Inversion of seismic attributes for petrophysical parameters and rock facies

Mohammad Sadegh Shahraeeni

Thesis Submitted for the degree of Doctorate of Philosophy

University of Edinburgh

September 2010

“Are those equal, those who know and those who do not know? It is those who are endued with understanding that receive admonition.”

Declaration

This thesis is the result of my own original work. Where it draws on the work of others, this is acknowledged at the appropriate points in the text.

Mohammad Sadegh Shahraeeni

Abstract

Prediction of rock and fluid properties such as porosity, clay content, and water saturation is essential for exploration and development of hydrocarbon reservoirs. Rock and fluid property maps obtained from such predictions can be used for optimal selection of well locations for reservoir development and production enhancement. Seismic data are usually the only source of information available throughout a field that can be used to predict the 3D distribution of properties with appropriate spatial resolution. The main challenge in inferring properties from seismic data is the ambiguous nature of geophysical information. Therefore, any estimate of rock and fluid property maps derived from seismic data must also represent its associated uncertainty.

In this study we develop a computationally efficient mathematical technique based on neural networks to integrate measured data and a priori information in order to reduce the uncertainty in rock and fluid properties in a reservoir. The post inversion (a posteriori) information about rock and fluid properties are represented by the joint probability density function (PDF) of porosity, clay content, and water saturation. In this technique the a posteriori PDF is modeled by a weighted sum of Gaussian PDF's. A so-called mixture density network (MDN) estimates the weights, mean vector, and covariance matrix of the Gaussians given any measured data set. We solve several inverse problems with the MDN and compare results with Monte Carlo (MC) sampling solution and show that the MDN inversion technique provides good estimate of the MC sampling solution. However, the computational cost of training and using the neural network is much lower than solution found by MC sampling (more than a factor of 10^4 in some cases). We also discuss the design, implementation, and training procedure of the MDN, and its limitations in estimating the solution of an inverse problem.

In this thesis we focus on data from a deep offshore field in Africa. Our goal is to apply the MDN inversion technique to obtain maps of petrophysical properties (i.e., porosity, clay content, water saturation), and *petrophysical facies* from 3D seismic data. Petrophysical facies (i.e., non-reservoir, oil- and brine-saturated reservoir facies) are defined probabilistically based on geological information and values of the petrophysical parameters.

First, we investigate the relationship (i.e., petrophysical forward function) between compressional- and shear-wave velocity and petrophysical parameters. The petrophysical forward function depends on different properties of rocks and varies from one rock type to another. Therefore, after acquisition of well logs or seismic data from a geological setting the petrophysical forward function must be calibrated with data and observations. The uncertainty of the petrophysical forward function comes from uncertainty in measurements and uncertainty about the type of facies. We present a method to construct the petrophysical forward function with its associated uncertainty from the both sources above. The results show that introducing uncertainty in facies improves the accuracy of the petrophysical forward function predictions.

Then, we apply the MDN inversion method to solve four different petrophysical inverse problems. In particular, we invert P- and S-wave impedance logs for the joint PDF of porosity, clay content, and water saturation using a calibrated petrophysical forward function. Results show that posterior PDF of the model parameters provides reasonable estimates of measured well logs. Errors in the posterior PDF are mainly due to errors in the petrophysical forward function.

Finally, we apply the MDN inversion method to predict 3D petrophysical properties from attributes of seismic data. In this application, the inversion objective is to estimate the joint PDF of porosity, clay content, and water saturation at each point in the reservoir, from the compressional- and shear-wave-impedance obtained from the inversion of AVO seismic data. Uncertainty in the a posteriori PDF of the model parameters are due to different sources such as variations in effective pressure, bulk modulus and density of hydrocarbon, uncertainty of the petrophysical forward function, and random noise in recorded data. Results show that the standard deviations of all model parameters are reduced after inversion, which shows that the inversion process provides information about all parameters. We also applied the result of the petrophysical inversion to estimate the 3D probability maps of non-reservoir facies, brine- and oil-saturated reservoir facies. The accuracy of the predicted oil-saturated facies at the well location is good, but due to errors in the petrophysical inversion the predicted non-reservoir and brine-saturated facies are ambiguous. Although the accuracy of results may vary due to different sources of error in different applications, the fast, probabilistic method of solving non-linear inverse problems developed in this study can be applied to invert well logs and large seismic data sets for petrophysical parameters in different applications.

Acknowledgements

First, I would like to thank my supervisor, Professor Andrew Curtis, for his invaluable scientific support, advice and encouragement throughout my work with him. He gave me the opportunity to explore my ideas and his constructive comments inspired me. Thank you Andrew, for your patience and support.

I would like to thank Professor Malcolm Sambridge and Dr. Mark Chapman the external and internal examiners of this thesis for their invaluable comments and suggestions, which improves the quality of the final version of this thesis a lot.

I thank Professor Ian Main for his support and guidance during the first year of my PhD. I must also thank all administrative staff at the Grant Institute and in particular Helena Sim, for her help and guidance during my PhD.

I thank all the members of the Geophysics group in the Grant Institute whom I had the pleasure of knowing the entire time here. Thank to Arash Gandomi, David Halliday, Thomas Guest, Heather Nicolson, Simon King, Angeleena Thomas, John Gonzalez, Craig Duguid, Travis Poole, Debbie Polson, and my officemate Gillian McCay. I also thank all my friends from Heriot Watt University whom made my stay in Edinburgh enjoyable.

I would like to acknowledge the sponsor of my PhD, Geoscince Research Centre (GRC) of TOTAL E&P UK. In particular I would like to appreciate the support and encouragement of Dr. Olivier Dubrule, the former head of GRC, who introduced Andrew to me and also accepted to sponsor this PhD. I never forget our exciting intellectual discussions. I would like to express my gratitude to Dr. Gilles Drullion for his kind help during and after my internship at GRC. I also would like to thank the support and invaluable discussions with Dr. Christian Deplante, the former head of the geophysics group of GRC. I would like to thank all members of geophysics team at GRC, Dr. Andrea Grandi, Dr. Darrel Cole, and Dr. Gracjan Lambert for their friendship. In particular, I must thank Dr. Gabriel Chao my industrial contact, for his support, friendship and invaluable scientific discussions.

Last, but by no means least, is my family. I would like to express my gratitude to the memory of my dad, whom spiritual existence and faith always guide me in my life. I would like to thank Mom, my wife Leila, Ebrahim, Leila, and Mehdi for their love, kindness, encouragement, and understanding. I dedicate this thesis, which I have worked very hard for, to them.

Table of Contents

List of Figures

List of Tables

1 Introduction

> *"I keep the subject of my inquiry constantly before me, and wait till the first dawning opens gradually by little and little into a full and clear light."*
>
> Isaac Newton

1.1 Objective

The overall objective of this thesis is to obtain information about petrophysical parameters from 3D seismic data in Akpo field, which is a deep offshore field in Africa. More specific goals in this research are: 1) to estimate petrophysical parameters (i.e., porosity, clay content, and water saturation) of hydrocarbon reservoirs from seismic data using a given calibrated petrophysical forward function; 2) to develop and test a mathematical tool for probabilistic inversion that is able to solve repeated, similar, multi-dimensional, nonlinear inverse problems in a computationally efficient and robust manner; and 3) to construct a probabilistic petrophysical forward function that links seismic data to the petrophysical parameters of different facies. The motivations for each of the above problems are described below.

1.2 Background and motivations

1.2.1 Petrophysical inversion of seismic data

3D seismic data acquired from a field can be used for various purposes in different phases of reservoir development. In the exploration phase, it can be used to detect hydrocarbon reservoirs. In the reservoir appraisal, it can be used to describe the shape, size, and extent of reservoirs. In the production phase, it can be used to characterize transport properties of reservoirs and connectivity between different reservoir compartments. It can also be used to forecast the performance of a reservoir during production.

Development of 3D seismic inversion techniques and the quantitative description of relations between seismic attributes and petrophysical parameters give us the opportunity to perform quantitative reservoir characterization more accurately. However, due to the ambiguous nature of geophysical information any estimate of the petrophysical parameters from seismic data must represent its associated uncertainty. Uncertainty of the estimated petrophysical parameters can have at least three sources: first, there is non-uniqueness in the inversion of (usually Amplitude Versus Offset (AVO)) seismic data for seismic attributes, second there is non-uniqueness in the petrophysical inversion of seismic attributes for the petrophysical parameters given a petrophysical relationship, and third there is ambiguity in these petrophysical relationships themselves.

There are two approaches to represent petrophysical relations for inversion (Avseth, 2000): 1) physical models, and 2) statistical methods. In physical models, theoretical petrophysical forward relations are used to link the petrophysical parameters to seismic attributes. In statistical methods on the other hand the petrophysical parameters are represented as empirical functions of seismic attributes; the coefficients of these empirical functions are estimated using well log data. Geostatistical cokriging (Doyen, 1988; Dubrule, 2003) is an example of a statistical method, where the empirical functions are linear. In a more complicated statistical approach the relationship between the petrophysical parameters and seismic attributes are assumed to be nonlinear and are modeled using a neural network (Saggaf et al., 2003). The parameters of the neural network are estimated using well log data. All statistical approaches suffer from the lack of a definite physical theory that links the petrophysical parameters to seismic attributes.

Many authors have studied physical links between the petrophysical parameters and seismic attributes. Mavko et al. (2009) and Avseth et al. (2005) provide a detailed summary of these theories. In particular, Avseth et al. (2005) provide a list of different rock physics models for different geological scenarios such as the unconsolidated sand model (Dvorkin and Nur, 1996), the shaley sand model (Marion et al., 1992; Dvorkin and Gutierrez, 2001), the cemented shaley sand model (Jizba, 1991; Dvorkin and Nur, 1996), the sandy shale model (Dvorkin and Gutierrez, 2001), and the laminated sand-shale model (Dvorkin and Gutierrez, 2001). In principle, these rock physics theories and models can be applied to predict the petrophysical parameters from seismic data with less uncertainty than just pure geostatistical relations because they incorporate information about the theoretical physical relationships. Successful examples include lithology and fluid identification from seismic data in North Sea turbidite systems (Avseth et al., 2001; Mukerji et al., 2001), applying geostatistical methods to decrease the uncertainty of reservoir characterization (Eidsvik et al., 2004; Larsen et al., 2006; Ulvmoen and Omre, 2010; Ulvmoen et al., 2010), and estimating petrophysical parameters (i.e., porosity, clay content, water saturation) from seismic data (Bachrach, 2006; Spikes et al., 2007; Bosch et al., 2009; Grana and Della Rossa, 2010). These examples demonstrate that the rock physics link is essential for rigorously relating seismic data to reservoir properties. Moreover, they all indicate that the physical link should be combined with statistics to account for uncertainties related to variability in the petrophysical parameters.

All of the above applications show that there is a need for fast computational methods to calculate uncertainty in the petrophysical parameter estimates obtained from the petrophysical inversion of seismic data. We also require new methods to account for uncertainty in the petrophysical forward relations due to uncertainty about different geological scenarios that can happen within a reservoir or even a well. The fast computational method is required because applying conventional Monte Carlo sampling based methods with large seismic datasets would be computationally demanding to the point of being generally impractical unless large parallel computing facilities are available. . For each point in the subsurface the forward problem (in this case the petrophysical forward function and approximations of Zoeppritz's equation) must be evaluated for a large number of model samples (typically at least of the order of thousands), and this process would have to be repeated for all data points in a seismic cube, which usually includes hundreds of millions of data points. Therefore, MC sampling usually applied for a small or sparse subset of the available data points (Eidsvik et al., 2004; Larsen et al., 2006; Bachrach, 2006; Spikes et al., 2007; Bosch et al., 2009). The uncertainty in the geological scenario must be included because variation in the physical link between petrophysical parameters and seismic attributes is an additional source of uncertainty in the petrophysical inversion of seismic data. In this thesis we propose and apply methods to address the above two needs.

1.2.2 Fast probabilistic inversion

Traditionally, nonlinear inverse problems are solved using iterative linearization algorithms (Tarantola, 2005, pp. 68-71), or sampling methods (Mosegaard and Tarantola, 2002; Sambridge and Mosegaard, 2002; Tarantola, 2005, pp. 41-55). Buland and Omre (2003), and Buland et al. (2003) applied a linearized forward model to invert AVO data for compressional- and shear-wave velocity, and density. Sen and Stoffa (1991) and Beaty et al. (2002) demonstrate applications of the simulated annealing sampling method in the inversion of seismic data for wave velocity and density. These methods have also been used to solve petrophysical inverse problems where seismic attributes are inverted to obtain petrophysical parameters. Bosch (2004) and Bosh et al. (2009) demonstrate applications of the iterative linearization method in the inversion of seismic amplitudes for P-wave impedance, porosity and water saturation. Avseth et al. (2001), Mukerji et al. (2001), Eidsvik et al. (2004), Bachrach (2006), Bosh et al. (2007; 2009) and Grana and Della Rossa (2010) demonstrate applications of sampling methods in petrophysical inversions.

The computational cost (i.e., processing time and required memory) is a critical factor in geophysical inverse problems. Neural networks have been used to solve nonlinear geophysical inverse problems in a computationally efficient manner. Roth and Tarantola (1994) demonstrate an application of neural networks in the inversion of synthetic common shot gathers for compressional velocity. Devilee et al. (1999) and Meier et al. (2007a,b) demonstrate other applications of neural networks in the probabilistic inversion of global and regional surface wave dispersion velocities for crustal thickness. Meier et al. (2009) demonstrate another application of neural networks in a probabilistic petrophysical inversion. The above applications of neural networks show that they can be used to solve repeated and similar inverse problems robustly and efficiently. In this thesis we extend the probabilistic neural network applied by Meier et al. (2007,2009) to solve petrophysical inverse problems with multi-dimensional model and data spaces.

1.2.3 Petrophysical forward relations

The petrophysical forward relation, which links the petrophysical parameters and seismic attributes, is a basic requirement for petrophysical inversion. Avseth (2000) explains that the petrophysical forward relations depend on the type of facies, and these relations can be used to improve the applicability of seismic data for reservoir characterization. In current applications of the petrophysical inversion, petrophysical forward functions are defined for a single facies and their uncertainties are modeled by Gaussian or truncated Gaussian probability density functions (PDF's) (Bachrach, 2006; Spikes et al., 2007; Bosch et al., 2009; Chao et al., 2009; Grana and Della Rossa, 2010). The assumption of a single facies might be a correct assumption for some depth intervals in a reservoir, but in general the heterogeneity of facies in a reservoir can be high and in many locations it will not be obvious which facies is present. Buland et al. (2008) and Grana and Della Rossa (2010) applied multi-facies a priori probability distribution function (PDF) of the petrophysical parameters to solve petrophysical inverse problems. In the same manner as the multi-facies assumption for the a priori PDF of the petrophysical parameters, we need to define a general multi-facies petrophysical forward function that can describe the link between seismic attributes and the petrophysical parameters for different types of facies within a reservoir. In this thesis we propose a method to construct such a petrophysical forward function.

1.3 Description of chapters

We address the aforementioned problems in different chapters of this thesis. In chapter 2 we present a method to construct a multi-facies petrophysical forward function. We show that this method can be used to address the facies uncertainty in the petrophysical forward function predictions. Results of this chapter show that introducing facies uncertainty improves the accuracy of the petrophysical forward function predictions. This work is going to be submitted to Geophysics.

In chapter 3 we present a mixture density neural network (MDN) as an efficient method to solve nonlinear inverse problems and apply it to solve four different petrophysical inverse problems. We show that the MDN can be applied in both of the previously mentioned approaches to represent petrophysical relations in a petrophysical inversion. We also show that the MDN inversion method can be used to estimate the Monte Carlo sampling solution of an inversion with an acceptable accuracy. Some parts of this chapter were published in Shahraeeni and Curtis (2009) and some other parts were submitted to Geophysics (Shahraeeni and Curtis, 2010).

