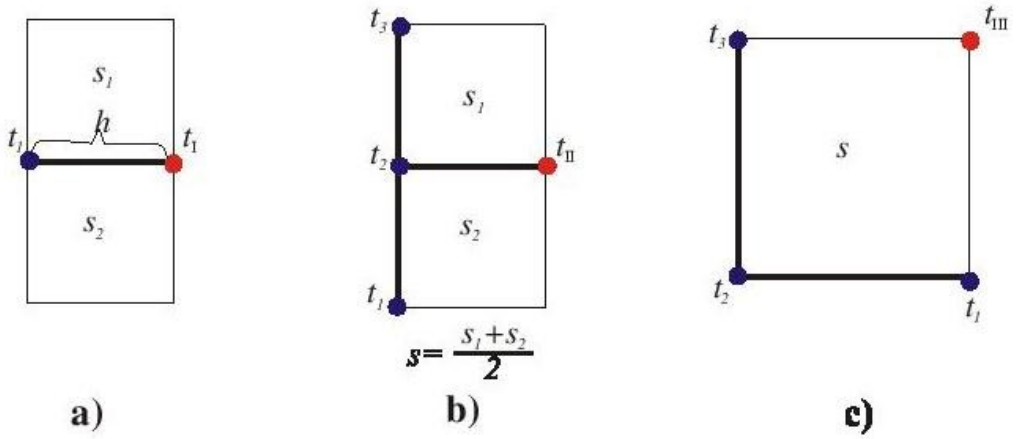


Figure 2.3 Stencils implemented for the forward finite difference scheme of the arrival times

The geometries of the grid points involved in each of the three methods are shown in Figure 2.3

a) Scheme type I

$$t_I = t_1 + min(s_1 h, s_2 h) \tag{2.09}$$

b) Scheme type II

$$t_{II} = t_2 + \sqrt{(2 s^2 h^2 - (t_3 - t_1)^2)} \tag{2.10}$$

c)Scheme type III

$$t_{III} = t_2 + \sqrt{(s^2 h^2 - 1/4 (t_3 - t_1))} \tag{2.11}$$

The scheme type I is preferred for direct waves, the scheme type II is for circular wavefronts and scheme type III is adequate for plane wavefronts.

The main drawback of Vidal's method is that it is not valid in the case of large velocity gradients. Hole and Zelt (1998) propose modification that solves this problem by :

1) the inclusion of head wave operators
2) reverse propagation through the model using a wall source initiated on the edge of the model

Travel-time inversion

The regularized inversio algorithm has been implemented for seismic refraction inversion by Zelt and Barton in their research (with jumping strategy)(Zelt and Barton,1998). The objective function to be minimized is

$$\Phi = \delta t^T C_d^{-1} \delta t + \lambda [\alpha (m^T C_h^{-1} m + s_z m^T C_v^{-1} m) + (1-\alpha) \delta m^T C_p^T C_p \delta m] \qquad (2.12)$$

where :

m = model vector

δm = perturbated model vector (δm= m-m0)

δt = data residual vector

λ =trade-off parameter

s_z = parameter that determines vertical versus horizontal smoothness

C_d = data covariance matrix

C_h ,C_v = horizontal and vertical roughening matrices

C_p = perturbation weighting matrix

Solution of the this inverse problem is non-unique because the data are finite in number and inaccurate , so regularization is necessary. Regularization deals with the ill-posedness of the inverse problem . At the next step constraints are applied, so the model has minimum structure because we are looking for models that include only structure that is required to fit the data according to noise level. Minimum structure is measured in term of model roughness -second spatial derivatives (Lees and Crosson,). In our case, regularized inversion minimizes objective function that includes norms that measure data misfit, model roughness and perturbation. The perturbation weighting matrix is a diagonal matrix of the reciprocal starting slowness values, measuring the relative perturbation of the current model from the starting model .

The three parameters decided to be given for inversion are λ, s_z and α.

The trade-off parameter λ controls the overall regularization and is systematically reduced, starting from λ_0, using a reduction factor. For large values of λ, model smoothness is emphasized over fitting the data. As the value of λ decreases, the relative importance of fitting the data increases. Generally, as the iteration progresses and a final solution is approached, the linearization errors become smaller, the required trade-off parameter λ becomes smaller and smaller model details are resolvable . In practice, a large initial value λ_0 is assigned, and due to the systematic reduction, the final tomographic results are not overly sensitive to the initial trade-off parameter value.

For the solution of the inverse problem, the sparseness of the large system of linear equations is exploited by using the fast and robust LSQR method (Paige & Saunders 1982; Nolet 1987). Due to regularization we have a huge number of linear equations but most of matrix elements are zero (often more than 99%), so computations are not time consuming. At the end of each iteration, the χ2 normalized misfit between observed and predicted data is evaluated (Zelt 1999):

$$\chi^2 = \frac{1}{N}\left[\frac{(t_i^o - t_i^p)}{(\sigma_i)}\right]^2 \qquad (2.13)$$

where N is the number of data points, t_i^o observed traveltime, t_i^o prdecicted travel time, σ picking uncertainty.

3. Magnetotellurics

In this chapter theory behind magnetotelluric exploration is presented.

3.1. The magnetotelluric method-basic physics

Magnetotelluric method firstly appeared in 1950's as a method that serves for determining distribution of the electrical properties of the earth. Originally proposed by Andrey Tikhonov and Louis Cagniard. For the purposes of considering electromagnetic induction in the Earth, a number of simplifying assumptions are considered applicable (Cagniard, 1953). In contrast to other geoelectrical methods, which are based on the study of the relationship between electromagnetic fields and their sources, the MT method is studying the ratios between the horizontal components of the electric and magnetic fields to investigate the electrical conductivity structure of the Earth. The energy for the magnetotelluric technique is from natural source of external origin. When this external energy, known as the primary electromagnetic field, comes to the earth's surface, part of it is reflected back and remaining part penetrates into the earth .

Basic assumptions of the method (Simpson and Bahr, 2005):

- Maxwell's general electromagnetic equations are obeyed
- The Earth does not generate energy, but only dissipates and absorbs it
- All fields may be treated as conservative and analytic away from their sources
- The natural electromagnetic source fields utilized, being generated by large-scale ionospheric current systems that are relatively far away from he Earth's surface, may be treated as uniform, plane-polarized electromagnetic waves impinging on the Earth at near-vertical incidence.
- No accumulation of free charges is expected to be sustained within a layered Earth . If charges are accumulated this leads to phenomenon called static shift. In this case correction for the static shift need to be applied.
- Charge is conserved and the Earth becomes an ohmic conductor
- The electric displacement field is quasi static for MT sounding periods,consequently electromagnetic induction in the Earth is treated as a diffusion process
- Any variations in the electrical permittivities and magnetic permeabilities of rocks are assumed negligible compared with variations in bulk rock conductivities.

In most cases the electromagnetic field arising from distant sources can be considered a plane wave and the source geometry does not need to be known . A plane wave is one that propagates normal to a plane in which the fields are constant . The plane wave assumption is fundamental to the MT method, because it implies time invariance of the exciting source.

The electromagnetic fields within can be described by Maxwell's equations. These can be expressed in a differential form with the International system of Units (SI) as:

$$\nabla \times E = \frac{-(\partial B)}{(\partial t)} \quad \text{Faraday's law} \qquad (3.01)$$

$$\nabla \times H = j + \frac{(\partial D)}{(\partial t)} \quad \text{Ampere's law} \qquad (3.02)$$

$$\nabla \cdot D = \rho \quad \text{Gauss's law} \qquad (3.03)$$

$$\nabla \cdot B = 0 \quad \text{Gauss's law for magnetism} \qquad (3.04)$$

where :

E = electical field intensity (V/m)
j = electric current density (A/m^2)
H = magnetic field intensity (A/m)
B = magnetic flux density (T)
D = electric displacement (C/m^2)

In a homogeneous isotropic medium the following scalar constitutive relations apply

$$D = \varepsilon E \qquad (3.05)$$

$$B = \mu H \qquad (3.06)$$

$$J = \sigma E \qquad (3.07)$$

where ε is the electrical permittivity, μ is the magnetic permeability , σ is the electrical conductivity

These three parameters describe intrinsic properties of the materials through which the electromagnetic fields propagate. These magnitudes are scalar quantities in isotropic media. In anisotropic materials they must be expressed in a tensorial formulation.

If equations for D, B and J (3.05, 3.06, 3.07) are introduced into Maxwell's equations, after a little mathematics

$$\nabla\times(\nabla\times E)=-\mu(\nabla\times(\frac{\partial H}{\partial t})) \qquad (3.08)$$

Now we use the vector identity

$$\nabla(\nabla\times E)=\nabla(\nabla E)-\nabla^2 E \qquad (3.09)$$

From the equation for the divergence of the electric field in a charge free region we get

$$(\nabla E)=0 \qquad (3.10)$$

then

$$\nabla\times(\nabla\times E)=-\nabla^2 E \qquad (3.11)$$

After substituting equations for J and D , the wave equation:

$$\nabla^2 E=\mu(\sigma\frac{\partial E}{\partial t}+\epsilon\frac{\partial^2 E}{\partial t^2}) \qquad (3.12)$$

If w is the angular frequency and e^{iwt} is time dependence for E,then

$$\frac{\partial E}{\partial t}=iwE \quad \text{and} \quad \frac{\partial^2 E}{\partial t^2}=-w^2 E \qquad (3.13\text{ a. },3.13\text{ b.})$$

Introducing above relations into the wave equation

$$(\nabla^2+k^2)E=0 \qquad (3.14)$$

$$k=\mu\omega(\varepsilon\omega-i\sigma) \quad , \quad \text{k=propagation constant} \qquad (3.15)$$

Quasi-stationary approximation claims that displacement currents can be neglected relative to conductivity currents, so $\sigma >> \varepsilon\omega$

Second term in wave equation can be neglected,so diffusion equation :

$$\nabla^2 E = \mu \sigma \frac{\partial E}{\partial t} \tag{3.16}$$

Same stands for the magnetic field

$$\nabla^2 H = \mu \sigma \frac{\partial H}{\partial t} \tag{3.17}$$

The diffusion equation describes the behavior of electric and magnetic fields in Earth at the frequencies used in MT .

Faraday's law in the frequency domain implies the following relationships between the E and H

$$\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = -i \mu w H_x \tag{3.18}$$

$$\frac{\partial E_x}{\partial z} - \frac{\partial E_Z}{\partial x} = -i \mu w H_y \tag{3.19}$$

$$\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -i \mu w H_z \tag{3.20}$$

For a uniform plane wave H propagating in the z direction in a half-space the E_z term and its derivative are zero, as are as the x and y derivatives of E_y and E_x

$$E = E_0 \mathrm{e}^{-ikz} \tag{3.21}$$

$$\frac{\partial E}{\partial z} = -i k E \ , \ \frac{\partial E_x}{\partial z} = -ikE_x \ , \ \frac{\partial E_y}{\partial z} = -ikE_y \tag{3.22 a. b. c.}$$

Then it comes

$$\frac{E_y}{H_x}=\frac{-\mu w}{k} \quad , \quad E_x H_y=\frac{\mu w}{k} \qquad \text{(3.23 a. b.)}$$

As in other electromagnetic methods it is usual to define apparent resistivity. In MT it is defined as :

$$\rho=\frac{1}{\mu w}\left|\frac{E_x}{H_y}\right|^2 \qquad \text{(3.24)}$$

In a 2-D earth with a strike along the horizontal x-axis (i.e., $\partial x = 0$) and conductivity σ(y, z) (z positive downwards), the Maxwell's equations are decoupled into two polarization modes. The decoupling is valid since the EM-fields are treated as plane waves, which means that the interaction between electric and magnetic fields are always orthogonal with each other and therefore the horizontal component of the magnetic field tangential to the conductivity strike does not depend on the magnetic field component perpendicular to it.

In this context, the so-called TE-polarization mode refers to the tangential electric field and the TM-polarization mode to the tangential magnetic field; both components are tangential with respect to the strike (x-axis) of the conductivity structure:

TE − polarization : Ex , By

$$\frac{\partial B_z}{\partial y}-\frac{\partial B_y}{\partial z}=\mu\sigma E_x \qquad \text{(3.25 a. b.)}$$

$$-\frac{\partial E_x}{\partial z}=i\omega B_y \qquad \text{(3.26 a. b.)}$$

$$-\frac{\partial E_x}{\partial y}=i\omega B_z \qquad \text{(3.27 a. b.)}$$

TM − polarization : Bx , Ey

$$-\frac{\partial E_z}{\partial y}+\frac{\partial E_y}{\partial z}=i\omega B_x \qquad \text{(3.25 a. b.)}$$

$$\frac{\partial B_x}{\partial z}=i\mu\sigma E_x \qquad \text{(3.26 a. b.)}$$

$$\frac{\partial B_x}{\partial y}=i\mu\sigma E_z \qquad \text{(3.27 a. b.)}$$

For 2D interpretation (we will use 2D in our research) both TE and TM mode are used, together or separately, depending on sensitivity for structure depth or robustness to the 3-D effects caused by real geological bodies.

Impedance tensor

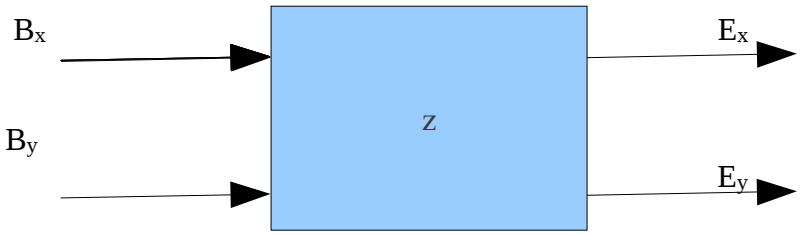


Figure 3.1 shows how components of magnetotelluric tensor are mutually connected

In the presence of multidimensional structures the apparent resistivity is dependent on the source field polarization and the measurement directions of the electric and magnetic fields. Here tensor formulation for electrical impedance will be introduced.
The electrical impedance Z [mV /T] is the ratio between the electric and magnetic field components, which comes from the matrix form relation: E = ZB.

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} Z_{xx} Z_{xy} \\ Z_{yx} Z_{yy} \end{pmatrix} \begin{pmatrix} B_x \\ B_y \end{pmatrix} \qquad (3.28)$$

This linear coupling between the different components is not proved but holds within measurement accuracy (Rokityansky,1982)

Each tensor element is $Z_{ij} = E_i / B_j$ (3.29)

In a 2-D earth the diagonal elements of Z vanish (in the 2-D strike coordinate system):

$$Z_{xx} = Z_{yy} = 0 \qquad (3.30)$$

In a 1-D layered earth, besides having vanishing diagonal elements, the off-diagonal elements are related in the form:

$$Z_{xy} = -Z_{yx} \tag{3.31}$$

Processing of the time series to yield impedance tensor estimates is referred to as data processing, and MT responses (impedances) output from the data processing are referred to as MT data . Advanced data analyses techniques are required in order to make use of the MT measurements in many cases contaminated by industrial noise . MT functions are used to construct earth conductivity models and are a critical link between the observed data and the inferred conductivity models. Magnetotelluric transfer functions are used to represent the frequency relationship between electric and magnetic data. Reliable estimates of the transfer functions are therefore essential and this depends on selecting the correct frequency relationship between the electric and magnetic data and the statistical method used to obtain unbiased estimates of the transfer function.

Data processing steps:

- time series preprocessing
- spectral analysis
- statistical analysis
- post processing

More details about this can be found in Smirnov and Varentsov(2004) and Egbert(2002)

3.2. Applications of MT

Due to the big range of penetration in MT, this method has found numerous applications in both academy and industry. Here we are giving some of them:

- Geothermal studies

  It has been used last 20-30 years in geothermal energy, but with high energy prices it is getting more popular so these studies are of interest to both academia and industry. MT is particularly suitable for geothermal areas because it is capable to detect both water reservoirs and magmatic bodies which produce conductive anomalies. There are several mechanism by which hydrothermal fluids can change the conductivity of rocks (altering rocks, increasing salinity) due to high temperature.

  In Iceland, LBNL researchers (Gasperikova *et al.*, 2011) have investigated two large geothermal fields with 3D MT, showing that this approach is very promising in imaging geothermal reservoirs in a single self-consistent manner . In Hungary, research was carried out with aim to delineate potential geothermal locations, after this cooperative inversion was performed with gravity. Successful results, interpreted two deep faults and one shallow fault, all of them potential geothermal targets (Strack and Vozoff, 2010).

  In northern Portugal (Chaves region), Monteiro Santos (Monteiro Santos *et al.* ,1995) have lead investigations of low enthalpy field. Magnetotelluric measurements have been carried out to determine the crustal electrical conductivity distribution. Interpreted data indicate the existence of a conductive layer and a first estimate of the regional geothermal gradient.

  Newman (Newman *et al.*, 2005) have characterized Coso geothermal field using three-dimensional magnetotelluric inversion .The initial 3D resisitivity model clearly shows the controlling structures that could possibly influence well production at Coso. The inverted model indicate geological structure that may correlate with a zone of higher permeability and fluid content.

- Deep crustal and mantle studies

  The aim of crustal and mantle studies is to improve knowledge of the internal structure and plate-tectonic processes in the regions of with complex geology. Numerous examples:

Booker and Egbert(1986) have dealt with long period MT to determine the mantle conductivity. Jones et al.,(2009) have carried out surveys at Archean Slave craton in northwestern Canada, which has become an international focus of broad geoscientific investigation following the discovery of economic diamondiferous kimberlite pipes . Their research has revealed anomaly in electrical conductivity collocated with kimberlite field.

Also group of researchers from IDL (University of Lisbon) has also conducted research in Portugal. Soundings have been made across main tectonic structures and detailed 3D model of graben area was constructed (Monteiro Santos *et al*. ,1999)

- Earthquake precursor studies

  Field measurements have documented electromagnetic signals which are attributed to precursory changes which were followed ultimately by earthquakes . Precursory electric field changes have been used to issue earthquake predictions. In Japan, Geographical Survey Institute of Japan has set up MT monitoring systems. Similar activities were performed by Swiss Seismological Service , to complement standard seismological procedures.

- Oil and mineral exploration

  Because of seismic method limitations, MT methods are becoming important supplementary tool to seismic surveys for petroleum exploration . Hydrocarbon exploration in the Papuan fold belt is made extremely difficult by mountainous terrain, equatorial jungle and thick karstified Miocene limestones at the surface. Mike Hoversten pointed out that the high-velocity karstified limestones at or near the surface often render the seismic technique useless for imaging the subsurface because they scatter the seismic energy. In such areas magnetotellurics (MT) provides a valuable capability for mapping subsurface structures (Hoversten, 2006).

  In marine exploration,MT is also becoming more and more important. Petroleum and natural gas are often trapped in pockets between rock and salt that were formed by the upward movement of salt towards the surface, penetrating and bending the existing rock. Internal salt structure geometry can be complicated as well, leading to ambiguities in the interpretation of seismic reflection data. With the resistivity of salt is often more than ten times greater than the surrounding sediments, MT could be used for solving these issues In areas where seismic imaging of the base of salt structures is difficult, electromagnetic techniques offer

complementary as well as independent structural information (Hoversten, 1998). He proved that the mapping of the base of the salt with marine MT is virtually unaffected by internal variation within the salt. Further research of Steven Constable and Kerry Key, with new broadband marine MT instrument (which makes frequency range much larger) at Gemini Prospect (Gulf of Mexico) showed that 2d inversions can recover the salt body remarkably well despite its complex 3D shape (Key, 2003)

- Environmental studies

  Numerous applications, Martin Unsworth(Unsworth et al., 1997) has been investigated using a variety of EM exploration techniques for potential radioactive waste disposal site at Sellafield (Great Britain). Controlled-source audio-frequency magnetotelluric (CSAMT) data have given the best subsurface information in an environment that has a high level of cultural noise. The resistivity models obtained after the inversion showed good agreement with well log data collected at the site .

  MT has been also useful for buried waste characterization when used at radio (RMT) frequencies. Buried waste sites represent a critical environmental problem .It often has an electrical conductivity much higher than the surrounding host material and therefore MT or any of other EM methods present good choice for investigating the extent of the waste. Further repetition of these measurements can help to monitor the changes in conductivity within the waste site and the surrounding . Newman *et al.* (2003) have done RMT over a buried waste site near Cologne, Germany. Inverting these data using a 3D MT inversion scheme they successfully recovered the shape and extent of the contamination, which agrees with the borehole data. They have incorporated prior data to get the sharper images. Also there work pointed out differences in 3D and 2D inversion process.

Apparent resistivity and phase data were modeled using regularized 2D MT inversion code OCCAM2DMT (UCSD).

MT data set s consists of a finite number of imprecise data, an infinity of solutions to the inverse problem exists. This code generates smooth models by restricting set of solutions. Smooth models are less likely to result in over interpretation of the data and reflect the true resolving power of the MT method. Since smooth inversions only output structures required by the data, they are often referred to as as minimum structure inversions. Thus, a structure not sensed by the data will not be imaged in the inversion models. Together with limiting the class of acceptable models, the smoothness requirement also increases the convergence rate .

Inversion is explained in chapter 2.

We start from forward problem

$$d = F[m] \tag{3.32}$$

where d is a vector of apparent resistivity and phase data and F is the forward operator that computes the MT response for a given set of model parameters m. The forward operator F is calculated using a standard 2D MT finite element modeling program (Wannamaker *et al.*, 1987)

The inversion is performed with Lagrange multiplier formulation by minimizing the following unconstrained functional

$$U[m] = \|\partial m\|^2 + \mu^{-1}\left\{\|Wd - WF[m]\|^2 - X\right\} \tag{3.33}$$

where

|| || - Euclidean norm

∂ - roughening matrix

1/μ - Lagrange multiplier

W - diagonal matrix of weights determined from the uncertainty of each data point

X - the acceptable level of data misfit

The functional is minimized at points where the gradient with respect to the model is zero. First term accounts for the model roughness,second how well the model fits the data. The Lagrange multiplier acts as a weighting term that is adjusted to strike a balance between minimizing the data misfit and minimizing the model roughness. Detailed description of minimization is given in Constable *et al.* (1987) and de Groot-Hedlin and Constable (1990).

4. Pattern recognition

Pattern recognition is the scientific discipline whose goal is the classification of objects into a number of categories or classes. Depending on the application, these objects can be images or signal waveforms or any type of measurements that need to be classified. (Koutroumbas and Theodoridis, 2009)

Pattern recognition arises from the need for automated decision-making based on a given set of parameters. Despite over half a century of productive research, pattern recognition continues to be an active area of research because of many unsolved theoretical problems as well as a rapidly increasing number of applications that can benefit from pattern recognition. As computers are getting more and more powerful, this has enabled faster processing of huge data sets and using diverse methods for data analysis and classification .At the same time, demands on pattern recognition systems are rising enormously due to the availability of large databases and stringent performance requirements especially regarding speed and accuracy.

Terminology

"A set of variables believed to carry discriminating and characterizing information about an object to be identified are called **features,** which are usually measurements or observations about the object. A collection of d such features, ordered in some meaningful way into d-dimensional column vector is the **feature vector**, denoted x, which represents the signature of the object to be identified. The d-dimensional space in which the feature vector lies is referred to as the feature space. A d-dimensional vector in a d-dimensional space constitutes a point in that space. The category to which a given object belongs is called the **class** . A collection of features of an object under consideration, along with the correct class information for that object, is then called a **pattern**. Any given sample pattern of an object is also referred to as an instance or an exemplar. The goal of a pattern recognition system is therefore to estimate the correct class corresponding to a given feature vector based on some prior knowledge obtained through training. **Training** is the procedure by which the pattern recognition system learns the mapping relationship between feature vectors and their corresponding class. This relationship forms the decision boundary in the d-dimensional feature space that separates patterns of different classes from each other. Therefore, we can equivalently state that the goal of a pattern recognition algorithm is to determine these decision boundaries, which are, in general, nonlinear functions " (Polikar, 2006).

A quantitative measure that represents the cost of making a classification error is called the **cost function.** The pattern recognition algorithm is specifically trained to minimize this function. A pattern recognition system that adjusts its parameters to find the correct decision boundaries, through a learning algorithm using a training dataset, such that a cost function is minimized, is usually referred to as the **classifier**. Incorrect labeling of the data by the classifier is an error and the cost of making a decision, in particular an incorrect one, is called the **cost of error** .

Components of pattern recognition system

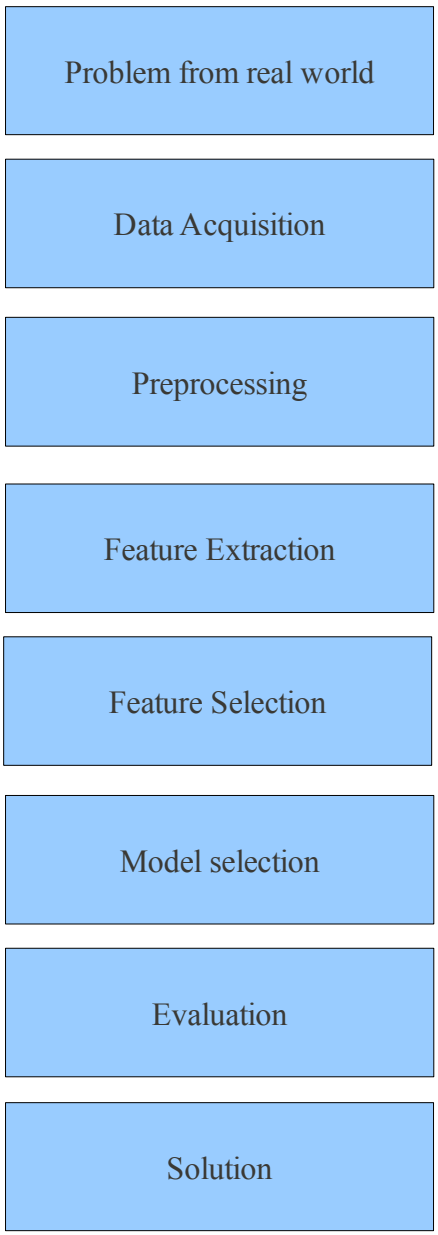


Figure 4.1 shows flowchart for pattern recognition (Robi Polikar, 2006)

Data Acquisition

"Having adequate and representative training and test datasets is important requirements for designing a successful pattern recognition system " (Polikar,2006). Adequacy ensures that a sufficient amount of data exists to learn the decision boundary as a functional mapping between the feature vectors and the correct class labels. Representative data ensures that all meaningful variations of field data instances that the system is likely to see are sampled by the training and test data . Meeting this requirement is quite difficult, because it is hard to determine whether the training data distribution adequately spans the entire space on which the problem is defined .

Preprocessing

This is important step with multiple goals. One goal is to condition the acquired data such that noise from various sources are removed to the extend that it is possible (filtering techniques can be employed). Conditioning may also include normalization, as classifiers are known to perform better with feature values that lie in a relatively smaller range. Normalization can be done with respect to the mean and variance of the feature values or with respect to the amplitude of the data. In FCM algorithm, instances are normalized such that the normalized data have zero mean and unit variance . The reason is that resistivities cells are in the range from 10-1200, and P wave velocities from 3-6km/s. Preprocessing should also include outlier removal when possible (not in this case because it would require much more complicated numerical computations)

Feature Extraction

The goal of feature extraction is to find preferably small number of features that are particularly distinguishing or informative for the classification process, and that are invariant to irrelevant transformations of the data .

Feature Selection

Selecting a subset of features from a set of features that have already been identified by a preceding feature extraction algorithm. There is more reasons to reduce the number of features to a sufficient minimum (e.g high mutual correlation between feature vectors).

Model selection

Only after acquiring and preprocessing data and extracting and selecting the most informative features is one finally ready to select a classifier and its corresponding training algorithm .

Performance evaluation

Evaluation depends on the problem we have (accuracy , 'detection rate', precision).

Cluster analysis is the art of finding groups in data (Kauffman and Rousseeuw, 2005). This simple definition is different from other stuff that can be found. The definition of clustering leads directly to the definition of a single “cluster” , so if know what cluster is then we know what clustering presents (circular way). Everitt defines vectors as points in the l-dimensional space, and the clusters are described as “continuous regions of this space containing a relatively high density of points, separated from other high density regions by regions of relatively low density of points” . This definition is closer to our visual perception of clusters in the two- and three-dimensional spaces.

Clustering is one of the most primitive mental activities of humans, used to handle the huge amount of information they receive every day . We are used to categorize objects that surrounds us into clusters , each cluster is then characterized by the common attributes. By learning about the world around him child learns to separate into groups (cars or trains or cats or dogs). Next time when child meets a cat, child does not expect cat to bite him (not from that category).

Historically, classification has played important role in science. In the last half century, new algorithms have led to widespread cluster analysis in both natural and social sciences (biology, medicine, psychology, machine learning, information retrieval, image analysis).

The aim of cluster analysis is to partition a given set of data or objects into clusters (subsets, groups, classes). This partition should have the following properties:

- Homogeneity within the clusters (data that belong to the same cluster should be as similar as possible)
- Heterogeneity between the clusters (data that belong to different clusters should be as different as possible)

“The greater the similarity (or homogeneity) within a group and the greater the difference between groups, the better or more distinct the clustering”(Kumar, 2006). The concept of "similarity" has to be specified according to the data. Various similarity measures have been used for discovering clusters. ” In metric spaces, similarity is often defined by means of a distance norm” (Duda *et al.*, 2000). Since the data are in most cases real-valued vectors, the Euclidean distance between data can be used as a

measure of the dissimilarity . One should consider that the individual variables (components of the vector) can be of different relevance. In particular, the range of values should be suitably scaled in order to obtain reasonable distance values . But the whole process from data collection to clusters is dependent on types of data, different types are processed and classified into clusters in different ways, usually following a trial-and-error strategy.

“The performance of most clustering algorithms is influenced not only by the geometrical shapes and densities of the individual clusters, but also by the spatial relations and distances among the clusters. Clusters can be well-separated, continuously connected to each other, or overlapping each other.” (Babuska, 2009)

Steps in clustering process (Koutroumbas ,Theodoridis):

- Feature selection-Features must be properly selected so as to encode as much information as possible concerning the task of interest. The aim is to identify a subset of features without deforming the original representation . Preprocessing of features may be necessary prior to their utilization in subsequent stages .
- Proximity measure-This measure quantifies how “similar”or “dissimilar”two feature vectors are. It is natural to ensure that all selected features contribute equally to the computation of the proximity measure and there are no features that dominate others. This must be taken care of during preprocessing.
- Cluster criterion-When do samples make a cluster?No straightforward answer, This criterion depends on the interpretation the expert gives to the term sensible, based on the type of clusters that are expected to underlie the data set. For example, a compact cluster of feature vectors in the l-dimensional space, may be sensible according to one criterion, whereas an elongated cluster may be sensible according to another. The clustering criterion may be expressed via a cost function or some other types of rules.
- Clustering algorithms-Having adopted a proximity measure and a clustering criterion, this step refers to the choice of a specific algorithmic scheme that unravels the clustering structure of the data set. The choice of a clustering algorithm depends both on the type of data available and on

particular purpose.

- Validation of the results- Once the results of the clustering algorithm have been obtained, we have to verify their correctness. This is usually carried out using appropriate tests. A fundamental drawback of all partitioning clustering techniques is that the number of clusters must be chosen ahead of the clustering process. Because the optimum number of clusters for a given data- base is seldom known a priori, the clustering procedure must be repeated for different values of c (e.g., Höppner 1999.)
- Interpretation of the results-In many cases,the expert in the application field must integrate the results of clustering with other experimental evidence and analysis in order to draw the right conclusions.
- But some questions always on our mind. Firstly we do not know if available data possess a clustering structure at all .Given data from a new problem, can we determine whether there exists a clean decision boundary between the classes? To what extent can this boundary be inferred by the algorithms? Which classifiers can do the best job?

As already mentioned in chapter 4, development of computers has enabled more complex numerical computations, which had a positive effect in geophysics. With aim to integrate different geophysical data sets, new ideas are progressing:

Urbat has used FCM algorithm for assessing behaviour of geochemical and magnetic properties. They have managed to identify boundaries of magnetic minerals, also to asses diagenetic processes(Urbat *et al.* 1999) .

Finol employed fuzzy rule-based systems to help predict permeability in sedimentary rocks using well-log responses . Numerical simulation results show that the fuzzy system is an improvement over conventional empirical methods in terms of predictive capability. The benefits of this approach are demonstrated with a case study in the Lake Maracaibo basin, Venezuela. (Finol *et al.,* 2002)

Vernieuwe (Vernieuwe *et al., 2007*) has described groundwater flow in the unsaturated zone with help of fuzzy models. Alternative to classical Richards equation approach, fuzzy models were evaluated and results compares for different fuzzy algorithms.

Guler and Thyne (2004) have classified hydrochemical data set (water samples and hydrochemical variables) from southeastern California by FCM and HCA clustering techniques to obtain hydrochemical facies delineation. Results from both FCM and HCA clustering produced cluster centers (prototypes) that can be used to identify the physical and chemical processes creating the variations in the water chemistries. Their results show that a fuzzy classification gives improved spatial definition and less error sensitive clusters than a classical hard classification .

Alvarez Grima (2000) proposed use of nerro-fuzzy methods for groundwater flow models and control of excavation machinery and also for risk assessment of complex engineering geology projects.

At the ETH Zurich Paache (Paache *et al.* , 2006) have combined information contained in the physical-property models that result from inverting the individual data sets (porosity, georadar velocity, georadar attenuation, seismic velocity) to derive multiparameter model which effectively outlines the major sedimentary units . In addition to determining physical-property models from the different geophysical data sets, it is often necessary to use the same data for estimating the distribution of various petrophysical parameters (gamma ray and hydraulic conductivity). An important outcome of this study is the development of a method for estimating the spatial distribution of petrophysical pa- rameters‘e.g., lithologies, porosities, and hydraulic conductivities’ that cannot be determined uniquely from the physical-property models or geophysical data

Paache and Eberle have employed fuzzy c-means cluster analysis for the rapid and largely automated integration of complementary geophysical datasets comprising airborne radiometric and magnetic as well as ground-based gravity data . FCM cluster analysis enabled automatic generation of a zoned integrated geophysical map delineating distinguished sub-surface units based upon the information contained by each dataset .Altogether with available geological information from the study area,this allowed assigning a sound geological meaning to the zones (clusters)(Paache *et al.* ,2009).

Fuzzy sets were introduced in 1965 by Lofti Zadeh. They were a generalization of conventional set theory,one of the basic structures underlying computational mathematics and models. The concept of fuzzy sets emerged when it was realized it may not be possible to model ill-defined systems with precise mathematical assumptions of classical methods. The basic logic of the fuzzy set theory is that it allows event to belong to more than one sample space where sharp boundaries between spaces are hardly found (Zadeh, 1965).

Children learn how to interpret fuzzy instructions presented to them (“go to bed about 10”) but adults also (“will get married before 40”). Somehow in our daily lives, we all use vague data and imprecise information from which we make decisions,sometimes very important ones. “Computational models of real systems should be able to recognize, represent, manipulate, interpret and use both fuzzy and statistical uncertainties” (Bezdek, 1993). Computational pattern recognition has played a central role in the development of fuzzy models because fuzzy interpretations of data structures are natural way to formulate and solve various problems.

Let X be space of points, with a generic element of X denoted by x. A fuzzy set (class) A in X is characterized by a membership(characteristic) function $f_A(x)$ which associates with each point in X a real number in the interval [0,1], with the value of $f_A(x)$ at x representing the grade of membership of x in A. The nearer value of $f_A(x)$ to unity,the higher the grade of membership of x in A. Membership function determines how much fuzziness a fuzzy set contains. Here analogy can be drawn with information theory, ”how much information is contained in message”. In fuzzy sets , quantification of the amount of imprecision captured depends on the extent to which the objects do or do not possess the concept or property represented by fuzzy set.

In many real situations, fuzzy clustering is more natural than hard clustering, as objects on the boundaries between several classes are not forced to fully belong to one of the classes, but rather are assigned membership degrees between 0 and 1 indicating their partial memberships.

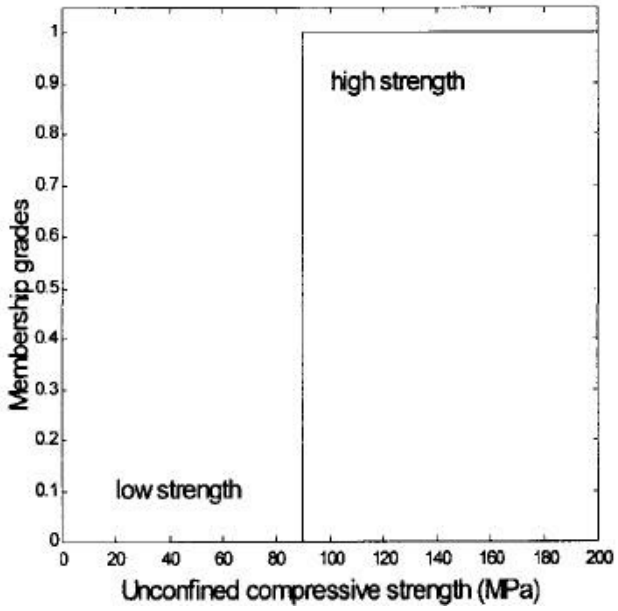


Figure 4.2

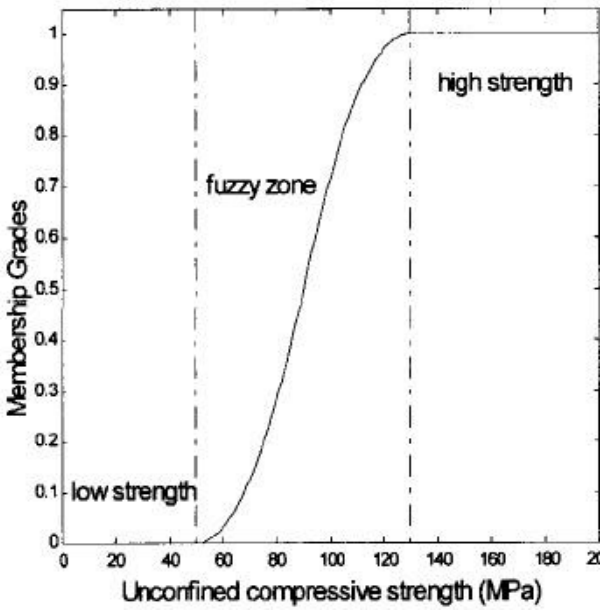


Figure 4.3

Figures 4.2 and 4.3 show how strength of the rock versus membership grades (Alvarez Grima, 2000)

Figures 4.2 and 4.3 illustrate strength of the rock fuzzy explanation. Lets assume the assumption 'the strength of the rock is high'. Using classical set theory, term high can be defined A={UCS | UCS > 90}, UCS unconfined compressive strength. In this case rock strength can be high or not. A problem arises for UCS values of about 100 MPa. For some applications such rocks have high strengh for others low and for some applications cane be regarded as a minimum. This type of uncertainty can be handled by using fuzzy sets. Lets assume now that a classical set high is modified to fuzzy set and is represented via a predefined membership function as indicated in Figure 2.2. Now we define rock strength as a continuous function that allows incorporation of expert knowledge and fuzziness in term high.(Alvarez Grima, 2000)

Fuzzy-set algorithms outperform other conventional methods when they are appropriately applied to solve the problems of fuzzy uncertainty. Main features of fuzzy methodologies (Bezdek, 1983) :

1. Fuzzy logic provides a systematic basis for quantifying uncertainty due to the vagueness and incompleteness of information
2. Classes with non-sharp boundaries can easily be modeled using fuzzy sets
3. Fuzzy reasoning is a formalism that allows the use of expert knowledge, and is able to process

this expertise in structured and consistent way

4. There is no broad assumption of complete independence of evidence to be combined using fuzzy logic, as required for other subjective probabilistic approaches
5. When the information is inadequate to support random definition,the use of probabilistic methods may be difficult,fuzzy methods are good alternative in such cases

C means algorithm

The fuzzy c-means (FCM) clustering algorithm is a set-partitioning method based on Picard iteration through necessary conditions for optimizing a weighted sum of squared errors objective function(J_m).

$$J_m(U, v) = \sum_{k=1}^{N} \sum_{i=1}^{c} (u_{ik})^m (\|y_k - v_i\|)_A^2 \qquad (4.01)$$

$Y=\{y_1, y_2, ..., y_N\}$ data
c=number of clusters in Y
m=weighting exponent
U=fuzzy c partition of Y
$v=(v_1, v_2, ..., v_c)$ vectors of centrers
$v_i=(v_{i1}, v_{i2}, ..., v_{ic})$ centers of clusters i
$\| \|_A$ induced A-norm
A positive definite(n x n) weight matrix

The squared distance between y_k and v_i shawn in equation 4.01 is computed as A-norm

$$d_{ik}^2 = \|y_k - v_i\|_A^2 = (y_k - v_i)^T A (y_k - v_i) \qquad (4.02)$$

The weight attached to each squared error is $(u_{ik})^m$,the m^{th} power of y_ks membership in cluster i. The vectors v_i are regarded as centers of mass of partitioning subsets. Weighting exponent m controls the relative weights placed on each of the squared errors d_{ik}^2 (if m is getting close to 1 partitions are

becoming increasingly hard,m=1 hard partition). On the other side,increasing m tend s to degrade(blur) membership towards the fuzziest state. Each choice for m defines, all other parameters being fixed, one FCM algorithm. No theoretical or computational evidence distinguishes an optimal m.

Weight matrix A deserves special discussion. This matrix controls the shape that optimal clusters assume. Norm is always defining a generalized notion of "distance" between two points/vectors. Norm is inner product via the formula:

$$<x,y>_A = x^T A_y \qquad (4.03)$$

The norm influences the clustering criterion by changing the measure of similarity(or dissimilarity). The norms of greatest interest for use with equation (4.01) correspond to

A=I ,what is euclidean norm

$A=C_y^{-1}$,Mahalonobis norm

The other will be explained more detail in next chapter.

Optimal fuzzy clustering of Y are defined as pairs (U,v) that locally minimize J_m

For m>1 ,if y_k =/ v_j for all j and k, (U,v) may be the local optimum for J_m only if

$$v_i = \frac{\left(\sum_{k=1}^{N} (u_{ik})^m y_k\right)}{\left(\sum_{k=1}^{N} u_{ik}^m\right)} \quad ; \quad 1 \leqslant i \leqslant c$$

$$u_{ik} = \left(\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}}\right)^{\left(\frac{2}{(m-1)}\right)}\right)^{-1} \quad ; \quad 1 \leqslant i \leqslant c, 1 \leqslant k \leqslant N$$

Conditions expressed in equations are necessary, but not sufficient; they provide means for optimizing Jm via a simple Picard iteration, by looping back and forth from equations until the iterate sequence shows but small changes in successive entries of U and v.

Basic algorithmic steps:

1. Fix c,m,A, $||k||_A$ .Choose an initial matrix $U^{(0)}$
2. Compute means $v^{(k)}$, i=1,2,....,c
3. Compute an updated membership matrix $U^{(k+1)}=u_{ik}^{(k+1)}$
4. Compare $U^{(k+1)}$ to U^k in any convinient matrix norm. If $||U^{(k+1)} - U^k|| < \varepsilon$,stop. Otherwise set $U^k= U^{(k+1)}$ and return to step 2.

Bezdek has established the convergence of a class of clustering procedures. He proved that the arbitrary sequences generated by these (Picard iteration) procedures always terminates at a local minimum or always contains a subsequence which converges to a local minimum of the generalised least squares objective function which defines the problem. Numerical experiments with real data in various applications have confirmed this theory.

Gustafson and Kessel (Gustafson and Kessel 1979) proposed that the matrix A in equation (4.01) be a third variable . They have modified the objective function by putting $A=(A_1,A_2,..,A_c)$, A_i being symmetric and positive-definite matrix

$$J_{m,GK}(U,v,A) = \sum\sum u_{ik}^{M} ||y_k - v_i||^2_{Ai} \quad (4.06)$$

The main disadvantage of algorithms with constant norm (FCM is one example) is that they search for clusters with the shape that may not occur in data. In comparison with fuzzy c means algorithm, in addition to the cluster centers each cluster is characterized by A. The variables estimated by Gustafson and Kessel model are the triplet (U,v,A), contrary to previous FCM where only (U,v) were used as optimization variables in the function. This allows the distance norm to adapt to the local topological structure of the data. The important idea here is that the i-th cluster in U might be best matched by a hyperellipsoidal shape generated by the eigenstructure of the variable matrix A_i.

The eigenstructure of the cluster covariance matrix A_i provides information about the shape and orientation of the cluster. The ratio of the lengths of the cluster's hyper- ellipsoid axes is given by the ratio of the square roots of the eigenvalues of A_i. The directions of the axes are given by the eigenvectors of A_i , as shown in Figure 4.4. The GK algorithm can be used to detect clusters along linear subspaces of the data space. These clusters are represented by flat hyperellipsoids, which can be regarded as hyperplanes. The eigenvector corresponding to the smallest eigenvalue determines the normal to the hyperplane, and can be used to compute optimal local linear models from the covariance matrix. (Babuska, 2009)

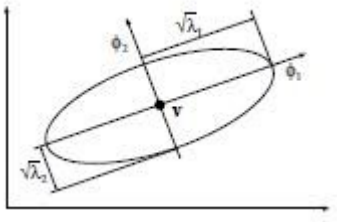


Figure 4.4 shows ellipsoid with axes (eigenvectors)

Equation $(y-v)^T A^{-1} (y-v) = 1$ defines a hyperellipsoid. The length of the jth axis of this hyperellipsoid is given by $\sqrt{\lambda_j}$ and its length is spanned by $\sqrt{\phi_j}$, where $\sqrt{\lambda_j}$ and $\sqrt{\phi_j}$ are the jth eigenvalue and the corresponding eigenvector of A.

J is linear in A_i and we have a singular problem. The cost J may be made as small as desired by simply making A_i less positive definite. What is important here to allow variations in the shape of each class induced by the metric but not let the metric grow without bounds. To get a feasible solution ,we must constrain A_i in some manner. The additional constraint that $\det(A_i) = \rho_i > 0$, guarantees that A is positive-definite; p. is a user defined constant for each cluster. Gustafson and Kessel showed that minimization of $J_{m,GK}$ with respect to A_i leads to the necessary condition.

$$A_i = [\rho_i \det(C_i)]^{(1/p)} C_i^{-1}; 1 \leqslant i \leqslant c \tag{4.07}$$

$$C_i = \frac{\left(\sum_{k=1}^{n} u_{ik}^m (y_k - v_i)(y_k - v_i)^T\right)}{\sum_{k=1}^{n} u_{ik}^m} \tag{4.08}$$

C_i coovariance matrix

v_i represent cluster centers

With this choice,norm for distances becomes

$$d_{ik}^2 = \det(C_i)^{(1/p)} \|y\_k - v\_i\|^2 \tag{4.09}$$

The minimization of the GK objective functional is achieved by using the alternating optimization (AO) method according to the algorithm which is similar to the one from previous chapter(4.1.1). The GK-AO algorithm is more sensitive to initialization than FCM-AO because its search space is much larger.

In order to obtain a satisfactory solution to the problems of large variability in cluster shapes and densities, a two layer clustering strategy has been developed. Gath-Geva (Gath and Geva, 1989) implements fuzzy partition based on maximum-likelihood estimation.

First step, modification of fuzzy K-means algorithm is performed. The fuzzy K-means algorithm is based on minimization of the following objective function

$$J_q(U,V)=\sum_{j=1}^{N}\sum_{i=1}^{K}(u_{ij})^q d(X_j,V_i)^2 \qquad (4.10)$$

X-the j-th m-dimensional feature vector

V_i-centroid of i-th cluster

d^2-inner product metric

N-number of data points

K-number of clusters

q -weighting exponent

Iterative optimization steps:

1. choose primary centroids Vi
2. compute the degree of membership of all feature vectors in all the clusters
3. compute new centroids
4. check if the termination requirement has been met (termination criterion ε), if yes stop,if not go to step 3

Computation of the degree of membership u_{ij} depends on the definition of the distance measure

$$d(X_j,V_i)^2=(X_j-V_i)^T A(X_j-V_i) \qquad (4.11)$$

If A equals to the identity matrix the distance is Euclidean(k-means algorithm).In the other case “exponential” distance measure,based on maximum likelihood estimation is defined. This distance will be used in calculation of h(i|Xj), the posterior probability(the probability of selecting i-th cluster given the jth feature vector).

$$h(i|X_j)=\frac{1/d_e^2(X_j,V_i)}{\sum_{k=1}^{K}1/d_e^2(X_j,V_k)} \tag{4.12}$$

$$d_e^2(X_j,V_i)=\frac{(det\,F_i)^{0.5}}{P_i}\exp((X_i-V_j)^T F_i^{-1}(X_i-V_j)/2) \tag{4.13}$$

where F_i is the fuzzy covariance matrix of the ith cluster, and P_i the a priori probability of selecting the cluster

The new algorithm(FMLE), modification of fuzzy K means. Step 3 of the FMLE algorithm includes,in addition to computation of the new centroid ,calculation of P_i , the apriori probability of selecting the i-th cluster :

$$P_i=\frac{1}{N}\sum_{j=1}^{N}h(i\,|\,X_j) \tag{4.14}$$

Fi fuzzy covariance matrix ith cluster

$$F_i=\frac{(\sum_{k=1}^{N}h(i|Xj)(X_j-V_i)(X_j-V_i)^T)}{(\sum_{k=1}^{N}h(i|X_j))} \tag{4.15}$$

unsupervised tracking of cluster prototypes

The algorithm from previous section starts with initial guesses of classification prototypes and the iterative process results in convergence of the cluster centroids to local optimum.

1. compute average and standard deviation of whole data set
2. choose the initial cluster prototype at the average location of all feature vectors
3. choose an additional classification prototype equally distant from all data points
4. calculate a new partition of the data set according to steps 1 and 2
5. if k,number of clusters is less than a given maximum,goto 3,otherwise stop

5. Geological setting and data processing

5.1. Geological setting

The present studies are illustrated using geophysical data acquired at the Odeleite test site. Odeleite is an area located in the Iberian Pyrite Belt (IPB), which is the part of the South Portuguese Zone geological terrain(SPZ) .

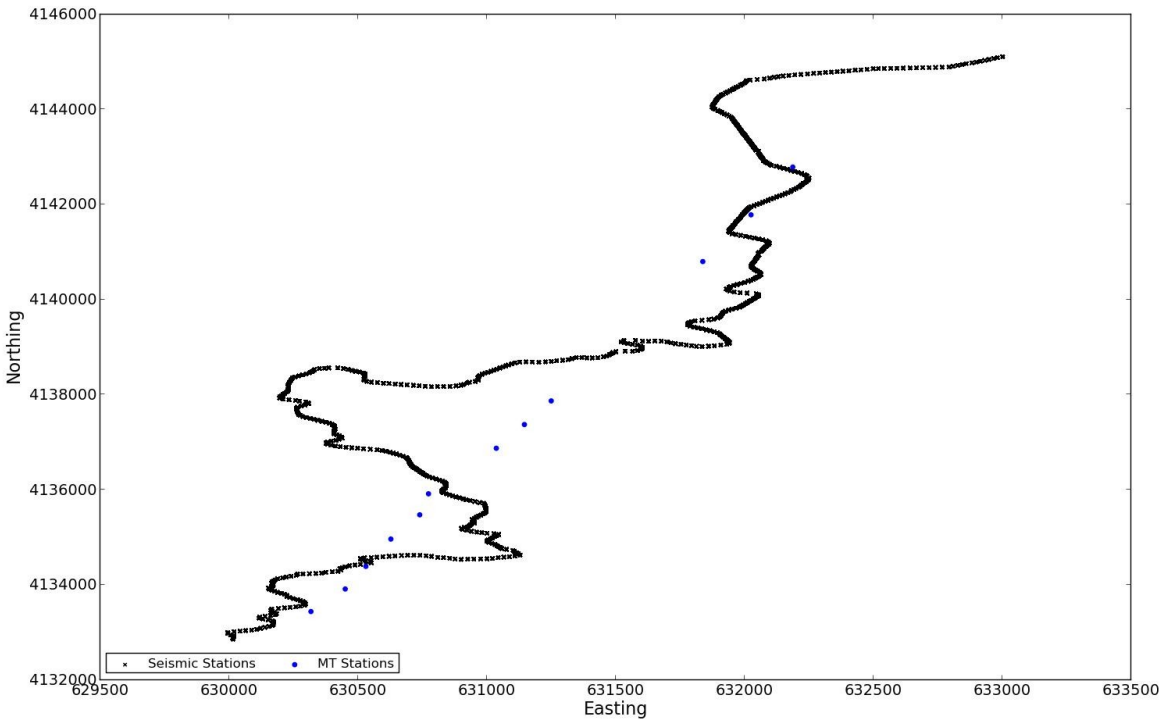


Figure 5.1 shows exact location where MT stations (blue points) and receivers from seismic experiment (presented with black points) were located

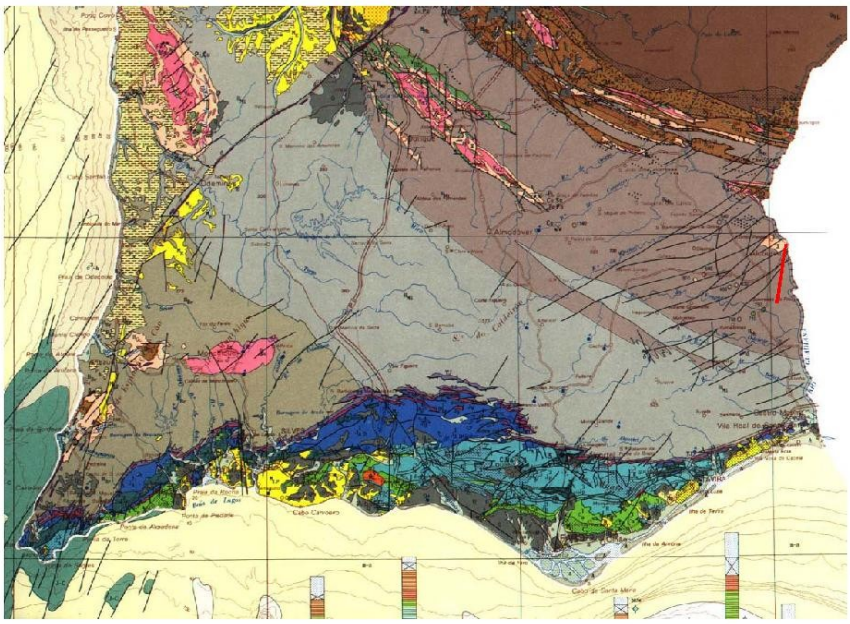


Figure 5.2 shows geological map of Algarve , red line indicates approximate profile location

SPZ is one of the SW Iberia geological units which hosts the Iberian pyrite belt (IPB), an important metallogenic province for volcanogenic massive sulphide (VMS) deposits (Mitjavila et al. 1997; Carvalho et al. 1999). IPB extends across an area in the south-western part of the Iberian Peninsula measuring 250 by 40 km from Atlantic to Guadalquivir basin. The Iberian pyrite belt volcanogenic sulfides are hosted in volcano-sedimentary successions of Upper Devonian to Lower Carboniferous age and represent the great concentration of large massive sulfide deposits of Late Devonian and Carboniferous folded and thrusted sediments, which experienced almost no metamorphism, are exposed within the IPB unit of the SPZ terrane. The stratigraphic succession of the IPB has been established according to the occurrence of volcanic rocks and low- and high-temperature hydrothermal deposits associated with volcanism . The sequence of the IPB rocks is generally divided, from oldest to youngest, into the following three major units: the phyllite-quartzite (PQ) group; the volcano-sedimentary complex (VSC) group and the flysch group (Schermerhorn 1971). Late Devonian sediments of the PQ group crop out only in the cores of anticlines ; they are of terrigenous origin and were deposited on the platform of a continental margin. The Early Carboniferous aged VSC group is placed above the PQ group. The VSC group lithology represents a heterogeneous assemblage of bimodal volcanic rocks inter-bedded with metasediments, hosting the VMS deposits (Carvalho et al. 1999). The third major lithostratigraphic unit in the IPB is the Middle to Late Carboniferous aged flysch group. Several hundred metres to few kilometre thick turbiditic sequences of shale, greywacke and local lenses of intraformal conglomerate and till-like rocks build up this synorogenic flysch unit (Oliveira 1990).

From a structural point of view, the SPZ is a foreland fold-and- thrust belt of the Variscan orogeny, characterized by a southwest vergent imbricated thrust system (Silva et al. 1990; Soriano & Casas 2002). The thin-skinned tectonic deformation of folding and thrust displacement is dated to Middle Carboniferous times.

5.2. Data acquisition and processing 56

5.2.1. Seismic profile

The seismic acquisition was a cooperation between LNEG (Laboratorio Nacional de Energia e Geologia, Portugal) and Prospectiuni (Romania).

The system used was the state of the art cable-less Sercel 428, E-Unite Full with 480 active channels, six in-line geophones per channel, a channel spacing of 25 m and a source interval of 50 m, resulting in 120 full fold data. The source used were three synchronised ION AHV-IV vibroseis trucks of 28.5 tons each. A linear up sweep 16 s long with a frequency varying between 10 to 90 Hz was used to allow recording information of shallow and deeper reflections. Two or three vibrations per vibration points were carried out. Finally, SG 10 Hz receivers were employed to record the seismic wavefield, which was sampled at 2 ms rate during six seconds.

From the line 6 (Figure 5.1), P-wave arrivals were picked along the entire line, then we have selected receiver locations ,which are located on our MT profile. Because sources and receivers of the line are not located a long one straight line, the crooked-line acquisition geometry has to be taken into account. Both source and receiver positions were projected on MT line with maintaining correct source–receiver offset . Schmelzbach (2008) have compared results with this type 2D and 2.5D (when the correct source and receiver geometry is maintained), and very similar results were obtained with both methods. The traveltime difference of reciprocal source–receiver pairs was used to estimate the total picking uncertainty.

shot	picks	uncertainty(ms)	model(km)
40	6579	12	10.6, 1.4

Table 1

2-D magnetotelluric (MT) survey has been carried out on a approximately 11 km long profile .

Measurements have been accomplished using the multichannel geophysical measurement (Metronix Inc). For field work, two data loggers were used : ADU- 06 (DC to 20,000 Hz frequency range, five input channels with a 24 bit A/D converter) and ADU-07 (DC to 250,000 Hz frequency range, ten input channels with a 24 bit A/D converter). Synchronization is accomplished by a GPS clock, that provided exact positioning and time . Three magnetometers and two pairs of non-polarizable electrodes are connected to the data logger. For the registration of magnetic field variations in the range from 10,000 to 0.1 Hz broadband induction coil magnetometers are used. The electric field variations are registered by measuring potential differences with Pb-PbCl-electrodes. The experimental set-up includes four electrodes, which are distributed at a distance of 50 m in north–south and east–west direction. They are buried at a depth of about 50-60 cm and coupling to the soil is improved using water. The ADU logger and magnetometers are located in the centre, while the same time three induction coils are oriented north–south, east–west and vertical at a distance of 10 m from the data logger and at least 1 m from electric field wires. The vertical coil was buried and covered by a plastic tube in order to prevent recordings from the influence of wind . Power was supplied by a external battery.

After timeseries have been collected , next step was data processing.

Processing the recorded time series includes filtering and re-sampling of the time series and the calculation of spectral matrices at determined frequencies. Then the impedance tensor is estimated from the spectra by statistical means. In this study, two independent processing softwares were employed, processing with Mapros (Metronix Inc.) and with WinGLink (Geosystems). The processing results were stored in Electrical Data Interchange (EDI) format.

Firstly cascade decimation was performed. After that we discard all low quality segments of time series, this was done manually by visual inspection. Complete time series are divided into overllapping segments. To reduce the spectral leakage each segment was window tapered. After that transformation into frequency domain was done with Fast Fourier Transformation. By averaging a certain bandwidth into evaluation frequency using a Parzen window, the signal to noise ratio was improved . Finally MT transfer functions were calculated. This procedure was performed for each frequency band.

At the very end we have rotated impedance tensor using Torquil-Smith method (also with WinGLink software).

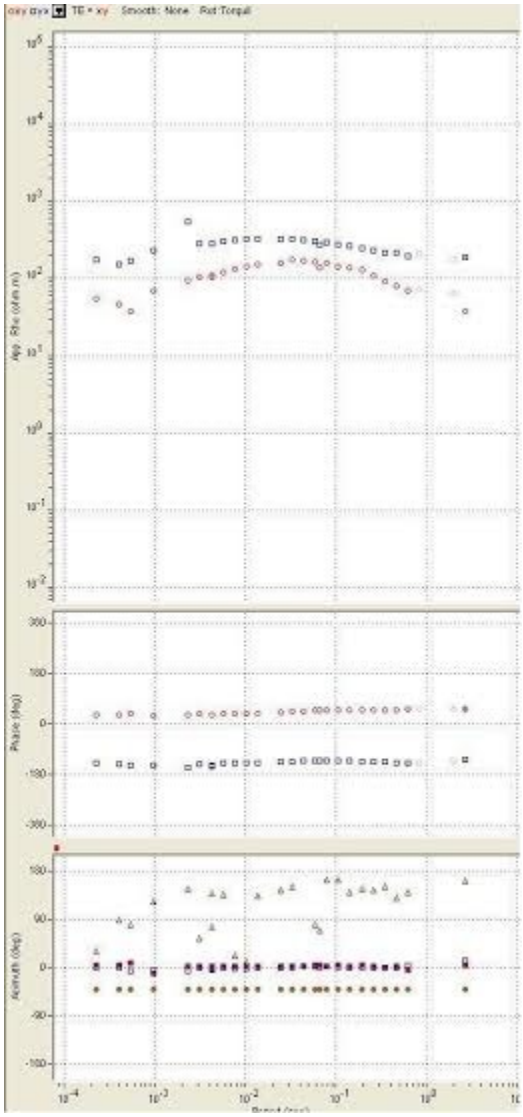


Figure 5.3 shows apparent resistivity,phase and azimuth for one site

6. PARAMETRIZATION AND METODOLOGY 59

6.1. PARAMETRIZATION

As Tarantola(2005) pointed out parameterization of the system equals to discovery of a minimal set of model parameters whose values completely characterize the system (from a given point of view).
In seismic tomography, a lot of research was dedicated to comparisons between different parametrizations, how particular model representations may favorably accomodate one or two boundary conditions and that different model respresentations potentially lead to different results, resolution, and reliability estimates (Haslinger and Kissling, 2001 and Husen and Kissling, 2001). One must consider not only data quality and numerical precision of forward and inverse solvers but also appropriateness of model parametrization . The quality of the forward solution, in particular, strongly depends on parametrization of the velocity field and is of great importance both for calculation of travel times and partial derivatives that characterize the inverse problem. Almost same stands for MT. Both forward and inverse problem are dependent on parametrization.

In order to combine two methods, choice of model parametrization was somehow compromise between seismic and MT inversion codes (the first one has velocity structure with nodes, the other conductivity one with cells).
Model dimensions are 10.6km(length), 1.4km(width). It has been divided it into grid with square cells (100m x 100m). For FAST , this leads to input model with 107 nodes in each of 15 rows.

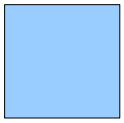


Figure 6.1 One grid element characterised by resistivity ρ and velocity v

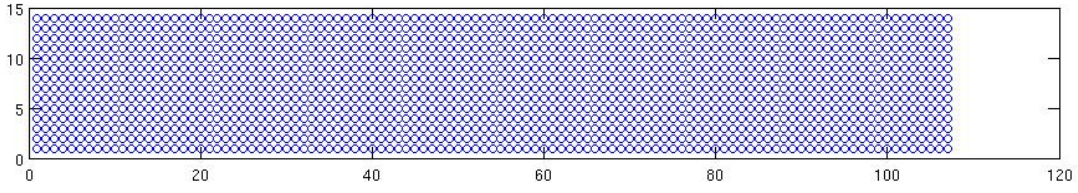


Figure 6.2 Square grid with a total of 1484 grid elements

After the seismic inversion is performed, 1605 node velocities were obtained. To go for clustering, cell model identical to that from Occam code was created, simply by taking the average value of 4 neighbouring nodes for each cell.

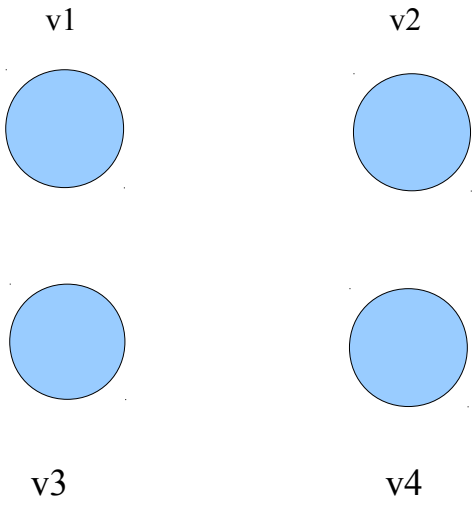


Figure 6.3 shows four node velocities that will produce one cell velocity

$$v_{sr} = (v_1+v_2+v_3+v_4)/4 \qquad (6.01)$$

After this model has 1484 cells. Pattern recognition process (explained in chapter 4) creates new models, but now again with seismic part we have problem how to go back from cells to nodes. Parameter value(velocity) from 1484 cells must be mapped into 1605 nodes. At this point it was inevitable (this will also contribute to error but there are no easy solutions) to make approximation of first row velocities by taking 90 % of value from the first row cells. For second row and all other node rows, values were simply calculated by taking the average from 4 neighbouring cells(Formula 6.01). First and last values in every row were obtained as averages of 2 cells (red circles in Figure 6.4).

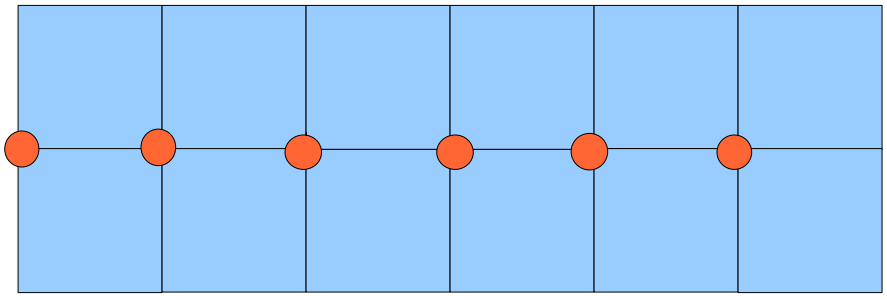


Figure 6.4 shows grid with cells and nodes

To double check validity of going from nodes to cells (and vice versa) calculations were performed using the Gauss - Seidel method. For a synthetic test initial model velocities were inverted again from nodes to cells, overall errors were less than 1% (acceptable level).

All calculation were latter performed with above defined cell values. Different parametrization (make cells 4 times smaller for example) could possibly lead to some improvements, but that is computationally expensive because all the calculation need to be repeated (both for synthetic and for Algarve data sets). Also changing cells to more complex geometries is impossible with these codes (FAST and OCCAM2D)

In this chapter, methodology is presented.

In the first approach, algorithm proposed by Paache and Tronicke (2007) for zonal cooperative inversion of georadar and P-wave velocities, was employed. The basic idea of cooperative inversion approach is to combine conventional single input data set inversion algorithms with FCM cluster analysis with FCM cluster analysis at each iteration. This ensures information exchange between different geophysical data sets and models during the inversion process.

Second and third approaches are similar, only different clustering algorithms were employed (GK and GG respectively), everything else was similar to the first algorithm. The idea behind this, is to try with different algorithms and see how finding groups in data in three different manners will influence inversion. Similarities between variables can arise in quite different ways, as a consequence of this starting models for inversion will not be the same.

In our work, two points similarity (or dissimilarity) is determined by calculating distance between them. In the first algorithm (FCM), distance is Euclidean and in other two Mahalonobis. Mahalonobis distances are based on both the mean and variance of the predictor variables, plus the covariance matrix of all the variables, and therefore take advantage of the covariance among variables. The region of constant Mahalanobis distance around the mean forms an ellipsoid in vector space. Mahalonobis distances are calculated as :

$$D^2 = (x-m)^T C^{-1} (x-m) \tag{6.02}$$

where:

D = mahalonobis distance

x = vector od data

m = vector of mean values

C^{-1} = Covariance matrix

Iterative procedure starts with two disparate data sets, that represent velocity and resistivity structure. With grid cell parametrization scheme, each cell has P wave velocity and resistivity. After this follows normalization of each data set.

Normalization is simply a case of getting all data we have on the same scale: if the scales for different features are wildly different, this can have a knock-on effect on algortihm's ability to extract clusters. If one of the features has a broad range of values, the distance will be governed by this particular feature . Ensuring standardized feature values implicitly weights all features equally in their representation. For the first algorithm scaling is zero mean variance one, for the rest normalization by scaling between 0 and 1. So after this ,we have 2-dimensional model that span an 2-dimensional parameter space. In practice it looks this way, if we have velocity values in range from (2925.25, 5549.5m/s), after zero mean variance one range will be (-2.0837067, 1.1910209), after scaling between 0 and 1 (0,1).

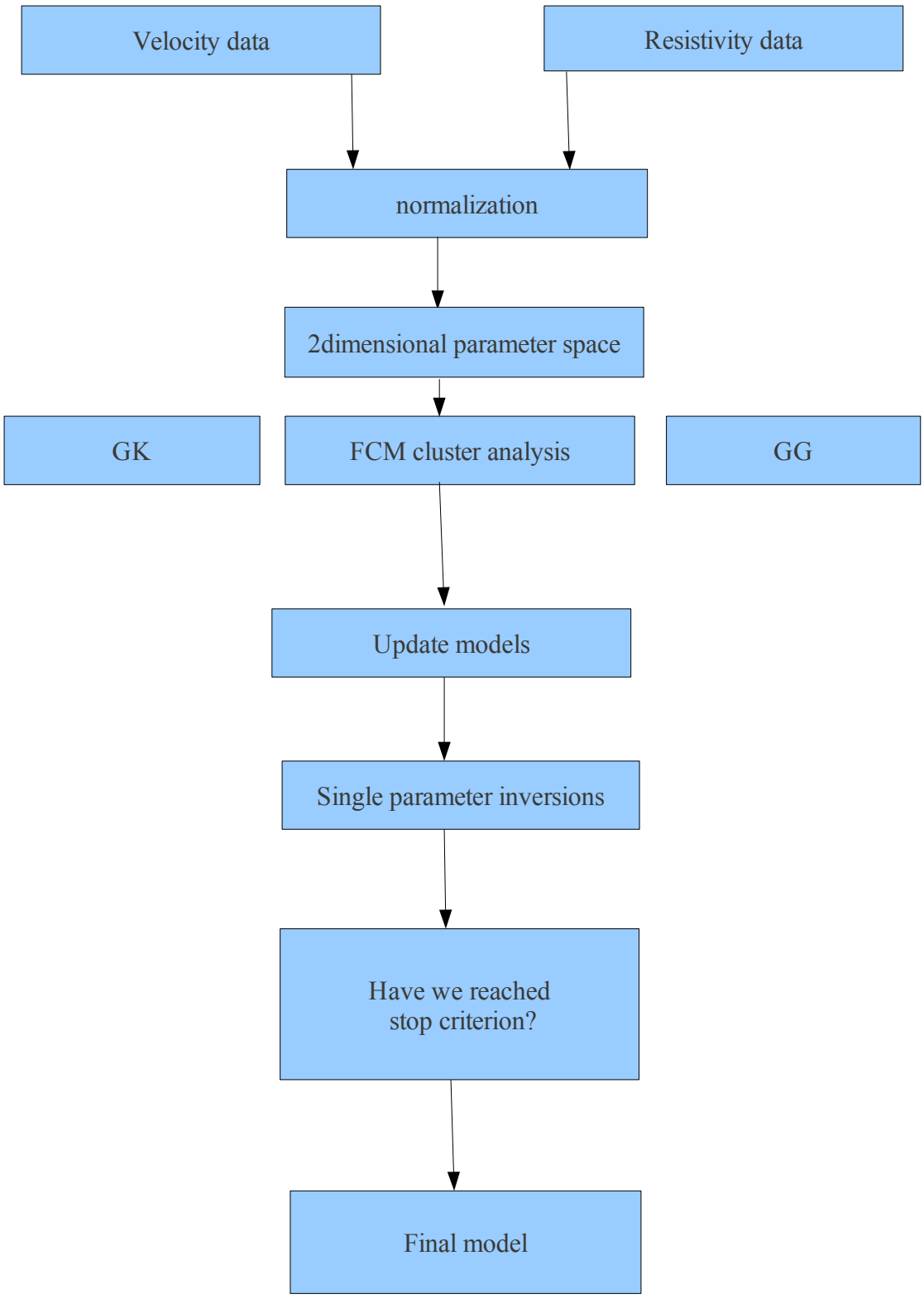


Figure 6.5 flow diagram illustrates cluster analysis based cooperative inversions

Important issue here is that all points in parametric space are going to be organized with respect to their numerical value, no matter how close or distant are grid cells in real geometrical structures. So two points can be close to cluster center in vector space even they represent cells from totally distant parts of the model.

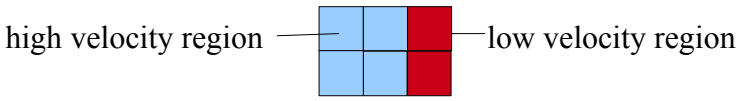


Figure 6.6 illustrates that algorithms do not posess any information about the spatial context.