In chapter 4 we apply the MDN method to invert 3D seismic attributes (compressional and shear impedance) for the petrophysical parameters in the Akpo field. Results show that the inversion method can be applied to invert 3D seismic datasets very efficiently. We also apply the results of the petrophysical inversion to derive probability of different facies defined in the petrophysical forward function. We show that the accuracy of the result is influenced by the accuracy of the petrophysical forward function and uncertainty in the inverted seismic attributes (seismic impedances). This work is also submitted to Geophysics (Shahraeeni et al., 2010).

In chapter 5 we summarize and discuss the methodology presented in the previous chapters and in chapter 6 we conclude and mention possible extensions and improvements to the methodology and results presented in this thesis.

In Appendix A, we present a detailed mathematical description of neural networks. In this appendix we develop a new type of MDN, which enables us to solve inverse problems with multi-dimensional model and data spaces more accurately than previous types of MDN. We apply this new MDN to solve all petrophysical inverse problems in the main chapters of this thesis. In Appendices B and C, we present two sets of petrophysical forward relations which are applied in different inverse problems addressed in the main chapters of the thesis.

1.4 Available data

We apply the proposed methods to data from Akpo field, a deep offshore field in Africa. The available data include measured well log datasets and facies logs from five different wells. The well log dataset includes effective porosity, clay content, water saturation, effective pressure, bulk modulus and density of hydrocarbon and brine, compressional-wave velocity, shear-wave velocity, and bulk density. We are provided with a calibrated petro-elastic model, which links the petrophysical parameters (i.e., effective porosity, clay content, water saturation, effective pressure, bulk modulus and density of hydrocarbon and brine) to seismic attributes (i.e., compressional- and shear-wave velocity, and density) for this field. We are also provided with the in-line and cross-line sections of the AVO-inverted seismic attributes (i.e., compressional- and shear-wave impedance) at one of the wells in the field. We also have the depth of some of the stratigraphic units applied in the AVO inversion of seismic data at the well locations.

1.5 References

Avseth, P., 2000, Combining rock physics and sedimentology for seismic reservoir characterization of north sea turbidite systems. PhD Thesis, Stanford University.

Avseth, P., T. Mukerji, A. Jorstad, G. Mavko and T. Veggeland, 2001, Seismic Reservoir Mapping from 3-D AVO in a North Sea Turbidite System, *Geophysics*, **66**, 1157-1176.

Avseth, P., T. Mukerji and G. Mavko, 2005, *Quantitative seismic interpretation: applying rock physics tools to reduce interpretation risk*. Cambridge ; New York: Cambridge University Press.

Bachrach, R., 2006, Joint estimation of porosity and saturation using stochastic rock-physics modeling, *Geophysics*, **71**, O53-O63.

Beaty, K. S., D. R. Schmitt and M. Sacchi, 2002, Simulated annealing inversion of multimode Rayleigh wave dispersion curves for geological structure, *Geophysical Journal International*, **151**(2), 622-631.

Bosch, M., 2004, The optimization approach to lithological tomography: Combining seismic data and petrophysics for porosity prediction, *Geophysics*, **69**(5), 1272-1282.

Bosch, M., L. Cara, J. Rodrigues, A. Navarro and M. Diaz, 2007, A Monte Carlo approach to the joint estimation of reservoir and elastic parameters from seismic amplitudes, *Geophysics*, **72**(6), O29-O39.

Bosch, M., C. Carvajal, J. Rodrigues, A. Torres, M. Aldana and J. Sierra, 2009, Petrophysical seismic inversion conditioned to well-log data: Methods and application to a gas reservoir, *Geophysics*, **74**(2), O1-O15.

Buland, A., O. Kolbjornsen, R. Hauge, O. Skjaeveland and K. Duffaut, 2008, Bayesian lithology and fluid prediction from seismic prestack data, *Geophysics*, **73**(3), C13-C21.

Buland, A., O. Kolbjornsen and H. Omre, 2003, Rapid spatially coupled AVO inversion in the Fourier domain, *Geophysics*, **68**(3), 824-836.

Buland, A. and H. Omre, 2003, Bayesian linearized AVO inversion, *Geophysics*, **68**(1), 185-198.

Chao, G., G. Lambert and H. Cumming, 2009, Analysis of Intrinsic Uncertainties of Petro-elastic Models Using Simulated Annealing, 71st EAGE Conference & Exhibition, Amsterdam.

Devilee, R. J. R., A. Curtis and K. Roy-Chowdhury, 1999, An efficient, probabilistic neural network approach to solving inverse problems: Inverting surface wave velocities for

Eurasian crustal thickness, *Journal of Geophysical Research*, **104(B12)**, 28841-28856.

Doyen, P. M., 1988, Porosity from seismic data: A geostatistical approach, *Geophysics*, **53**, 1263-1275.

Dubrule, O., 2003, *Geostatistics for seismic data integration in earth models*: Society of Exploration Geophysicists.

Dvorkin, J. and M. A. Gutierrez, 2001, Grain sorting, porosity, and elasticity, from pangea.stanford.edu/~jack/Bimodal.pdf.

Dvorkin, J. and A. Nur, 1996, Elasticity of high-porosity sandstones: Theory for two North Sea data sets, *Geophysics*, **61**(5), 1363-1370.

Eidsvik, J., P. Avseth, H. Omre, T. Mukerji and G. Mavko, 2004, Stochastic Reservoir Characterization using Prestack Seismic Data, *Geophysics*, **69**, 780-993.

Grana, D. and E. Della Rossa, 2010, Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion, *Geophysics*, **75**(3), O21-O37.

Jizba, D., 1991, Mechanical and acoustical properties of sandstones and shales. Thesis, Stanford University.

Larsen, A. L., M. Ulvmoen, H. Omre and A. Buland, 2006, Bayesian lithology/fluid prediction and simulation on the basis of a Markov-chain prior model, *Geophysics*, **71**(5), R69-R78.

Marion, D., A. Nur, H. Yin and D. Han, 1992, Compressional Velocity and Porosity in Sand-clay Mixture, *Geophysics*, **57**, 554-563.

Mavko, G., T. Mukerji and J. Dvorkin, 2009, *The rock physics handbook: tools for seismic analysis of porous media*. Cambridge ; New York: Cambridge University Press.

Meier, U., A. Curtis and J. Trampert, 2007a, A global crustal model constrained by non-linearised inversion of fundamental mode surface waves, *Geophysical Research Letters*, **34**, L16304-L16304.

Meier, U., A. Curtis and J. Trampert, 2007b, Global crustal thickness from neural network inversion of surface wave data, *Geophysical Journal International*, **169**, 706-722.

Meier, U., J. Trampert and A. Curtis, 2009, Global variations of temperature and water content in the mantle transition zone from higher mode surface waves, *Earth Planetary Science Letters*, **282**, 91-101.

Mosegaard, K. and A. Tarantola, 2002, Probabilistic approach to inverse problems, *International handbook of earthquake and engineering seismology*. W. H. K. Lee, H. Kanamori, P. C. Jennings and C. Kisslinger: Academic Press Inc.**:** 237-265.

Mukerji, T., A. Jorstad, P. Avseth, G. Mavko and J. R. Granli, 2001, Mapping lithofacies and pore-fluid probabilities in a North Sea reservoir: Seismic inversions and statistical rock physics, *Geophysics*, **66**, 988-1001.

Roth, G. and A. Tarantola, 1994, Neural Networks and Inversion of Seismic data, *Journal of Geophysical Research*, **99**, 6753-6768.

Saggaf, M. M., M. N. Toksoz and H. M. Mustafa, 2003, Estimation of reservoir properties from seismic data by smooth neural networks, *Geophysics*, **68**(6), 1969-1983.

Sambridge, M. and K. Mosegaard, 2002, Monte Carlo methods in geophysical inverse problems, *Reviews of Geophysics*, **40**(3), 1009.

Sen, M. K. and P. L. Stoffa, 1991, Nonlinear one-dimensional seismic waveform inversion using simulated annealing, *Geophysics*, **56**(10), 1624-1638.

Shahraeeni, M., A. Curtis and G. Chao, 2010, Fast probabilistic petrophysical mapping of reservoirs from 3D seismic data, *Geophysics*, **Submitted**.

Shahraeeni, M. S. and A. Curtis, 2009, Nonlinear petrophysical inversion with mixture density neural networks, *71st EAGE Conference & Exhibition, Amsterdam*.

Shahraeeni, M. S. and A. Curtis, 2010, Fast probabilistic nonlinear petrophysical inversion, *Geophysics*, **submitted**.

Spikes, K., T. Mukerji, J. Dvorkin and G. Mavko, 2007, Probabilistic seismic inversion based on rock-physics models, *Geophysics*, **72**(5), R87-R97.

Tarantola, A., 2005, *Inverse Problem Theory and Methods for Model Parameter Estimation*: SIAM.

Ulvmoen, M. and H. Omre, 2010, Improved resolution in Bayesian lithology/fluid inversion from prestack seismic data and well observations: Part 1 --- Methodology, *Geophysics*, **75**(2), R21-R35.

Ulvmoen, M., H. Omre and A. Buland, 2010, Improved resolution in Bayesian lithology/fluid inversion from prestack seismic data and well observations: Part 2 --- Real case study, *Geophysics*, **75**(2), B73-B82.

2 Facies uncertainty in petrophysical forward function

> *"The wrong view of science betrays itself in the craving to be right; for it is not his possession of knowledge, of irrefutable truth, that makes the man of science, but his persistent and recklessly critical quest for truth."*
>
> Karl Popper

2.1 Abstract

In the context of geophysical inversion, a petrophysical forward function provides the link between properties that are directly geophysically detectable (e.g., seismic velocities and attenuation) and rock and fluid properties. Due to heterogeneity of rock properties uncertainty of the petrophysical forward function can be high. Uncertainty in the petrophysical forward function comes from two different sources: (1) measurement uncertainty, and (2) theoretical uncertainty. Measurement uncertainty accounts for errors in the acquisition and processing of data. Theoretical uncertainty on the other hand accounts for lack of knowledge about the rock type. We present a method to construct the petrophysical forward function with its associated uncertainty from the both sources indicated above. The uncertainty of the petrophysical forward function is represented by a weighted sum of Gaussian distributions. Each Gaussian represents the measurement uncertainty of one facies, and the weight of the Gaussian represents the probability of the associated facies being the correct rock type to model the elastic behavior of samples. We first apply the method by assuming no theoretical uncertainty and show that the predictions of the petrophysical forward function are biased under this assumption. Then we apply the method by considering theoretical uncertainty about the type of the facies. The results show that introducing uncertainty in the facies reduces uncertainty in the final probabilistic petrophysical forward function and removes biases from its predictions.

2.2 Introduction

Seismic velocities are sensitive to rock properties such as porosity, lithofacies (i.e., clay content, grain size, sorting, etc.), pore fluid type, saturation, pore pressure, depth, age, cementation, compaction, and temperature (Gutierrez, 2001, p. 67; Avseth et al., 2005, ch. 2). Petrophysical forward functions or models describe relations between seismic velocities and the rock properties. In order to construct the petrophysical forward function of a mixture of grains and pores, we need to specify three parameters: (1) the volume fraction of the various components (2) the elastic moduli of various components, and (3) the geometric details of how components are arranged relative to each other (Mavko et al., 2009, p. 169). Information about the last two parameters of the above list is usually obtained from a limited number of cores and thin sections (Avseth et al., 2005, pp. 73-81), and because of the heterogeneity of rock properties in a field or even in a well, their uncertainty is high. In practice the above parameters are seldom measured (Dvorkin and Gutierrez, 2001) and instead well observations and geological information are used to select an appropriate form for the petrophysical forward function, and the selected function is calibrated with well log measurements (Gutierrez, 2001, pp. 97-100; Avseth et al., 2005, pp. 96-101; Chao et al., 2009).

Petrophysical forward models fall into one of two general categories: theoretical and empirical (Avseth et al., 2005, p. 43-47). Theoretical models approximate the elastic behavior of rocks by assuming a simplified geometry of pore space and grains (Eshelby, 1957; Kuster and Toksoz, 1974; Dvorkin and Nur, 1996; Mavko et al., 2009, ch. 5). Due to idealized (and unrealistic) assumptions about the geometry of components, theoretical models tend to predict incorrect relations between the elastic moduli and pressure, and give poor estimates of the bulk to shear moduli ratio (Avseth et al., 2005, pp. 43-45; Mavko et al., 2009, pp. 261-263).

In empirical models some petrophysical functional form is assumed and then parameters of the selected function are determined by calibrating the function against available data. Simple forms of empirical models are defined between two or three parameters such as velocity-porosity models (Wyllie et al., 1956; Raymer et al., 1980), velocity-effective pressure models (Shapiro, 2003), and velocity-porosity-clay content models (Han et al., 1986). More complicated empirical models (Goldberg and Gurevich, 1998; Dvorkin and Gutierrez, 2001; Keys and Xu, 2002; Spikes et al., 2007; Bachrach and Avseth, 2008) take the form of the petrophysical forward function from a theoretical model and calibrate it with available data. This type of models can account for the effect of more parameters such as fluid saturation, effective pressure, or depth, and because the petrophysical forward model is constructed based on physical theories, it can be used to predict the elastic behavior of rocks.

The grand challenge of petrophysical modeling is to provide a general petrophysical forward function that can explain the elastic properties of different types of rocks in an oil field (Bachrach, 2006). Some theoretical laws, such as an increase of elastic velocities with depth due to the decrease in porosity (Avseth et al., 2005, pp. 98-101) or the exponential decay of the derivative of elastic moduli with respect to effective pressure (Shapiro, 2003), seem to be universal. However, it is difficult to describe the elastic behavior of all rock types in one geological setting with a single law or petrophysical function. As a result any petrophysical forward function needs to include uncertainty in its predictions.

Information about uncertainty of petrophysical forward function is of great importance if we want to invert wave-impedances (or velocities) for rock properties such as porosity, clay content, and water saturation (Bachrach, 2006; Bosch et al., 2009). In current applications of petrophysical inversion, where seismic compressional-wave and shear-wave impedance are inverted to obtain porosity, water saturation, and possibly clay content, petrophysical forward functions are defined for a single facies and their uncertainties are modeled by a Gaussian or truncated Gaussian probability density function (PDF) (Bachrach, 2006; Spikes et al., 2007; Bosch et al., 2009). In statistical rock physics applications, Monte Carlo sampling method is used to estimate uncertainty of seismic attributes for different types of facies by using petrophysical forward functions of different facies; the estimated uncertainty can then be applied to obtain the probability of occurrence of different facies given seismic data (Avseth et al., 2001; Mukerji et al., 2001; Eidsvik et al., 2004; Gonzalez et al., 2008). However, none of the above applications, which consider different types of facies, attempts to produce a general petrophysical forward function that can describe the elastic behavior of different types of facies with its uncertainty in a geological setting.

In the context of the petrophysical inversion, Grana and Della Rossa (2010) demonstrate a method based on the Gaussian Mixture Models (GMM) to represent the a priori probability density function (PDF) of the petrophysical parameters (i.e., porosity, clay content, and water saturation) for different types of facies. In a GMM a given PDF is modeled as a summation or mixture of Gaussian distributions. For each facies a Gaussian distribution represents the a priori information about the petrophysical parameters. The weight of each

Gaussian represents the frequency of occurrence of the corresponding facies. This application shows that the GMM can be applied to represent a priori facies uncertainties.

In this chapter we propose a method based on the GMM to construct the general petrophysical forward function and apply it to data from one well in Akpo field. In the GMM, each Gaussian distribution represents the uncertainty of the petrophysical forward function of one facies, and the weight of each Gaussian in the summation represent the probability of the associated facies being the correct rock type to model the elastic behavior of data. We first introduce the petrophysical forward functions we used to model the elastic behavior of different facies, then we introduce the method to estimate the uncertainty of the petrophysical forward function for each facies, and finally we present the mixture of Gaussians model for the petrophysical forward function. Then follows the results of applying this method with two different assumptions about facies. Finally, we discuss the results and conclude.