Ater the cluster centers and memberships are obtained, next step is defuzzifaction. We have allocated points to clusters with maximum memberships (maximum membership degree principle). This step is inevitable but could cause problems especially with large number of clusters, simply because some points have almost same membership values. (for example two points have memberships 0.242 and 0.239)

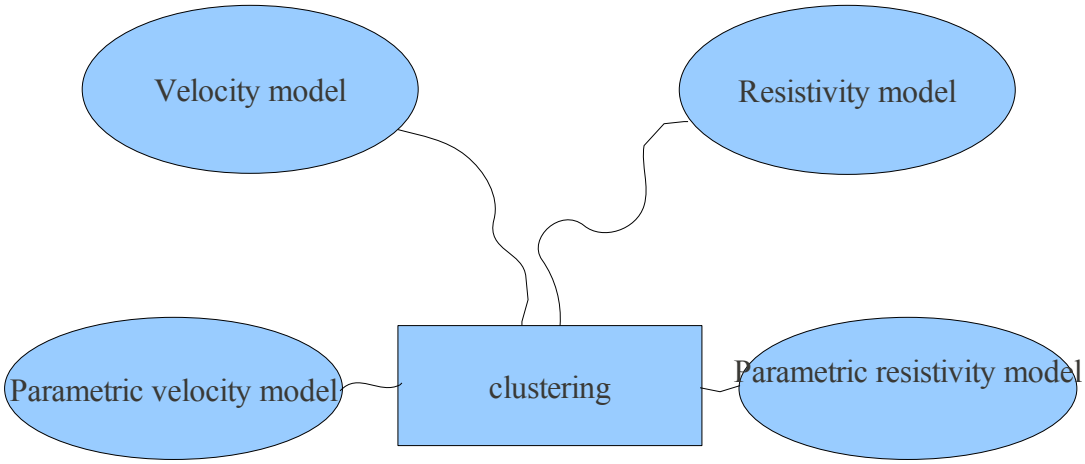


Figure 6.7 illustrates how from initial models, after clustering and defuzzification, new parametric models are obtained(they serve as initial models for next iteration)

Then we go for update of model parameter both for P-wave velocities and resisitivities. The main idea is that new parameter value is a mixture of the corresponding cluster center values weighted with the degree of membership of the model cell to the c clusters.

Formula:

$$np_{ij} = \sum mb_{ij} \times avg_{ki} \qquad (6.03)$$

mb_{ij} memberships

np_{ij} new model parameters

avg_{ki} mean value of parameter for the cluster

For all models, we compute rms errors (in ms for travel times, dimensionless units for MT functions) from the corresponding forward calculations and the model updates according to the involved single input data set inversion routines. Updated models (new velocity structure and new resistivity structure) are then used in the following iteration, and we repeat this central inversion loop until a stopping criterion is met; that is, until the data misfits cannot be improved by further iterations. This procedure results in a zonal model in which each of the zones is characterized by similar or consistent relationships between the individual model parameters. The basic idea behind this method is to develop a process through repeated cycles which would at the end produce consistent zonation models. So the aim is to have a lot of iterations,with improvements at each step.

The procedure is repeated for different number of clusters(for 3 algorithms).

Further calculations were done regarding efficiency of optimization in fuzzy c-means. In our algorithm we use MATLAB fcm function,so we wanted the compare results for cluster centers obtained this way and with hybrid methods (AO+ Quasi-Newton method)(Hu and Hathaway, 2002) . Results obtained give cluster centers and objective function converging to same values, cluster center coordinates differences were on the order 10^{-3}. This confirms that MATLAB function fcm is accurate enough for our problem,no need to worry about numerical precision.

To try how different algorithms work, synthetic data sets were generated. Data has not been contaminated with additional noise. The algorithms were designed for the integration of two geophysical methods which are mathematically described with different set of equation (diffusion and wave equation). The multi-step process with inverse problems and clustering at each step, should guarantee that we get results that are within expected range of the parametric values and these result could be integrated into geological model. The objective of work with synthetic data is to see how reliable our algorithms are, does it make sense at all to exchange the information with pattern recognition algorithms and if the structural features can be detected.

It was necessary to create data as geological structure is known. It is important that synthetic data are realistic. Therefore, the values for the parameters of the initial models correspond to real values (of geological structures). Initial models are with 6 regions, they are presented in Figures 7.1 and 7.2.

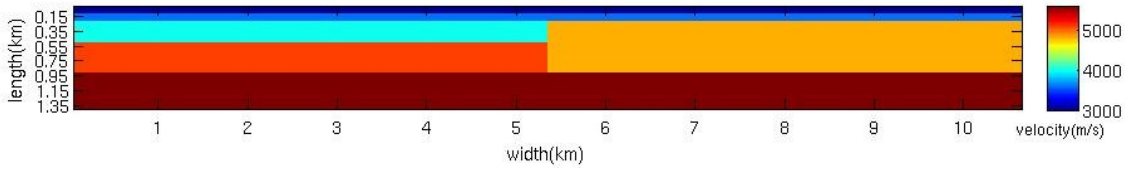


Figure 7.1 shows velocity model

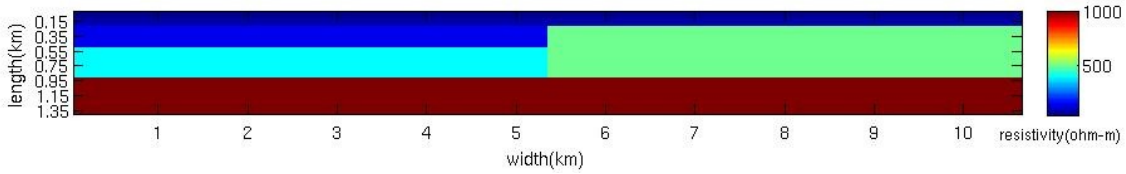


Figure 7.2 shows resistivity model

cell	velocity(m/s)	resistivity(Ωm)
1	3000	50
2	3600	80
3	4000	150
4	4800	500
5	5100	400
6	5600	1000

Table 2 shows values of each region(used to create synthetic data)

Both data sets were generated with forward modeling schemes (explained in chapters two and three), so at this point we have travel-times and MT functions that are representative for above model. For MT, MT functions were calculated at 8 sites with frequency range 1-1200Hz (18 frequencies). Seismic response was generated with 30 sources with 4936 source-receiver pairs. Forward model calculates observed traveltimes(of course they are not observed but synthetic, but inversion algorithm will regard them that way).

First step was to invert synthetic data sets separately. Regularization parameters will be kept the same for latter cooperative inversion research.

Starting velocity model for seismic profile is shown in Figure 7.4

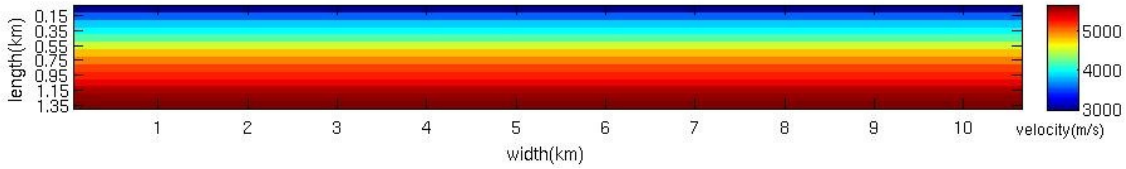


Figure 7.4 shows starting model for inversion

Velocity model obtained.

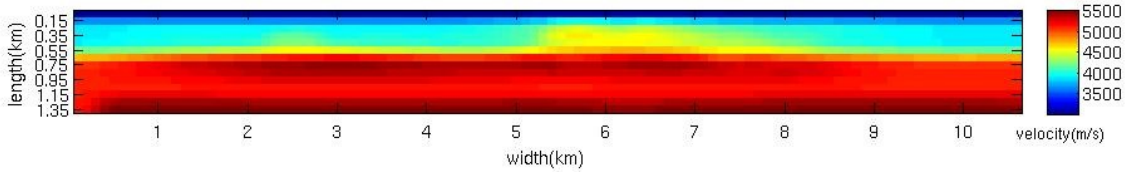


Figure 7.5 velocity model after inversion

iteration	1	2	3	4
RMS(ms)	28.14	16.37	11.28	10.92

Table 3 shows traveltime RMS for every iteration

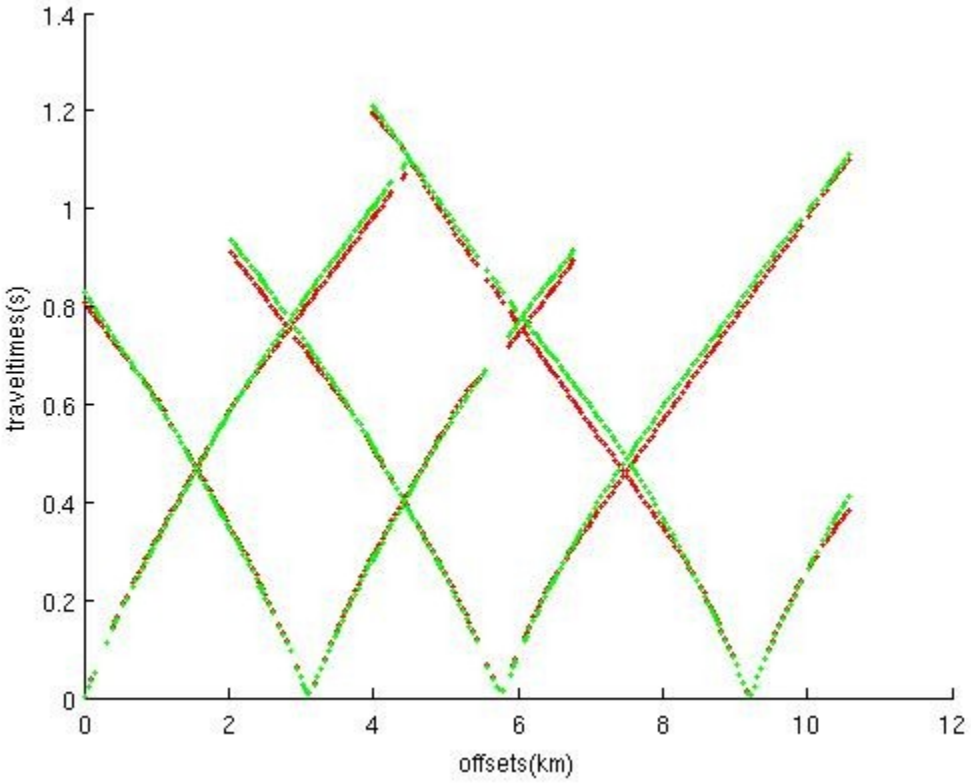


Figure 7.6 Seismic traveltime data. Red points represent original velocity model(Fig. 7.1) and green the inverted velocity model (Fig. 7.4)

From Fig 7.5 it is obvious that velocity structure is poorly resolved although misfits are decreased. This can be explained with non-uniqueness of inverse problem. There is no delineation between regions, for cells 4 and 5 , there are just outlines with geometry that does not fully represent the original model. Theoretical explanation for this lies in the choice of starting model and inversion parameters. In seismic tomography initial model is regarded as a 'holy grail', something of huge importance. Inversion mechanism does not have capability to detect structure with totally different velocity values, only small perturbations to the starting model could be determined. Same traveltime misfits could be obtained

with different initial model and with different regularizations, but that would lead to different velocity models after inversion.

For MT, inversion of the magnetotelurrics data (synthetic data created with model from 7.2) was performed with the starting uniform model (100Ωm).

RMS(for MT functions) of 1 was reached after three iterations.

Inversion results are presented in Fig. 7.7

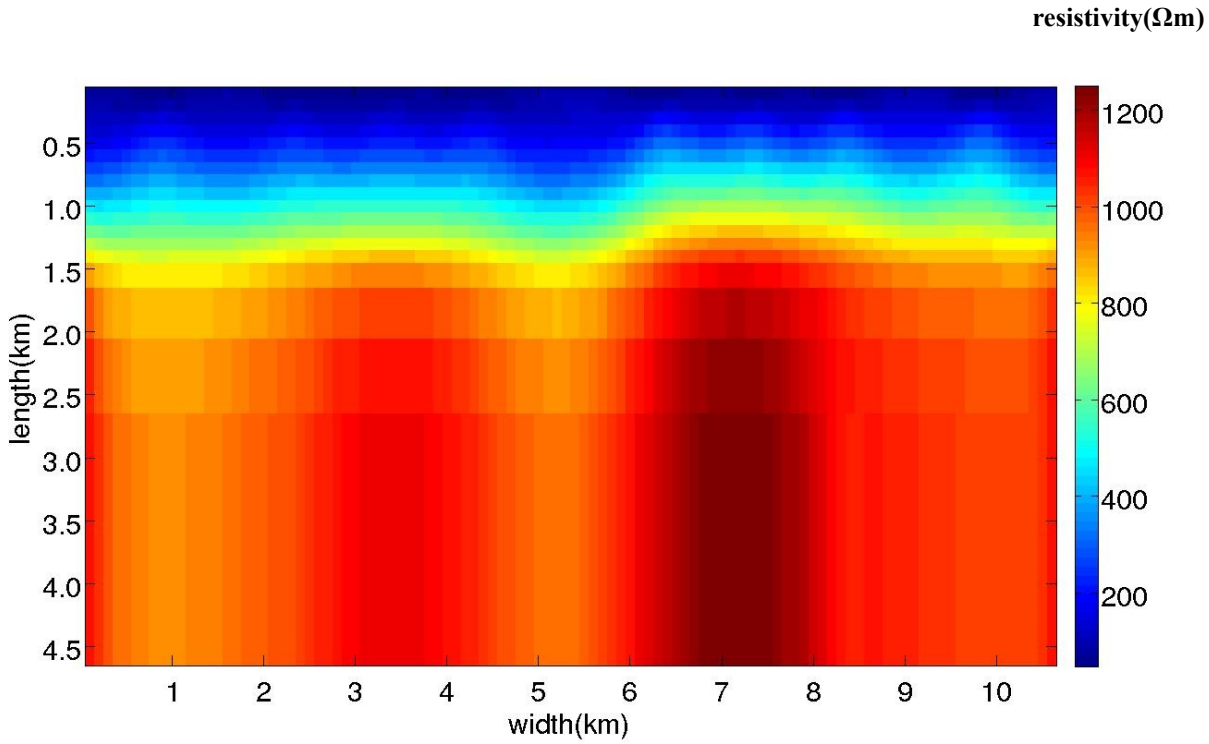


Figure 7.7 resistivity structure obtained after inversion

We are interested in upper 1.4km of this model,this part will be examined in further work. From above picture,outlines of structures are visible, but we are aiming to get better resolution. MT functions are presented in Figures 7.8-7.14 for three sites (1,3,8). For apparent resistivity, there is excellent match between data and model, for phase also good match.

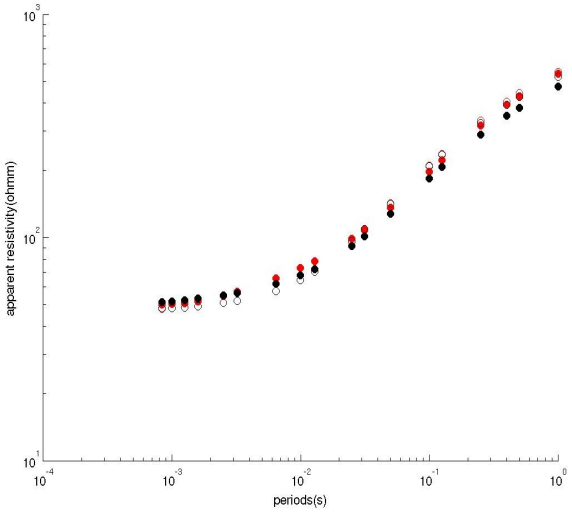


Figure 7.8

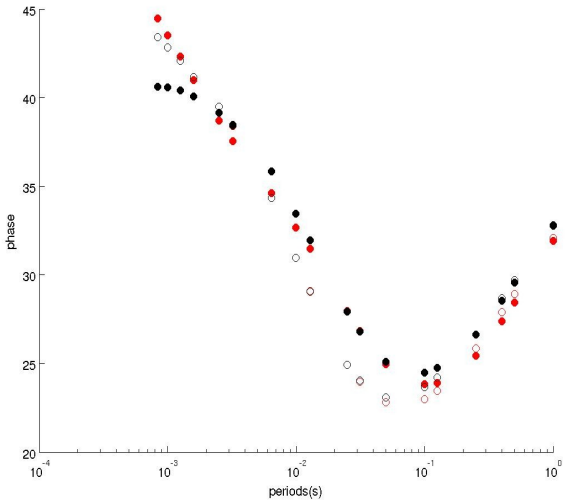


Figure 7.9

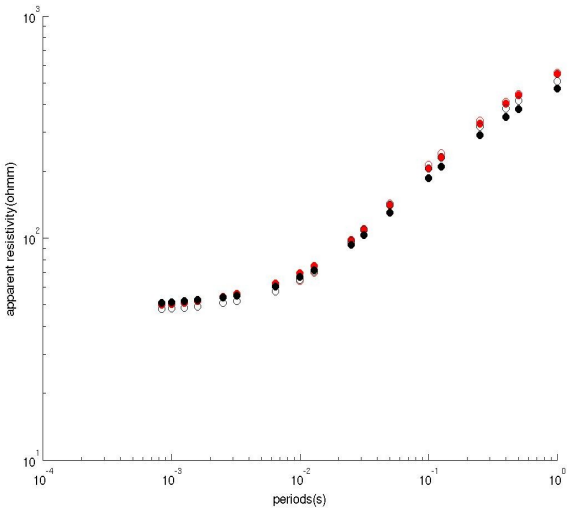


Figure 7.11

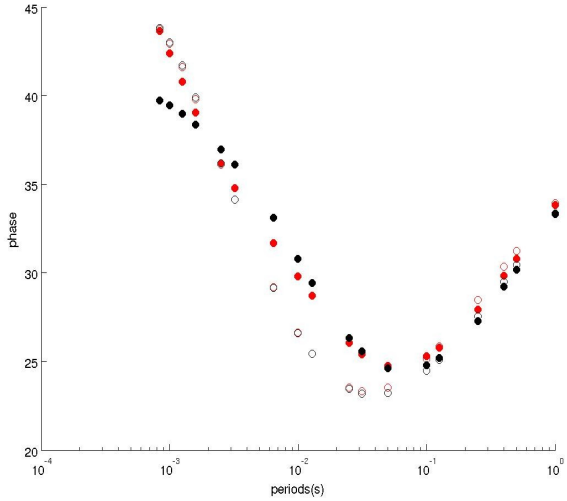


Figure 7.12

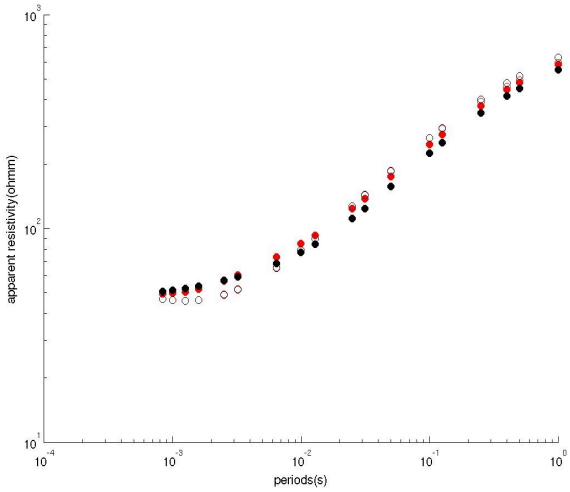


Figure 7.13

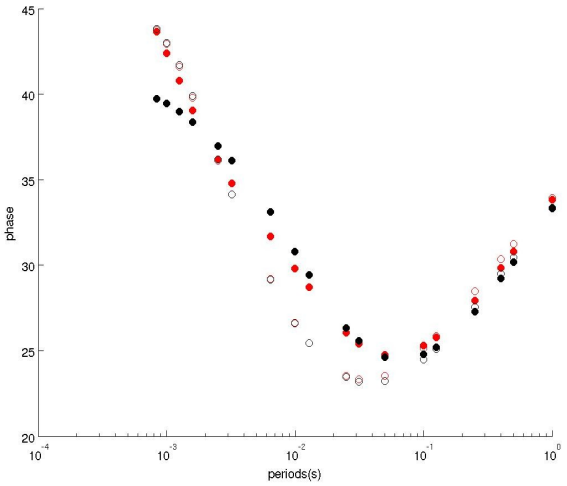


Figure 7.14

Fig. 7.8-7.14 show apparent resistivity and phase for E mode(red) and H mode(black) for three sites (1,3,8). Empty circles denote synthetic data, filled circles inversion response.

Using methodology explained in chapter 6, cooperative inversion was performed with 5,6,7 clusters. Parameters the 'fuzziness' exponent (m=2) and the termination tolerance (ε=0.001) were kept the same while we changed cluster number. Below are presented results of 6 cluster solution obtained after 3 iterations(at this point misfit was reached for seismic inversion). Adding or subtracting one cluster had not improved results (the pattern has not changed too much)

Figures 7.15, 7.16 and 7.17 present clusters, resistivity and velocity parametric models

Fig 7.18 shows traveltimes obtained after inversion(green) and original velocity model(red). Original velocity model is the one that produced the synthetic data,so the goal is that the inverted model matches first one as good as possible.

Figures 7.19-7.24 present MT function for sites 1,3,8 respectively

For apparent resistivity there is good matching, for phase there is some misfit. The MT functions are almost the same with one we got with separate MT inversion which makes sense because resistivity models are almost the same.

Table 4 shows how RMS changes in 3 iterations.

iteration	1	2	3
RMS(ms)	28.15	16.26	11.61

Table 4

Black rectangle is drawn to represent original cell 3, to compare original model with our inversion results (Fig 7.1 and Fig 7.2). Although there is no distinct borders, there is good matching to original model. Also we can recognize first two cells. In other parts results are not so good. For example, the lowest cell, has smeared resistivity structure (in original model 1000Ωm) and quite deformed velocity structure (in original model 5600 m/s).

Stopping criterion was reached too fast, but there is a problem how points are clustered and latter parametric models constructed. FCM has inherent limitations, due to the fact that it searches for sphere clusters with similar number of points. We know how original model looks like, so first choice for cluster numbers was 6 (this can be very tricky issue with all clustering techniques). The whole procedure was repeated with 5 and 7 clusters but without improvement. More clusters does not necessarily mean better results because, because there will still be points with low memberships. This could implicate parametric models that do not describe the properties as it would be expected.

Possible solution would be to weight each cell specifically to membership values and thus somehow prevent outliers to influence the clustering.

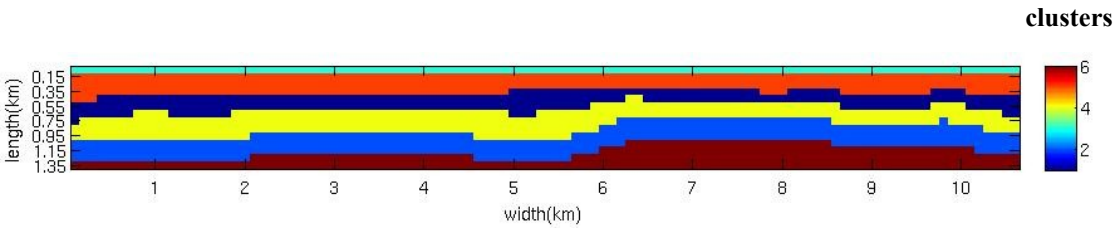


Figure 7.15 shows clusters

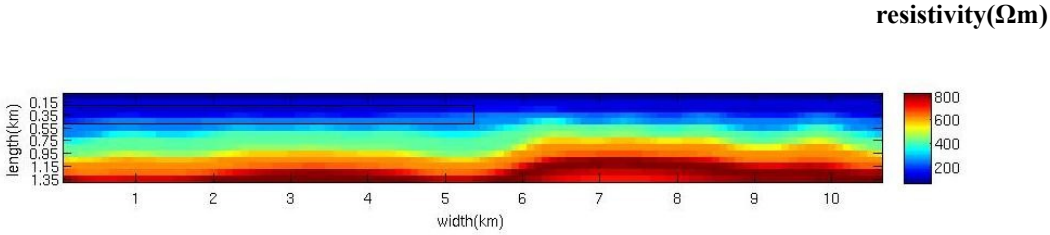


Figure 7.16 shows resistivity model

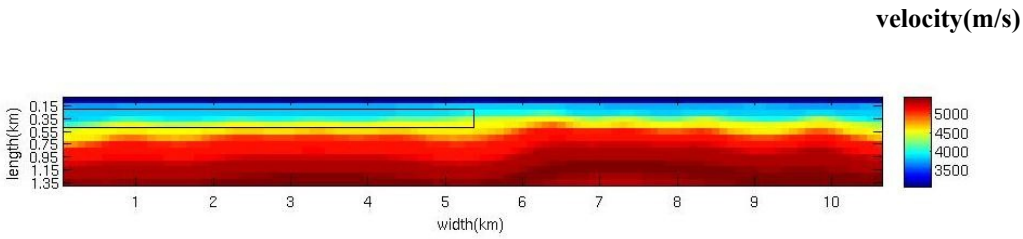


Figure 7.17 shows velocity model

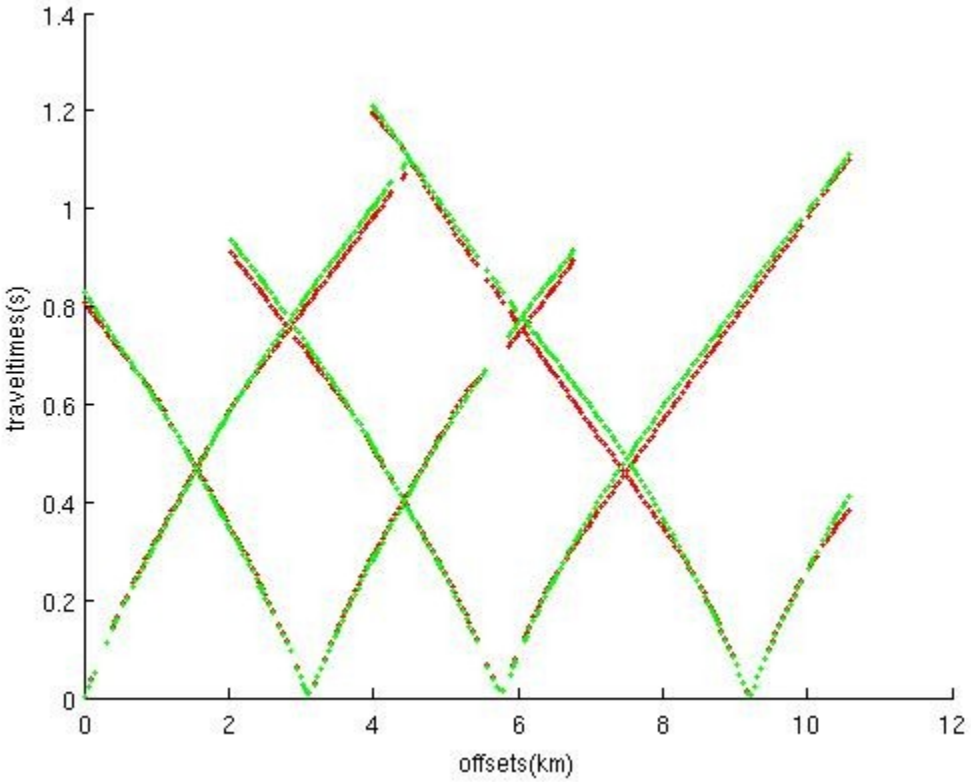


Figure 7.18 Seismic traveltime data. Red points represent original velocity model(Fig. 7.1) and green the inverted velocity model (Fig. 7.16)

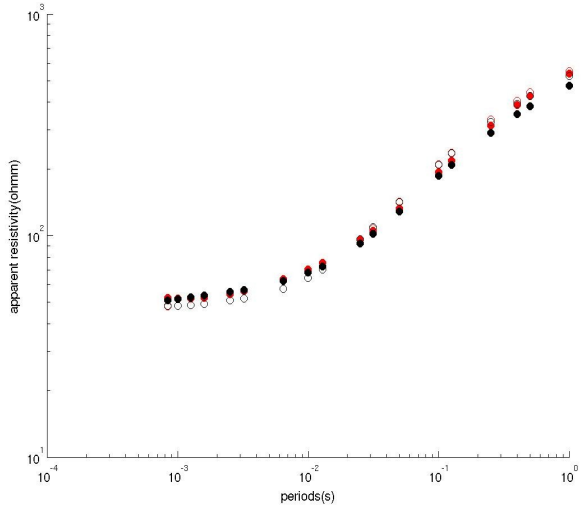


Figure 7.19

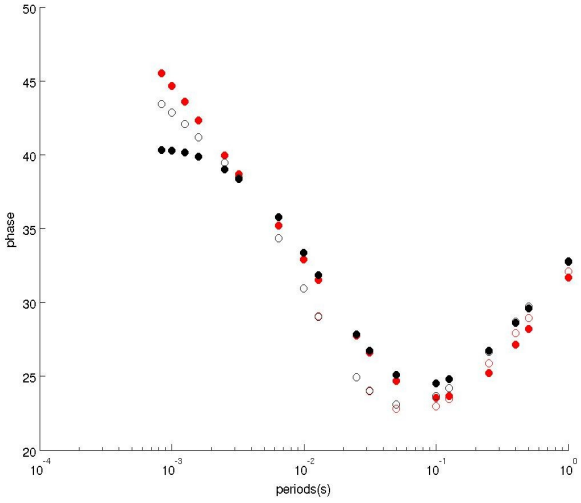


Figure 7.20

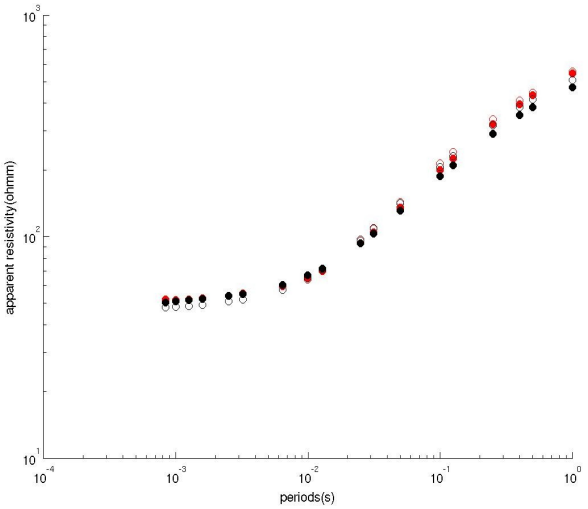


Figure 7.21

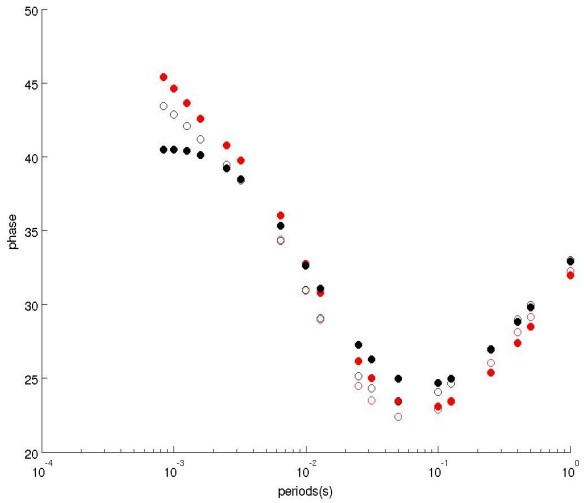


Figure 7.22

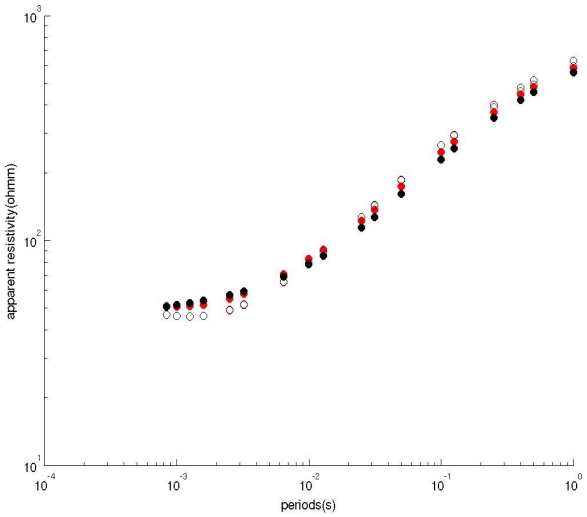


Figure 7.23

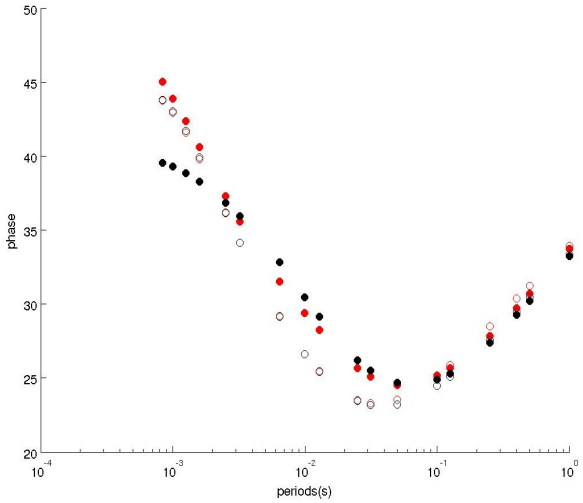


Figure 7.24

Fig. 7.19-7.24 show apparent resistivity and phase for E mode(red) and H mode(black) for three sites (1,3,8). Empty circles denote synthetic data, filled circles inversion response.

Using methodology explained in chapter 6, cooperative inversion was performed with 5,6,7 clusters. Here are presented results of 7 cluster solution obtained after four iterations.

Beside the parameters we had in FCM(the number of clusters,the 'fuzziness' exponent(m=2), the termination tolerance(ε=0.001)) we have additional parameter cluster volumes - ρ_i(formula).We have tried for ρ_i=1 and 2 , second gives gives possibility for unequal number of points in clusters.

Figures 7.25, 7.26 and 7.27 present clusters, resistivity and velocity parametric models.(first case)

Figures 7.28, 7.29 and 7.30 present clusters, resistivity and velocity parametric models.(second case)

Fig. 7.31 illustrates how clusters are organized with this algorithm in 2dimensional parametric space(ρ_i=2)

Table 5 below shows how RMS changes in 4 iterations.

iteration	1	2	3	4
RMS(ms) , ρ_i=1	28.15	15.27	13.82	12.84
RMS(ms) , ρ_i=2	28.15	14.88	13.19	11.29

Table 5

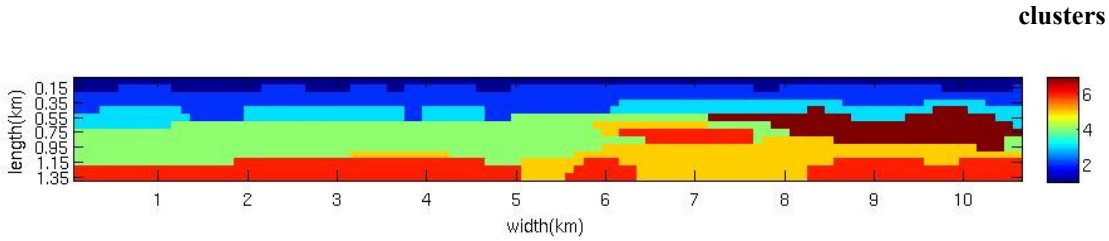


Figure 7.25 shows clusters

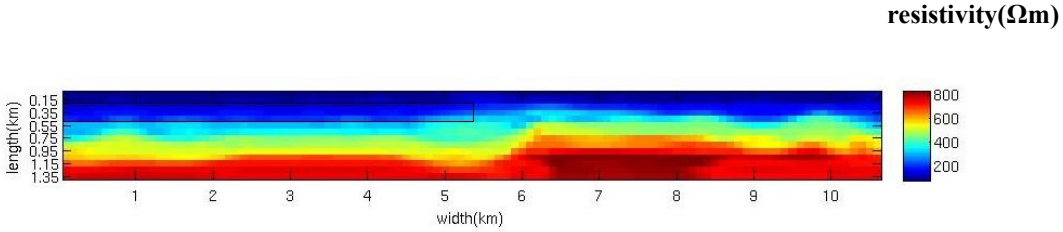


Figure 7.26 shows resistivity model

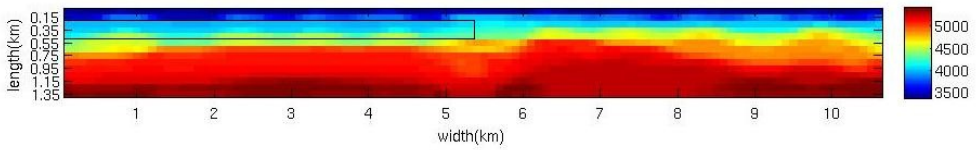


Figure 7.27 shows velocity model

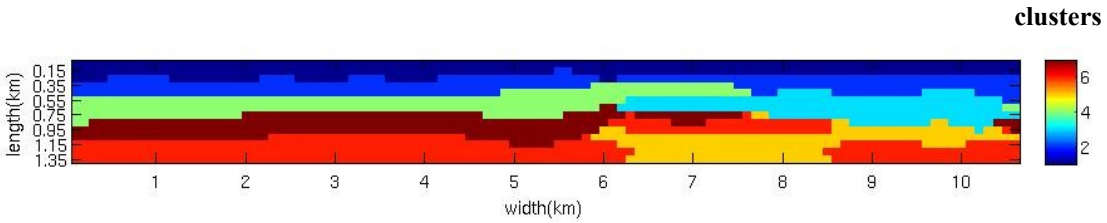


Figure 7.28 shows clusters

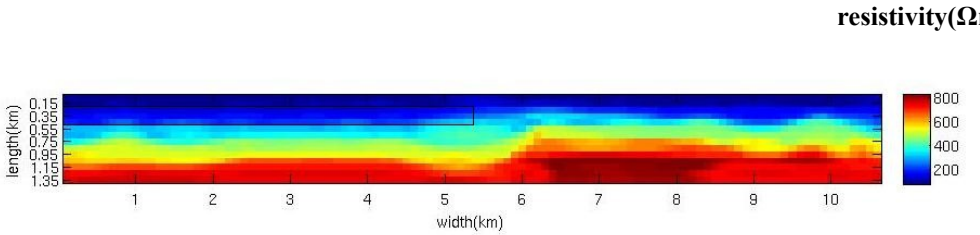


Figure 7.29 shows resistivity model

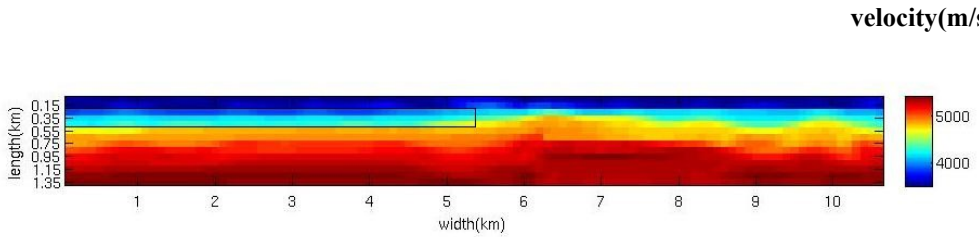


Figure 7.30 shows velocity model

If we look at Fig 7.31 and 7.32, it can be seen that cluster centers are drawn along with points (after normalization). GK is looking for ellipsoid clusters (metrics is not Euclidean any more). Comparison of these two indicates how ρ_i parameter has influenced calculations. Resistivity models are very similar, but for velocity models we got better results in second case (better misfit and better region separation, similarity to original model). Again rectangle is put over same cell (now we have better result than with FCM). Also lowest velocity cell is much more homogeneous as well as resistivity one though values from original model have not been reached.

Regarding velocity, problem occurs regarding the issue of velocity range. In final parametric model range is much lower than in the synthetic velocity model (Fig 7.1). This is a consequence of hidden complexity of the parametric models. There are couple of factors that influence this: numerical values of both parameters, membership functions and also defuzzification.

One important lesson that we could learn from here is how cluster organization at one step will influence parametric models not just on that iteration but also in further iterations. In seismic inverse problem, linearization assumption leads to that only small perturbations of the starting model . But here starting model itself is result of unsupervised learning (clustering), algorithm is grouping points and after that parameter models are created (detail explanation given in chapter 6). Extremely resistive or conductive structure will produce one distinct cluster, but will also determine other cluster centers and memberships. So if we have some structures at the second iteration, they can be slightly modified but will also influence pattern recognition process. In the iteration that comes next, this structure will be present again. It also depends on how both structures look like(cluster arrangement), but in most of the cases this will be the happen. Drawback of this method is that the number of iterations can hardly be very large (which is generally desirable because we want as lower as possible misfits but even changing starting model could not improve this,to have for example 15-20 iteration with improvement at every step)

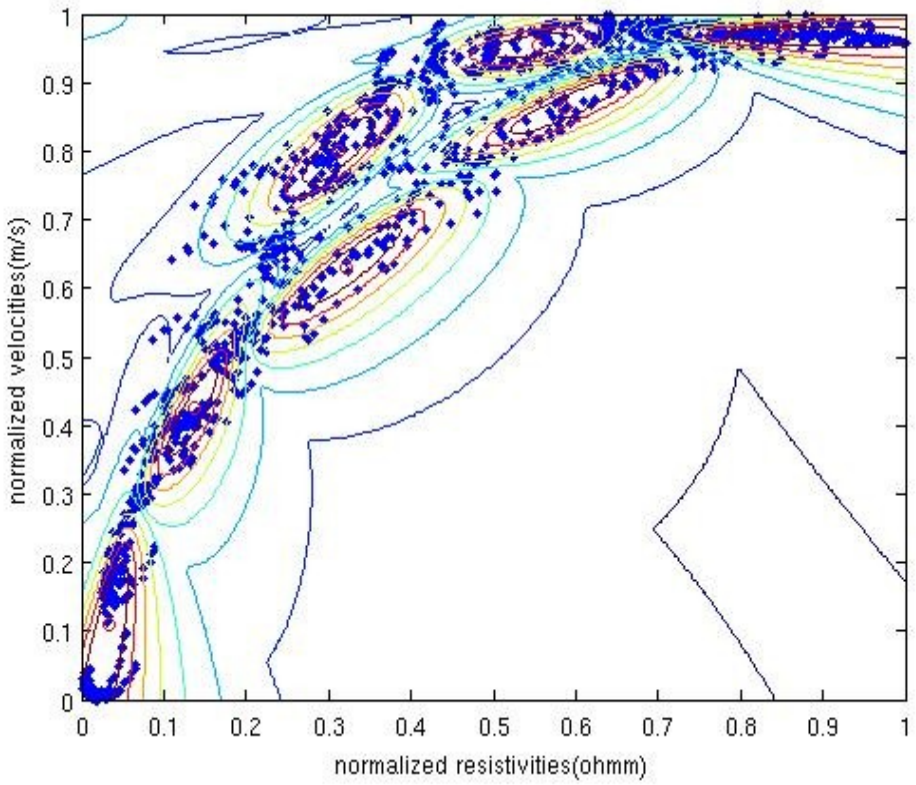


Figure 7.31 shows clusters and all points after normalization

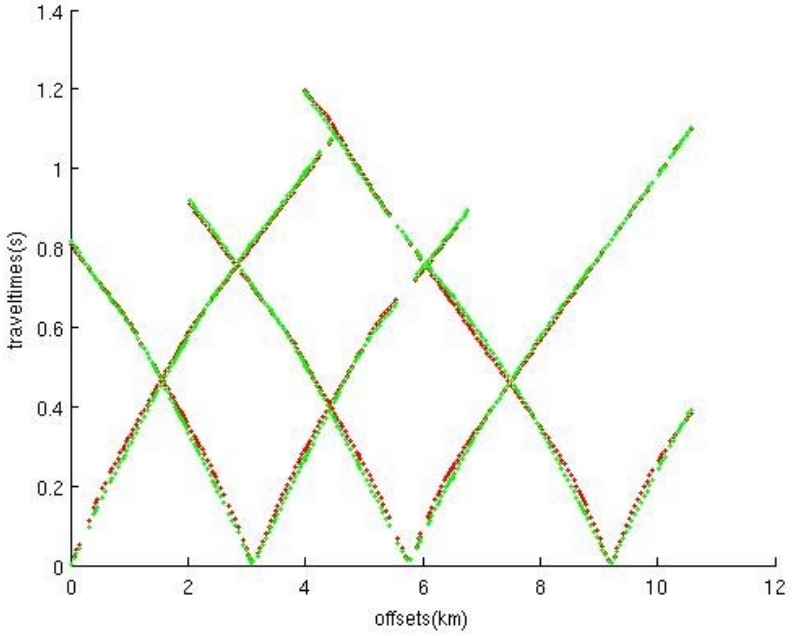


Figure 7.32 Seismic traveltime data. Red points represent original velocity model and green the inverted velocity model

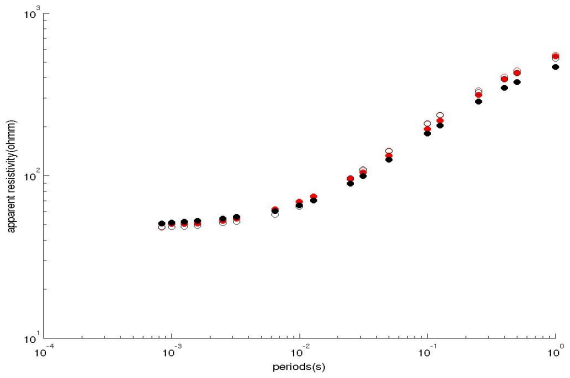


Figure 7.33

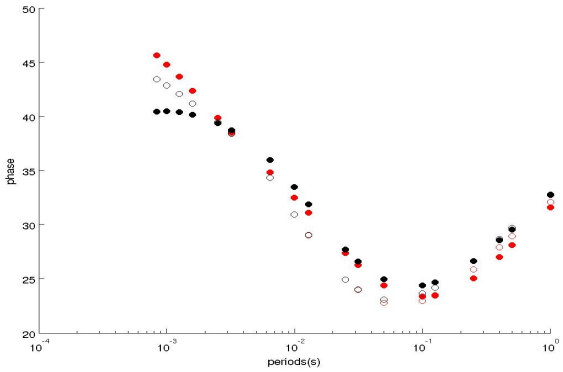


Figure 7.34

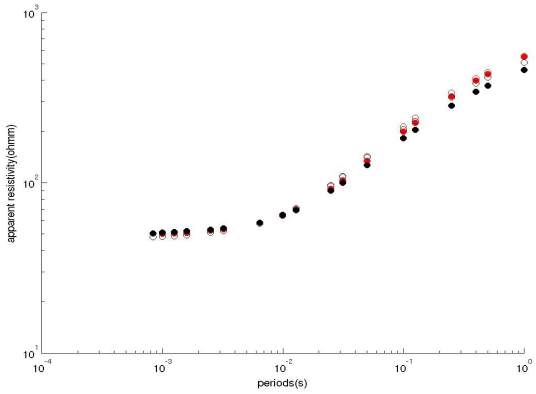


Figure 7.35

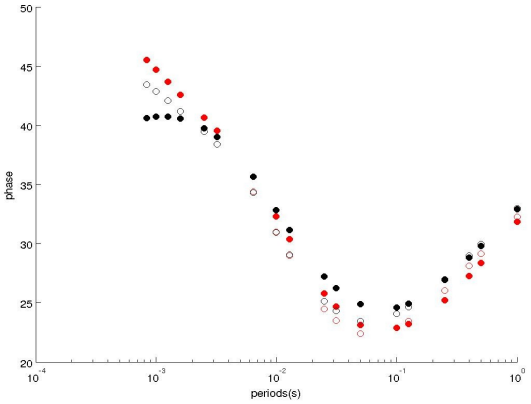


Figure 7.36

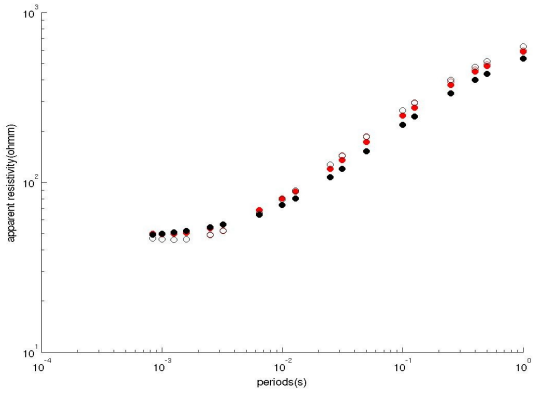


Figure 7.37

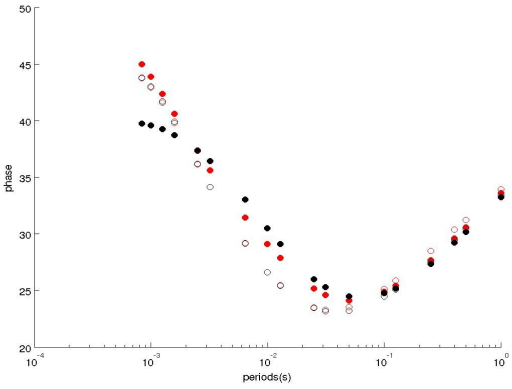


Figure 7.38

Fig. 7.33-7.38 show apparent resistivity and phase for E mode(red) and H mode(black) for three sites (1,3,8). Empty circles denote synthetic data, filled circles inversion response.

Gath-Geva clustering was performed for 5,6 and 7 clusters.

Solution for 7 clusters gave best results and is presented below.

Figures 7.39, 7.40 and 7.41 present clusters, resistivity and velocity parametric models.

Figures 7.42 shows cluster centers and all points (after normalization).

Figures 7.43-7.48 present MT function for sites 1,3,8 respectively.

Fig 7.49 shows seismic traveltime data.

Table 6 below shows how RMS changes in four iterations.

iteration	1	2	3	4
RMS	28.15	21.05	14.64	11.57

Table 6

Table 7 shows average value of resistivity and velocity for each cluster

Cluster	v(m/s)	ρ
1	2976	69
2	3538	92
3	4094	145
4	5202	471
5	5563	819
6	5364	635
7	4727	311

Table 7

The Gath-Geva algorithm is an extension of Gustafson-Kessel algorithm, but distant function is calculated in way that small distance means a high probability and a large distance means a low probability of membership. For Gath-Geva, most of the points in parameter space have very high memberships with respect to belonging cluster.

If we start again with cell 3 from original model, velocity matches quite well the original model, with resistivity there are some problems with low resistivity region being pushed-down. Lowest cell in original model is represented here with 3 clusters (4,6,5), not homogeneous as we would like to, but with high resistivities and velocities to be interpreted as a separate structure. Two upper cells are quite good matching. Only problem is the somehow striped structure,so it does not differentiate between the left side and right side cells in original model. The same can be observed at Fig 7.48, where high residuals occur exactly at he central part of the model.

MT functions response is very similar to original,only H mode at some period we have larger misfits.

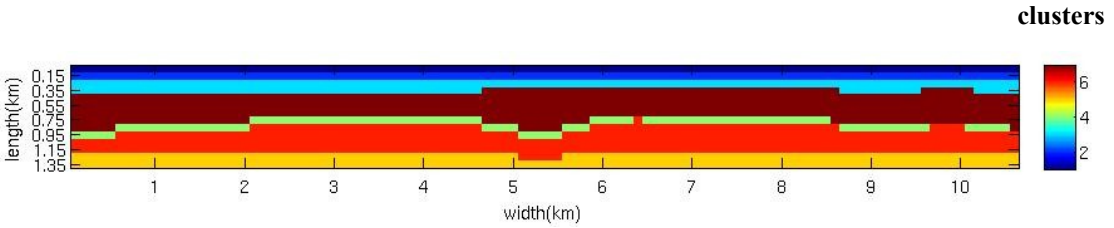


Figure 7.39 shows clusters

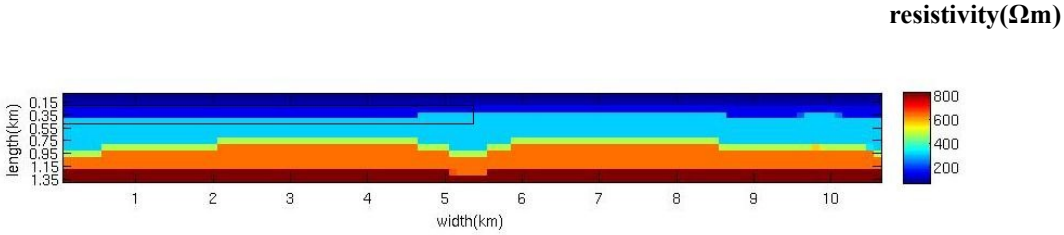


Figure 7.40 shows resistivity model

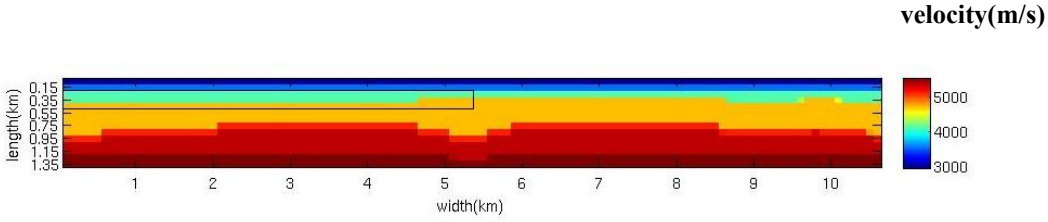


Figure 7.41shows velocity model

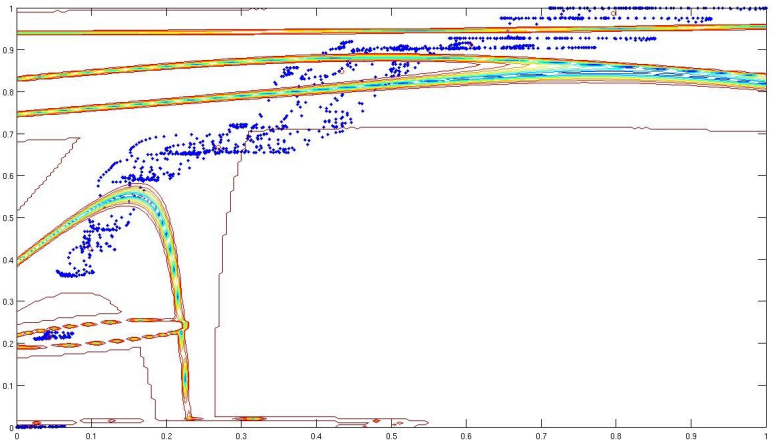


Figure 7.42 clusters and points(normalized values of parameters)

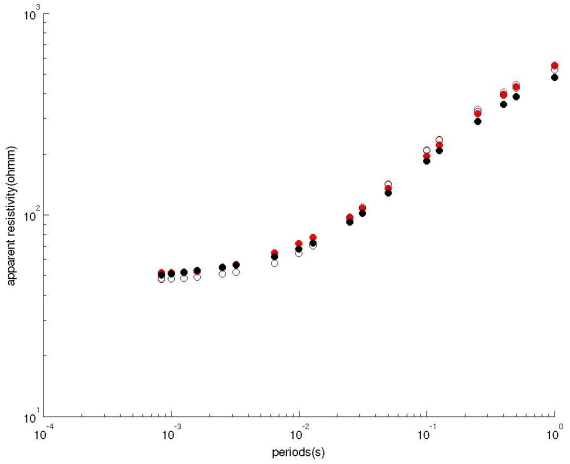


Figure 7.43

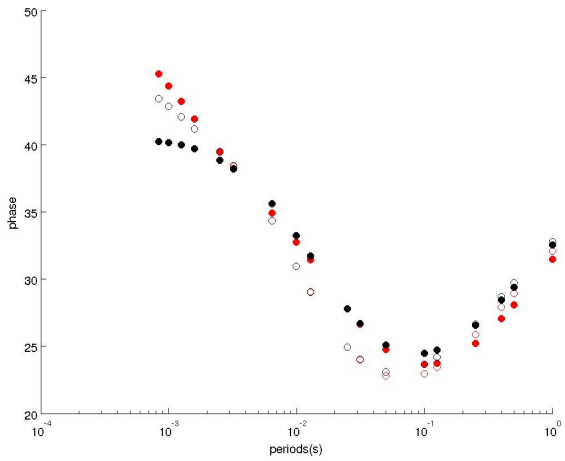


Figure 7.44

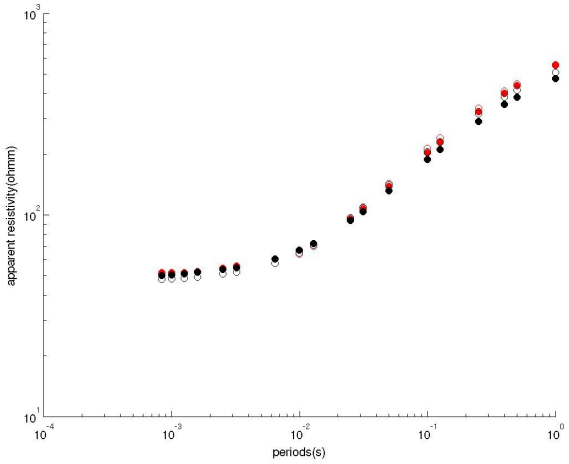


Figure 7.45

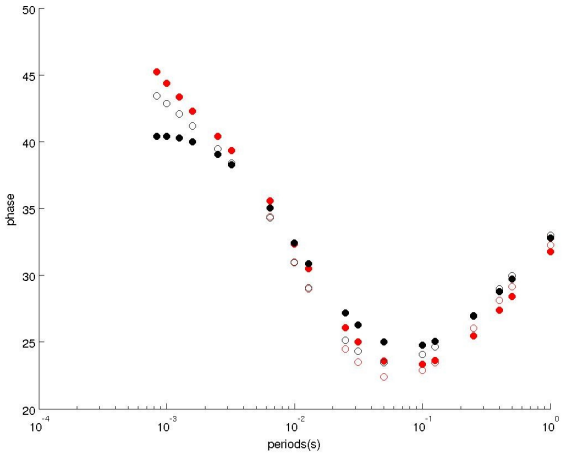
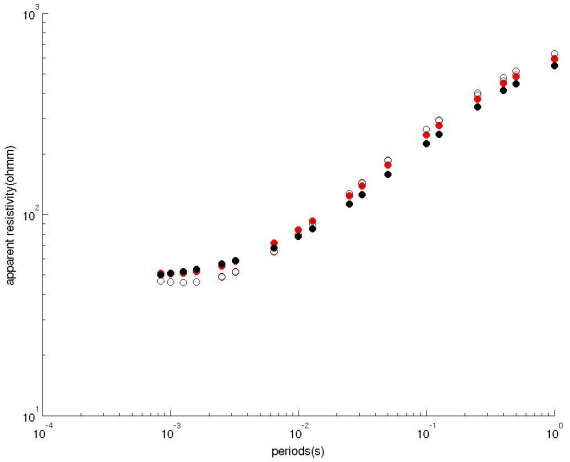


Figure 7.46



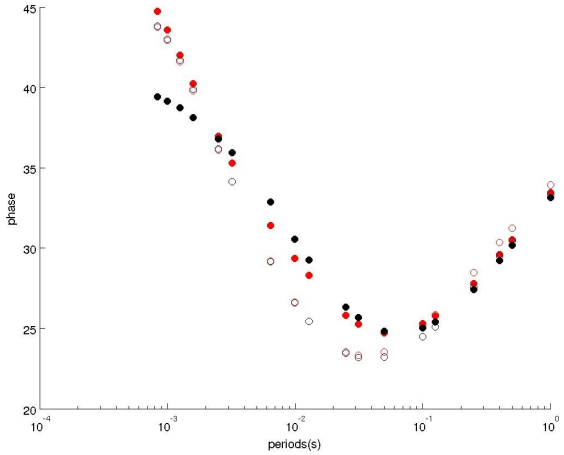


Figure 7.47 Figure 7.48

Fig. 7.43-7.48 show apparent resistivity and phase for E mode(red) and H mode(black) for three sites (1,3,8). Empty circles denote synthetic data, filled circles inversion response.

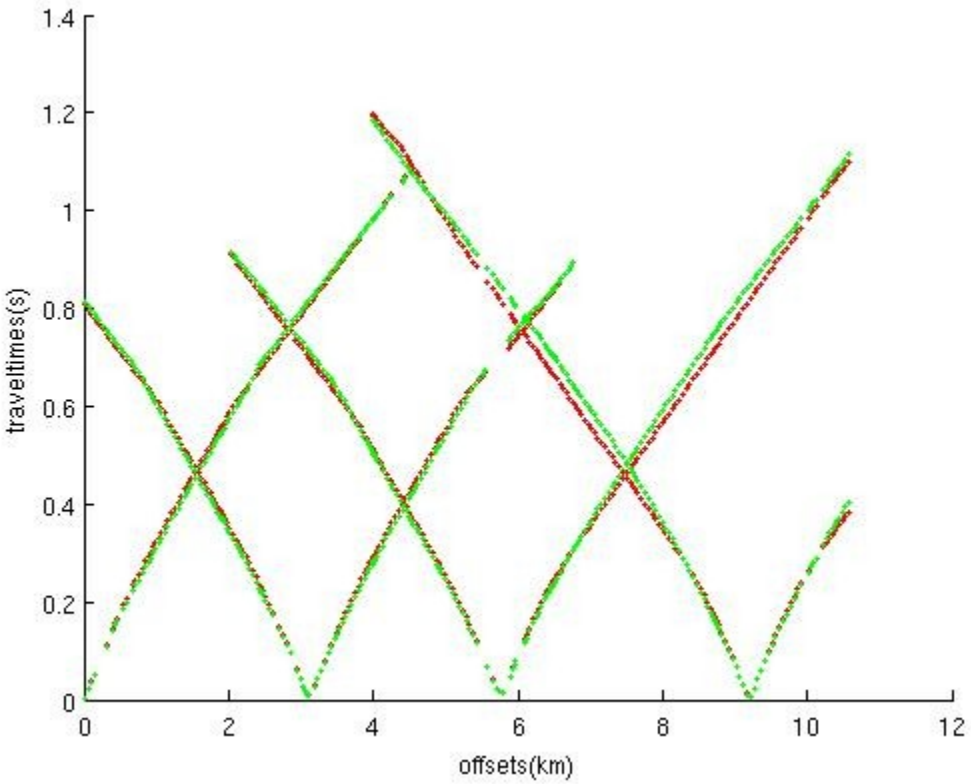


Figure 7.49 Seismic traveltime data. Red points represent original velocity model and green the inverted velocity model

8.1. SEPARATE INVERSION RESULTS

After data processing, we have inverted the data and the following work is related to obtained results. We have inverted separately both data sets. Results presented here will be latter compared with models we got from all three cooperative inversion algorithms.

In case of MT, OCCAM2D inverse code was employed. Starting model was chose to be homogeneous , each cell with resistivity 316.22 ohmm. (log ρ= 2.5). We have tried also different starting models, changing log ρ(2.0, 2.3, 2.6) but keeping all the other inversion parameters the same. After five or more iterations, they also lead to very similar results, there is no possibility to substantially improve R.M.S, almost the same conductivity models were obtained.

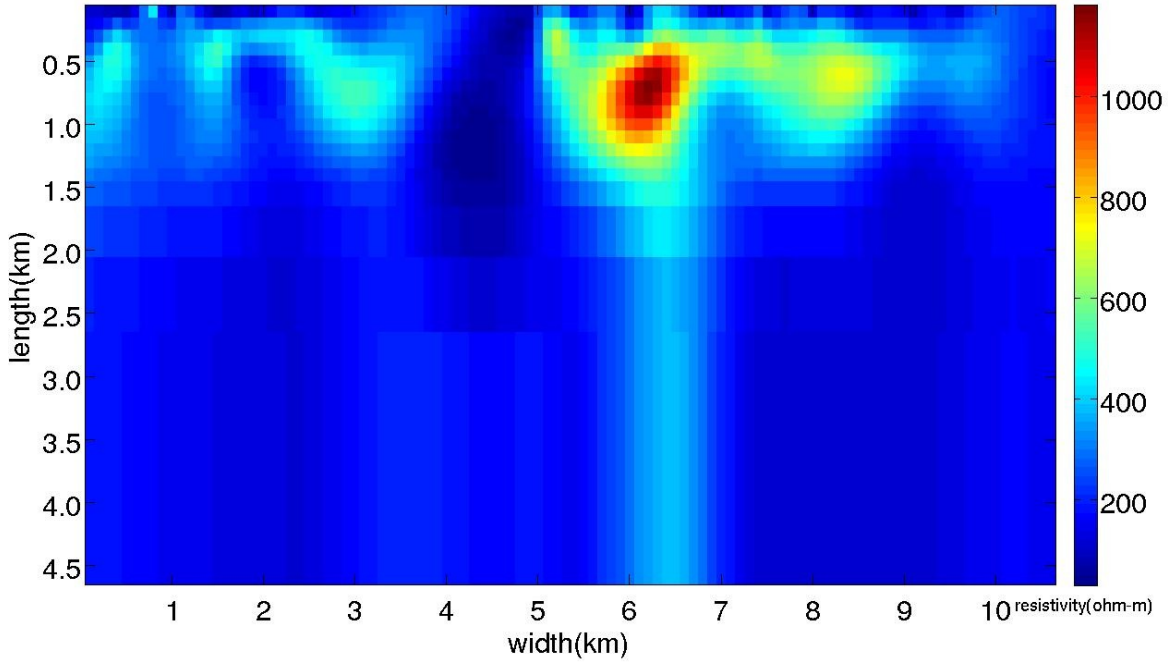


Figure 8.1 resistivity model after 5 iterations

We are interested in the first 1.4 km, upper part of the model. These 1484 cells will be used for clustering with the respective cells from seismic model (methodology was explained in chapter 6).

	start. model	1	2	3	4	5
R.M.S	3	2.76	2.43	2.42	2.41	2.4
Lagrange	4	2.09	1.32	1.31	1.18	1.18
roughness	10000000	8.66	21.48	20.02	23.03	24.8

Table 8 shows how R.M.S, Lagrange multiplier and model roughness change at every iteration

site 3

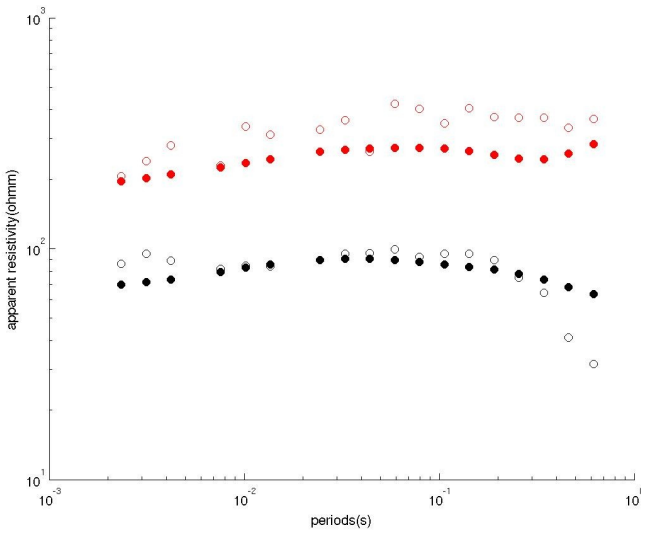


Figure 8.2

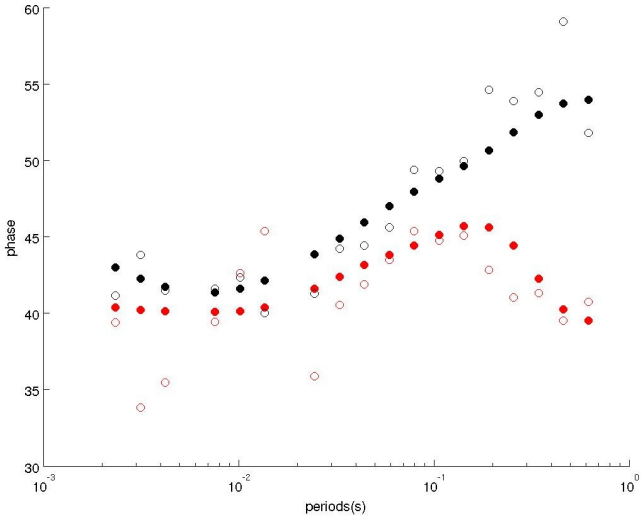


Figure 8.3

site 5

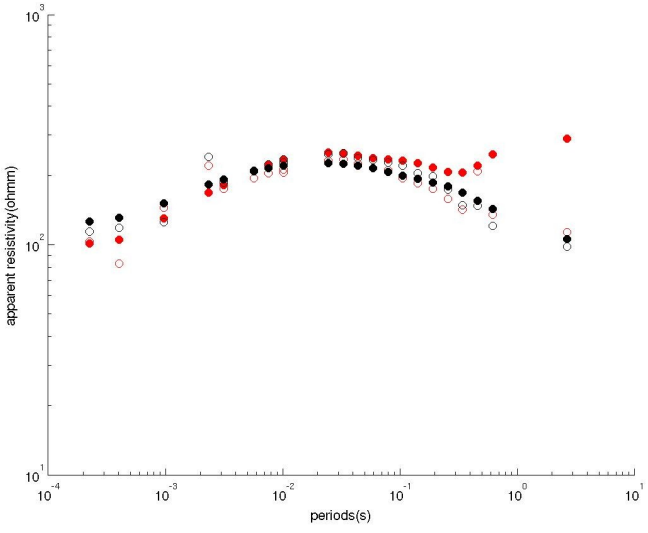


Figure 8.4

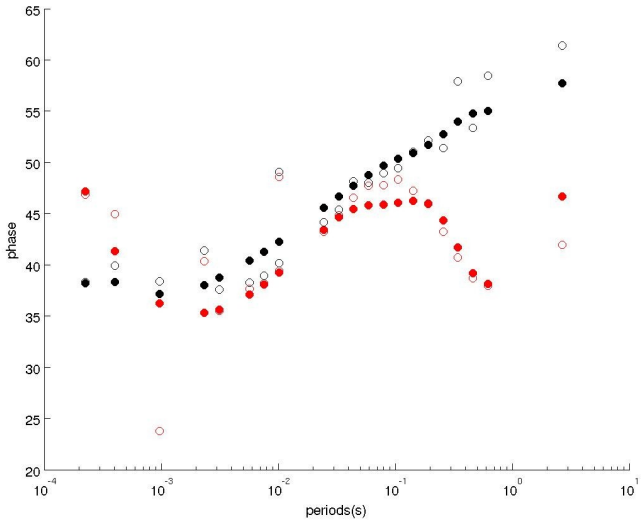


Figure 8.5

site 12

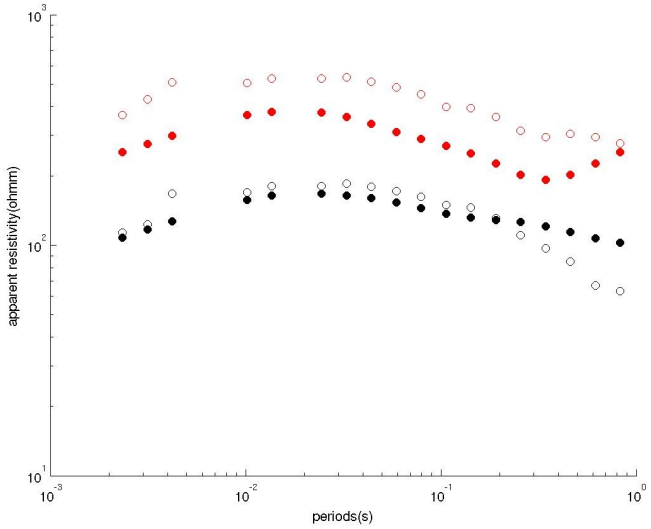


Figure 8.6

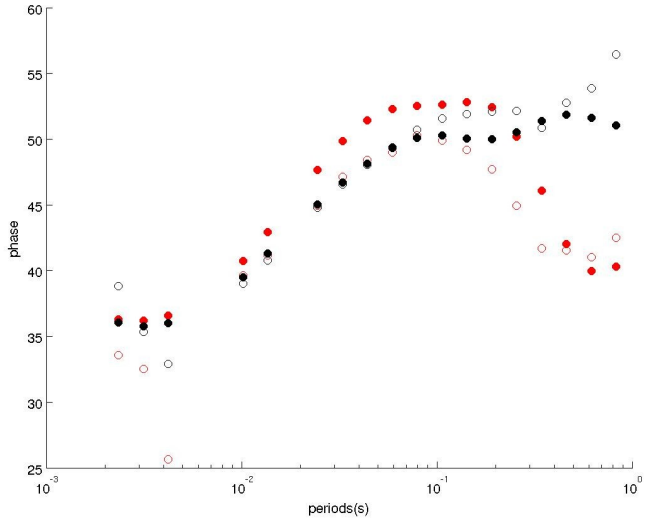


Figure 8.7

site 14

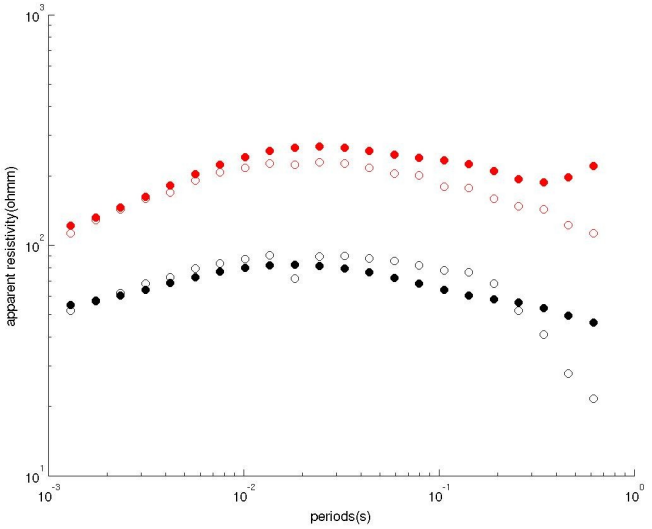


Figure 8.8

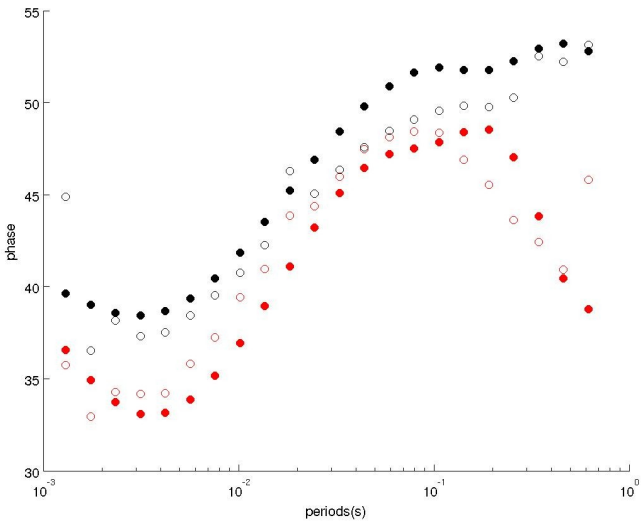


Figure 8.9

Figures 8.2-8.9 shows apparent resistivity and phase for E mode(red) and H mode(black) for sites 3,5,12,14. Empty circles denote fieldwork data, filled circles inversion response

If we look at Figures 8.2-8.9, generally apparent resistivity was better estimated with inversion than phase. At higher periods, some phase points are scattered, that influences inversion process. Although algorithm is looking for the smoothest model which fits the data to within an expected tolerance ,huge residuals (between data and model) will definitely have negative impact. This explains why we can not get better RMS with more iterations. From Fig. 8.1, we can see resistive structures between 6-7km and 8-9km, along the profile (the full length of the profile is 10.6km).