2.3 Method

2.3.1 Well log data and geological information

The well log data used in this study are from one well (i.e., well No. 3) in Akpo field, which is an offshore field in Africa. Well-log data include P- and S-wave velocity, bulk density, porosity, clay content, water saturation, and effective pressure. Bulk modulus and density of brine and hydrocarbon are also given as functions of depth in this well. In addition to these measured logs, a facies log is also available in this well (Figure 2-1).

The geological context differs from the top to the bottom of the well. Structural and stratigraphic interpretations divide the well interval into 8 sub-intervals with different geological contexts (Seismic inversion report of the Akpo field, 2009). In figure 2-1 thick blue lines show the boundaries between these sub-intervals. The first five intervals mainly include channel complexes, which are characterized by thick sequences of superimposed sand channels and levees. The last three intervals include laterally wide and relatively thin sand lobes within a matrix of shale. The facies log in figure 2-1 shows that five different types of sedimentary clastic facies can be seen in this well. The majority of samples are mixtures of sand and shale with different values of clay content, and a small number of them can be either coal or limestone, which are shown by white color in figure 2-1. We focus on the construction of the petrophysical forward function for the sand-shale mixture.

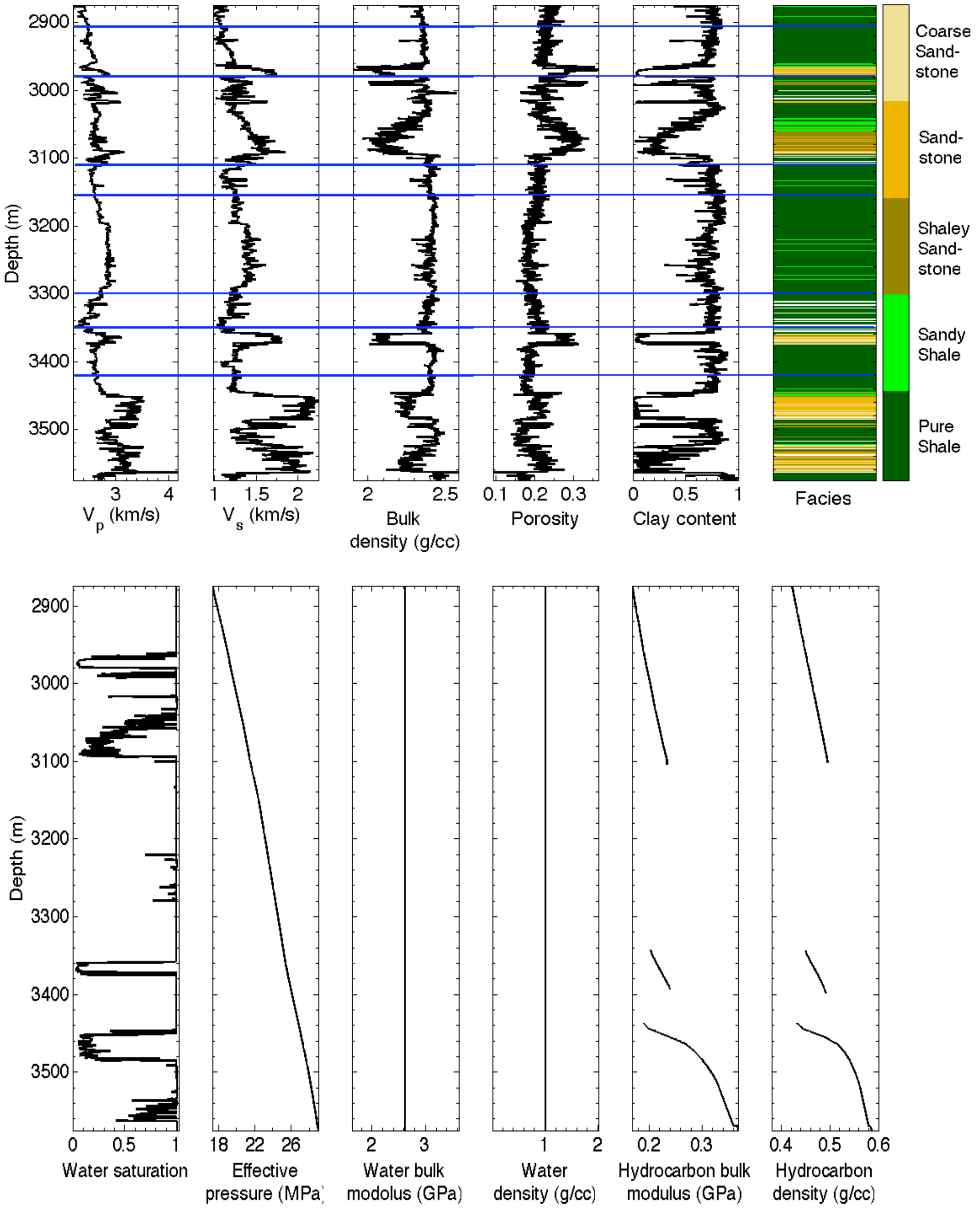


Figure 2-1: Well-log dataset from well No.3 in Akpo field. The bulk modulus and density of hydrocarbon are shown for the intervals with non-zero hydrocarbon saturation. The horizontal blue lines define different stratigraphic units with different geological contexts.

2.3.2 Theoretical models for sand-shale mixture

The mixture of sand and shale can be either laminated or dispersed (Dvorkin and Gutierrez, 2001). In a dispersed sand-shale mixture, packs of relatively small clay particles fill the pore space of the matrix of larger sand particles. By increasing the amount of clay in the mixture the pore space of the sand matrix decreases and eventually packs of clay particles fill all the pore space of the sand matrix, while sand particles in the matrix are still in contact with each other. Up to this point sand particles are the load-bearing phase in the sand-shale mixture and the mixture is usually referred to as shaley sand. By increasing clay content after this point sand particles become suspended in the clay particles framework, which gradually becomes the load-bearing phase in the mixture. The mixture in this case is referred to as sandy shale. Due to change of the load-bearing material in shaley sands and sandy shales, the elastic behavior of a dispersed sand-shale mixture varies for these two types of mixtures.

Marion et al. (1992) and Dvorkin and Gutierrez (2001) developed a theory to describe the elastic behavior of dispersed sand-shale mixture. A detailed description of Dvorkin and Gutierrez (2001) model for dispersed sand-shale is given in Appendix B. In this model, the relationship between clay content and porosity is not differentiable when clay content is equal to the initial porosity of the sand matrix, ϕ_0 (figure 2-2(a)). For sand and shaley sand, where clay content increases from zero to the initial porosity of the sand matrix, porosity decreases linearly as clay content increases. For shale and sandy shale, where clay content is larger than the initial porosity of the sand matrix, porosity increases linearly as clay content increases. In the same way, the relationship between clay content and P-wave velocity is also non-differentiable when clay content is equal to ϕ_0 (figure 2-2(b)). The Dvorkin and Gutierrez model predicts that P-wave velocity as a function of porosity has two distinct branches for sandy shales and shaley sand mixtures. This behavior is usually referred to as a V-shaped trend in velocity-porosity relation and is shown in figure 2-2(c).

In a laminated sand-shale mixture, clay particles are distributed as laminas between sand units. Dvorkin and Gutierrez (2001) developed a theory to describe the porosity-clay content-velocity relation for laminated sand-shale mixtures. P-wave velocity changes continuously by increasing clay content in laminated sand-shale mixtures. Therefore, there is a gradual spread from sand via shaley sand and sandy shale to shale in the cross plot of P-wave velocity versus porosity of the laminated sand-shale mixture, and the V-shaped trend cannot be seen in this case (figure 2-2(d)). This gradual spread in the velocity-porosity cross plot of the laminated sand-shale mixture in conjunction with the V-shaped trend in the velocity-porosity, porosity-clay content, and velocity-clay content cross plots of the dispersed sand-shale mixture can be used to diagnose the type of the sand-shale mixture from measured logs (Avseth et al., 2005, pp. 71, 258-263).

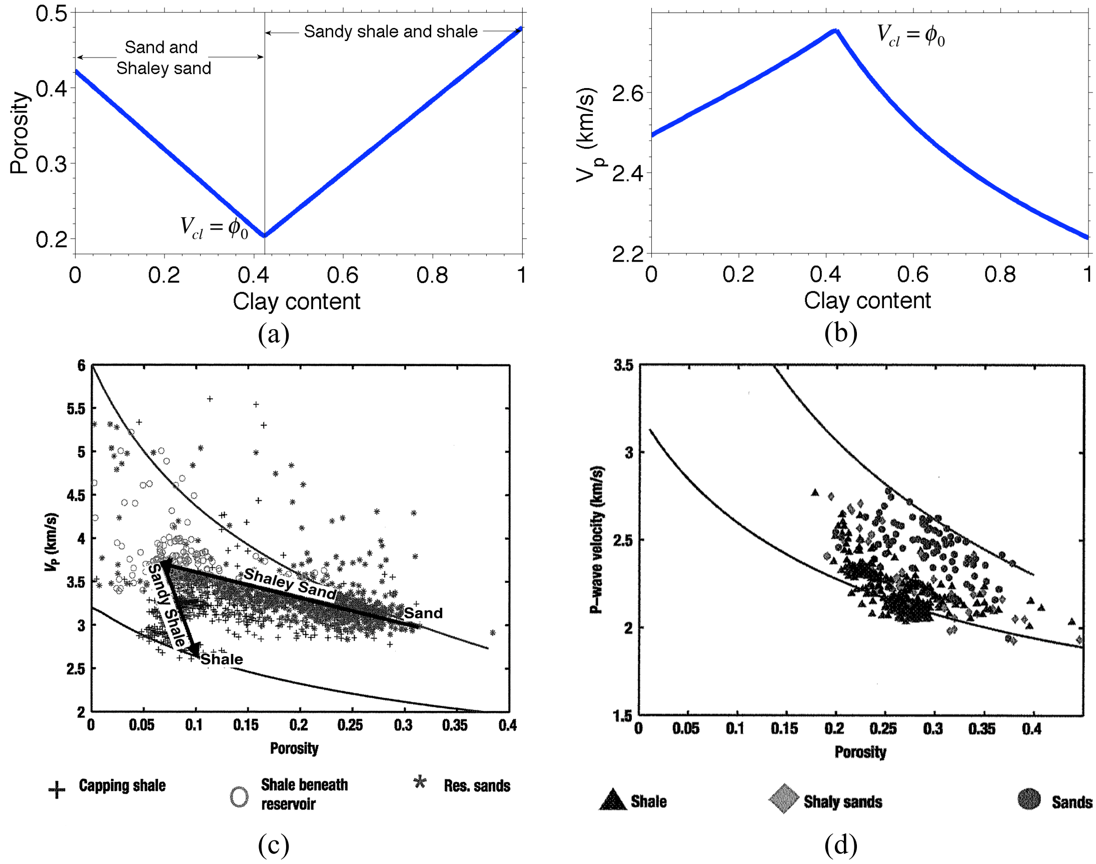


Figure 2-2: Different types of sand-shale mixture. (a) Dispersed sand-shale mixture. Porosity versus clay content from Dvorkin and Gutierrez (2001) model (Appendix B). (b) Dispersed sand-shale mixture. P-wave velocity versus clay content from Dvorkin and Gutierrez (2001) model (Appendix B). (c) Dispersed sand-shale mixture. P-wave velocity versus porosity for a Norwegian sea reservoir zone. Obvious transition from clean sand to shale via shaley sand and sandy shale. (d) Laminated sand-shale mixture. P-wave velocity versus porosity for an offshore Angola turbidite reservoir. Gradual decrease of V_p from clean sand via shaley sand to shale. (Figures (c) and (d) are from Avseth et al. (2005).).

The petrophysical forward model for a laminated sand-shale mixture is derived by assuming that clay laminas are arranged perpendicular to the direction of wave propagation (Dvorkin and Gutierrez, 2001). In this case the effective bulk and shear moduli are derived by Backus averaging and the bulk density of rock is estimated by arithmetic averaging of the bulk density of sand and clay laminas:

$$\frac{1}{K_{mix}} = \frac{V_{cl}}{K_{clay}} + \frac{1-V_{cl}}{K_{sand}} \tag{2-1}$$

$$\frac{1}{G_{mix}} = \frac{V_{cl}}{G_{clay}} + \frac{1-V_{cl}}{G_{sand}} \tag{2-2}$$

$$\rho_{mix} = V_{cl}\,\rho_{clay} + \left(1 - V_{cl}\right)\rho_{sand} \tag{2-3}$$

In the above equations V_{cl} is the clay content, K_{mix}, G_{mix}, and ρ_{mix} are the bulk and shear moduli, and the bulk density of the laminated sand-shale mixture, K_{clay}, G_{clay}, and ρ_{clay} are the equivalent parameters for clay laminas, and K_{sand}, G_{sand}, and ρ_{sand} are equivalents for sand units. Note that equations (2-1) and (2-2) are the Reuss average (Mavko et al., 2009) of the elastic moduli of sand and shale laminas. The Reuss average is a lower bound of the elastic moduli of a mixture of rocks. Therefore, if the perpendicular wave incident assumption is violated in the laminated sand-shale mixture, application of the above formulas result in an underestimate of the equivalent elastic moduli of the mixture.

We assume that the earth is isotropic and therefore the P- and S-wave velocity are derived from the bulk and shear moduli, and the bulk density of the sand-shale mixture as:

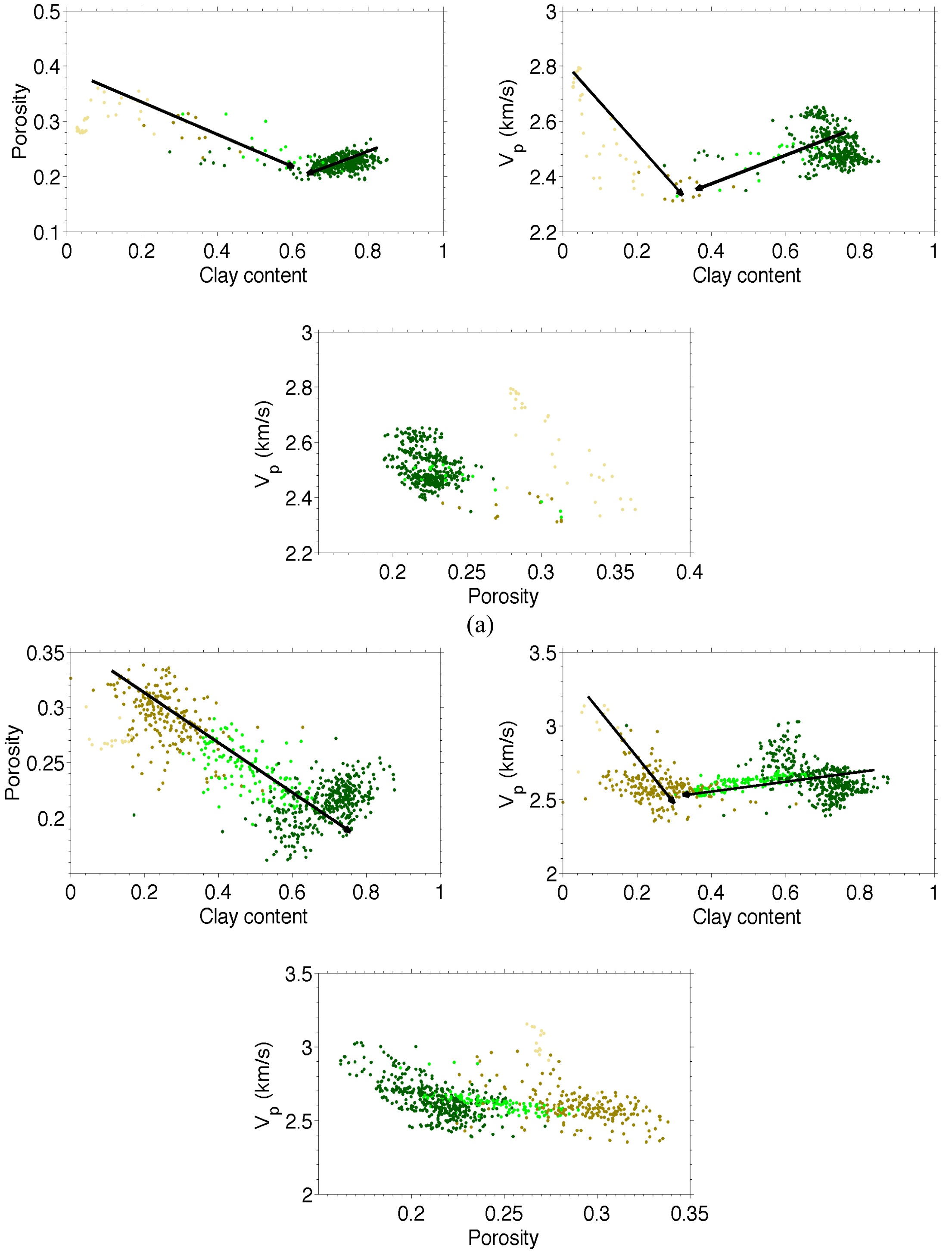
$$V_p = \sqrt{\frac{K_{mix} + \left(4/3\right)G_{mix}}{\rho_{mix}}} \tag{2-4}$$

$$V_s = \sqrt{\frac{G_{mix}}{\rho_{mix}}} \tag{2-5}$$

For the laminated sand-shale mixture, K_{mix}, G_{mix}, and ρ_{mix} in the above equations are obtained from equations (2-1) to (2-3). The model for the dispersed sand-shale mixture is described in detail in Appendix B.