For traveltime inversion, it was necessary to find 1-D starting model that reasonably fits the field data. In seismic tomography, initial model is very important. Tomographic images resulting from linearized inversion are dependent on starting models, only small perturbations to starting model could be determined. By using trial-and-error strategy, we have constructed starting model, to suit observed high velocity gradient in first 300m (Table3, Figure 8.10).

depth(km)	velocity(km/s)
0.000	2.700
0.100	3.100
0.200	3.500
0.300	3.750
0.400	4.000
0.500	4.200
0.600	4.400
0.700	4.700
0.800	5.000
0.900	5.100
1.000	5.200
1.100	5.300
1.200	5.400
1.300	5.500
1.400	5.600

Table 9 1-D starting velocity model

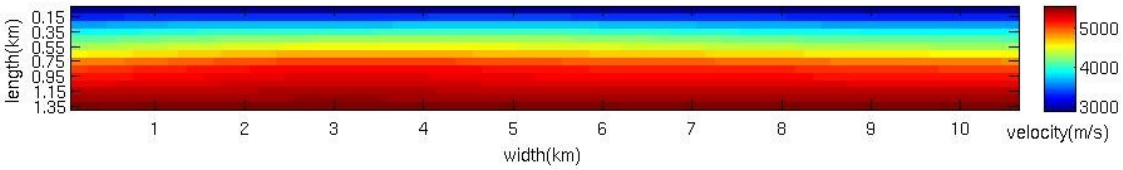


Figure 8.10 shows 1-D starting velocity model

Traveltime inversion parameters λ, s_z and α are crucial, because they control how objective function looks like (explained in detail in chapter 2). The preferred parameter set should provide a minimum-structure model in accordance with Occam's principle (Constable et al. 1987) to avoid interpreting small-scale features on the final models that are not warranted by the data.

$$\Phi = \delta t^T C_d^{-1} \delta t + \lambda [\alpha (m^T C_h^{-1} m + s_z m^T C_v^{-1} m) + (1-\alpha) \delta m^T C_p^T C_p \delta m] \qquad (8.01)$$

looks like and thus inversion results We have tested these parameters proving that different parameter set could also lead inversion process with similar results. So by changing inversion parameters (λ , s_z and α) we would still have iteration leading χ^2 to 1, but with slightly different traveltime rms and chi-squared. So alternative solutions exist, because we are dealing with non-unique and ill-posed problem.

λ	s_z	α
300	0.13	0.99

Table 10 shows inversion parametres (will use these parameter set in our cooperative inversion alghorithms)

	1	2	3	4	5	6	7	8	9	10	11	12
Rms (ms)	29.7	18.92	17.47	16.63	15.97	15.24	14.52	13.94	13.11	12.78	12.28	12.04
χ2	6.13	2.48	2.12	1.92	1.77	1.61	1.46	1.35	1.26	1.13	1.08	1.01

Table 11 shows how traveltime rms and chi-squared(χ^2) change during inversion

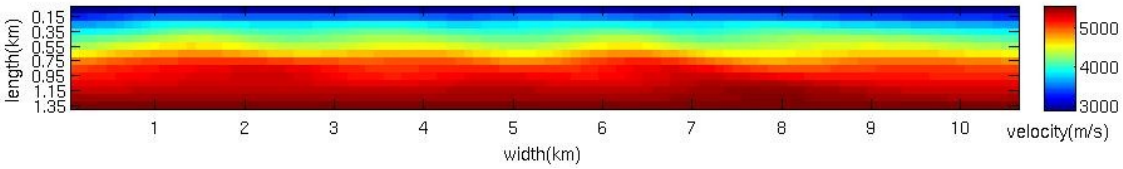


Figure 8.11 shows velocity model after inversion

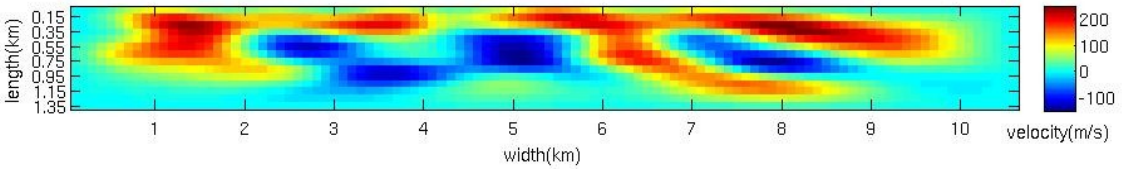


Figure 8.12 shows velocity difference between final and starting model

Figure 8.12 shows that the regions of higher and lower velocities are quite distinctive. Cells close to the edge of the model have practically same values, due to the low ray coverage in this part of the model. This could be a serious problem because these cells (their numerical values) will also enter into the pattern recognition algorithms, possibly pulling cluster centers away from ideal solution (whole model perfectly covered with rays).

Explanation about stopping criteria. As mentioned in 6.2, algorithm stops when stopping criterion is met. In case of traveltime inversion, stopping criterion is rms (root mean square) error of traveltimes

$$rms = \sqrt{(1/n(\sum (t_{obs} - t_{calc})^2))} \qquad (8.02)$$

where:
n = number of offsets
t_{obs} = observed traveltime (data)
t_{calc} =calculated traveltime (predicted by model)

Data acquisition and processing were explained in 5.2.1. With 40 shots and 6579 picks, n=6579. For 6579 source receiver pairs at every inversion, traveltime difference (residual) between observed and calculated traveltimes provides information how close model is to data. Smaller value indicates better model performance. With synthetic data obtained 12.04 ms. Note at Figure 8.13 that some residuals have very high values, which indicate that some parts of the model are inadequate and poorly resolved.

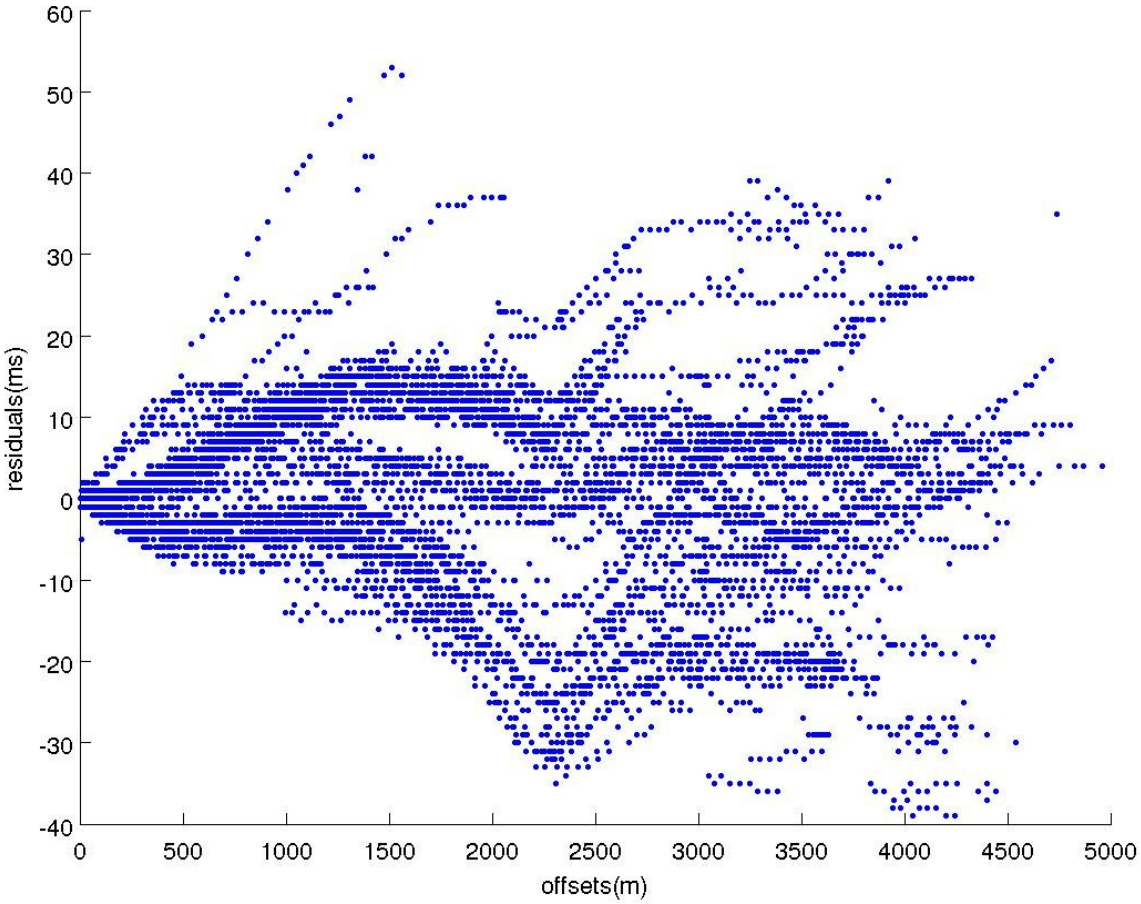


Figure 8.13 illustrates traveltime residuals for all shot receiver offsets

To fully reconstruct P-wave structure, resolution analysis should be performed and anisotropy taken in account (Baumann-Wilke *et. al.*, 2012), but that is out of the scope of our work (we will try to exchange information with MT and thus improve knowledge about velocity structure)

Fuzzy c means cooperative inversion was performed for different number of clusters (6,7,8,9,10) and fuzzifers (2,3). Results are presented here, that are unsatisfactory in our case (data from Pyrite belt). This algorithm was chosen initially because it achieved success for seismic P-wave and ground penetrating radar cooperative inversion (Paache *et. al.*, 2006). The result of their inversion procedure was a multiparameter model outlining the major zonation in which each zone is characterized by similar relationships between the individual model parameters. The same aim was here, just with MT and with expectation of consistent models that will improve at every step. We apply cluster analysis to the available data set, and the resulting clusters are characterized based on the characteristics of the patterns by which they are formed . With MT and seismic, neither changing number of clusters nor fuzzifier didn't help too much, stopping criterion was met too fast.

With synthetic data, results were much better, but we should bear in mind that for velocity we have started from models very similar to real structures, it was not strictly 'guessing in the dark'.

The fcm algorithm (explained in 4.1) is looking for sphere clusters with approximately similar number of points, which illustrates figure below, normalized values of resistivity and velocity (8 cluster solution, after 2nd iteration, RMS 13.76ms, next step gives larger misfit).

In this algorithm, zero mean variance one normalization was used. With normalization we want to give equal weight to both variables. But smaller units means larger range for both velocity and resistivity.

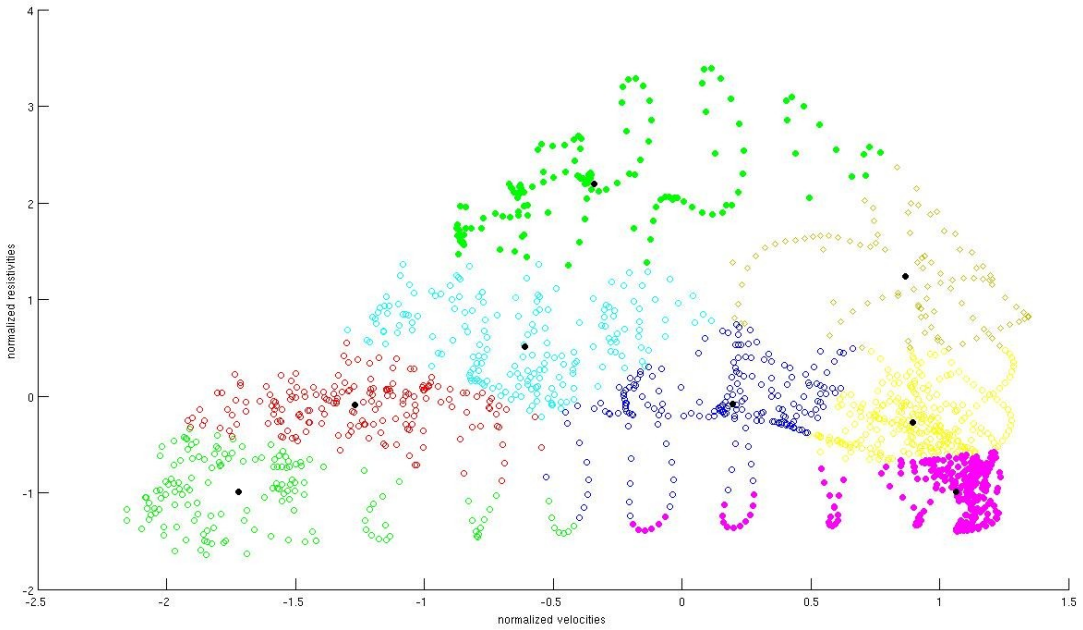


Figure 8.14 shows how normalized velocities and resistivities were clustered. Each cluster is represented by one color. With FCM, undesirable effects were produced (dissimilar elements fall into same cluster)

Cluster centers are shown with black points in Figure 8.14. If we analyze clusters it is obvious that velocity spread is quite high, resistivities tend to be in shorter range. The velocity values in the most resistive cluster (green) are in range of 3946-5096m/s (0.772 - 0.873), which is too wide range. From Figure 8.14 it is obvious that clusters are quite heterogeneous. New parameter models will be calculated using formula:

$$np_{ij} = \sum mb_{ij} \times avg_{ki} \qquad (8.03)$$

The value of a fuzzy membership function is a mathematical characterization of cluster. Multiplying it with average cluster value practically means that cluster heterogeneity does not generally lead to dissimilarity of neighbouring cells in new models. But still some parts of the velocity model are with quite dissimilar elements. As a consequence of this, new velocity structure is with lower velocities pushed-down from 6-8km.

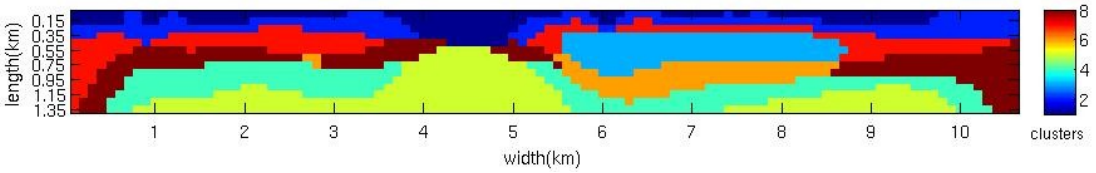


Figure 8.15

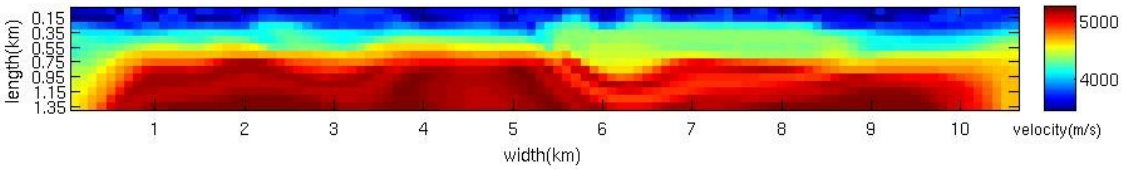


Figure 8.16

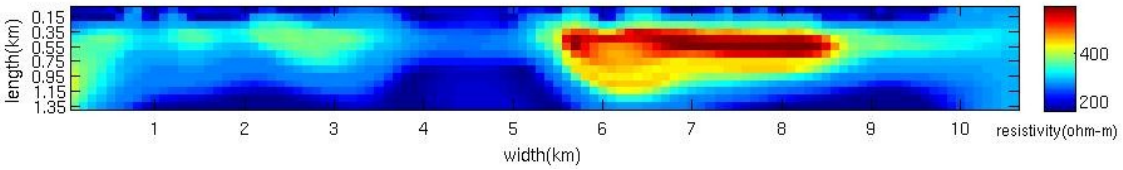


Figure 8.17

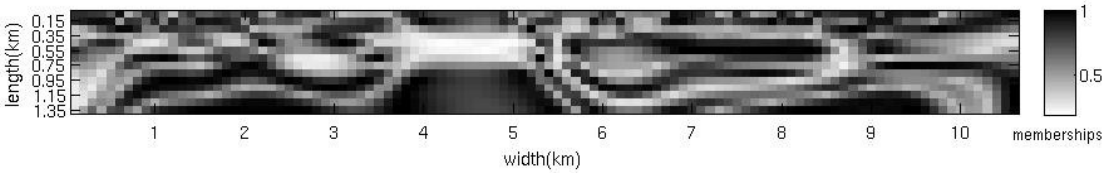


Figure 8.18

Figures 8.15-8.18 show 8 clusters after defuzzifications, parametric models after cooperative inversion with FCM (velocity and resistivity) and membership of each cell (degree to which it belongs to cluster)

Slightly better results were obtained if we change fuzzifier from 2 to 3 (all other parameters were kept the same). Fuziffier m is an important parameter for FCM. It controls how much clusters are allowed to overlap. Misfits for traveltimes were reached after 3rd iteration (13.68ms). With MT inversion, RMS has lower value at every iteration (2.41).

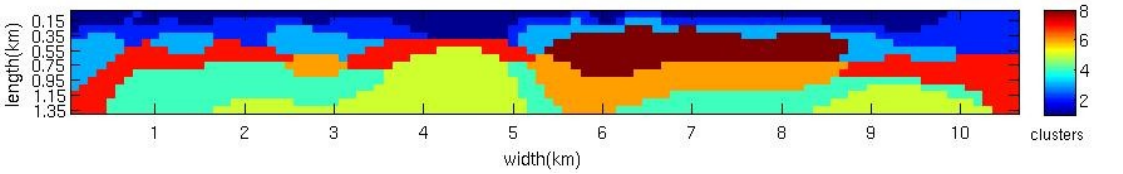


Figure 8.19

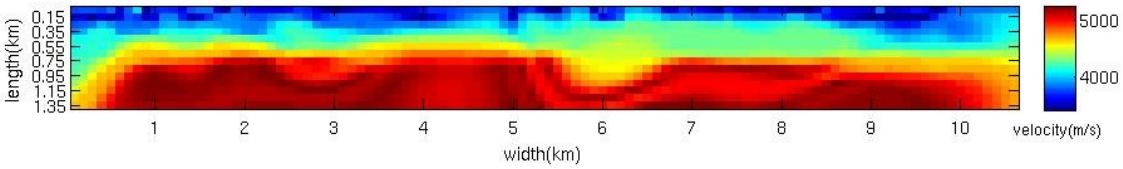


Figure 8.20

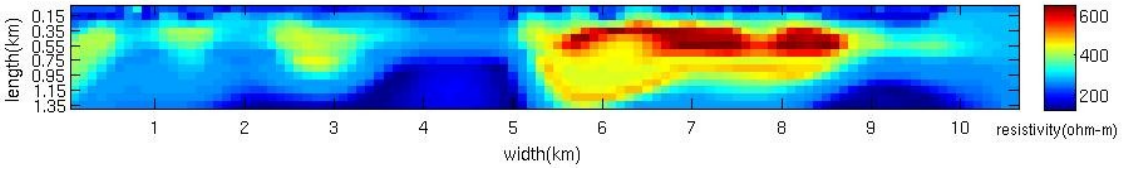


Figure 8.21

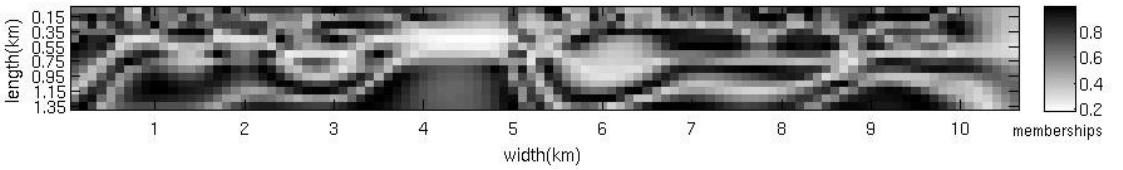


Figure 8.22

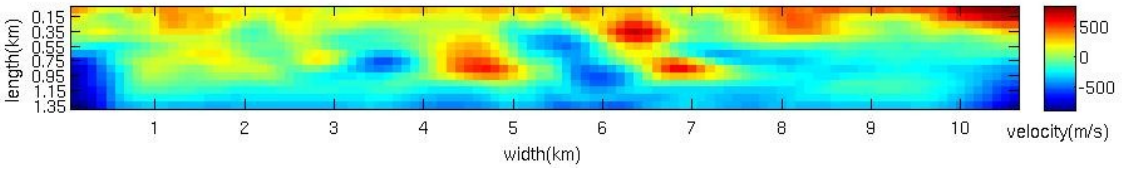
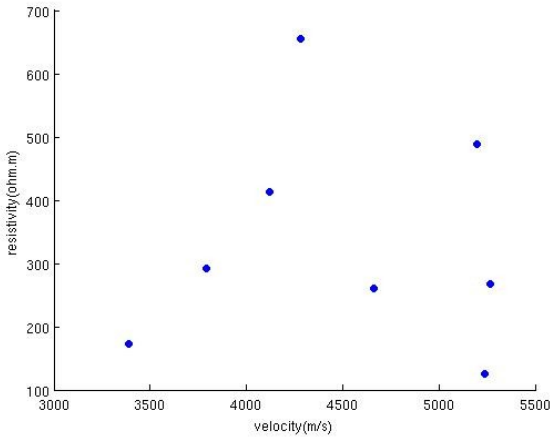


Figure 8.23

Figure 8.19 - 8.23 show 8 clusters after defuzzifications, parametric models after cooperative inversion with FCM (velocity and resistivity), membership of each cell (degree to which it belongs to cluster) and velocity perturbation (with respect to the starting model)

Figure 8.24 shows average value of resistivity and velocity for each cluster



If we compare parametric resistivity model (Fig. 8.20) with the one we got after the separate inversion (Fig. 8.1), it is quite clear that most resistive structure is not represented in a same way. One reason is that misfit criterion was reached too fast (with more iterations range of resistivities gets higher), but also fcm process itself forces clusters to have the shape in a way that should account for both data sets. On the other side what was already been noticed in previous example (fuzzifier equals 2 , Fig 7.14) that velocity structure is deformed between 6-8 km (along the profile). This can be interpreted as a geological layer that has a lower velocity than that of the overlying layer. But there is no strong evidence for this (low quality of data is also problem). Two most resistive clusters are located in this part of the model, this arrangement disturbs velocity model.

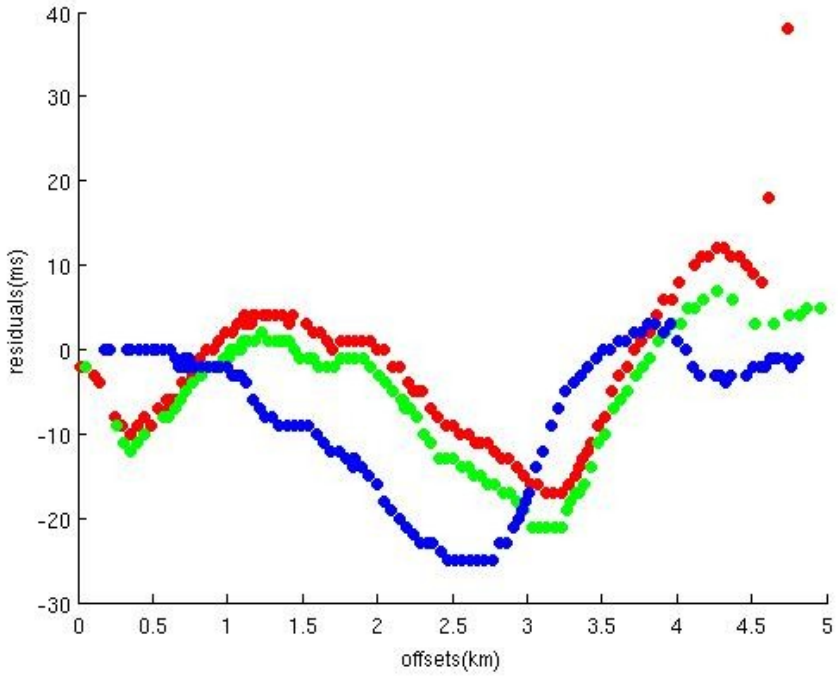


Figure 8.25 shows traveltime residuals for 3 sources , located at 5.35km, 5.42km and 5.75km respectively (they have good ray coverage for this specific 6-8 km region)

Fig. 8.24 provides further information about this region. As we go further away from sources, residuals are bigger and bigger. Residuals for these three sites (that cover properly this part of the model) indicate quite high negative residuals exactly in region we are analyzing.

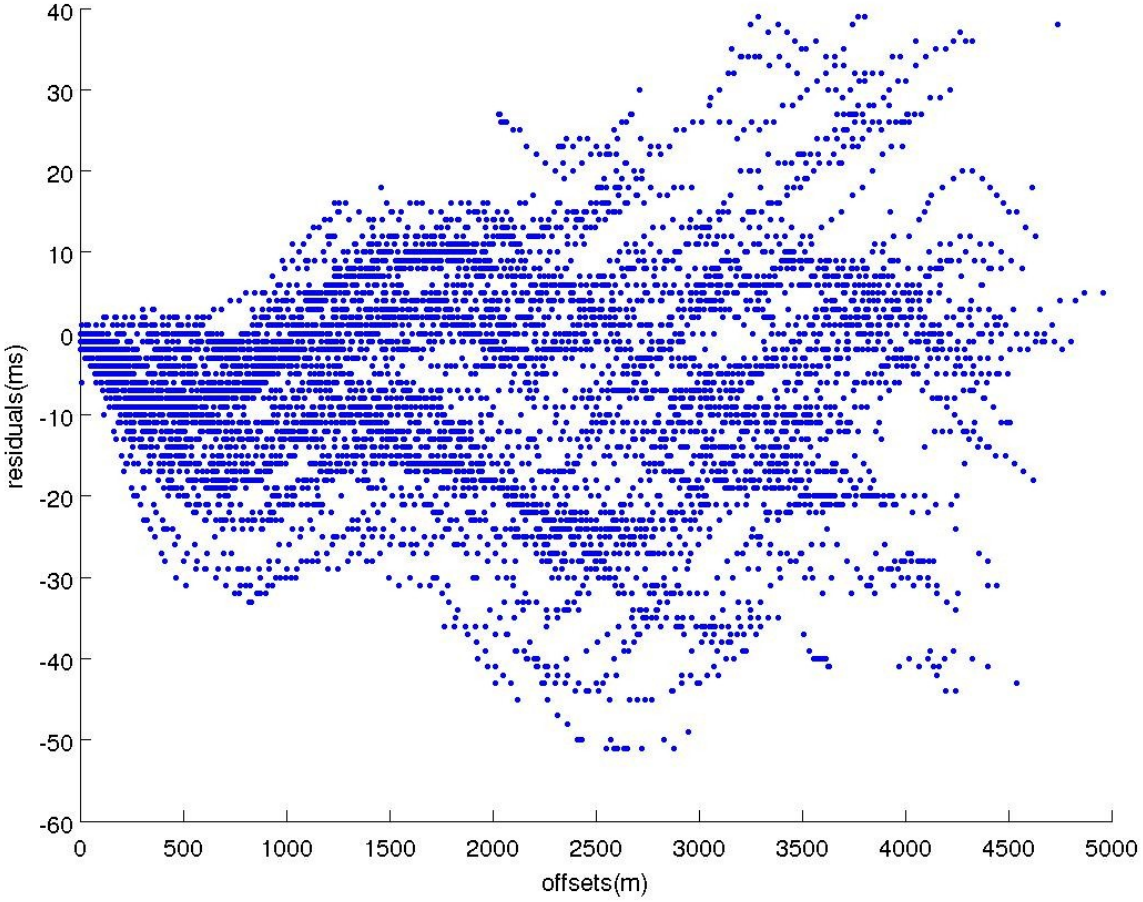


Figure 8.26 illustrates traveltime residuals for all shot receiver offsets

Figure 8.26 shows all 5736 residuals. It can be seen that at some offsets residuals have quite high values (-50 ms or 40ms on the positive side)

Notice that FCM clustering produces that three high-velocity clusters have different resistivity values (Fig. 8.19). From Fig 8.22 we can see that some elements have low membership values which could indicate dissimilarity between elements (the value of a fuzzy membership function is a mathematical characterization of cluster.). Because of different nature of MT and seismic forward problems, poorly resolved region in velocity model will add large traveltime residuals to the functional and have bigger RMS then previous step (stop the algorithm). On the other side MT 'speaks volume', so small poorly resolved regions here will not have such damaging impact on inversion process.

Perturbation analysis (Fig. 8.23) shows as velocity with respect to starting model. It is obvious that first layers now have too high velocities, we have chosen starting model(1-D) as a decent approximation of field data results, so we know that these high velocities are quite unrealistic. Ray coverage in lower corners of the model is quite low, so we know these parts are not constrained well.

The reasons why it does not perform well here are due to inherent limitations of fuzzy c-means algorithm. FCM is very sensitive to noise, also it does not consider any information about spatial context, two neighbouring cells could have very similar or dissimilar values, they will be treated as all others . FCM does not work well when clusters' s sizes or densities are different (in our case densities could be quite different, especially because of the broadening of resistivity range with every iteration) Also cluster size can not be predicted in advance.

In some pattern recognition algorithms , step known as clustering tendency should be involved. This includes various tests that indicate whether or not the available data possess a clustering structure (Theodoris and Koutroumbas, 2009). FCM could be used for this, to say if there is substructure in data at all (in case of seismic refraction and MT). Latter when we present other two algorithms and compare the results, it will be shown that fcm parametric models are very similar to other two.

MT FUNCTIONS

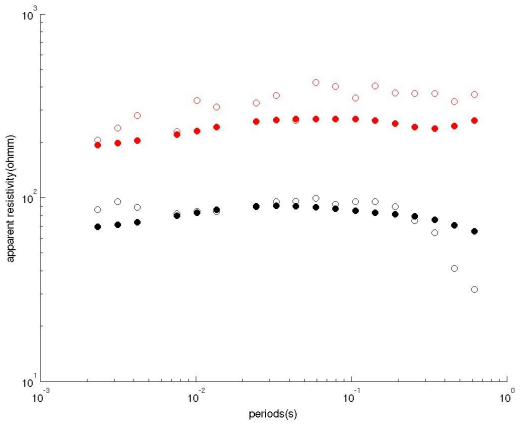


Figure 8.27

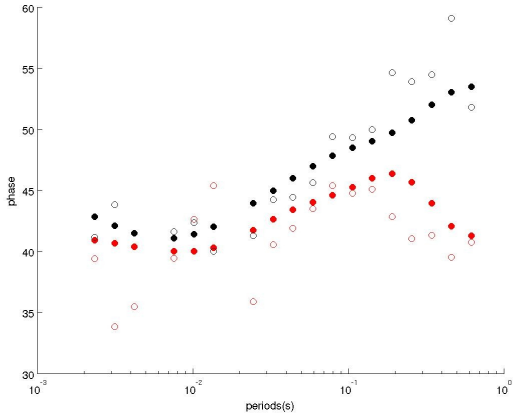


Figure 8.28

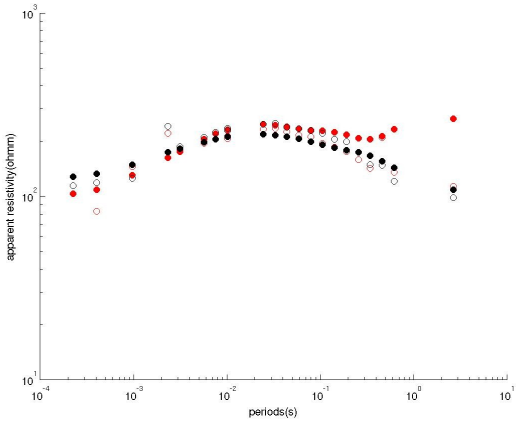


Figure 8.29

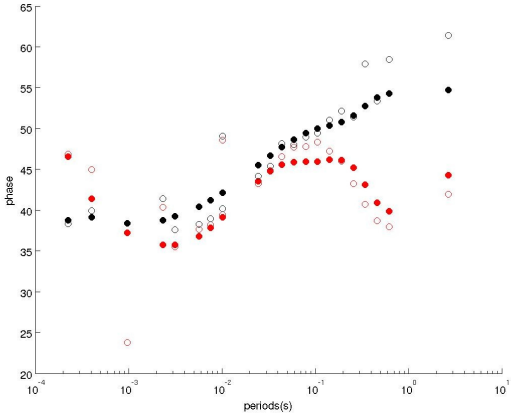


Figure 8.30

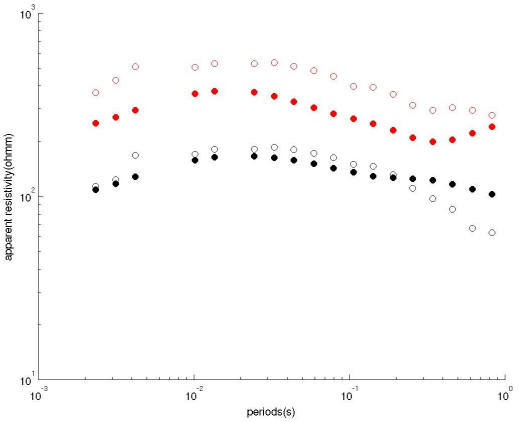


Figure 8.31

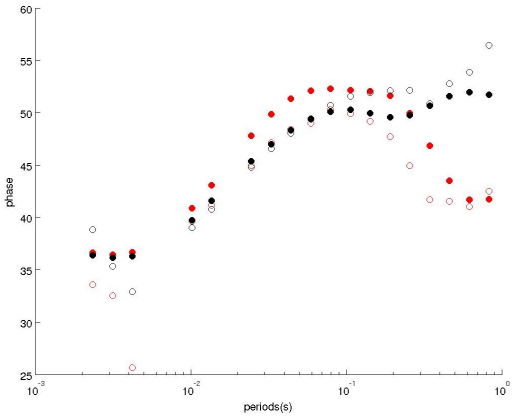


Figure 8.32

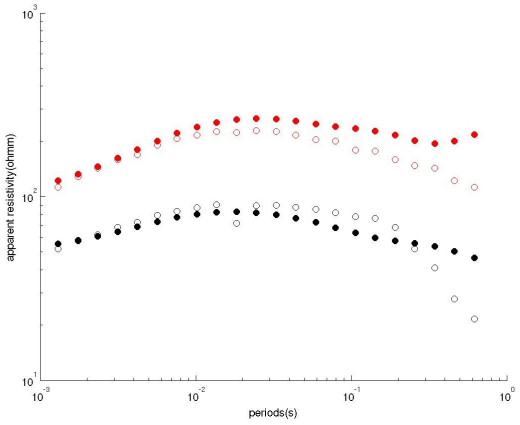


Figure 8.33

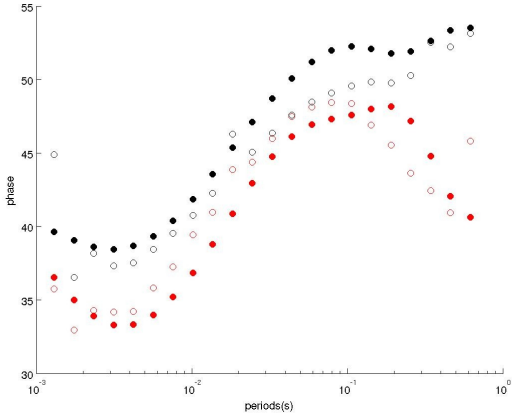


Figure 8.34

Fig. 8.27- 8.34 shows apparent resistivity and phase for E mode(red) and H mode(black) for sites 3,5,12,14. Empty circles denote fieldwork data, filled circles inversion response

In the previous chapter, the research conducted with fcm algorithm was shown, and the weaknesses of that approach were discussed. This chapter will build on this, trying to improve this, by changing the distance metric (now Mahalanobis distances will be used instead of Eucleadean) with all other steps in integration of two models being the same.