Types of sand-shale mixture in well No.3

We applied the above diagnostic test to identify the type of the sand-shale mixture from P-wave velocity, clay content, and porosity logs. The cross-plots of porosity versus clay content, P-wave velocity versus clay content, and P-wave velocity versus porosity for samples from reservoir sub-intervals are shown in figure 2-3. The change of effective pressure in all sub-intervals is less than 3 MPa, and therefore we neglect the possible effect of effective pressure on the trends in the cross plots. For three out of the four sub-intervals (figure 2-3(a), (b), and (d)), the general trend in the cross plot of P-wave velocity versus clay content is the opposite of the behavior of a dispersed sand-shale mixture (figure 2-2(b)). However, in very small intervals we can observe a behavior consistent with the dispersed sand-shale mixture model; an example of such intervals can be seen in figure 2-4. Figure 2-3 suggests that for the majority of sandy-shale and shaley-sand facies in this well the elastic behavior of the sand-shale mixture is consistent with the laminated model. However, as

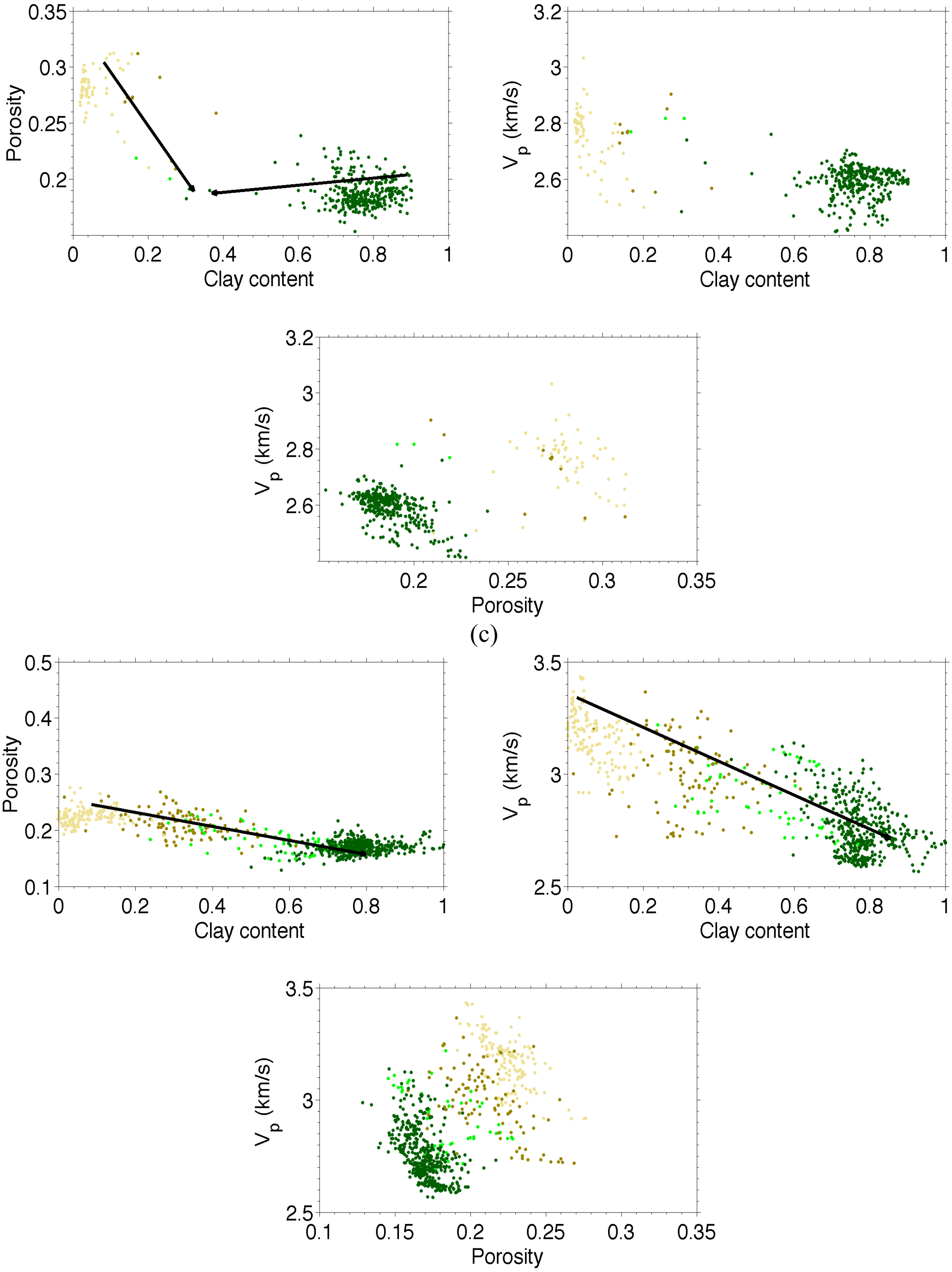


(a)

(b)

Figure 2-3: Cross-plots of porosity, clay content, and P-wave velocity, for samples from reservoir sub-intervals of well No. 3. (a) Cross-plots for 2905 m-2979 m interval. (b) Cross-plots for 2979 m-3112 m interval. (Continued on next page)

figure 2-4 suggests for small intervals the dispersed sand-shale model is more appropriate. For each of depth sub-intervals, we select the appropriate sand-shale mixture model using

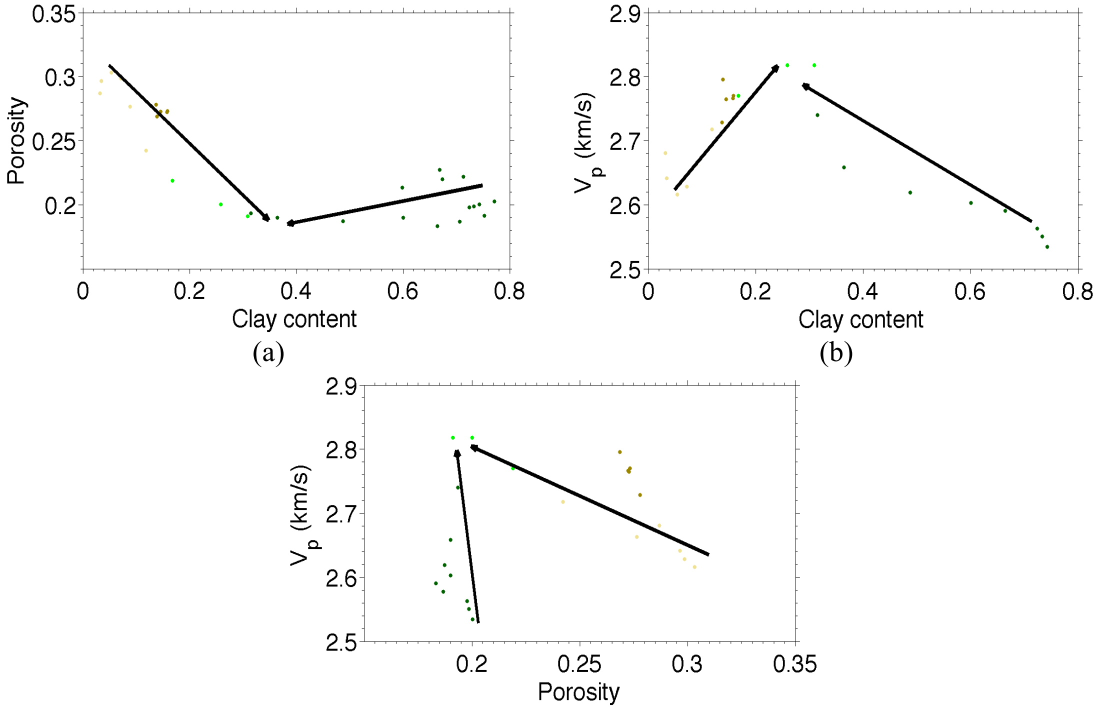


(c)

(d)

Figure 2-3: Continued from previous page. Cross-plots of porosity, clay content, and P-wave velocity, for samples from reservoir sub-intervals of well No. 3. (c) Cross-plots for 3349 m-3420 m interval. (d) Cross-plots for 3420 m-3580 m interval.

the above cross plots.



(a) (b)

(c)

Figure 2-4: Cross-plots of porosity, clay content, and P-wave velocity for 3357 m-3361 m interval. (a) Porosity-clay content cross plot. (b) P-wave velocity-clay content cross plot. (c) P-wave velocity-porosity cross plot. In this interval the behavior of porosity, clay content, and P-wave velocity suggests that the sand-shale mixture is dispersed.

In the next section we introduce a method to estimate the bulk and shear moduli, and the bulk density of sand and shale units in the dispersed and laminated sand-shale mixture.

2.3.3 Rock physics theories

In order to describe the dependency of the bulk and shear moduli of sand and clay components of a mixture on the porosity, effective pressure, and water saturation, we need to model the elastic behavior of packs of granular materials. Several rock physics model are proposed with different assumptions about normal and shear deformation of grains (Mavko et al., 2009, ch. 5). For a particular geological scenario one of the rock physics models may be more appropriate and hence be preferred (Avseth et al., 2005, ch. 2).

Bulk and shear moduli of sand units

The sand units of the Akpo field are uncemented (Personal Communication, 2010, Christian Deplante, manager of Geophysics research group TOTAL E&P UK). In uncemented sands, cement is deposited away from grain contacts (Mavko et al., 2009, p. 258). Bachrach and Avseth (2008) developed a theory to describe the elastic behavior of uncemented sand. In this theory, it is assumed that the starting framework of uncemented sand is a dense random pack of spherical grains, which are not necessarily identical, with an initial porosity ϕ_0. At the initial porosity, under the hydrostatic loading assumption, the effective bulk and shear moduli, K_{eff} and G_{eff}, are given as (Bachrach and Avseth, 2008):

$$K_{eff} = \left(\frac{(1-\phi_0)^2 G^2}{18\pi^2 (1-\nu)^2} \right)^{1/3} c\, p^{1/3} \tag{2-6}$$

$$G_{eff} = \frac{2+3f-\nu(1+3f)}{5(2-\nu)} \left(\frac{3(1-\phi_0)^2 G^2}{2\pi^2 (1-\nu)^2} \right)^{1/3} c\, p^{1/3} \tag{2-7}$$

Here, p is the effective pressure at grain contacts, G is the shear modulus of grains, ν is the Poisson's ratio of grains, f is the fraction of no-slip grain contacts (contacts with finite tangential stiffness), and c is a calibration parameter. Bachrach and Avseth (2008) pointed out that c can capture some geometric aspects of the contacts between grains. The bulk and shear moduli of grains are estimated from the bulk and shear moduli, and volume fraction of sand and clay grains using the Voigt-Reuss-Hill average moduli (Hill, 1952; Mavko et al., 2009,pp. 177-178).

Equations (2-6) and (2-7) describe the effective elastic moduli of sand at the initial porosity ϕ_0. In order to find the bulk and shear moduli of uncemented sand at a different porosity, ϕ, Dvorkin and Nur (1996) proposed a heuristic Hashin-Shtrikman lower bound based on the original Hashin-Shtrikman lower bound (Hashin and Shtrikman, 1963). This model connects two end members–one has zeros porosity and moduli of solid end members, and the other has the initial porosity and pressure dependent moduli as given by equations (2-6) and (2-7). One heuristic argument for this model is that poorly sorted sand (with smaller porosity than the initial porosity) can be modeled as an original frame with large grains with the initial porosity ϕ_0 enveloped by soft shells of fine-grained sands(Avseth et al., 2005, p. 56).

At porosity ϕ, the concentration of the original sphere-pack frame is ϕ/ϕ_0, and that of the pure solid phase added to the sphere-pack frame to decrease porosity is $1-\phi/\phi_0$. The bulk and shear moduli of the mixture is then given by the Hashin-Shtrikman lower bound as:

$$K_{dry} = \left[\frac{\phi/\phi_0}{K_{eff} + 4G_{eff}/3} + \frac{1-\phi/\phi_0}{K + 4G_{eff}/3} \right]^{-1} - \frac{4}{3}G_{eff} \tag{2-8}$$

$$G_{dry} = \left[\frac{\phi/\phi_0}{G_{eff} + z} + \frac{1-\phi/\phi_0}{G + z} \right]^{-1} - z \tag{2-9}$$

where

$$z = \frac{G_{eff}}{6}\left(\frac{9K_{eff} + 8G_{eff}}{K_{eff} + 2G_{eff}} \right)$$

Here, K and G are the bulk and shear moduli of grains, respectively; K_{eff} and G_{eff} are the bulk and shear moduli of the original sphere pack with the initial porosity ϕ_0 (equations (2-6) and (2-7)).

The bulk and shear moduli of saturated sands are derived from the moduli of dry sands (equations (2-8) and (2-9)) using Gassmann's equation (Gassmann, 1951; Avseth et al., 2005, pp. 15-24), and the bulk density is obtained by arithmetic averaging:

$$K_{sand} = K_{dry} + \frac{\left(1 - K_{dry}/K\right)^2}{\phi/K_{fl} + (1-\phi)/K - K_{dry}/K^2} \tag{2-10}$$

$$G_{sand} = G_{dry} \tag{2-11}$$

$$\rho_{sand} = (1-\phi)\rho + \phi\rho_{fl} \tag{2-12}$$

Here, K_{fl} and ρ_{fl} are the effective bulk modulus and density of pore fluid, and K and ρ are the bulk modulus and density of grains. By assuming uniformly mixed fluid phases, in which all pore fluid phases have the same wave-induced pore pressure, the effective bulk modulus of fluid in a two-phase mixture of oil and brine is derived as:

$$K_{fl} = \left(\frac{s_w}{K_{br}} + \frac{(1-s_w)}{K_{hc}} \right)^{-1} \tag{2-13}$$

$$\rho_{fl} = s_w\rho_{br} + (1-s_w)\rho_{hc} \tag{2-14}$$

Here, s_w is the brine saturation, K_{br} and ρ_{br} are the bulk modulus and density of brine, and K_{hc} and ρ_{hc} are the bulk modulus and density of hydrocarbon, which is oil in our example. If saturation is heterogeneous over scales larger than the characteristic diffusion length, i.e. patchy saturation, elastic velocity will be higher than what is obtained using the uniform mixing assumption above (Mavko and Mukerji, 1998). For low seismic frequencies, patchy and uniform assumptions about fluid distribution in rocks result in upper and lower bounds on elastic velocity.