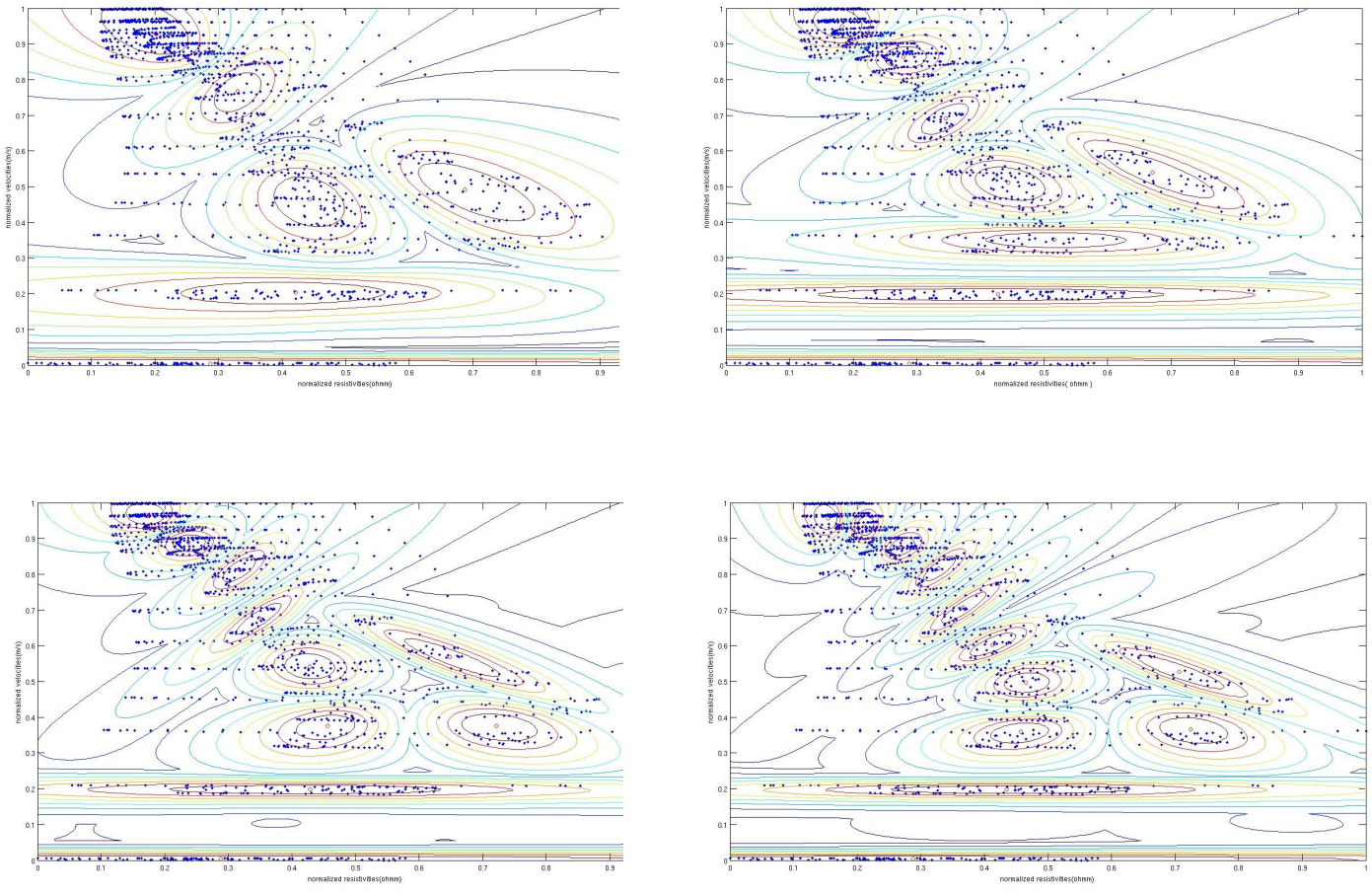
Parameters:
number of clusters: 6,7,8,10,12
weighting exponent 2
termination toleration $\varepsilon=0.001$
cluster volumes $\rho_i=1$,2

We will change cluster numbers and cluster volumes, to get the best integration. What we really want is cluster arrangement in that way, that enables both parametric models to give a smaller misfit at each iteration. It is desired to have as many iterations as possible until a stopping criterion is met. If we have gradual development of common features then this can show that models are consistent with data.

Two- dimensional data enables visual inspection of points in data space, so we can compare cluster organization and how it works with GK algorithm. Figures 7.33-7.36 show clusters for four different numbers of clusters, GK is always looking for ellipsoid shape of clusters (more details about theory in 4.1.2). In the previous chapter, it was clearly shown that sphere arrangement of points(fcm) is limiting factor that prevents iterative procedure to repeat in numerous steps. Also fcm is not invariant to linear transformations, GK is. Thus it is more robust concerning the scaling of our data sets.

First iteration is the same, start with initial models. Figures below show what clusters (6,8,10,12) look like after, for normalized velocities and resistivities . It is intended to show how GK creates new clusters when we increase their number. Before normalization velocity is in the range of 2849-5552 m/s, while the range of resistivity is 66.701 - 569.457 Ωm.

From 8.35-8.38 it is quite obvious that if the number of clusters is increased, there are still some outliers, points far away from cluster centers. Although not so numerous they will play role in creation of clusters and consequently the initial model for both inversions. The very fact that we are combining pattern recognition with two inverse problems makes situation much more complicated. Clustering precipitates inversions (both MT and seismic), so the presumption is that points that are similar to our criterion will lead to adequate starting models for inversion(this needs to be proved).



Figures 8.35-8.38 show cluster centers and clusters for 6,8,10,12 solution after first iteration. Both velocities and resistivities are normalized (in a range zero-one)

c=10

m=2

ε=0.001

ρ_i=1

Below are 10 cluster solution results obtained after 4 iterations (next step gives larger RMS misfit for traveltimes).

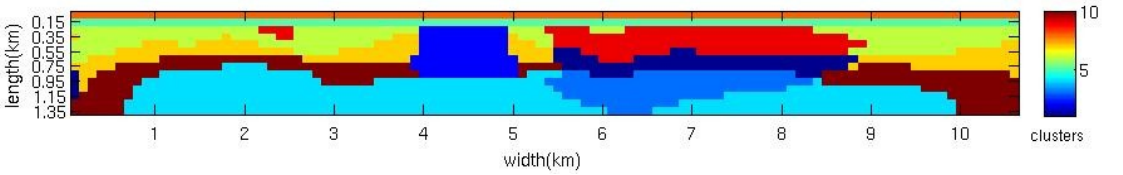


Figure 8.39

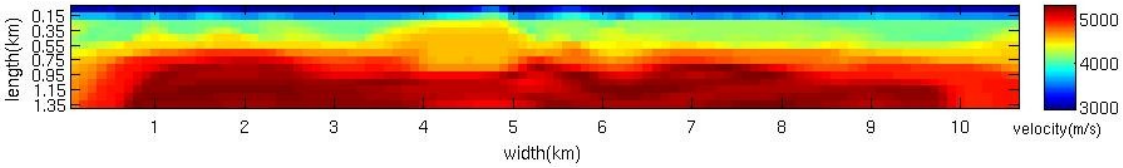


Figure 8.40

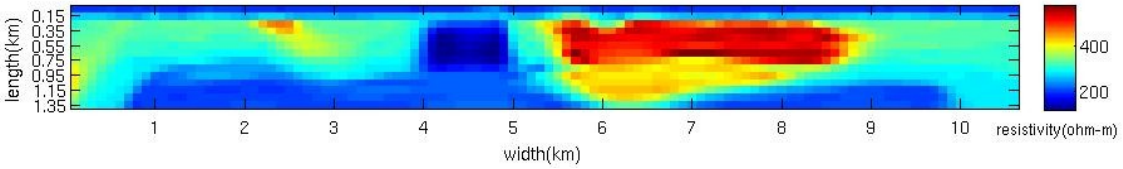


Figure 8.41

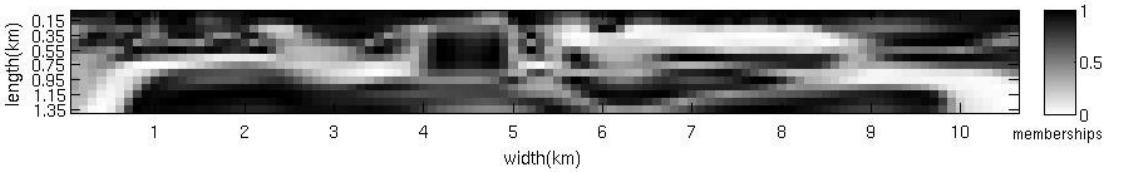


Figure 8.42

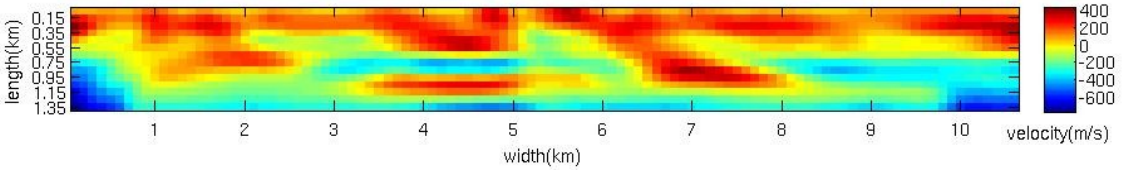


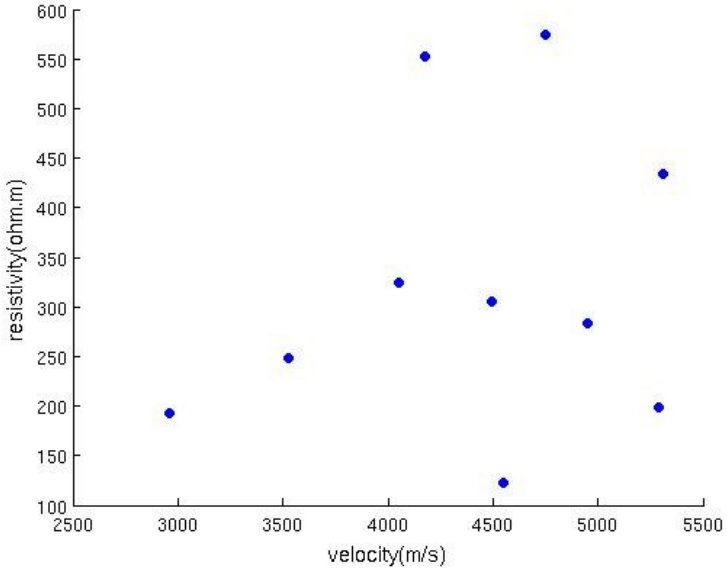
Figure 8.43

Figure 8.39-8.43 show 10 clusters after defuzzification, parametric models after cooperative inversion with GK(velocity and resistivity), membership of each cell (degree to which it belongs to cluster) and velocity perturbation (with respect to the starting model)

Iteration	1	2	3	4
RMS traveltime	18.92	13.58	13.24	13.01
RMS MT	2.76	2.61	2.54	2.48

Table 12 shows how RMS misfit change for traveltime and MT functions at every iteration

Figure 8.44 shows the average value of resistivity and velocity for each cluster



Figures 8.39 and 8.41 have a couple of striking features that will have repercussion also on the parametric velocity model. The most resistive cluster creates a rectangle close to the center of the model. Two highly-resistive clusters are located one above the other (unlike separate MT inversion, Fig 8.1), similar to what we have already seen in previous chapter (fcm). Also a resistive 'island ' in the upper 2-3 km (along the profile), is a little smeared but creates a distinct formation. From 8.42 a couple of regions can be seen with a very low membership function value which is a consequence of having 10 clusters (quite a big part of the model). Low memberships does not mean that these model elements will be matched in incorrect way, but it still adds uncertainty. Velocity perturbation comparing to the initial model also shows positive gradient in the upper part of the model. Note that the RMS after the 4^{th} iteration is still quite high, so this implies that this velocity structure is not perfect. Both lower corners are with high discrepancies with the starting model, but that part of model is not covered with rays properly. From Figure 8.44, it can been seen that some clusters have same velocities but the average resistivity is quite different (this could be related to different geological structures)

As has already been noted (research with synthetic data, chapter 7), once there is dominant structure it is very likely that it will appear in the next iterations. We should bear in mind that starting models in traveltime tomography is important and the chances for successful estimation of true model are improved if the starting model is selected in the neighborhood of the true model. Range of resistivity elements tends to be larger with every iteration, but this will not eradicate dominant structures. The rectangle-shaped cluster, formed after the second step, keeps appearing in following iterations with more less the same elements. This cluster is with high values of membership function. Implications of this are visible also for its velocity counterpart. Note that resistive structure will dominate the cluster creation (MT inversion is dominated by field data, at least with this set of optimization parameters).

Now the same region (6-8km), is analyzed again with same three sources, as was done in the previous chapter. Two most resistive clusters will affect the velocity model.

As Figure 8.45 shows that good results are obtained, residuals here are much smaller, which confirms that velocity model from fcm was not the adequate one.

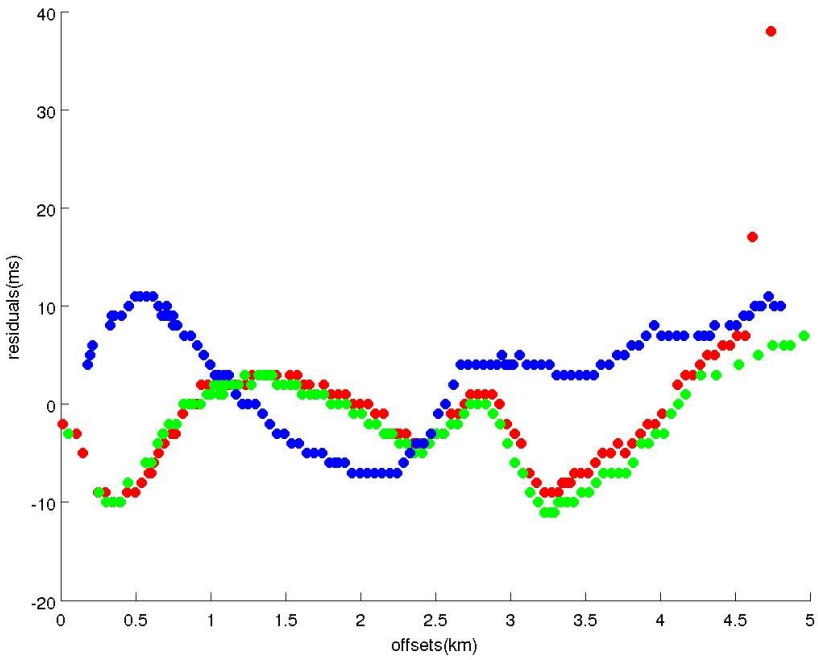


Figure 8.45 shows traveltime residuals for 3 sources , located at 5.35km,5.42km and 5.75km respectively (they have god ray coverage for this specific 6-8 km region)

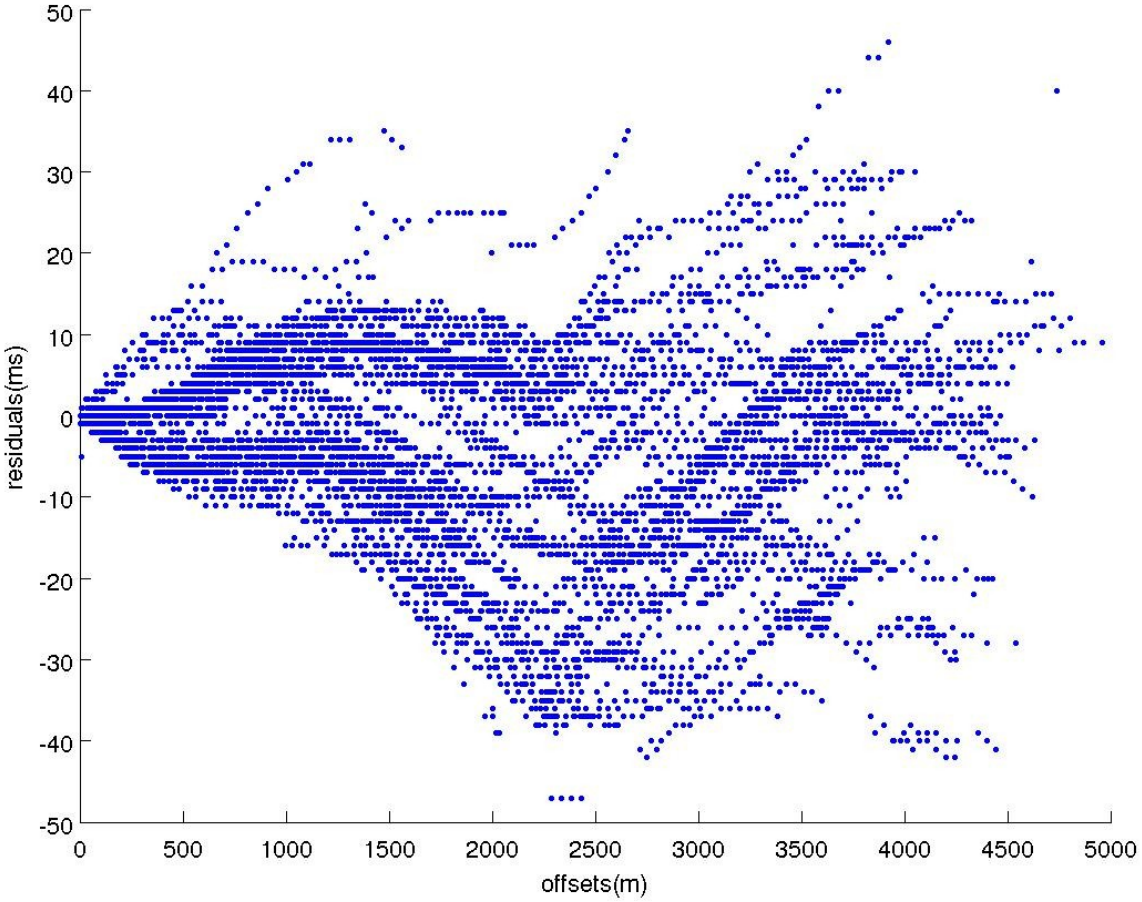


Figure 8.46 illustrates traveltime residuals for all shot receiver offsets

The majority of residuals (Fig 8.46) for smaller offsets(up to 1500m) are quite small, but as offsets become larger, huge negative residuals appear

c=12

m=2

ε=0.001

ρ_i=1

Two more clusters have been introduced to see if this will improve integration of data.

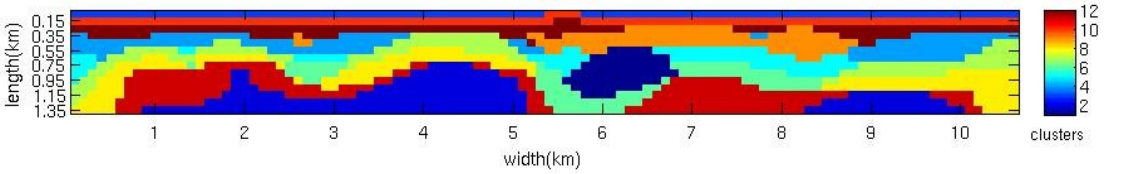


Figure 8.47

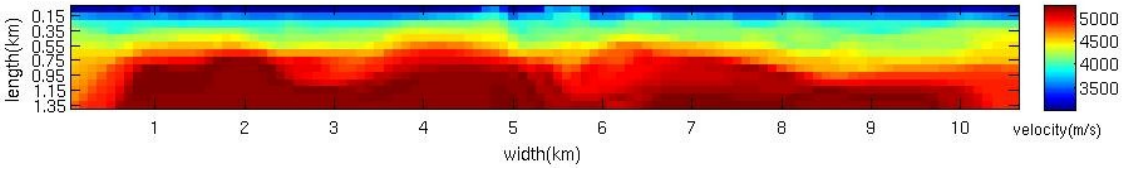


Figure 8.48

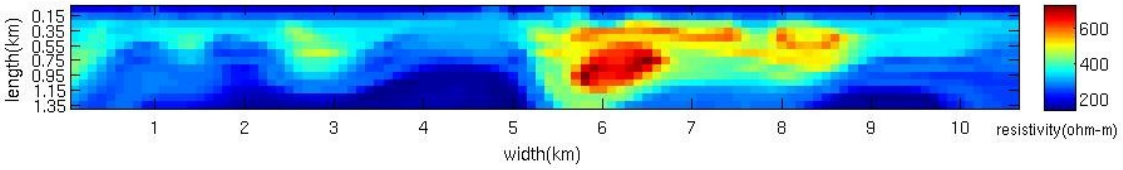


Figure 8.49

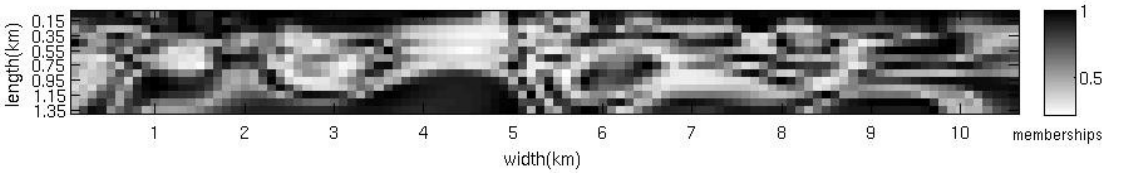


Figure 8.50

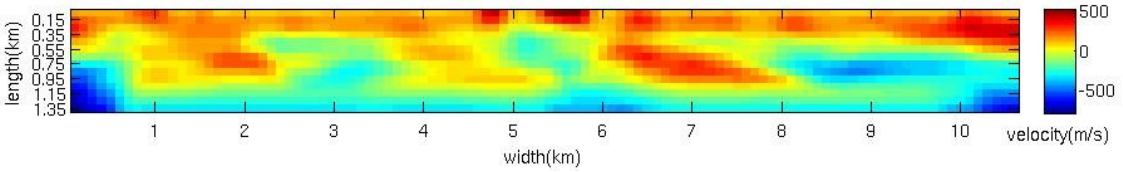


Figure 8.51

Figures 8.47-8.51 show 12 clusters after defuzzification, parametric models after cooperative inversion with GK(velocity and resistivity), membership of each cell (degree to which it belongs to cluster) and velocity perturbation (with respect to the starting model)

iteration	1	2	3	4	5
RMS traveltime	18.92	13.01	12.62	12.22	11.88
RMS MT	2.76	2.61	2.54	2.44	2.32

Table 13 shows how RMS misfit change for traveltime and MT functions at every iteration
In this case RMS was reached after one more iteration (5), with better RMS for both inversions (as shown in Table 5). One iteration more in case of MT, broadened the resistivity range which will also affect the layout of points in two-dimensional parametric space and inversion process.

Note that at 8.49, resistive parametric model starts to look like the separately inverted (Figure 8.1). Unlike two previous cases where a big resistivity structure was present with a duplicate below it, now more iterations allows allow the model to have separately one structure at 6-7 km and another structure at approximately 8-9 km(along the profile). MT curves are shown in 8.56-8.63 and still apparent resistivity is better estimated than phase . Some data points are distant from the curve(bad data), but these points can not be improved . Cluster layout implies a huge number of cells with low membership values in upper part of the model. Velocity perturbation is very similar to the previous velocity perturbation (from 10-cluster solution).

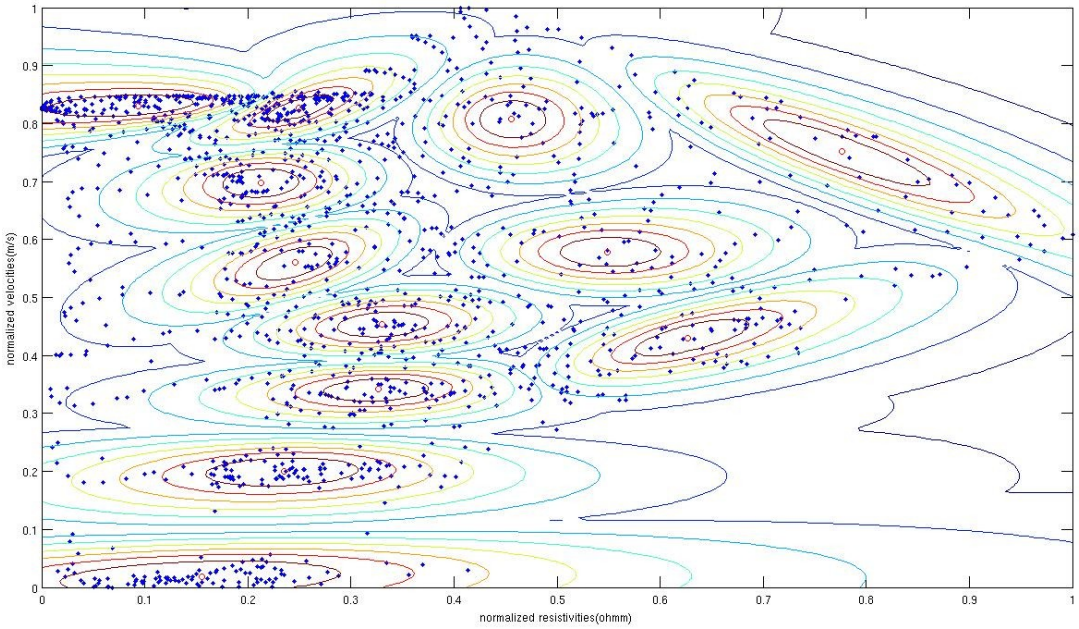


Figure 8.52 shows clusters after 5 iteration

Comparing Figures 8.52 and 8.38, it is quite visible that clusters were much more homogeneous after 1st iteration than after the 5th iteration. Resistivity range was 83.6-454 in the first and 50.3-931.6 in the last iteration. For example, for two cells with resistivity values of(100 Ωm ,200Ωm) correspond to normalized values in the first (0.044,0.313) and (0.057,0.169) in the last iteration, in normalized space now they are much closer .

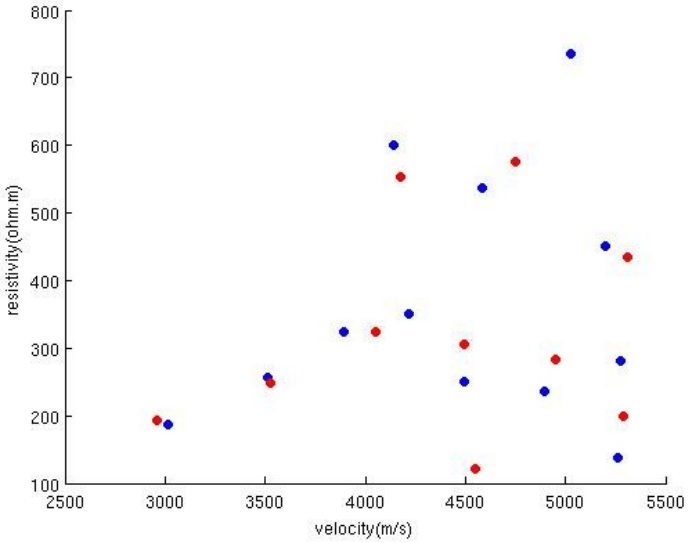


Figure 8.53 shows the average value of resistivity and velocity for each cluster (blue circles for 12 clusters). Also previous 10 cluster results were added with red circles to compare the results.

Figure 8.53 gives us opportunity to compare average cluster values for both solutions. Except the most resistive cluster, we could say that there is similarity between models

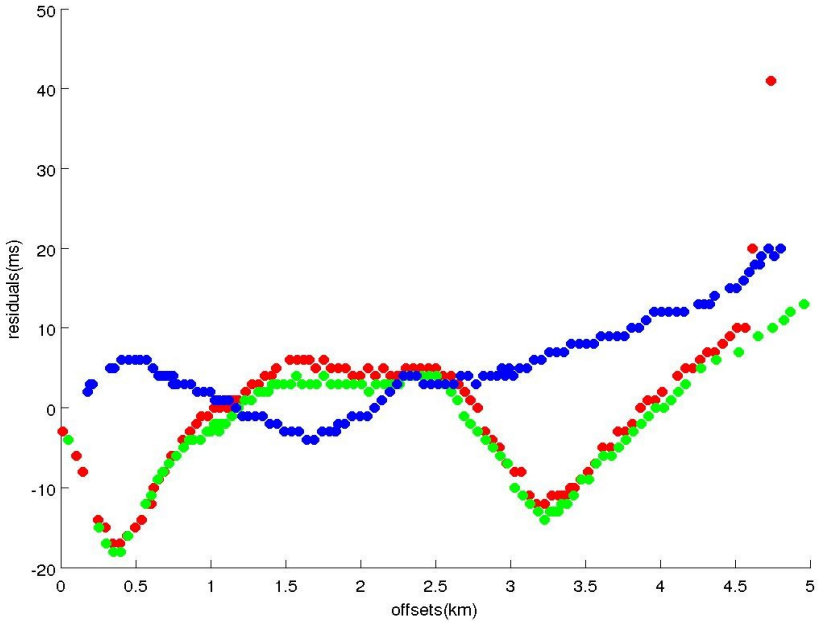


Figure 8.54 shows that good results are obtained, although negative residuals here are larger then for the previous solution(at the same offsets

Figure 8.54 shows traveltime residuals for 3 sources ,located at 5.35km,5.42km and 5.75km

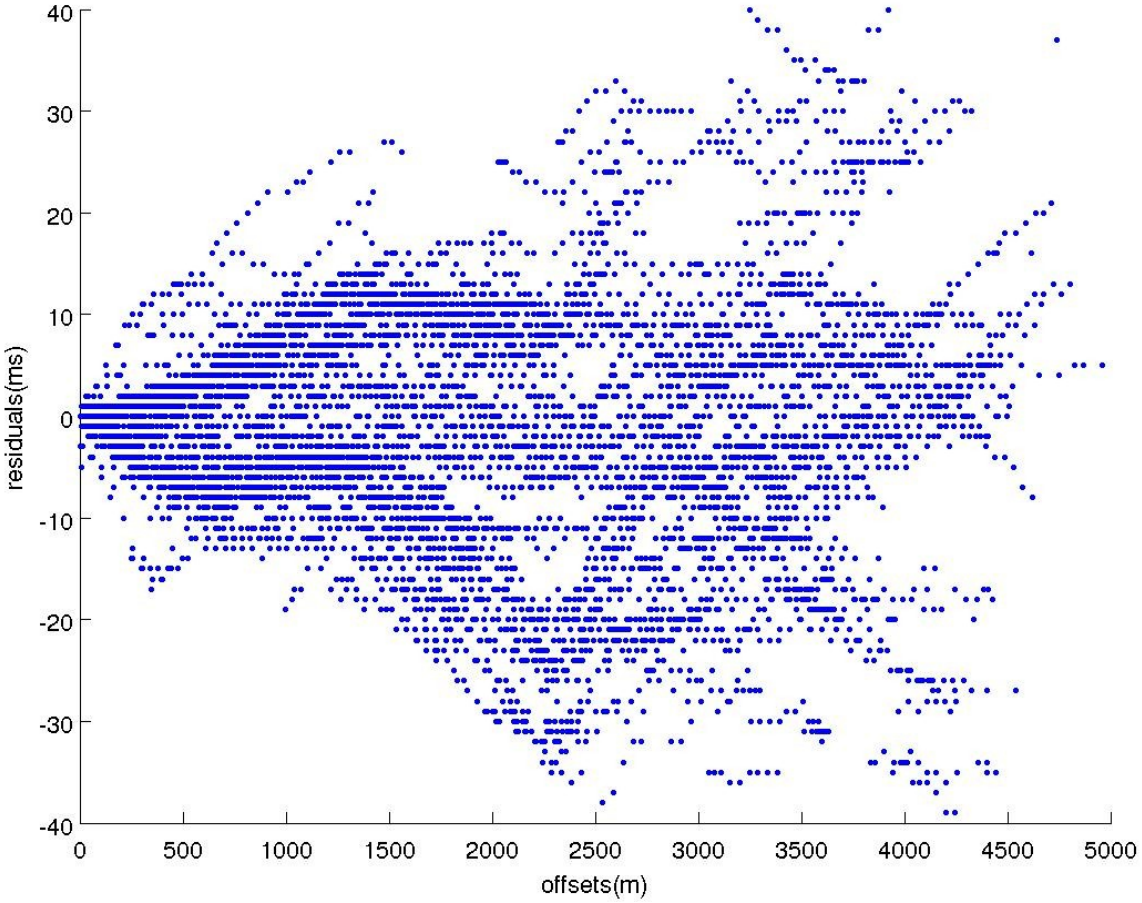


Figure 8.55 illustrates traveltime residuals for all shot receiver offsets

Note (Figure 8.55) that the range of the residuals is now smaller comparing to fcm results (Figure 8.26). Still there are some offsets with large residuals. Although they are not so numerous (most of the points are up to 10ms, on one or other side), they are limiting factor in getting better results.