In order to use equations (2-10) and (2-11) to estimate the bulk and shear moduli of sand units, we need to specify values of different parameters in equations (2-6) to (2-9), and (2-13) and (2-14). Some of these parameters such as porosity, effective pressure, and water saturation can be obtained from measured well logs. Some other parameters can be estimated as a function of measured parameters, for example the bulk modulus and density of fluids can be represented as a function of pressure (Batzle and Wang, 1992). The remaining parameters, which include bulk and shear moduli and density of grains K, G, and ρ, respectively, initial porosity ϕ_0, fraction of no-slip contacts f, and calibration parameter c, are not known and must be assigned either from core measurement or via calibration. Values of these parameters depend on the mineralogy and internal structure of rock, and sorting and angularity of grains (Bachrach and Avseth, 2008; Mavko et al., 2009, p. 349). Therefore, due to the change of the properties of rocks, uncertainty of the unknown parameters can be high even within one well. In the next section, we describe a method to estimate the unknown parameters and the uncertainty in the above equations.

Bulk and shear moduli of shale units

Normally, shale units are not cemented and therefore the effective dry bulk and shear moduli at the initial porosity can be modeled using equations (2-6) and (2-7) (Dvorkin and Gutierrez, 2001; Avseth et al., 2005, p. 60). In equations (2-6) and (2-7) for shale units the bulk and shear moduli of grains at the intial porosity are equal to the bulk and shear moduli of clay grains. Gassmann’s equations (equations (2-13) and (2-14)) are used to find the elastic moduli of saturated shale at the initial porosity. We apply the Dvorkin-Guttierez model (Appendix B) to estimate the elastic moduli of shale units with different values of clay content. In the shale units the pack of shale particles envelops sand grains thus the mixture is a composite of two elastic elements–the softer element is porous fluid saturated shale and the stiffer element is the sand grain material. The softer element envelops the stiffer element thus the Hashin-Shtrikman lower bound is used to model the elastic moduli:

$$K_{shale}=\left[\frac{V_{cl}}{K_{eff}+4G_{eff}/3}+\frac{1-V_{cl}}{K_q+4G_{eff}/3}\right]^{-1}-\frac{4}{3}G_{eff} \tag{2-15}$$

$$G_{shale}=\left[\frac{V_{cl}}{G_{eff}+z}+\frac{1-V_{cl}}{G_q+z}\right]^{-1}-z \tag{2-16}$$

where $$z=\frac{G_{eff}}{6}\left(\frac{9K_{eff}+8G_{eff}}{K_{eff}+2G_{eff}}\right)$$

In the above equations K_q and G_q are the bulk and shear moduli of sand particles. The density of shale units is obtained from the density of grains and fluid (equation (2-12)). For shale units also, bulk and shear moduli and density of grains, initial porosity of clay matrix, and the calibration parameter in equations (2-6) and (2-7) are unknown and must be assigned from core measurement or via a calibration process.

Note that, in the derivation of equations (2-6) and (2-7) it is assumed that the grain particles are spherical. The applicability of this assumption in clay matrices can be questioned, because clay particles are usually plate-like (Dvorkin and Gutierrez, 2001). However, the calibration parameter *c* in equations (2-6) and (2-7) takes into account the effect of the shape of grains (Bachrach and Avseth, 2008). Practical applications of the Hertz-Mindlin theory (Mindlin, 1949; Mavko et al., 2009, pp. 246-248), which is a simplified version of the above model, with clay laminas show that the above model can be applied with plate-like clay grains (Dvorkin and Gutierrez, 2001; Avseth et al., 2005, p. 60, p. 95, p.100). In the next section we explain the procedure for estimation of the parameters in equations (2-6) to (2-13) for shale units.

2.3.4 Estimation of uncertainty of the petrophysical forward function

Estimation of uncertainty in the petrophysical forward function for each facies

For a given facies, $\Im_i$, the petrophysical forward function can be represented as:

$$\mathbf{d}=\mathrm{G}_{\Im_i}(\mathbf{m})+\mathbf{e} \tag{2-17}$$

Here, $\mathbf{d}=(V_p, V_s)$, $\mathbf{m}=(\phi, V_{cl}, s_w, p_e, K_{hc}, \rho_{hc}, K_{br}, \rho_{br})$, $\mathrm{G}_{\Im i}$ is the petrophysical forward function for facies $\Im_i$, which is described by the above equations, and **e** is a random vector representing the uncertainty in the petrophysical forward function. The components of the model vector **m** are the measured or estimated parameters of the petrophysical forward function as explained above. At the outset, the other parameters of the petrophysical forward function are not known and must be estimated by using information in a measured dataset (e.g., well-logs).

The probability distribution of the uncertainty vector **e**, which we refer to as measurement uncertainty, is not known and can have different characteristics for different facies. For each facies, we assume that the error of petrophysical forward function prediction is not systematic and can be represented by a symmetric 2×2 covariance matrix Σ with three independent elements. The probability distribution that estimates the maximum uncertainty under the above assumptions is Gaussian and represented as (Malinverno and Parker, 2006):

$$p(\mathbf{e}|\Im_i) = p\left(\mathbf{d}|\mathbf{m},\Im_i\right) = \tag{2-18}$$
$$= \frac{1}{\left((2\pi)^2|\Sigma|\right)^{0.5}} \exp\left(-\frac{1}{2}\left(\mathbf{d}-\mathrm{G}_{\Im_i}(\mathbf{m})\right)^T \Sigma^{-1} \left(\mathbf{d}-\mathrm{G}_{\Im_i}(\mathbf{m})\right)\right)$$

where

$$\Sigma = \begin{bmatrix} \sigma_{V_p}^2 & r\sigma_{V_p}\sigma_{V_s} \\ r\sigma_{V_p}\sigma_{V_s} & \sigma_{V_s}^2 \end{bmatrix}$$

Here, σ_{Vp} and σ_{Vs} are the standard deviation of V_p and V_s, respectively, and r is the correlation coefficient. The non-systematic assumption about the error in the petrophysical forward function can be violated if for example the vertical incident assumption in equations (2-1) and (2-2) or uniform assumption about the fluid mixture in equation (2-14) are violated. We will discuss about the effect of the non-systematic error assumption on the uncertainty of the petrophysical forward function in more details in Discussion section below. For each facies we group together all unknown parameters in the probability distribution of the uncertainty (equation (2-18)) into a vector $\mathbf{h}_{\Im i}$. For example, for sand facies this vector is represented as:

$$\mathbf{h}_{\Im_i} = \left(c,\phi_0,K,G,\rho,f,\Sigma\right) \tag{2-19}$$

Where c, ϕ_0, K, G, ρ, and f are the parameters of the forward petrophysical function (equations (2-1) to (2-14)) and Σ is the covariance matrix in equation (2-18). By using the above representation we replace $\mathrm{G}_{\Im i}(\mathbf{m})$ in equation (2-18) by $\mathrm{G}(\mathbf{m}, \mathbf{h}_{\Im i})$ and rewrite equation (2-18) as:

$$
\begin{aligned}
p(\mathbf{e}|\mathbf{h}_{\Im_i}) &= p\left(\mathbf{d}\middle|\mathbf{m},\mathbf{h}_{\Im_i}\right) = \\
&= \frac{1}{\left((2\pi)^2|\Sigma|\right)^{0.5}} \exp\left(-\frac{1}{2}\left(\mathbf{d}-\mathrm{G}\left(\mathbf{m},\mathbf{h}_{\Im_i}\right)\right)^T \Sigma^{-1} \left(\mathbf{d}-\mathrm{G}\left(\mathbf{m},\mathbf{h}_{\Im_i}\right)\right)\right)
\end{aligned}
\tag{2-20}
$$

In order to estimate the unknown parameters of the petrophysical forward function, $\mathbf{h}_{\Im i}$, and the covariance matrix of the uncertainty vector, Σ, simultaneously, we apply the Empirical Bayes method (Malinverno and Parker, 2006). In the Empirical Bayes approach, for a given facies $\Im_i$, we approximate the distribution of **e** for all values of $\mathbf{h}_{\Im i}$ by the distribution of **e** given $\hat{\mathbf{h}}_{\Im i}$, the most probable value of the a posteriori unknown parameters of the equation (2-20). In mathematical terms we use the following approximation:

$$
p\left(\mathbf{e}\middle|\mathbf{h}_{\Im_i}\right) \approx p\left(\mathbf{e}\middle|\hat{\mathbf{h}}_{\Im_i}\right)
\tag{2-21}
$$

For data samples (**m**, **d**) of a given facies $\Im_i$, the most probable value $\hat{\mathbf{h}}_{\Im i}$ maximizes the posterior probability of the unknown parameter vector $\mathbf{h}_{\Im i}$ in light of the data. Using Bayes' rule, the posterior probability of $\mathbf{h}_{\Im i}$ given the data samples, can be represented as:

$$
\begin{aligned}
p\left(\mathbf{h}_{\Im_i}\middle|\mathbf{e}\right) &\propto p\left(\mathbf{e}\middle|\mathbf{h}_{\Im_i}\right)p\left(\mathbf{h}_{\Im_i}\right) \\
&= p\left(\mathbf{d}\middle|\mathbf{m},\mathbf{h}_{\Im_i}\right)p\left(\mathbf{h}_{\Im_i}\right)
\end{aligned}
\tag{2-22}
$$

In the second line of the above equation, we substitute $p(\mathbf{e}|\mathbf{h}_{\Im i})$ with its definition in the equation (2-20). Therefore, by assuming a uniform a priori distribution for unknown parameters, $p(\mathbf{h}_{\Im i}) = \mathrm{U}[a, b]$, for a given dataset of the facies $\Im_i$ with n independent samples, $(\mathbf{m}_j, \mathbf{d}_j)$, $j = 1,\ldots,n$, the most probable value of unknown parameters, $\hat{\mathbf{h}}_{\Im i}$, can be estimated by maximizing the following function:

$$
\xi\left(\mathbf{h}_{\Im_i}\right) = \prod_{j=1}^{n} p\left(\mathbf{d}_j\middle|\mathbf{m}_j,\mathbf{h}_{\Im_i}\right)
\tag{2-23}
$$

In the above equation, $p(\mathbf{d}_j| \mathbf{m}_j, \mathbf{h}_{\Im i})$ is calculated from equation (2-20). It is usually more convenient to minimize $-\log(\xi(\mathbf{h}_{\Im i}))$ instead of maximizing equation (2-23) to find the optimum value of unknown parameters $\hat{\mathbf{h}}_{\Im i}$. In this work we applied the simplex algorithm (Lagarias et al., 1998) to find the minimum of the above function.

We apply several assumptions in the above derivation of the petrophysical forward function with its associated uncertainty for a given rock facies. The main assumption is the Gaussian assumption about uncertainty. This assumption implies that the error of predictions of the petrophysical forward function is non-systematic. However, in the derivation of the petrophysical forward function we observe that the vertical incident and uniform mixing fluid assumptions can result in an underestimation of elastic moduli, which is a systematic error. The non-systematic error assumption is applicable only if strong a priori geological information about the orientation of earth layers and fluid distribution are available, which is not usually the case. As a result the proposed method above is prone to uncertainty in such a priori geological information.

Estimation of uncertainty of the global petrophysical forward function

The global petrophysical forward function must predict the value of $\mathbf{d}=(V_p, V_s)$ for any given model vector $\mathbf{m}=(\phi, V_{cl}, s_w, p_e, K_{hc}, \rho_{hc}, K_{br}, \rho_{br})$ and for all facies presented in the well. In general, we do not have any explicit information about facies and therefore we cannot select a proper petrophysical function to derive the value of **d** for a given value of **m**. The model vector **m** cannot be used to distinguish one proper facies among different possible facies, uniquely. In fact, it is possible that two samples from different facies have exactly the same values of the model parameters. The lack of information about the type of facies acts as an additional source of uncertainty in the global petrophysical forward function, which we refer to as theoretical uncertainty. As we indicated above, the vector of the unknown parameters of the petrophysical forward function, **h** (equation (2-19)), depends on the type of facies. Using **h**, theoretical uncertainty can be represented as:

$$p(\mathbf{d}|\mathbf{m}) = \int_{\mathbf{h}} p(\mathbf{d}|\mathbf{m},\mathbf{h})\, p(\mathbf{h}|\mathbf{m})\ \mathrm{d}\,\mathbf{h} \qquad (2\text{-}24)$$

If m different facies exist in the dataset we can represent $p(\mathbf{h}|\mathbf{m})$ as:

$$p(\mathbf{h}|\mathbf{m}) = \sum_{i=1}^{m} p(\mathbf{h}|\mathbf{m},\Im_i)\, p(\Im_i|\mathbf{m}) \qquad (2\text{-}25)$$

We assume that for a given facies $\Im_i$, uncertainty about the unknown parameter vector $\mathbf{h}_{\Im i}$ is negligible and simplify the above equation as:

$$p(\mathbf{h}|\mathbf{m}) \approx \sum_{i=1}^{m} \delta\left(\mathbf{h}-\hat{\mathbf{h}}_{\Im_i}\right) p\left(\Im_i|\mathbf{m}\right) \tag{2-26}$$

In the above equation, $\delta(\mathbf{h})$ is the Dirac delta function, which is equal to zero everywhere except at the origin and has integral equal to one. In the above equation we approximate the probability distribution of **h** for a given facies $\Im_i$, by the most probable value of **h** for that facies, i.e., $\hat{\mathbf{h}}_{\Im_i}$. We will discuss about the possible limitations and errors due to the above approximation in the Discussion section below.

We can substitute equation (2-26) into equation (2-24) to obtain the following form for the global petrophysical forward function with its associated uncertainty:

$$p(\mathbf{d}|\mathbf{m}) = \sum_{i=1}^{m} p\left(\mathbf{d}\middle|\mathbf{m},\hat{\mathbf{h}}_{\Im_i}\right) p\left(\Im_i|\mathbf{m}\right) \tag{2-27}$$

In the above equation $p(\mathbf{d}|\mathbf{m}, \hat{\mathbf{h}}_{\Im_i})$ is derived from equation (2-20) for each facies.

The probability of each facies, $\Im_i$, conditioned on the value of the model vector **m**, $p(\Im_i|\mathbf{m})$, must be derived from well-log data and geological prior information. The geological prior information must be obtained independently of the log data used to estimate $\mathbf{h}_{\Im_i}$. For example this information could come from a geologist's direct inspection of cores, e.g. facies log. Samples in the facies log can be considered as realizations of $p(\Im_i|\mathbf{m})$, and therefore can be used to estimate that conditional probability. The histogram of the samples could be used to infer the above conditional probability density function. Other methods of density estimation are discussed in Bishop (1995, ch. 2). Here we will apply the K-nearest neighbor algorithm for density estimation (Cover and Hart, 1967; Bishop, 1995, pp. 55-57) to estimate $p(\Im_i|\mathbf{m})$.

In the K-nearest neighbor algorithm, a set of samples of the model vectors and their corresponding facies label is given: $(\mathbf{m}_r, \Im_{i_r})$. For any given value of **m**, the first K nearest neighbors of **m** in the sample set are considered. The probability of facies *i* conditioned on **m**, $p(\Im_i|\mathbf{m})$, is obtained as the frequency of the samples of facies *i* in the considered K nearest neighbors of **m**. We used the Mahalanobis distance (Bishop, 1995, p. 35), which is a scale-independent measure of distance, to select the K-nearest neighbor samples of **m**.

2.4 Results

We now present the results of the application of the above method to derive a closed-form for the probabilistic petrophysical forward function in well No.3 of the Akpo field. In this section we present two different applications of the above method. In the first application we assume that just a single sand and shale model can represent all the facies at the well and show the limitations of such an assumption. In the second application we specify possible sub-facies for sand and shale and apply the above method to construct the petrophysical forward function, showing the improvement in the accuracy caused by this approach.

2.4.1 First application: single facies petrophysical forward function

In this application we assumed that the unknown parameters (i.e., initial porosity, compressional and shear moduli and density of grains, and the calibration parameter) of sands with $V_{cl} < 0.20$ (and shales with $V_{cl} > 0.60$) in equations (2-4) to (2-14) are the same for all samples in the well No.3 of the Akpo field. We also assumed that all the sand-shale mixtures with $0.20 < V_{cl} < 0.60$ in this well are laminated (figure 2-5). These are significant assumptions because as we explained above the geologic context differs from the top to the bottom of this well. However, such simplifying assumptions are commonly used in oil industry to construct the petrophysical forward function (Chao et al., 2009). We applied the above method to estimate the vector of the unknown parameters, $\hat{\mathbf{h}}_{\Im_i}$ for sand and shale.

The a priori information about the range of the parameters of the unknown vector $\mathbf{h}_{\Im_i}$ is given in table 2-1. For initial porosity, elastic moduli, and density these intervals are selected by using information provided by Avseth et al. (2005, pp. 61, 307) and Mavko et al. (2009, pp. 458-460). We do not have any a priori information about the fraction of no-slip grains, f, thus it can take any value between zero and one. The only a priori information about the calibration parameter is that it is non-negative (Bachrach and Avseth, 2008). The standard deviation of the V_p and V_s are assumed to change within the 10% of the mean value of the measured logs. The value of the correlation between V_p and V_s can take any value between -1 and 1. In the simplex algorithm, which is applied to find the most probable value of the unknown parameters $\hat{\mathbf{h}}_{\Im_i}$, we limit the search of these parameters to the above intervals.

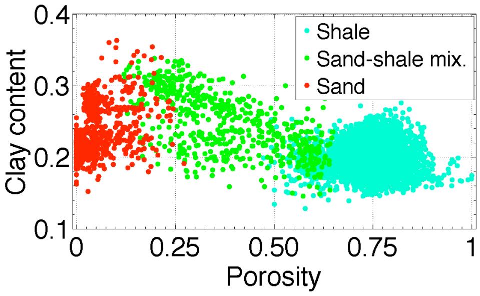


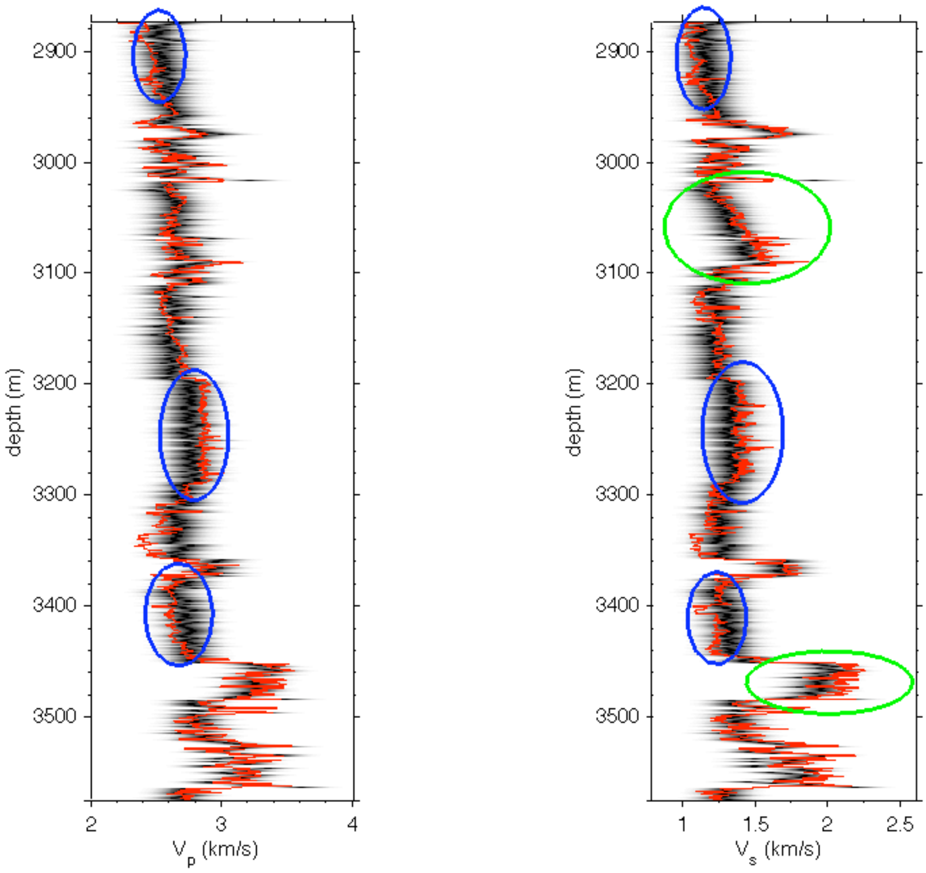
Figure 2-5: Cross-plot of porosity and clay content for different facies in well No. 3.

Parameters		
Sand		
	ϕ_0	0.3 – 0.45
	K (GPa)	30 – 40
	G (GPa)	40 – 50
	ρ (g/cc)	2.5 – 2.7
	f	0 – 1
	c	positive
Clay		
	ϕ_0	0.6 – 0.9
	K (GPa)	1.5 – 95
	G (GPa)	1.4 – 20
	ρ (g/cc)	1.58 – 2.6
	f	0 – 1
	c	positive
Covariance matrix		
	σ_p (m/s)	0 – 270
	σ_s (m/s)	0 – 140
	r	-1 – 1

Table 2-1: A priori intervals of the parameters of the unknowns vector $\mathbf{h}_{\mathfrak{I}_i}$.

Figure 2-6 shows the predicted probability distribution of V_p and V_s at the well location. In that figure the measured logs are also shown on red color. Note that even though the predicted values of V_p and V_s are very close to the measured values, the final estimate seems to be biased in some intervals of the well, indicated by colored ovals. The biased intervals for shales (blue ovals) coincide on both V_p and V_s logs. The predicted values of V_s underestimate the measured log values for sands (green ovals). The optimum values of the parameters of the vector of unknown parameters, $\hat{\mathbf{h}}_{fa_i}$ is presented in table 2-2. The percent error between the maximum a posteriori (MAP) point of the probability distribution of V_p and V_s, and measured value of these parameters are 3.2% and 5.8%, respectively.

The source of biases in the predictions of the petrophysical forward function may be the difference between elastic properties of sediments in different intervals, or the difference between the types of the sand-shale mixture with the assumed laminated model. In the next



(a) (b)

Figure 2-6: The probability distribution of P- and S-wave velocity derived by using the petrophysical forward function with single facies (the black color shows the high probability zones). (a) P-wave velocity. (b) S-wave velocity. The red curve is the measured well log. The blue ovals show the biased intervals for shales. The green ovals show the biased intervals for sands.

section, we further investigate this hypothesis and apply our method to construct the petrophysical forward model by assuming more than one facies for sands and shales.

Parameters		
Sand		
	ϕ_0	0.43
	K (GPa)	30.80
	G (GPa)	49.20
	ρ (g/cc)	2.66
	f	1.00
	c	45.05
Clay		
	ϕ_0	0.76
	K (GPa)	21.05
	G (GPa)	12.98
	ρ (g/cc)	2.60
	f	0.99
	c	55.86
Covariance matrix		
	σ_p (m/s)	107
	σ_s (m/s)	99
	r	0.61

Table 2-2: Optimum values of the parameters of the unknown vector $\mathbf{h}_{\mathfrak{I}i}$ for the single facies application.

2.4.2 Second application: multi facies petrophysical forward function

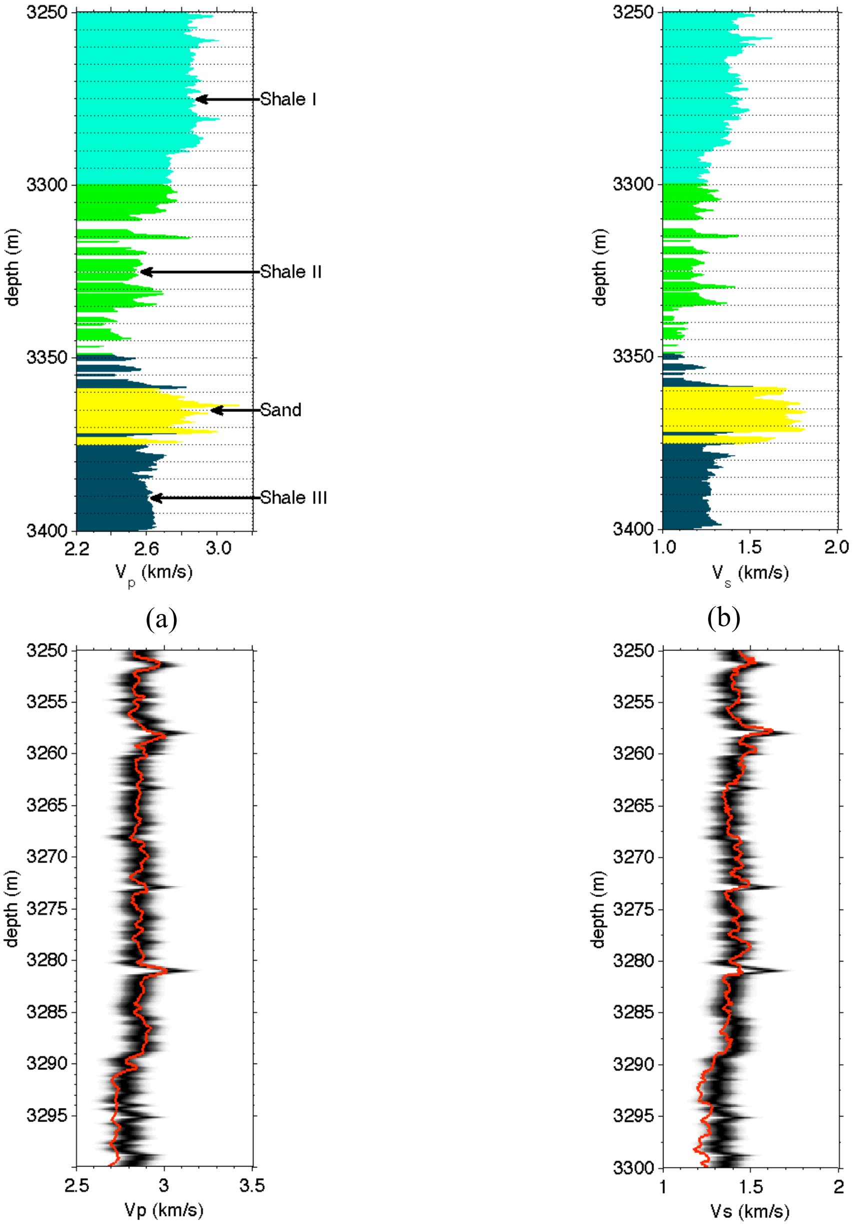
Estimation of the petrophysical forward function for each facies $p(\mathbf{d}\,|\,\mathbf{m}, \mathbf{h}_{\mathfrak{I}i})$

We applied the above method to model the elastic velocities and density for each of the 8 sub-intervals of the well in figure 2-1 with different geological context. Prior geological knowledge is commonly used in geophysical applications to assign different rock properties to different sub-intervals or layers of the earth model. For example, in the AVO inversion of seismic data for P- and S-wave impedance in Akpo field, prior information about model parameters in each of the above sub-intervals were different (Seismic inversion report of the Akpo field, 2009). We applied the elastic model for dispersed and laminated sand-shale mixture to model elastic properties of sandy shale and shaley sand facies in theses intervals. The cross-plots in figure 2-3 and figure 2-4 are used to select the appropriate type of sand-shale mixtures. We applied the elastic model for sand units, presented in Method section above, to model elastic properties of coarse sandstone and sandstone facies. The Voigt-Reuss-Hill average moduli were applied to estimate the elastic moduli of the grains in the sand units. We also applied the elastic model for shale units, presented above to model the elastic properties of pure shale facies.

Figure 2-7 and figure 2-8 show the prediction of the petrophysical forward functions for three sub-intervals of the well. The first sub-interval from 3250 m to 3300 m is a shale interval. The predicted PDF of P- and S-wave velocity for this intervals, with the measured well logs are shown in figure 2-7(c) and (d), respectively. In this interval, the standard deviations of the Gaussian PDF of V_p and V_s are equal to 49 m/s and 47 m/s, respectively. The bias in the prediction is in general small; however, it can be observed in the last 10 m of this interval.

Figure 2-8(a) and (b) show the predicted PDF of V_p and V_s in the second sub-interval from 3300 m to 3349.3 m. The unpredicted (white) intervals in this sub-interval correspond to coal facies, which is not included in our petrophysical forward function. The standard deviations of the Gaussian PDF of V_p and V_s are equal to 97 m/s and 66 m/s, respectively. Bias can be seen in the first 10 m and last 15 m of this interval. The larger standard deviation and bias in this interval may indicate that the petro-elastic model for shale units is inappropriate for predicting the elastic properties of shales in this interval. It may also indicate that this interval needs to be divided into more sub-intervals with more homogeneous properties, which are modeled more accurately using the above petro-elastic model.

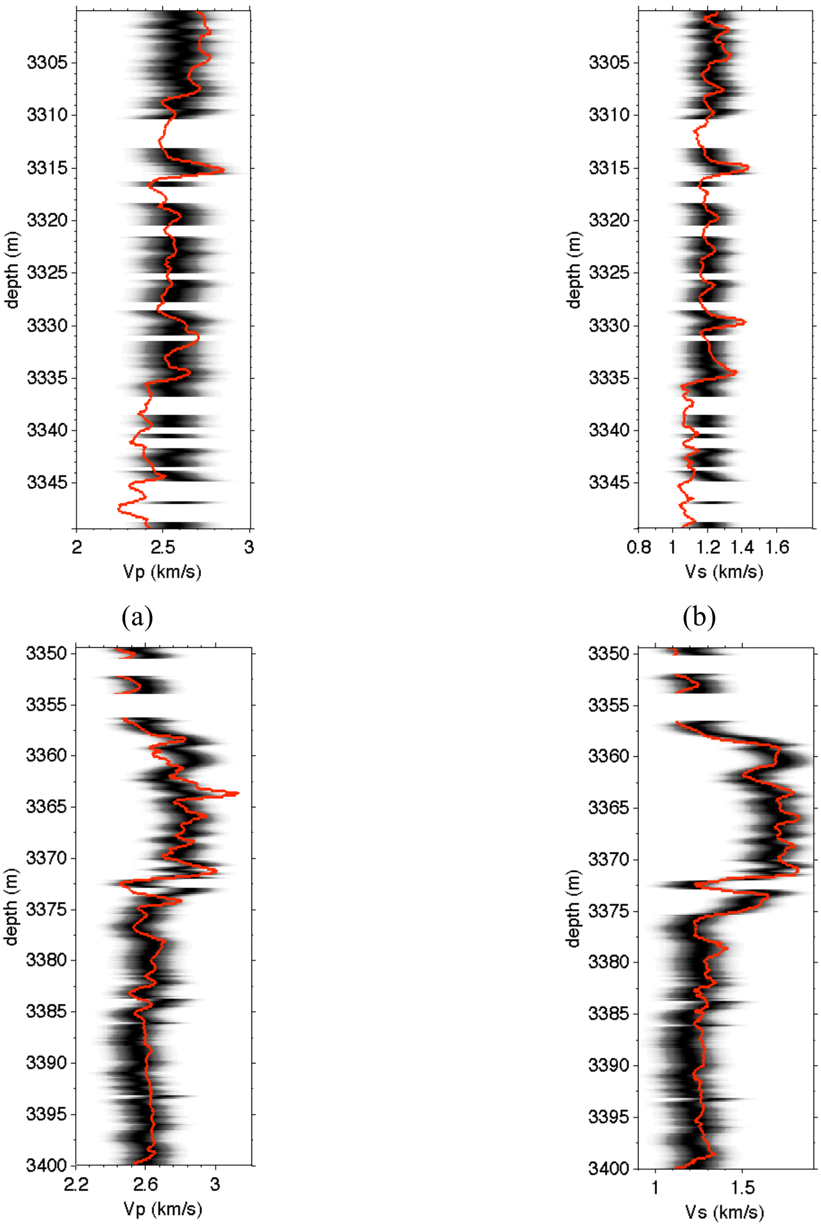
Figure 2-8(c) and (d) show the predicted PDF of V_p and V_s in the third sub-interval from 3349.3 m to 3400 m. Again there are some coal intervals in this sub-interval. The standard deviations of the Gaussian PDF of V_p and V_s are equal to 75 m/s and 82 m/s, respectively. In this interval also we observe some biases. However, in general the model prediction is an acceptable estimate of the measured well log. Note that on a short interval from 3357 m to 3361 m we have 7 samples of dispersed sandy shale and shaley sand facies (Figure 2-4). We apply the model for the sand and shale units above to model the elastic behavior of these samples instead the proper dispersed sand-shale mixture model. Due to the small number of these samples (7 sample) the error caused above does not influence the petrophysical forward function largely.