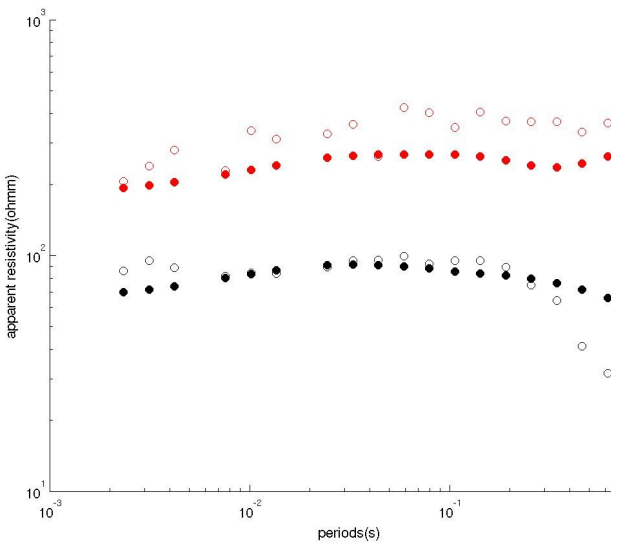


Figure 8.56

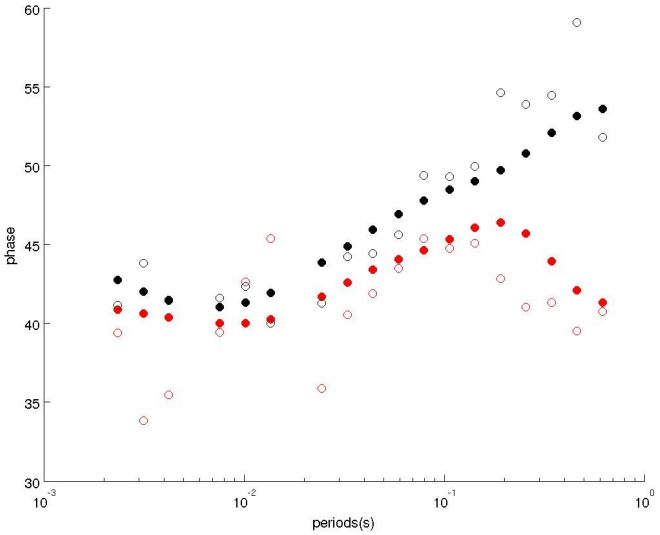


Figure 8.57

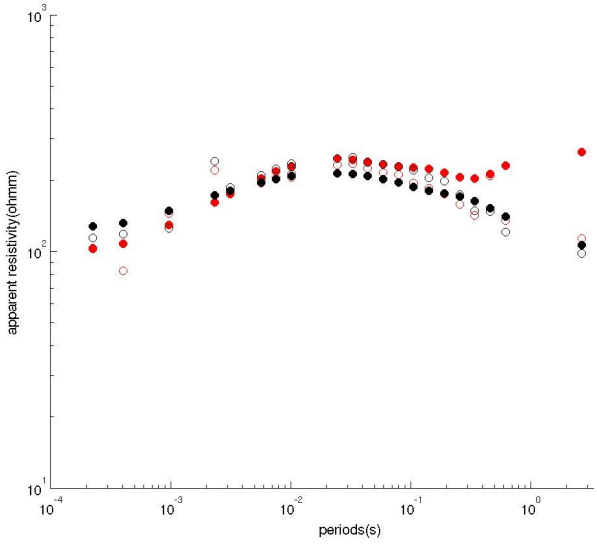


Figure 8.58

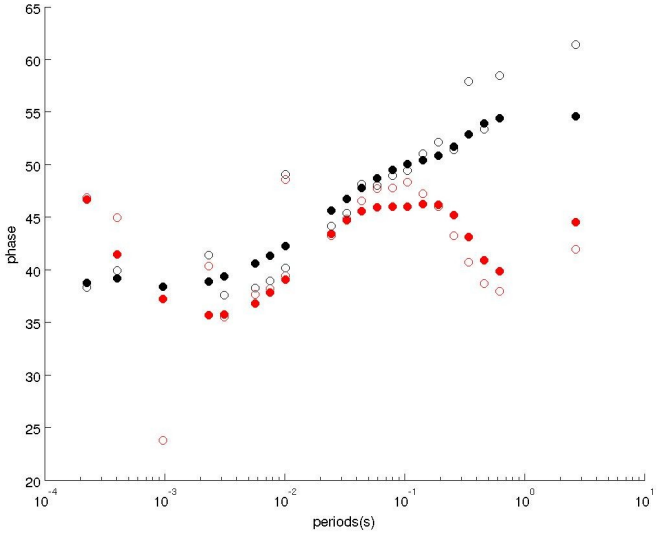


Figure 8.59

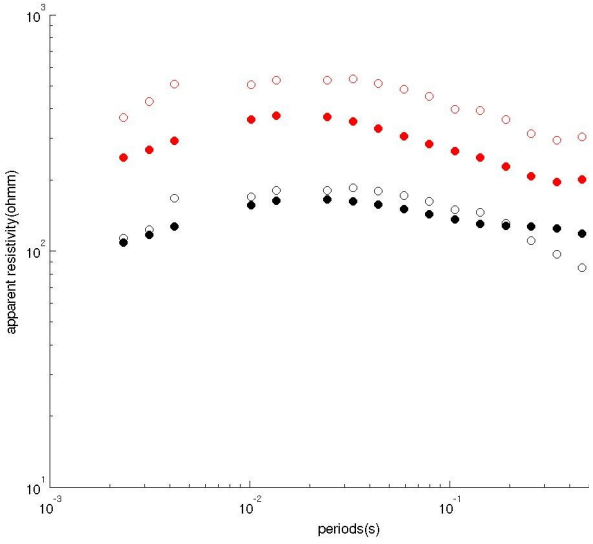


Figure 8.60

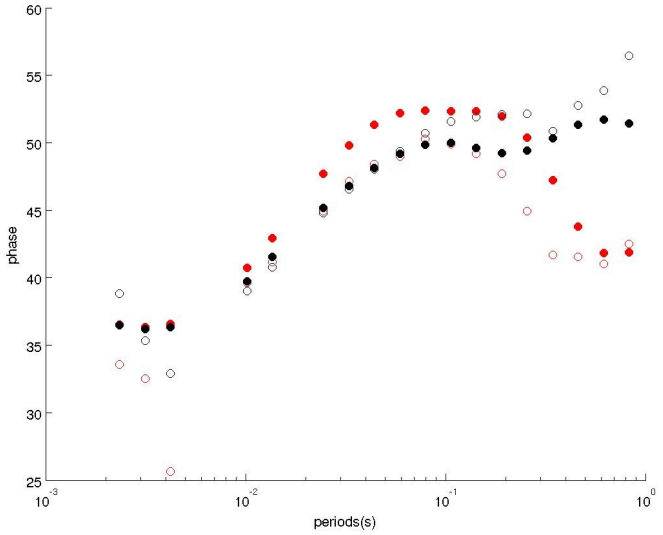


Figure 8.61

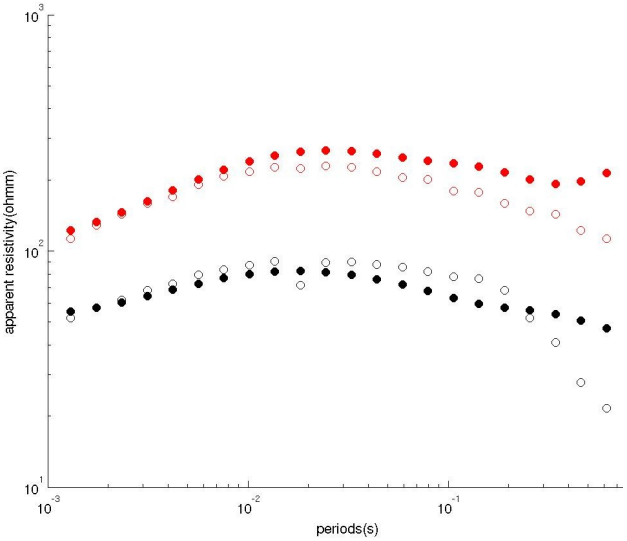


Figure 8.62

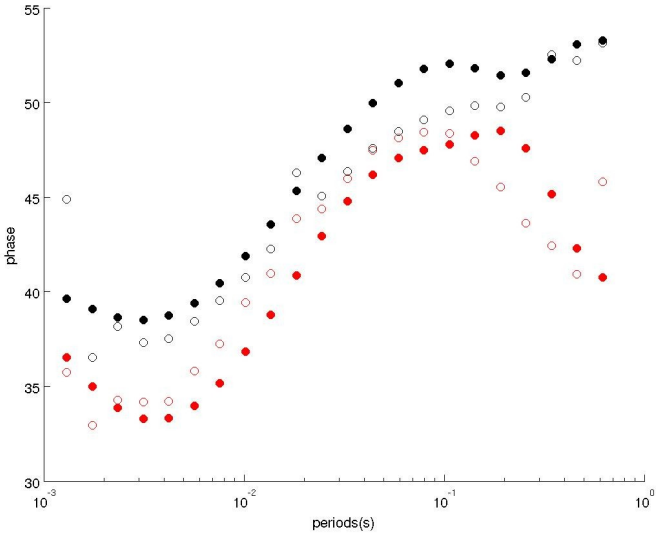


Figure 8.63

Fig. 8.56-8.63 show apparent resistivity and phase for E mode(red) and H mode(black) for sites 3,5,12,14. Empty circles denote fieldwork data, filled circles inversion response

In order to overcome limitations of previous two algorithms, we have researched possible integration of two geophysical data sets with the Gath-Geva algorithm. In contrast to the FCM algorithm and the Gustafson-Kessel algorithm , this one is not based on an objective function, but is a fuzzification of statistical estimators. This algorithm is able to detect clusters of varying shapes, sizes and densities. A detailed explanation of this algorithm was given in chapter 4 (4.1.3).

The memberships represent a somewhat strange picture because of the occurrence of the exponential function within the distance, all distances are more or less divided into two ranges:close and remote. This is not the preferred option for kind of a problem we are dealing with, because it implicates that all cluster elements have the same or very similar parameter values. It will be shown that for a small number of clusters, integration of two disparate data sets (MT and seismic refraction) is practically impossible. For a small number of clusters, the algorithm reaches a stopping criterion very fast(after 2 or 3 iterations). But with more clusters situation gets better and gives much more space for research.

Another important issue was initializations of prototypes ,FCM was used for initializations (this was part of the MATLAB code). We do not use results obtained with cooperative inversion with the fuzz y c means algorithm, rather we calculate at each step FCM clustering and then later go on with Gath-Geva algorithm.

Parameters

number of clusters: 6,7,8,10,12
weighting exponent 2
termination toleration ε=0.001

c=10
m=2
ε=0.001

Presented below are 10 cluster solution results obtained after 4 iterations (next step gives larger RMS traveltime misfit).

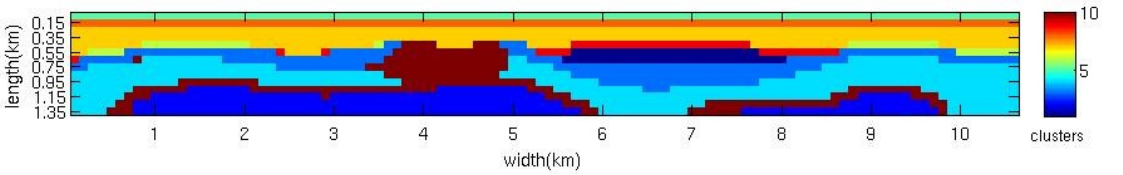


Figure 8.64

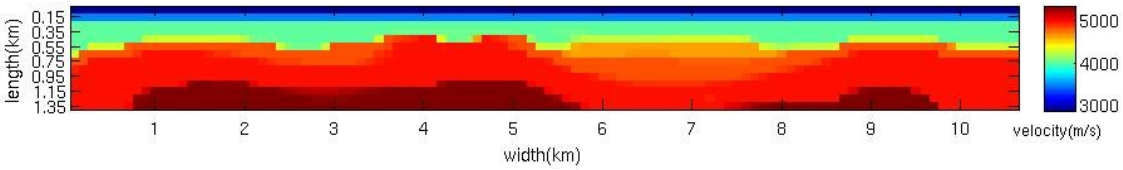


Figure 8.65

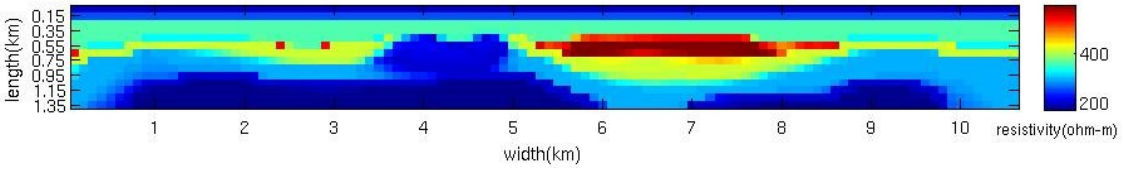


Figure 8.66

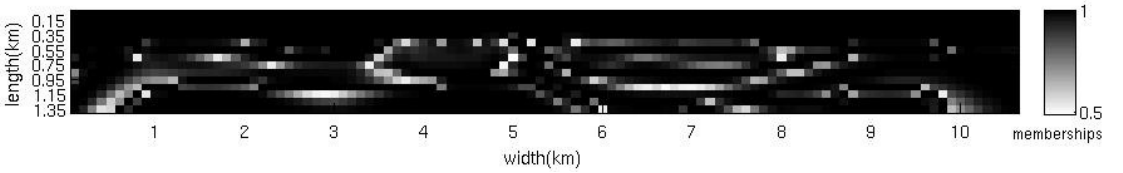


Figure 8.67

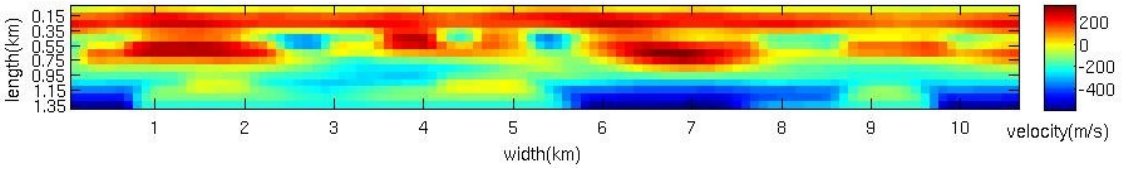


Figure 8.68

Figures 8.64-8.68 show 10 clusters after defuzzification, parametric models after cooperative inversion with GG(velocity and resistivity), membership of each cell (degree to which it belongs to cluster) and velocity perturbation (with respect the to starting model)

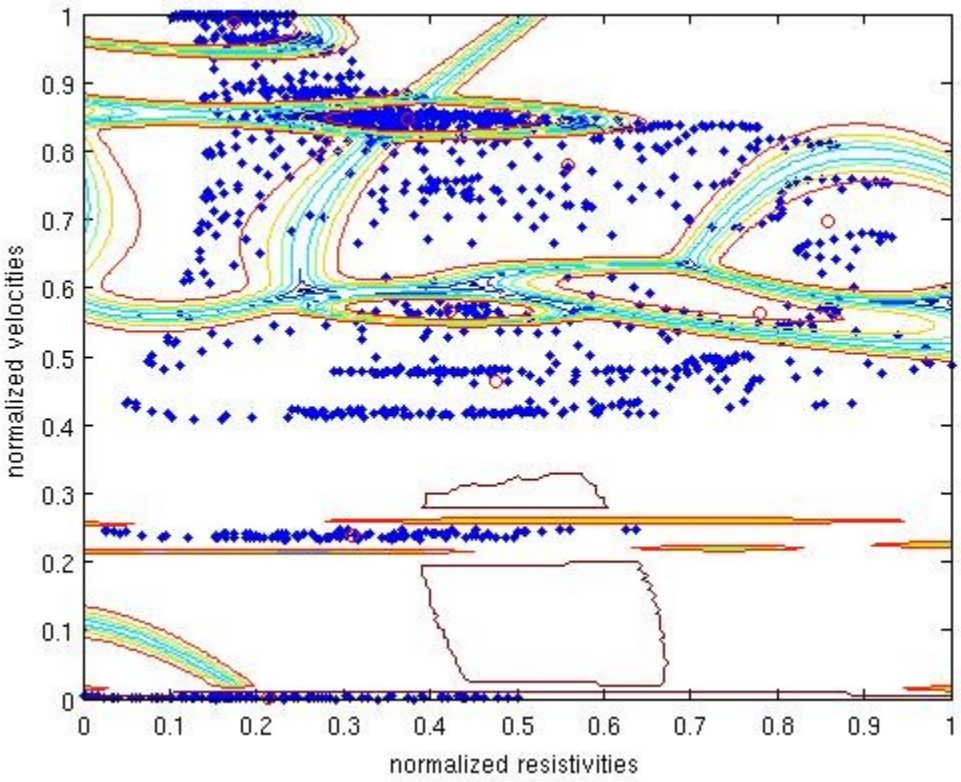


Figure 8.69 shows clusters after 4th iteration

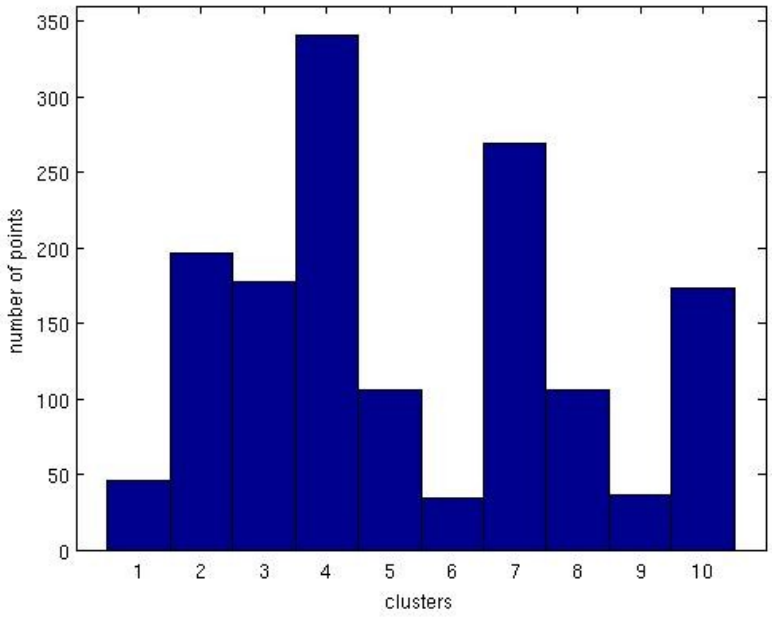


Figure 8.70 shows a distribution of points after clustering. There are clusters with different number of points, unlike the previous two algorithms where the clusters had similar dimensions.

iteration	1	2	3	4
RMS traveltime	18.92	14.62	14.12	13.08
RMS MT	2.76	2.54	2.48	2.44

Table 14 shows how RMS misift change for traveltime and MT functions at every iteration

From figures 8.63-8.68, it can be observed how changes were introduced comparing to previous work. Clusters now have different shapes and sizes. Having only points that are close to the cluster center or distant from it, each cell in velocity and resistivity parametric models will be almost equal to others from same cluster and quite distinctive from other cluster cells. This was not the case in previous algorithms. The similarity or dissimilarity of cells depends on layout of points in two dimensional space, which will latter be converted to cell values of parametric models. This was the main reason why this algorithm did not perform well for a small number of clusters. In that case it tends to put a large number of dissimilar elements into one cluster, inversion stops.

Memberships of most of the points is more than 99%, which leads to homogeneous structures. What is problematic, is the incapability of the algorithm to discern between two very similar structures. If points (of these structures) were allocated to the same cluster in vector space (of normalized velocities and resistivities) as a result of mathematical abstraction, two similar structures will be merged. Note in figure 8.69, that in some clusters velocities are grouped around value, but in others they are in still broad range. Introducing more clusters does not solve the issue. Twelve cluster solution will be presented later. Number of clusters is a really tricky issue and can not be determined with cluster validity criterion because each step has a link between pattern recognition and two inverse problems.

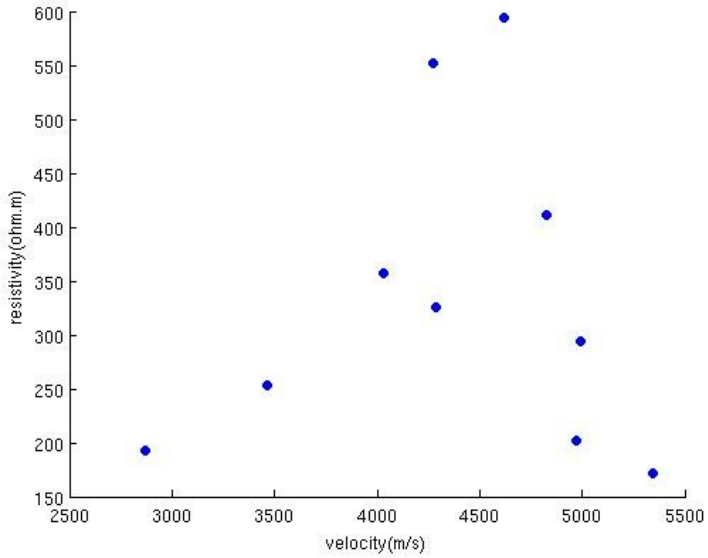


Figure 8.71 shows average value of resistivity and velocity for each cluster

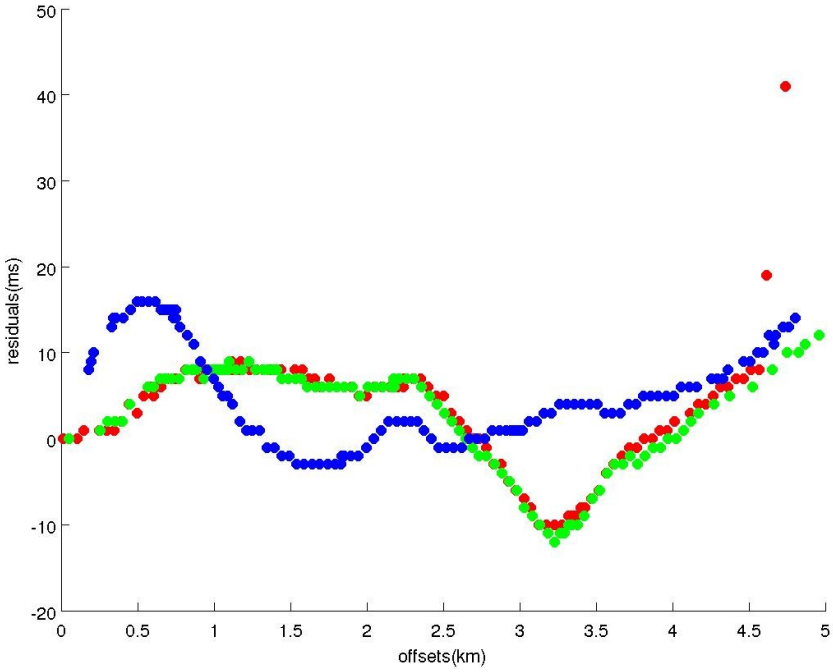


Figure 8.72 shows traveltime residuals for 3 sources,located at 5.35km,5.42km and 5.75km respectively

From Figure 8.72, the region of special interest (where high resistivity structures are) is well resolved. Residuals are very small,even better than with GK. On the other side RMS misfit is still high, and better results were expected. This indicates that some other parts of the model were not properly matched, and this falls in line with what was discussed above about limitations of the algorithm.

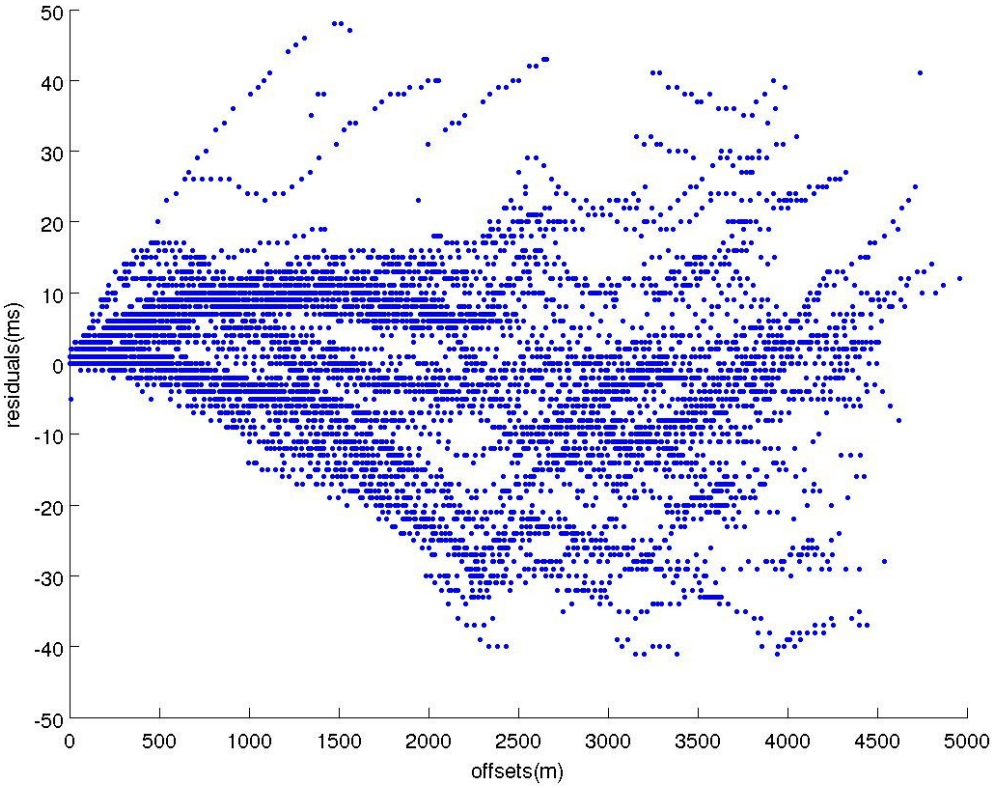


Figure 8.73 illustrates traveltime residuals for all shot receiver offsets

c=12

m=2

ε=0.001

Presented below are 12 cluster solution results obtained after 4 iterations (next step gives larger RMS traveltime misfit).

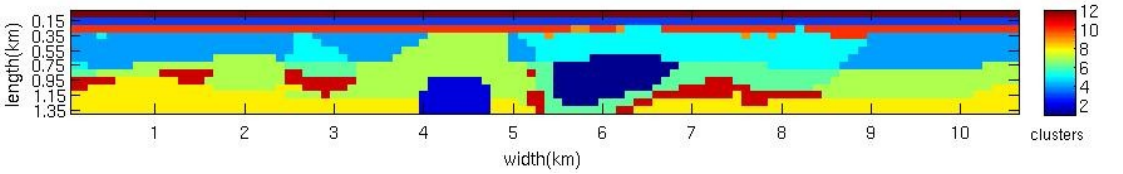


Figure 8.74

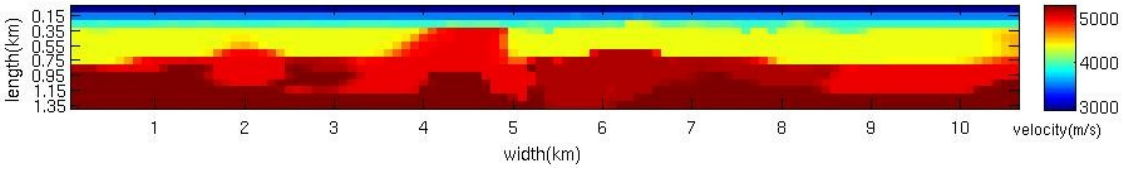


Figure 8.75

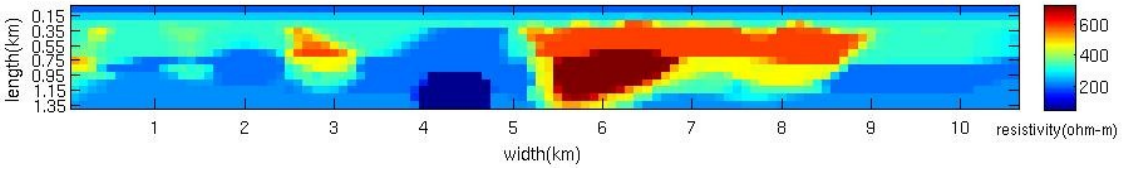


Figure 8.76

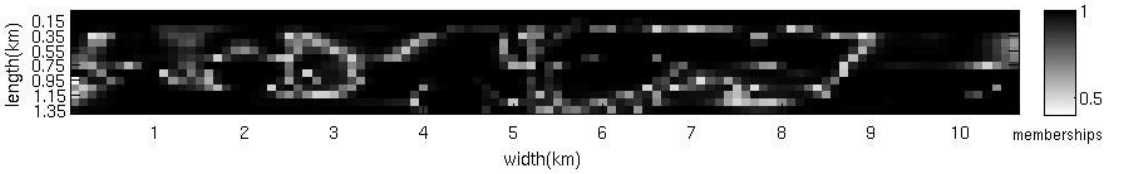


Figure 8.77

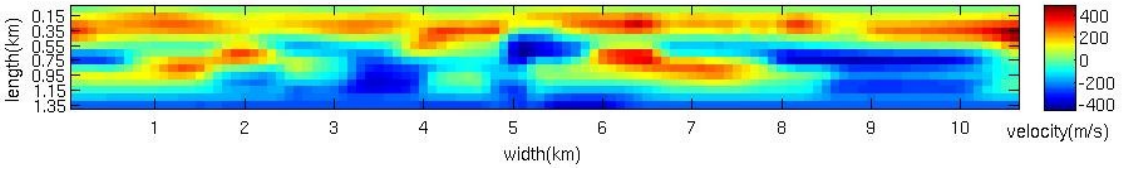


Figure 8.78

Figures 8.74-8.78 show 12 clusters after defuzzification, parametric models after cooperative inversion with GG(velocity and resistivity), membership of each cell (degree to which it belongs to cluster) and velocity perturbation (with respect to the starting model)

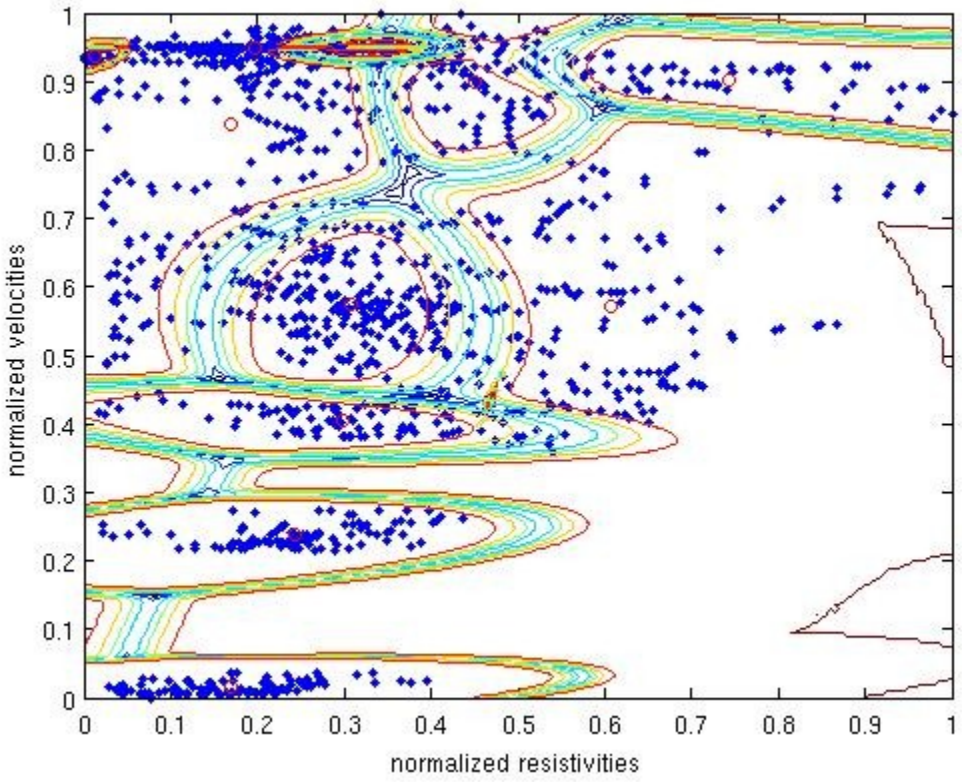


Figure 8.79 shows clusters after the 5th iteration

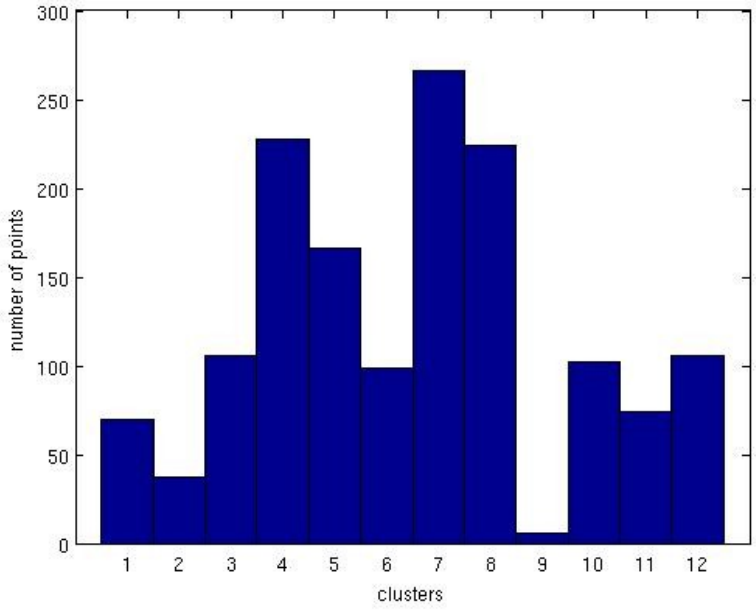


Figure 8.80 shows number of points per cluster

Again most cells are with high membership function (Figure 8.77). Detected clusters with cluster centers were shown at Figure 8.79. As expected, cluster sizes and densities are differing. With more cluster inevitably, there are some with small number of points, cluster 9 has only 6 points. Note in Fig. 8.77 that there are some clusters with velocities in a broad range. As mentioned in previous example, GG algorithm will calculate almost the same velocity and resistivity values for each cluster (due to high membership function). But big number of clusters will force parametric models to become heterogeneous, this especially affects the velocity model (note that in 10 cluster solution this was not the case). Between 4-5 km(along the profile), a structure appears which was not present with 10 clusters. The resistive structure is represented with 3 clusters, with the most resistive one being located almost at same location where it was in the case of a separate inversion (Figure 8.1) .

iteration	1	2	3	4
RMS traveltime	18.92	14.21	13.76	13.04
RMS MT	2.76	2.58	2.51	2.45

Table 15 shows how RMS misfit change for traveltime and MT functions at every iteration

Note that practically there has been no improvement in both RMS misfits. This confirms the assumption that adding more clusters does not necessarily guarantee better results.

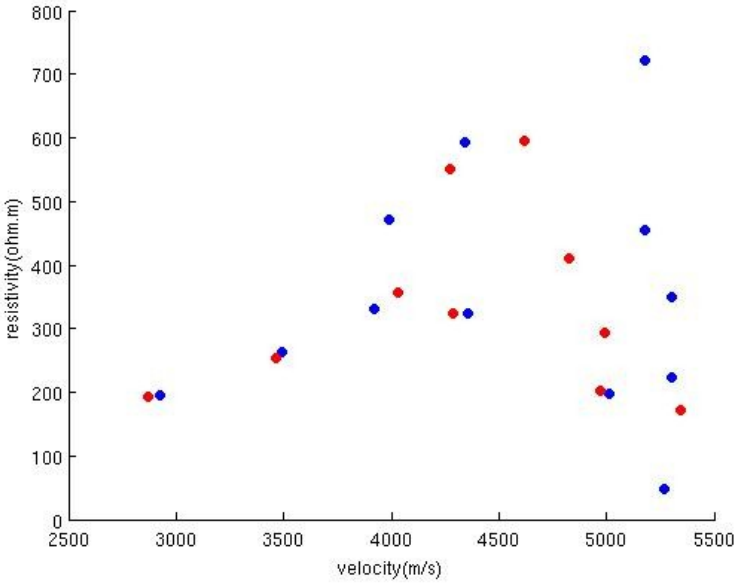


Figure 8.81 shows the average value of resistivity and velocity for each cluster (blue circles for 12 clusters). Also previous 10 cluster results were added with red circles to compare the results.