(a) (b)

(c) (d)

Figure 2-7: (a) P-wave velocity of an interval of the well with three different sub-intervals from 3250 m-3300 m, 3300 m-3349.3 m, and 3349.3 m-3400 m. (b) S-wave velocity for the corresponding interval. (c) The predicted PDF of the P-wave velocity obtained using the petrophysical forward model for sub-interval 3250 m-3300 m. (d) The corresponding PDF of the S-wave velocity. Darker areas show higher probabilities. The red curve is the measured log.



(a) (b)

(c) (d)

Figure 2-8: Prediction of the PDF of P- and S-wave velocity for sub-intervals of the well. (a) The predicted PDF of the P-wave velocity obtained using the petrophysical forward model for sub-interval 3300 m-3349.3 m. (b) S-wave velocity for the corresponding interval. (c) The predicted PDF of the P-wave velocity for sub-interval 3349.3 m-3400 m. (d) The corresponding PDF of the S-wave velocity. Darker areas show higher probabilities. The red curve is the measured logs.

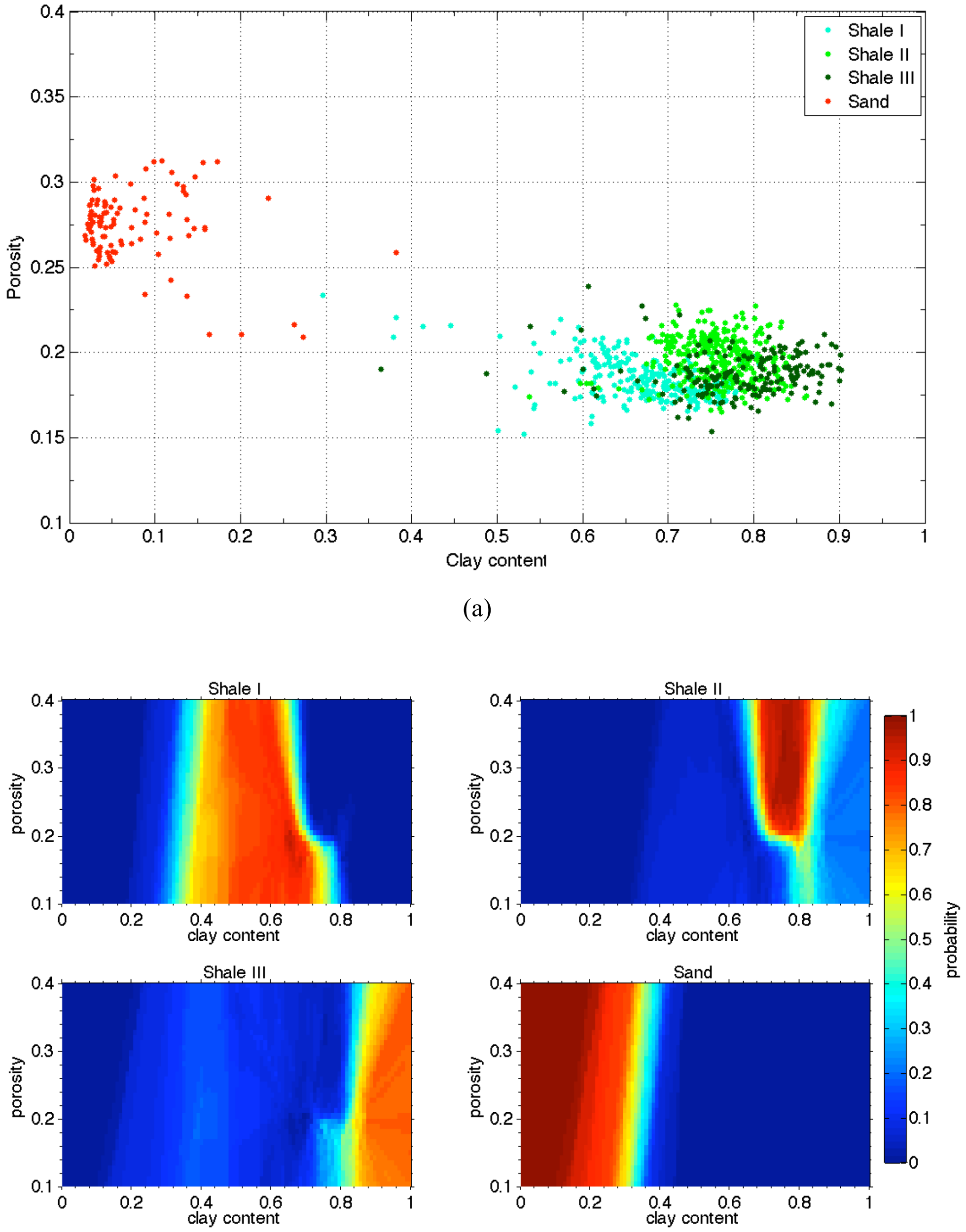
Estimation of facies probability from model vector $p(\Im_i | \mathbf{m})$

Above we apply the petrophysical forward function to estimate the PDF of data vector $\mathbf{d} = (V_p, V_s)$ conditioned on the model vector $\mathbf{m} = (\phi, V_{cl}, s_w, p_e, K_{hc}, \rho_{hc}, K_{br}, \rho_{br})$, for each facies $\Im_i$. This PDF is represented by $p(\mathbf{d}|\mathbf{m}, \hat{\mathbf{h}}_{\Im i})$ in equation(2-27). In the above interval, we observe three different shales and one sand facies. The probability of each of these facies conditioned on the model vector, $p(\Im_i|\mathbf{m})$ (equation (2-27)), is obtained using well log and facies log samples. We assume that the type of the facies is independent of water saturation, effective pressure, bulk modulus and density of fluid, and depends only on the value of porosity and clay content, therefore, we have $p(\Im_i|\mathbf{m}) = p(\Im_i|\phi, V_{cl})$.

The cross-plot of porosity versus clay content for the samples in Figure 2-7(a) are shown in figure 2-9(a). There is a clear overlap between Shale II and Shale III in the porosity-clay content plane, and the transition between different shales and even sand and shales is not abrupt. Figure 2-9(b) shows the estimate of $p(\Im_i|\phi, V_{cl})$ for the samples in figure 2-9(a) found by using the K-nearest neighbor algorithm. The probability of different types of shale facies in the overlap zone is non-zero. Moreover, assigning non-zero probabilities to different facies involved in a transition zone results in smooth transition between facies. For example, the probability of sand facies decreases smoothly from 0.9 to 0.1 as clay content increases from 0.3 to 0.4.

Multi facies petrophysical forward function

We apply the same procedure as above to estimate $p(\Im_i|\phi, V_{cl})$ and $p(\mathbf{d}|\mathbf{m}, \hat{\mathbf{h}}_{\Im_i})$ for the 8 sub-intervals of well Akpo-3 and use equation (2-27) to construct the petrophysical forward function. Figure 2-10 shows the predicted probability distribution of V_p and V_s at the well location, the measured logs are shown by the red color in that figure. We observe an improvement in the prediction of V_p and V_s over the biased intervals of the single facies model predictions (figure 2-6). The multi-modality of the prediction, which is a direct result of the application of the multi-facies petrophysical forward function, can be seen on several intervals in figure 2-10. For example the 3200 m- 3300 m, and 3380 m- 3460 m intervals, are two shale intervals, where two different petrophysical forward functions contribute to the predicted PDF of V_p and V_s. This multi-modality will be seen wherever we cannot definitively select the best petrophysical forward model due to the lack of information about the type of the facies.

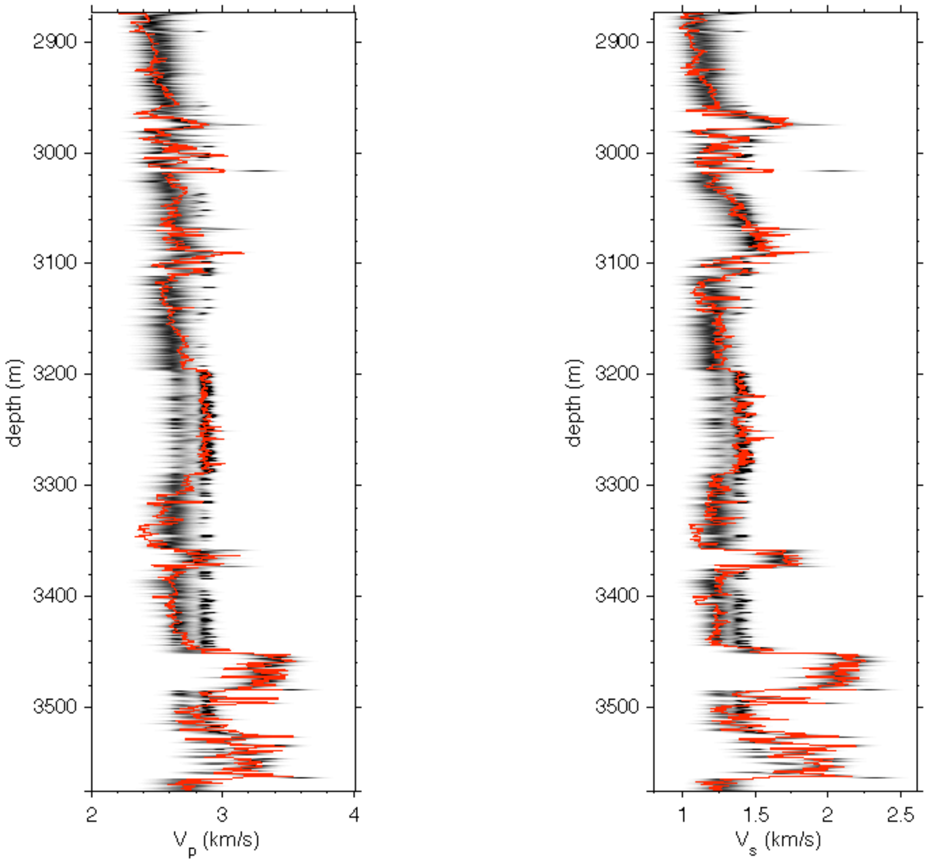


(a)

(b)

Figure 2-9: (a) Cross plot of clay content versus porosity for samples from the three selected sub-intervlas. (b) Probability of different facies conditioned on porosity and clay content, $p(\Im_i|\phi, V_{cl})$.

Figure 2-11 shows the estimated PDF of V_p for three different samples of the well log. This figure shows that for some samples (mainly the sand samples) the predicted PDF has a single mode (figure 2-11(a)). Figure 2-11(c) and figure 2-11(e) show two shale samples with two different modes in the predicted value of V_p. In figure 2-11(c) the measured log value is near the maximum of the estimated PDF, and in figure 2-11(e) the measured log is near the smaller peak of the PDF. This case shows that the maximum a posteriori of the estimated PDF of V_p can be a biased representative of the predicted PDF. Figures 2-10(b), 2-10(d), and



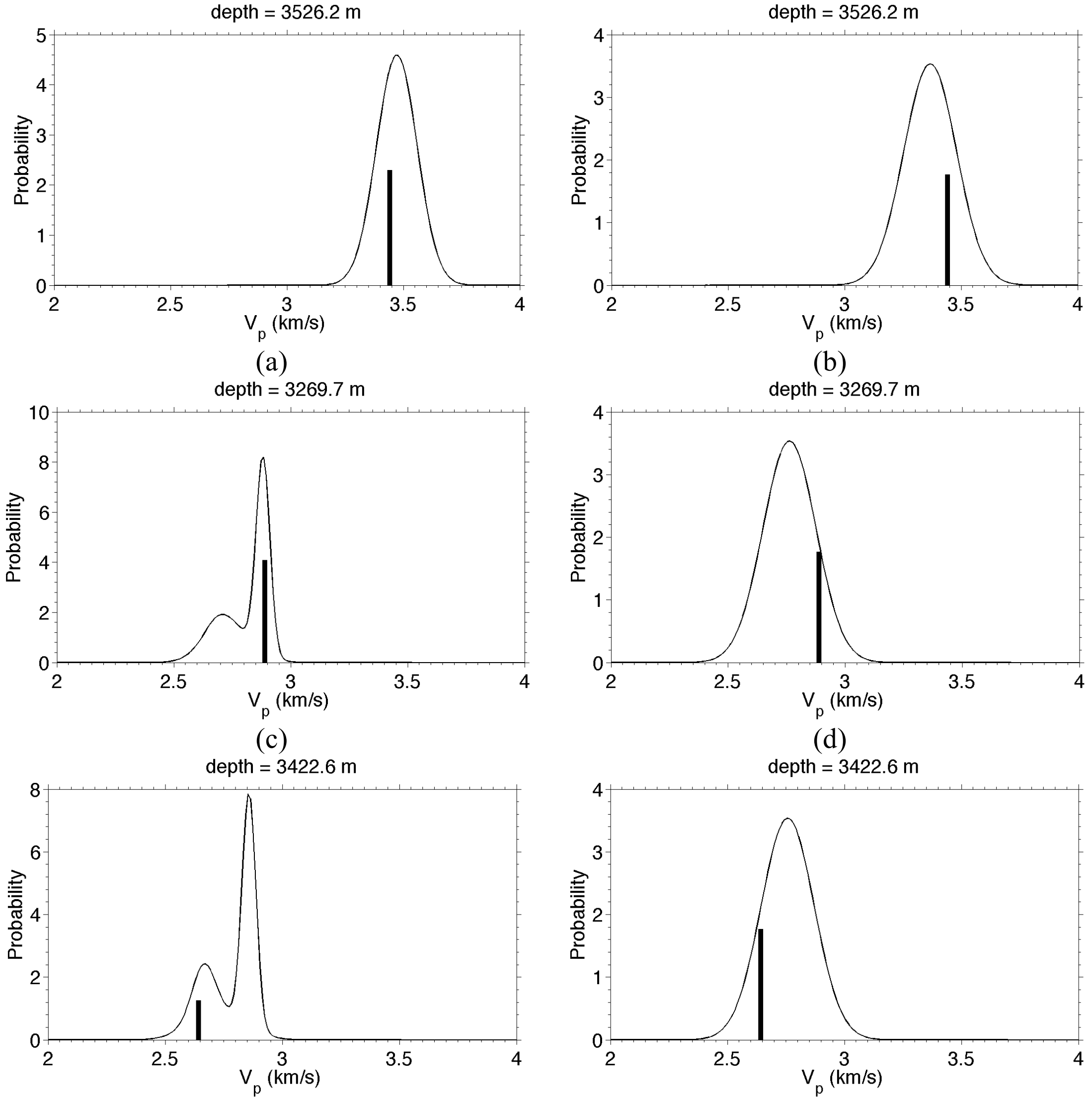
(a) (b)

Figure 2-10: The probability distribution of P- and S-wave velocity derived by using the petrophysical forward function with several facies (the darker color shows the high probability zones). (a) P-wave velocity. (b) S-wave velocity. The red curve is the measured well log.