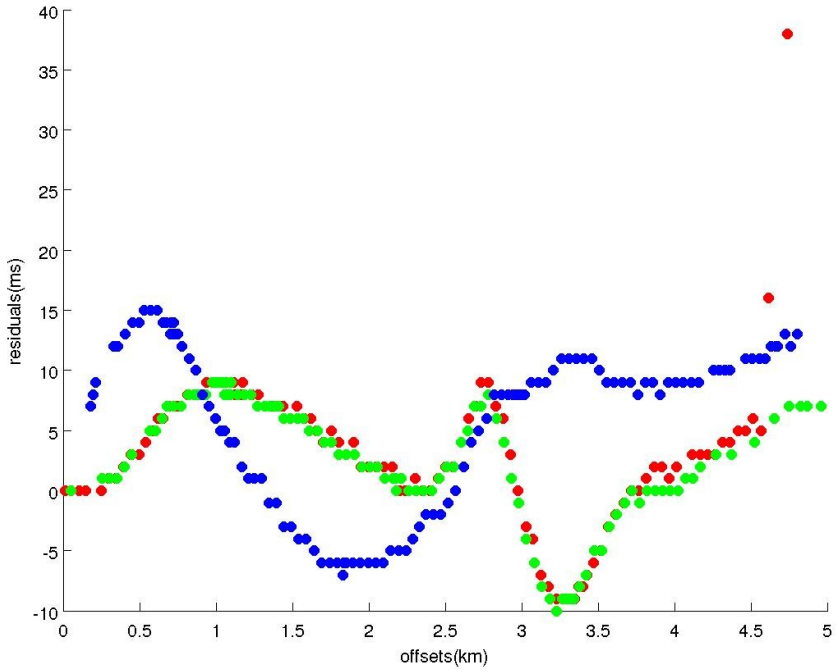


Figure 8.82 shows traveltime residuals for 3 sources,located at 5.35km,5.42km and 5.75km respectively

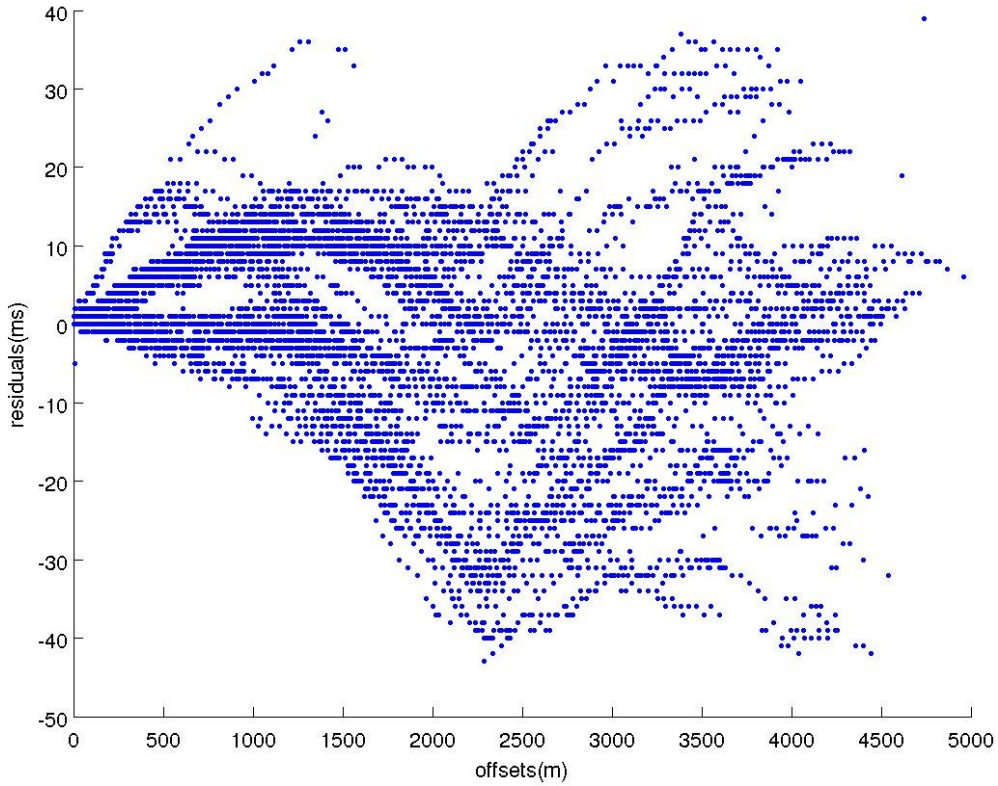


Figure 8.83 illustrates traveltime residuals for all shot receiver offsets

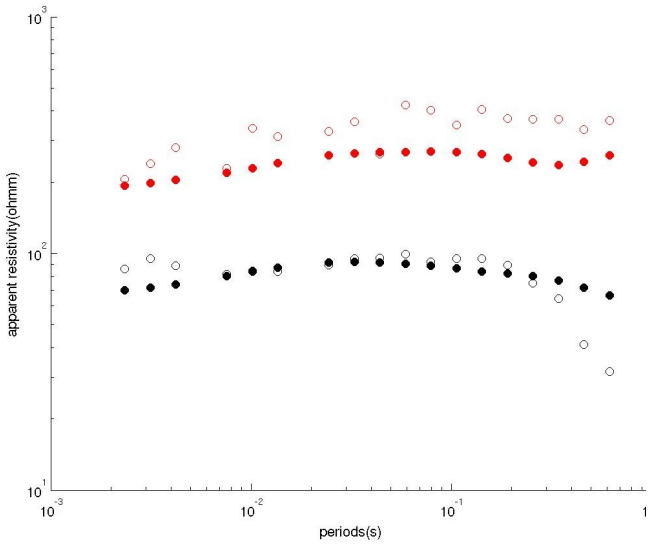


Figure 8.84

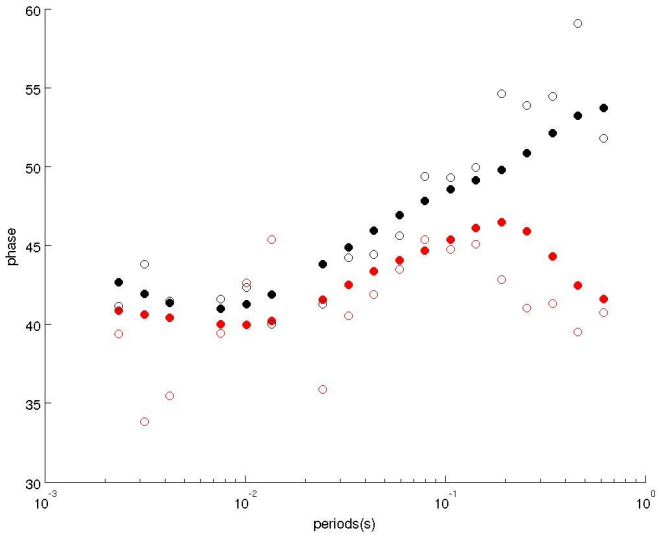


Figure 8.85

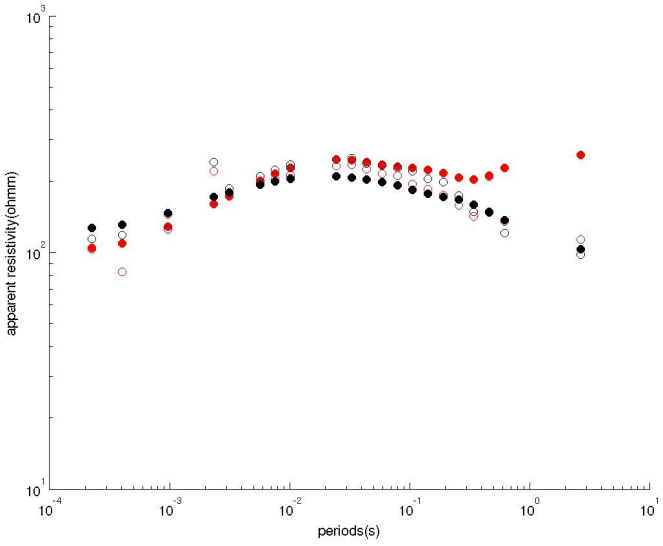


Figure 8.86

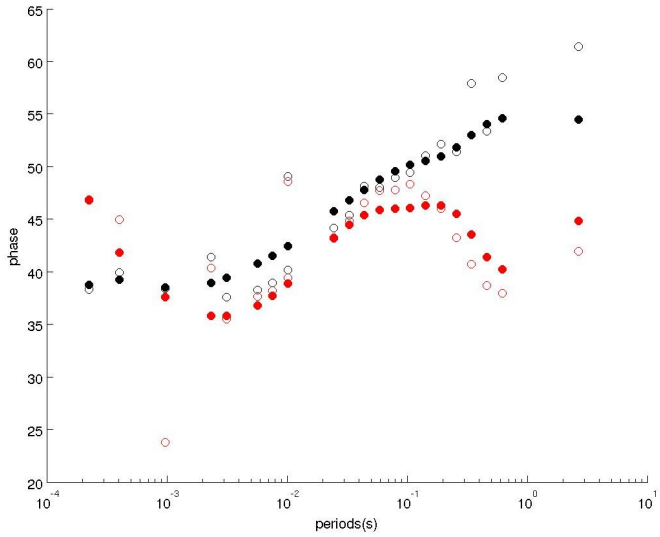


Figure 8.87

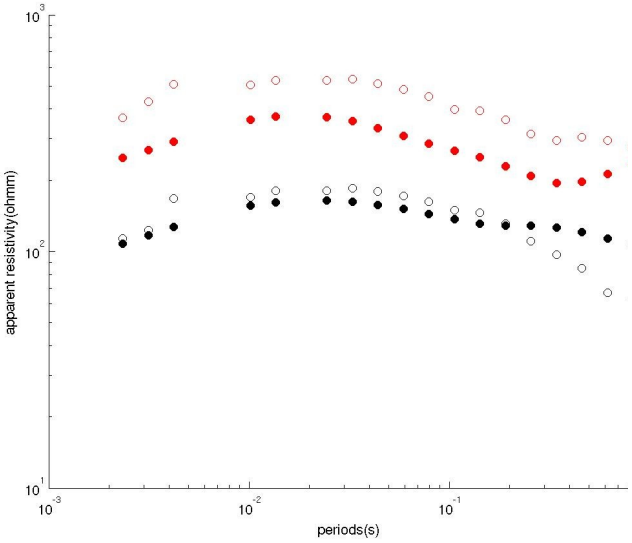


Figure 8.88

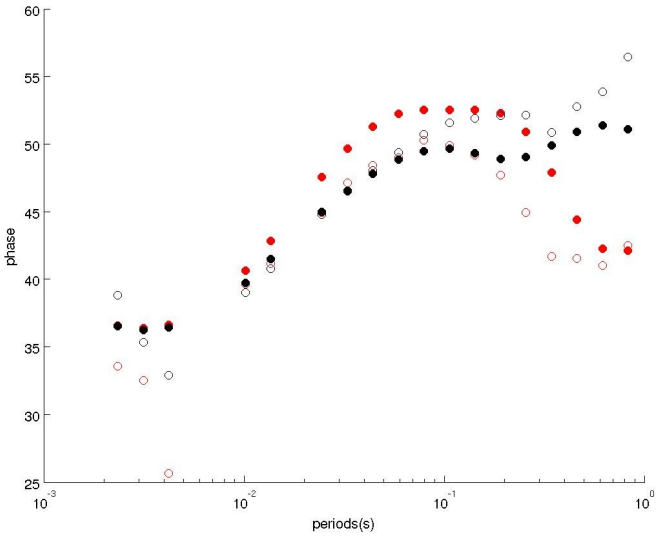


Figure 8.89

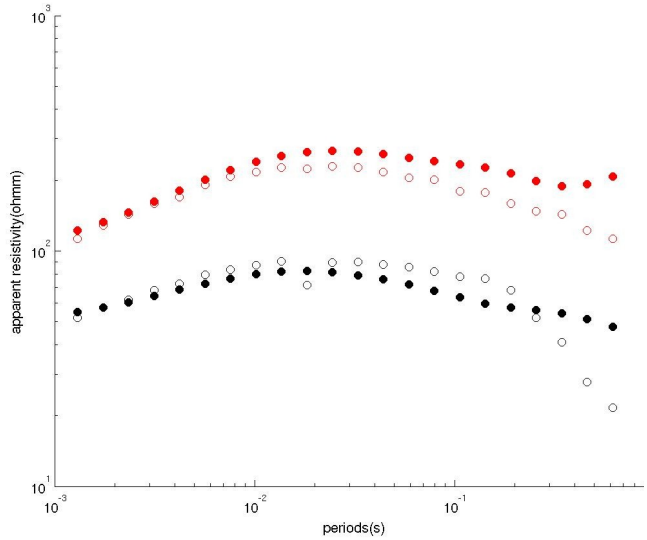


Figure 8.90

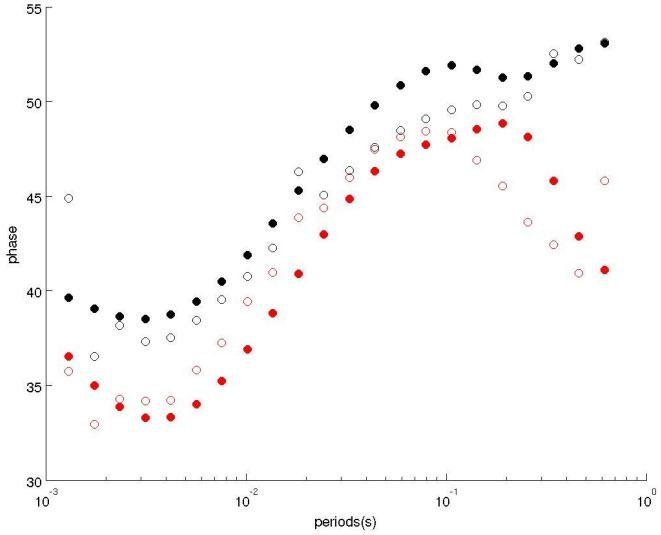


Figure 8.91

Figures 8.84-8.91 show apparent resistivity and phase for E mode(red) and H mode(black) for sites 3,5,12,14. Empty circles denote fieldwork data, filled circles inversion response

Research has shown that improvements were made with integration of two disparate data sets with these three algorithms (FCM, GK, GG).

1. With the fuzzy c-means algorithm, RMS misfit was reached after two or three iterations (depending on number of clusters). In vector space of normalized velocities and resistivities FCM tends to look for spheres. This causes clusters to be heterogeneous and does not lead traveltime inversion results into right direction. The algorithm could be used to give initial estimates about the structure.

2. The Gustafsson-Kessel algorithm showed its full potential. With clusters shaped into ellipsoids, at least one parameter does not vary too much, thus allowing parametric models to have homogeneous regions (for the invariable parameter). It was also shown that once there is a dominant structure (especially resistivity) it will keep appearing during the whole iterative process. Cluster formation was dictated by MT inversion, because of the travel-time inversions incapability to move far away from starting models. A large number of iterations gradually formed structures, with improving RMS misfit of both inversion processes. By changing the number of clusters, after iterative procedure, cooperative inversion will result in similar (not identical) models. Thus, the algorithm is confirming its robustness.

3. Gath-Geva allowed clusters to have different sizes, shapes and densities. Because of the way it calculates distances, all distances are very close to or very far away to cluster center. This causes sharp boundaries in the model. With more clusters this obstacle is surmountable and good results were obtained (similar parametric models to GK). It was observed that this algorithm is very sensitive to the number of clusters and starting models, adding or subtracting one cluster can lead cooperative inversions to different results.

9. Summary

This chapter will summarize what has been done in four years of research and new ideas will be proposed.

The research that has been performed in this thesis contributes to integrating two different geophysical methods (magnetotellurics and seismic refraction). Pattern recognition has not been used for the cooperative inversion of these two methods before. This study demonstrates potential of soft clustering methods for integration of two data sets. Gathering information from both inversions, a multi-parameter model was made. This multi-parameter model contains regions, each of which is characterized by a consistent relationship between the model parameters. Both synthetic and field data results were used to discuss an assumption that mutual exchange of information can lead to better results of inversions in both cases.

Having proven that exchange of information with these two algorithms can improve knowledge about the subsurface, some ideas for further improvement are presented next.

1. Introducing different parametrization (by making cells bigger in number), will probably affect the resolution of both models (especially the velocity one).
2. By calculating ray coverage and then changing the model dimensions (note that the cells in both lower corners are practically not covered by rays, if they are excluded, the model would be more plausible). If we want to exclude some cells, then the question is which cells are properly covered with rays. What is the threshold that must be exceeded to regard cells as sufficiently covered? Further research is required in that direction.
3. When calculating a new parametric model, membership function was multiplied with average cell value of velocity/resistivity. Weighting can be introduced to stress similarity for some cells and closeness to center of the cluster they belong to. This way the influence of outliers will be decreased.
4. There are other fuzzy algorithms which could possibly contribute to research, for example the entropy-based Gustafson-Kessel algorithm or fuzzy clustering with volume prototypes.
5. Research in the context of integration of methods should not be limited to fuzzy algorithms, but also use some other methods, for example spectral methods or neural networks.

10. Bibliography

1. Charles Groetsch, Inverse problems, The Mathematical Association of America (1999)
2. Trampert J., Global seismic tomography: the inverse problem and beyond, Inverse problems,1998
3. J. Barhen, J.G. Berryman, L. Borcea, J. Dennis Optimization and Geophysical Inverse Problems 2000, LBNL-46959
4. Lomax A. and R. Snieder, (1995), The contrast in upper mantle shear-wave velocity between the East European Platform and Tectonic Europe obtained with genetic algorithm inversion of Rayleigh wave group dispersion, Gephys. J. Int., 123, 169-182.
5. Alumbaugh D., Newman G. Image appraisal for 2-D and 3-D electromagnetic inversion, 2000, Geophysics 65, 1455-1467.
6. Pek, J. and Monteiro Santos, F.A**.**, 2002. Magnetotelluric impedances and parametric sensitives for 1-D generally anisotropic layer media. Computer and Geociences, 28, 939-950
7. Scales J. and R. Snieder, To Bayes or not to Bayes? GEOPHYSICS, VOL. 62, NO. 4 (JULY-AUGUST 1997); P. 1045–1046
8. Gouveia W. and Scales J. , Bayesian seismic waveform inversion : parameter estimation and uncertainty analysis, 1998, Journal of Geophysical Research: Solid Earth
9. Buland Arild, Bayesian seismic AVO inversion. PhD thesis,Trondheim, 2002
10. Parker, R. L., Geophysical Inverse Theory, 386pp, Princeton Univ. Press, May 1994.
11. Dobrin, M .B., 1976, Introduction to geophysical prospecting (3d ed .) New York, McGraw-Hill,
12. Portniaguine O. nad Zhdanov , M.S 1998a. Parameter Estimation for 3-D Geoelectromagnetic Inverse Problems,SEG
13. Portniaguine O. nad Zhdanov , M.S 1998b., Focusing of inversion images : SEG68
14. Lines, L. R., Schultz, A. K., Treitel, S., 1988, Cooperative inversion of geophysical data. Geophysics 53, 8-20
15. Nath S., Shahid S. 1999, Joint aplication of seismic refraction and vertical electrical sounding for delineation of shallow aquifers , Current Science Vol. 77
16. Scott, J. B. T., Barker, R. D., Peacock, S. ,2000, Combined seismic refraction and electrical imaging. EAGE meeting

17. Jean Beaujean, Frederic Nguyen, Andreas Kemna and Peter Engensgaard(2010) Joint and sequential inversion of geophysical and hydrogeological data to characterize seawater intrusion models , SWIM21 - 21st Salt Water Intrusion Meeting
18. Vozoff, K., Jupp, D. L. B. (1975) Joint inversion of geophysical data. Geophys. J. R. astr. Soc. 42, 977-991
19. Zhang, J., Morgan, F. D. (1996) Joint seismic and electrical tomography. Paper presented at EEGS Symposium on Applications of Geophysics to Engineering and Environmental Problems, Environ. and Eng. Geophys. Soc., Keystone,
20. Haber, E., Oldenburg, D. (1997) Joint inversion: A structural approach. Inverse Problems 13, 63-77.
21. Berge, P. A., Berryman, J. G., Bertete-Aguirre, H., Bonner, P., Roberts, J. J., Wildenschild, D. (2000) Joint inversion of geophysical data for site characterization and restoration monitoring, LLNL Rep. UCRL-ID-128343, Proj. 55411, Lawrence Livermore Natl. Lab., Livermore, Calif.
22. Gallardo L.,Meju M. Joint two-dimensional DC resistivity and seismic travel time inversion with cross-gradients constraints , 2003, JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 109, B03311
23. de Natale, G., Troise, C., Trigila, R., Chiarabba, C. (2004) Seismicity and 3-D substructure at Somma-Vesuvius volcano: evidence for magma quenching. Earth and Planet. Science Lett., 221
24. Roecker, S., Thurber, C. and McPhee, D. (2004) Joint inversion of gravity and arrival time data from Parkfield: New constraints on structure and hypocenter locations near the SAFOD drill site. Geophys. Res. Lett. 31(12), L12S04.
25. Christensen, N. L., Mooney, W. D (1995) Seismic velocity structure and composition of the continental crust: A global view. J. Geophys. Res. 100,
26. Marquis G, Hyndman R., Geophysical support for aqueous fluids in the deep crust: seismic and electrical relationships, Geophysical Journal International,Vol 110
27. Kozlovskaya, E. (2001) Theory and application of joint interpretation of multimethod geophysical data, Ph.D. dissertation, Univ. of Oulu, Oulu, Finland.
28. Luis Alonso Gallardo , Joint two-dimensional inversion of geoelectromagnetic and seismic refraction data with cross-gradients constraint , 2004, Lancaster University
29. Paache H. and Tronicke J.,,Cooperative inversion of 2D geophysical data sets: A zonal approach based on fuzzy c-means cluster analysis, 2007, GEOPHYSICS, vol. 72, no. 3 ‘may-june 2007; p. A35–A39

30. Aki, K., Christoffersson, A., and Husebye, E.S. Determination of the three-dimensional seismic structure of the lithosphere. J. Geophys. Res., 82:277–296, 1977.
31. Guust Nolet, A breviary of seismic tomography , 2008. , Cambridge University Press
32. Vidale J., Finite difference calculation of traveltimes in three dimensions, 1990, Geophysics, Volume 55
33. Hole J.A and B.Z Zelt , Three-dimensional finite difference reflection traveltimes, Geophys J. Int. 121, 1995
34. Aki K., Richards P., Quantitative seismology, University Science Books, 2002
35. Zhang, J. and Toksöz, M. N. (1998) Nonlinear refraction traveltime tomography. ,Geophysics 63,
36. Zelt, C. A., Barton, P. J. (1998) 3D seismic refraction tomography: A comparison of two methods applied to data from the Faeroe Basin. J. Geophys. Res. 103
37. Qin, F., Luo, Y., Olsen, K. B., Cai, W., Schuster, G. T. (1992) Finite-difference solution of the eikonal equation along expanding wavefronts. Geophysics 57
38. Korenaga, J., Holbrook, W. S., Kent, G. M., Kelemen, P. B., Detrick, R. S., Larsen, H. C., Hopper, J. R., Dahl-Jensen, T. (2000) Crustal structure of the southeast Greenland margin from joint refraction and reflection seismic tomography. J. Geophys. Res. 105,
39. Joseph R. Matarese , Nonlinear Traveltime Tomography , 1993, Phd thesis MIT
40. Lees J.M and R.S Crosson, Tomographic inversion for three-dimensional velocity structure at Mount. St Helens using earthquake data, J.Geophys Res. 94, 1989
41. Paige CC. and M.A Nolet ,LSQR : An algorithm for sparse linear equations and sparse least squares ,1982, Assos. Comp. Trans. Math. Software
42. Nolet G., Seismic wave propagation and seismic tomography, 1987, Modern Approaches in Geophysics
43. Zelt, C. A., A. M. Hojka, E. R. Flueh, and K. D. McIntosh, 3D simultaneous seismic refraction and reflection tomography of wide-angle data from the central Chilean margin, Geophys. Res. Lett., 26, 2577-2580, 1999.
44. Zelt, C. A., Barton, P. J. (1998) 3D seismic refraction tomography: A comparison of two methods applied to data from the Faeroe Basin. J. Geophys. Res. 103
45. Simpson F. and Bahr K. , Practical Magnetotellurics , 2005, Cambridge University Press
46. Cagniard, L. (1953) Basic theory of the magneto-telluric method of geophysical prospecting, Geophysics, **18**

47. Erika Gasperikova, Gregory Newman, Danny Feucht, and Knutur Arnason , 3D MT Characterization of Two Geothermal Fields in Iceland , 2011 , GRC Transactions, Vol. 35,
48. Strack, K. M., Tulinius, H., Vozoff, K., and Yu, G 2010 , Case histories of using magnetotellurics for geothermal exploration , 21 ASEG Conference and Exhibition
49. Monteiro Santos, F.A., Dupis, A., Andrade Afonso, A.R e Mendes Victor, L.A., 1995 Magnetotelluric Observations over the Chaves geothermal field (NE Portugal)-Preliminary results. Physics of the Earth and Planetary Interiors, 91: 203-211.
50. Monteiro Santos, F.A., Dupis, A., Andrade Afonso, A.R e Mendes Victor, L.A., 1999. Three-Dimensional modeling of a magnetotelluric survey over Chaves graben in northeast Portugal. Artigo incluido no livro 3-D Electromagnetic Methods (Oristaglio, M. and Spies, B. Ed.), Society of Exploration Geophysicists, USA.
51. Newman G., Hoversten M., Gasperikova E. and Wannamaker P. (2005) 3D magnetotelluric characterization of the coso geothermal field , PROCEEDINGS, Thirtieth Workshop on Geothermal Reservoir Engineering
52. Jones A., and Evans R.Velocity-conductivity relationships for mantle mineral assemblages in Archean cratonic lithosphere based on a review of laboratory data and Hashin-Shtrikman extremal bounds.,2009, *Lithos*, **109**,
53. Egbert, G. D. and Booker, J. R. (1986). Robust estimation of geomagnetic transfer functions. Geophysical Journal of the Royal Astronomical Society, 87(2)
54. M. Yu. Smirnov (2003) , Magnetotelluric data processing with a robust statistical procedure having a high breakdown point , Geophysical journal international, Volume 152
55. Hoversten M. Papua New Guinea MT: Looking where seismic is blind, 2006, Geophysical Prospecting Volume 44 Issue 6
56. Hoversten M. , Morrison F., Constable S. Marine magnetotellurics for petroleum exploration, Part II: Numerical analysis of subsalt resolution , GEOPHYSICS, VOL. 63, NO. 3, 1998
57. Key Kerry, Phd thesis: Application of Broadband Marine Magnetotelluric Exploration to a 3D Salt Structure and a Fast-Spreading Ridge , 2003 UCSD
58. Unsworth MJ, X Lu and MD Watts, Site characterization for radioactive waste disposal using CSAMT, Expanded abstracts 67th Ann. Internat. Mtg., 358-361, Society of Exploration Geophysicists, 1997
59. Newman G. , Recher S. , Tezkan B. , Nuebauer F. 3D inversion of a scalar radio magnetotelluric field data set ,2003, GEOPHYSICS, VOL. 68, NO. 3 (MAY-JUNE 2003);

60. Wannamaker, P. E., Stodt, J. A., and Rijo, L., 1987, A stable finite-element solution for two-dimensional magnetotelluric modeling: Geophysical Journal of the Royal Astronomical Society, 88, 277–296.
61. Constable, S.C., Parker, R.L., and Constable, C.G., 1987, Occam's inversion: A practical algorithm for generating smooth models from EM sounding data: Geophysics, 52, 289–300
62. Rokityansky, I. I., 1982, Geoelectomagnetic investigation of the Earth's crust and mantle: Springer-Verlag, New York, 381 p
63. Urbat M., Dekkers M.,Vriend S., The isolation of diagenetic groups in marine sediments using fuzzy *c*-means cluster analyses,, Geological Society, London, Special Publications 1999, v. 151
64. Finol J., X.D. Jing, Permeability prediction in shaly formations: The fuzzy modeling approach,2002,GEOPHYSICS,volume. 67 no. 3 p. 817-829
65. Hilde Vernieuwe, NEC Verhoest, Bernard De Baets, Rudi Hoeben, FP De Troch ; Cluster-based fuzzy models for groundwater flow in the unsaturated zone,Advances in water Resources, Volume 30, Issue 4, April 2007, Pages 701–714
66. C. Güler , Thyne G., Delineation of hydrochemical facies distribution in a regional groundwater system by means of fuzzy *c*-means clustering, 2004, Water Resources Research, Volume 40
67. Alvarez Grima, M.; Neuro-fuzzy modelling in engineering geology. Applications to rock strength estimation, machine excavation and geological mapping. Delft University of Technology. Promotor(s): Prof.dr. A.K. Turner, Prof.dr.ir. P. van der Veer. Uitgave: Delft University of Technology, Delft, The Netherlands, 2000, 244 p.
68. Paache H. and Tronicke J. , Cooperative inversion of 2D geophysical data sets: A zonal approach based on fuzzy c-means cluster analysis ; GEOPHYSICS, VOL. 72, NO.3 2007; . A35–A39,
69. Paache H., Tronicke J., Holliger K., Green A., Maurer H., Integration of diverse physical-property models: Subsurface zonation and petrophysical parameter estimation based on fuzzy c-means cluster analyses , GEOPHYSICS, VOL. 71, NO.3 2006;
70. Theodoridis S., K. Koutroumbas , Pattern Recognition(fourth edition), 2009, Elsevier Inc.
71. Kaufman Leonard and Rousseeuw Peter, (2005) Finding groups in data, John Wiley & Sons
72. Hopner F., Klawonn F., Kruse R., Runkler T. (2000) Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition, John Wiley & Sons
73. R. Polikar , (2006) Pattern recognition Tutorial,Rowan University , Glassboro, New Jersey

74. Baumann-Wilke M., Bauer K. , Schovsbo N. , Stiller M. , P-wave traveltime tomography for a seismic characterization of black shales at shallow depth on Bornholm, Denmark ; GEOPHYSICS, VOL. 77, NO.5 2012
75. Anderson, M., Matti, J., Jachens, R., 2004, Structural model of the San Bernardino Basin, California from analysis of gravity, aeromagnetic, and seismicity data: Journal of Geophysical Research, v. 109, B04404.
76. Paache H., Eberle D., Rapid integration of large airborne geophysical datasuites using a fuzzy partitioning cluster algorithm: a tool for geological mapping and mineral exploration targeting, Exploration Geophysics, 2009, 40, 277–287
77. Zadeh L., 1965, Fuzzy sets, Information and Control 8, 338-353
78. Y. Hu, R. Hathaway On Efficiency of Optimization in Fuzzy c-Means, 2002, Fuzzy Systems
79. J.C. Bezdek, Intro to Fuzzy Sets, 1993, IEEE Transactions on Fuzzy Systems, Vol. 1, No. 1,
80. Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum, New York.
81. D. Gustafson, W. Kessel, 1978, Fuzzy clustering with a fuzzy covariance matrix,IEEE
82. I. Gath, A.Geva , 1989, Fuzzy clustering for the estimation of the parameters of the components of mixtures of normaldistributions, Pattern Recognition Letters 9
83. Carvalho, D., Barriga, F.J.A.S. & Munh´J., 1999. Bimodal siliciclastic sysatems: the case of the Iberian Pyrite Belt,,Reviews in Economic Geology,Vol. 8,Society of Economic Geologists.
84. Mitjavila, J., Marti, J. & Soriano, C., 1997. Magmatic evolution and tectonic setting of the Iberian Pyrite Belt volcanism, J. Petrol., 38(6)
85. Schermerhorn, J.L.G., 1971. An outline of the stratigraphy of the Iberian Pyrite Belt, Boletin Geologico y Minero de Espana, 82
86. Silva, J.B., Oliveira, J.T. & Ribeiro, A., 1990. Structural outline of the South Portuguese Zone, in Pre-Mesozoic Geology of the Iberia Peninsula, Springer Verlag
87. Soriano, C. & Casas, J.M., 2002. Variscan tectonics in the Iberian Pyrite Belt, South Portuguese Zone, Int. J. Earth Sci., 91, 882–896
88. C. Schmelzbach, C. A. Zelt, C. Juhlin, R. Carbonell, 2008, P- and S V -velocity structure of the South Portuguese Zone fold-and-thrust belt, SW Iberia, from traveltime tomography, Geophys. J. Int. (2008) 175, 689–712
89. E. Kissling, S. Husen, F. Haslinger, 2001, Model parametrization in seismic tomography: a choice of consequence for the solution quality, Physics of the Earth and Planetary Interiors 123

RESUMO ALARGADO

A motivação para a integração de diferentes tipos de dados geofísicos surge do facto observável, de que existe uma grande variedade de objectos geológicos que podem causar anomalias semelhantes dos campos geofísicos. Com os problemas de fraca resolução e de ambiguidade a tarefa de interpretar os objectos geológicos torna-se assim muito complicada. Embora haja uma grande variedade de métodos para a prospecção da estrutura da terra, cada um deles tem os seus pontos fortes e também pontos fracos, parecendo inevitável a combinação de diferentes métodos com o objectivo de melhorar a investigação.

Nesta tese usaram-se os métodos MT e sísmica de refracção, dois métodos largamente usados pela comunidade académica e pela indústria. A integração destes métodos deverá dar mais informação sobre a estrutura da subsuperfície. Os dados de ambos os métodos são geralmente sujeitos a inversão. A física de ambos os métodos é descrita por duas equações (de difusão e onda, respectivamente) o que tem como consequência que os métodos têm difererente resolução (de facto, o método sísmico tem maior resolução). Como resultado das inversões obtêm-se modelos de distribuição de resistividade e de velocidade sísmica. O problema matemático em questão é: como combinar os dois tipos de dados. Trabalhos anteriores usando estes dois métodos, tiveram como objectivo minimizar uma "função objectivo" sujeita a constrangimentos: gradiente cruzado (Gallardo) ou estrutural (Haber e Oldenberg). Aqui uma nova aproximação é proposta: duas funções objectivo serão minimizadas separadamente e um novo modelo será gerado a cada iteração, combinando os modelos obtidos das inversões individuais. O nome de "inversão cooperativa" refere-se exactamente a este facto da informação ser trocada entre os modelos em cada iteração. Com este método criam-se modelos multi-zonais. Cada zona é caracterizada por relações consistentes entre os parâmetros. Com este método não há necessidade de um constrangimento estrutural, o que representa um vantagem em relação aos métodos usados anteriormente.

Nesta investigação, o problema inverso e o reconhecimento de padrões são usados num processo multi-etapas para a estimativa dos parâmteros dos modelos de subsuperfície. Um aspecto importante é que não se usam relações entre os parâmetros geofísicos. Foram empregues três algoritmos de análise de padrões (clustering): fuzzy C,

Gustafson-Kessel e Gath-Geva, sendo os resultados comparados. Nesta investigação, é a primeira vez que estes métodos são aplicados à inversão de dados de MT e sísmica.

Os dados experimentais para testar o método foram colhidos em Odeleite (Alentejo). Os dados de sísmica foram adquiridos pela Prospectiuni e pelo LNEG. O LNEG disponibilizou os dados da linha 6. O processamento dos dados e a interpretação foram feitas no âmbito desta tese. Porque as fontes e os receptores não eram colineares procedeu-se à projecção de ambos numa linha, mantendo a distância correcta entre os receptores e as fontes. Posteriormente foram adquiridos, no âmbito desta tese, os dados de MT, sobre a mesma linha. A partir da aquisição das variações dos campos elétrico e magnético obtiveram-se as funções de transferência MT. Os dados foram invertidos usando um modelo 2D. Infelizmente algumas das sondagens MT estavam contaminadas com ruído electromagnético, pelo que alguns dos dados são de baixa qualidade. Como consequência, o erro nas inversões, para alguns dos períodos, é elevado.

Para a combinação dos dois métodos a escolha da parametrização é feita com um compromisso entre os códigos de inversão de sísmica e MT (no primeiro caso a estrutura de velocidades na subsuperfície é feita por nós mas no segundo caso a resistiviade é atribuída a células). As dimensões do modelo são 10.6 km (comprimento) e 1.4 km de profundidade. Cada modelo foi dividido em células de 100x100 m. Foi demonstrado que esta subdivisão é satisfatória para este problema. A cada célula á atribuída um valor de velocidade e um de resistividade, constituindo, estes valores, o espaço de parâmetros no qual se agrupam pontos para formarem os “clusters”.

Nesta investigação foram empregues três algoritmos diferentes baseados na lógica “fuzzy”. A teoria básica da lógica “fuzzy” é a de permitir que um determinado evento possa pertencer a mais que um conjunto de eventos. A lógica “fuzzy” é uma alternativa à noção tradicional de “membro de um conjunto”. Em muitas situações reais a lógica “fuzzy” é mais natural que a aproximação clássica, dado que os eventos (ou objectos) na fronteira das diferentes classes não são forçados a pertencer só a uma delas mas, pelo contrário, podem pertencer a diferentes classes sendo o seu nível de participação, numa clase, traduzida por um valor compreendido entre 0 e 1.

Os algoritmos foram inicialmente aplicados a dados sintéticos com o fim de investigar como o processo de inversão é afectado pelo diferentes algoritmos, provando que se podem obter resultados em que se pode confiar. Os algoritmos foram depois aplicados a dados experimentais colhidos na Faixa Piritosa.

Na primeira parte da aplicação dos algoritmos, os valores de velocidade e de resistividade são normalizados. A normalização consiste na transformação de todos os dados para a mesma escala: se os valores dos diferentes conjuntos de dados forem de escalas muito diferentes, isto porá problemas na definição dos "clusters". Se um conjunto de dados possuir uma grande variedade de valores, a definição de "clusters" será dominada por este conjunto de dados. Para o primeiro algoritmo, a normalização é caracterizada por ter uma variância média de zero. Para os outros algoritmos a normalização apenas reescala os valores entre 0 e 1. O segundo passo da aplicação é o de calcular os centros dos "clusters" e organizar os "pontos" (parâmetros) pelos "clusters". A cada ponto será atribuído uma valor de uma função designada aqui por "membro do cluster" (membership em inglês) entre 0 e 1, sendo que para todos os "clusters" a soma dos "membro do cluster" deve ser 1. Após o cálculo dos "membro do cluster" de cada um dos pontos faz-se a distribuição dos pontos pelos diferentes "clusters" tendo em atenção os valores máximos da função "membro do cluster". Este processo é conhecido pelo termo inglês "defuzzification". Calculam-se ainda os valores médios da velocidade e da resistividade de cada "cluster". A actualização dos modelos é feita multiplicando-se os valores do "membro do cluster" pelos valores médios dos parâmetros (velocidade/resistividade). Estes modelos são utilizados como modelos iniciais da iteração seguinte dos processos individuais de inversão. O processo iterativo continua até que o valor do RMS seja superior ao da iteração anterior.

Como não se sabe antecipadamente qual é o melhor número de "clusters", o processo tem de ser repetido para vários "clusters" escolhendo-se no fim, a melhor solução, tendo em atenção um critério previamente estabelecido. No caso desta tese o critério adoptado foi o de menor valor de RMS.

Dos resultados obtidos concluiu-se o seguinte:

O algoritmo "Fuzzy C" produz "clusters" esféricos que não são aceitáveis para o caso em tratamento nesta tese. A distância entre os centros dos "clusters" e os pontos é Euclediana. O critério de paragem do processo iterativo é atingido ao fim de um par de iterações e o aumento do número de "clusters" não melhora os resultados. O algoritmo de Gustafson-Kesel produz "clusters" eleipsoidais e conduziu a bons resultados. A distância é neste caso designada por distância de Mahalonobis. A distribuição dos "clusters" é controlada pela inversão dos dados de MT devido, sobretudo, à incapacidade da inversão dos dados sísmicos de se afastar muito do modelo inicial. Ao fim de várias iterações observa-se a formação de estruturas coerentes com uma

diminuição do RMS para ambas as inversões. O aumento do número de “clusters” produz resultados semelhantes mas não iguais. O algoritmo Gath-Geva permite ter “clusters” com tamanhos, formas e densidades diferentes. O modo como as distâncias são calculadas (função exponencial) faz com que os pontos estejam ou muito próximos ou muito afastados dos centros dos “clusters”. O aumento do número de “clusters” minimiza este problema obtendo-se, então, bons resultados. Notou-se que este algoritmo é muito sensível ao número de “clusters” e também aos modelos iniciais.

Os códigos usados nesta tese para a inversãoconjunta (cooperativa) foram escritos pelo autor, tendo-se usado, também, códigos amplamente utilizados em MT e sísmica: FAST (sísmica de refracção) e OCCAM2D (para a MT).

Este trabalho mostrou que a utilização dos algoritmos de reconhecimento de padrões melhoram a integração dos dados dos diferentes métodos geofísicos. A tese termina com a sugestão para futuros desenvolvimentos nesta área.