2-10(f) show the predicted PDF of V_p with the single facies petrophysical forward function in three previous cases. Comparison between the results of the single facies and multi facies cases shows that the uncertainty of the estimate of the multi facies forward function is smaller. It also shows that the mean of the predicted PDF of V_p by single facies petrophysical forward function is clearly biased for two different shale samples, which means that the mean of the estimated PDF of V_p can also be biased as a representation of the predicted PDF.

2.5 Discussion

The result of the single facies petrophysical forward function (figure 2-6) shows that the single facies assumption results in biases in the prediction of V_p and V_s in several depth intervals of the well. This observation is in agreement with previous published works about the construction of the petrophysical forward function for Akpo field (Chao et al., 2009). Previous applications of rock physics theories for seismic interpretation indicated that



(a) (b) (c) (d) (e) (f)

Figure 2-11: The estimated PDF of P-wave velocity by assuming several facies (left), and single facies (right). (a) Multi facies estimate for the sand sample at 3526.2 m. (b) Single facies estimate for the sand sample at 3526.2 m. (c) Multi facies estimate for the shale sample at 3269.7 m. (d) Single facies estimate for the shale sample at 3269.7 m. (e) Multi facies estimate for the shale sample at 3422.6 m. (f) Single facies estimate for the shale sample at 3422.6 m. The black vertical line shows the value of measured log.

different petrophysical forward functions would apply for different burial depths, even for the same basin and the same stratigraphic level (Avseth et al., 2005, p. 107). Therefore, in order to apply rock physics theories for quantitative interpretation of seismic data we must construct a global petrophysical forward function that can predict possible values of V_p and V_s for the various different facies in a field, or in our case at least those encountered by one well.

Figure 2-7(a) shows a relatively short interval, with almost constant effective pressure, where we can see three different types of shale. All of these different types of shale are labeled as pure shale in the facies log (figure 2-1), and they are concentrated in a relatively small region of the porosity and clay content plane (figure 2-9(a)). However, the typical values of V_p and V_s for these samples are different and we can distinguish them on the V_p and V_s logs (figure 2-7(a)). Because the effective pressure, porosity, clay content, and fluid properties are nearly similar for all of these samples, we can relate the difference between petro-elastic properties to differences between petrophysical forward functions in each case, which in terms can be due to differences of depositional and sedimentary properties (e.g., initial depositional porosity, grain elastic moduli, etc.). We applied information provided by geologists to distinguish the above 3 sub-intervals with different types of shales and obtain the petrophysical forward function with its associated uncertainty for each facies (figure 2-7 and figure 2-8). More geological a priori information may change the boundaries between these three intervals, or may even introduce new facies in these intervals, and consequently reduce the error of the petrophysical forward function for each facies.

Equation (2-27) demonstrates how the method represents the uncertainty in V_p and V_s due to lack of information about facies type. Figure 2-9(b) shows that for some values of porosity and clay content, where the probability of different facies are non-zero, the prediction of the petrophysical forward function is a mixture of the result of the different facies petrophysical forward functions. Figure 2-10 and figure 2-11 show this mixing effect can result in multi-modal PDF's of V_p and V_s. However, counter-intuitively, figure 2-11 shows that introducing uncertainty in the facies actually reduces uncertainty in the final probabilistic predictions. This is because without this additional variations in facies, an incorrect petrophysical model is fit to this well interval; increasing estimated uncertainty is the only way the inference methodology can make the predictions compatible with the well data.

In this case study very little information about geology of the region and facies types at the well was available and used to construct the petrophysical forward function. Geological information and core observations can improve our understanding of facies type, which in terms can result in a better selection of the petrophysical forward function, and therefore reduce the uncertainty of the petrophysical forward function. For example, the clay content value used in equations (2-15) and (2-16) of the petrophysical model can vary for sand and shale laminas in a finely laminated sand-shale mixture. If the vertical resolution of the well log is larger than the height of sand and shale laminas then the measured clay content will be the average clay content of sand and shale laminas. Because we do not have any information about clay content in each lamina, we assumed that both of them are equal to the measured clay content. This assumption results in errors in the prediction of the petrophysical forward function, which might be reduced using core observation data.

Although, in general the result of a multi facies petrophysical forward function is better than a single facies function, in three intervals we cannot see any obvious improvements: 2875 m- 2900 m, 3150 m- 3200 m, and 3340 m-3360 m. Error on these intervals may be related to improper selection of the form of the petrophysical forward function in equations (2-1) to (2-16). Core observations in addition to petrophysical interpretation tools such as rock physics templates (Avseth et al., 2001; Odegaard and Avseth, 2004; Avseth et al., 2005) can be used to distinguish these different facies and might reduce the uncertainty of the petrophysical forward function further.

The main assumption in the methodology above is that in equation (2-27) the petrophysical forward function for each facies is represented with a single optimum vector of unknowns $\hat{\mathbf{h}}_{\Im_i}$. This assumption is acceptable if $\mathbf{h}_{\Im i}$ has a Gaussian PDF with small variance. In principle, such an assumption does not impose major limitations on the methodology, because any facies with non-Gaussian (e.g., multi-modal PDF) uncertainty in the vector of unknown parameters $\mathbf{h}_{\Im i}$ can be divided into several sub-facies with Gaussian uncertainty in $\mathbf{h}_{\Im i}$ based on geological and petrophysical prior information. However, in cases with weak prior geological information distinction between different sub-facies might be impossible and this assumption can result in large uncertainties and biases in the predictions of the petrophysical forward function. In our dataset, interval 3310 m-3360 m, which is a succession of shale and coal facies, is such an example.

It is worth to indicate that the above methodology accounts for two separate sets of uncertainty. First, the uncertainty in the facies that is represented by $p(\Im_i|\mathbf{m})$ in equation (2-27). Second, the uncertainty in the petrophysical forward function that is represented by $p(\mathbf{d}|\mathbf{m}, \hat{\mathbf{h}}_{\Im i})$ in equation (2-27). In the derivation of $p(\mathbf{d}|\mathbf{m}, \hat{\mathbf{h}}_{\Im i})$ we assume it can be represented by a Gaussian PDF, which implies that the error in the petrophysical forward function is non-systematic. We explained that in order to apply the non-systematic assumption, strong geological a priori information about orientation of earth layers and fluid distribution within rock are required. Such a priori information is usually not available and therefore the set of petrophysical forward relations used above are expected to predict underestimate values of the elastic moduli, as they can be interpreted as Reuss average of rock constituents.

The method presented above can be used in conjunction with any petrophysical forward function for different types of rock facies. Such petrophysical forward functions can be constructed based on rock physics theories (Marion et al., 1992; Dvorkin and Nur, 1996; Dvorkin and Gutierrez, 2001; Avseth et al., 2005, ch. 2) or from empirical correlations (Shapiro, 2003). In particular, the integration of state-of-the art shale models such as those proposed by Hornby et al. (1994) and Jakobsen et al. (2003) can result in a better estimate of shale properties, where the applied shale model performed poorly such as 2875 m-2960 m, 3100 m-3200 m, and 3375 m-3400 m. Also further information about properties of grains can improve the accuracy of the petrophysical forward function. In our application, we have very limited information about the properties of grains and as a result in the calibration process we estimate all of the unknown parameters by including only weak a priori information.

We present the result of the application of the method to data from one well. The petrophysical forward function can be updated with data from a new well in the same field or geological setting. In the first step, the same analysis as above must be used to identify different facies and construct the petrophysical forward function of each facies at the new well, and then the probability of each facies conditioned on the value of porosity and clay content must be reassigned by using additional information from the new well. The uncertainty of the updated petrophysical forward function may be higher than the previous function, however we have shown that this is certainly not always the case, and importantly the updated model will be more general in that it can be used for the different rock types of both wells.

The accuracy and applicability of the petrophysical forward function is critical in the petrophysical inversion of seismic data. Biases in the petrophysical forward function (Chao et al., 2009) can result in large uncertainty and biases in the petrophysical inversion results (Shahraeeni et al., 2010; Shahraeeni and Curtis, 2010). The method presented in this chapter can reduce the uncertainty and bias of the petrophysical forward function and therefore may improve results of the petrophysical inversion of seismic data.

2.6 Conclusions

We present a method to estimate the uncertainty in a petrophysical forward function. The results of the application of the method with a single-facies assumption show that a Gaussian assumption about the uncertainty of the petrophysical forward function can result in biases in its predictions. The second application shows that the multi-facies assumption can remove the biases from the predictions of the petrophysical forward function. The predicted PDF of the multi-facies petrophysical forward function may be multi-modal, and the uncertainty represented by the PDF may actually be smaller than that using a single-facies assumption. The comparison between the result of the single-facies and multi-facies petrophysical forward function shows that we need to consider the uncertainty of the forward rock physics model due to heterogeneity of rock properties even in only one well. A proper petrophysical forward function is an important a priori requirement in the petrophysical inversion, where seismic attributes are inverted for porosity, clay content, and water saturation. The method described above can be applied with data from several wells to construct a general petrophysical forward function that describes the elastic behavior of different rock facies in a geological setting, which can be used for the petrophysical inversion (Shahraeeni et al., 2010; Shahraeeni and Curtis, 2010).

2.7 References

Avseth, P., T. Mukerji, A. Jorstad, G. Mavko and T. Veggeland, 2001, Seismic Reservoir Mapping from 3-D AVO in a North Sea Turbidite System, *Geophysics*, **66**, 1157-1176.

Avseth, P., T. Mukerji and G. Mavko, 2005, *Quantitative seismic interpretation: applying rock physics tools to reduce interpretation risk*. Cambridge ; New York: Cambridge University Press.

Bachrach, R., 2006, Joint estimation of porosity and saturation using stochastic rock-physics modeling, *Geophysics*, **71**, O53-O63.

Bachrach, R. and P. Avseth, 2008, Rock physics modeling of unconsolidated sands: Accounting for nonuniform contacts and heterogeneous stress fields in the effective media approximation with applications to hydrocarbon exploration, *Geophysics*, **73**(6), E197-E209.

Batzle, M. and Z. Wang, 1992, Seismic Properties of Pore Fluids, *Geophysics*, **57**, 1396-1408.

Bishop, C. M., 1995, *Neural Networks for Pattern Recognition*: Oxford University Press.

Bosch, M., C. Carvajal, J. Rodrigues, A. Torres, M. Aldana and J. Sierra, 2009, Petrophysical seismic inversion conditioned to well-log data: Methods and application to a gas reservoir, *Geophysics*, **74**(2), O1-O15.

Chao, G., G. Lambert and H. Cumming, 2009, Analysis of Intrinsic Uncertainties of Petro-elastic Models Using Simulated Annealing, 71st EAGE Conference & Exhibition, Amsterdam.

Cover, T. and P. Hart, 1967, Nearest neighbor pattern classification, *Information Theory, IEEE Transactions on*, **13**(1), 21-27.

Dvorkin, J. and M. A. Gutierrez, 2001, Grain sorting, porosity, and elasticity, from pangea.stanford.edu/~jack/Bimodal.pdf.

Dvorkin, J. and A. Nur, 1996, Elasticity of high-porosity sandstones: Theory for two North Sea data sets, *Geophysics*, **61**(5), 1363-1370.

Eidsvik, J., P. Avseth, H. Omre, T. Mukerji and G. Mavko, 2004, Stochastic Reservoir Characterization using Prestack Seismic Data, *Geophysics*, **69**, 780-993.

Eshelby, J. D., 1957, The Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, **241**(1226), 376-396.

Gassmann, F., 1951, Uber die elastizitat poroser medien, *Vier. Natur Gesellschaft*, **96**, 1-23.

Goldberg, I. and B. Gurevich, 1998, A Semi-empirical Velocity-Porosity-Clay Model for Petrophysical Interpretation of P- and S-velocities, *Geophysical Prospecting*, **46**, 271-285.

Gonzalez, E. F., T. Mukerji and G. Mavko, 2008, Seismic inversion combining rock physics and multiple-point geostatistics, *Geophysics*, **73**(1), R11-R21.

Grana, D. and E. Della Rossa, 2010, Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion, *Geophysics*, **75**(3), O21-O37.

Gutierrez, M. A., 2001, Rock physics and 3-D seismic characterization of reservoir heterogeneities to improve recovery efficiency. Ph.D. Thesis, Stanford University.

Han, D.-h., A. Nur and D. Morgan, 1986, Effects of porosity and clay content on wave velocities in sandstones, *Geophysics*, **51**(11), 2093-2107.

Hashin, Z. and S. Shtrikman, 1963, A variational approach to the theory of the elastic behaviour of multiphase materials, *Journal of the Mechanics and Physics of Solids*, **11**(2), 127-140.

Hill, R., 1952, The Elastic Behaviour of a Crystalline Aggregate, *Proceedings of the Physical Society. Section A*, **65**(5), 349.

Hornby, B. E., L. M. Schwartz and J. A. Hudson, 1994, Anisotropic effective-medium modeling of the elastic properties of shales, *Geophysics*, **59**(10), 1570-1583.

Jakobsen, M., J. A. Hudson and T. A. Johansen, 2003, T-matrix approach to shale acoustics, *Geophysical Journal International*, **154**(2), 533-558.

Keys, R. G. and S. Xu, 2002, An approximation for the Xu-White velocity model, *Geophysics*, **67**(5), 1406-1414.

Kuster, G. T. and M. N. Toksoz, 1974, Velocity and attenuation of seismic waves in two-phase media: part I. theoretical formulation, *Geophysics*, **39**(5), 587-606.

Lagarias, J. C., J. A. Reeds, M. H. Wright and P. E. Wright, 1998, Convergence Properties of the Nelder--Mead Simplex Method in Low Dimensions, *SIAM J. on Optimization*, **9**(1), 112-147.

Malinverno, A. and R. L. Parker, 2006, Two ways to quantify uncertainty in geophysical inverse problems, *Geophysics*, **71**(3), W15-W27.

Marion, D., A. Nur, H. Yin and D. Han, 1992, Compressional Velocity and Porosity in Sand-clay Mixture, *Geophysics*, **57**, 554-563.

Mavko, G., T. Mukerji and J. Dvorkin, 2009, *The rock physics handbook: tools for seismic analysis of porous media*. Cambridge ; New York: Cambridge University Press.

Mindlin, R. D., 1949, Compliance of bodies in contact, *Journal of Applied Mechanics*, **16**, 259-268.

Mukerji, T., A. Jorstad, P. Avseth, G. Mavko and J. R. Granli, 2001, Mapping lithofacies and pore-fluid probabilities in a North Sea reservoir: Seismic inversions and statistical rock physics, *Geophysics*, **66**, 988-1001.

Odegaard, E. and P. Avseth, 2004, Well log and seismic data analysis using rock physics templates, *First Break*, **22**, 37-43.

Raymer, L. L., E. R. Hunt and J. S. Gardner, 1980, An improved sonic transit time-to-porosity transform, SPWLA 21st Annual Logging Symposium, Paper P.

Shahraeeni, M., A. Curtis and G. Chao, 2010, Fast probabilistic petrophysical mapping of reservoirs from 3D seismic data, *Geophysics*, **Submitted**.

Shahraeeni, M. S. and A. Curtis, 2010, Fast probabilistic nonlinear petrophysical inversion, *Geophysics*, **submitted**.

Shapiro, S. A., 2003, Elastic piezosensitivity of porous and fractured rocks, *Geophysics*, **68**(2), 482-486.

Spikes, K., T. Mukerji, J. Dvorkin and G. Mavko, 2007, Probabilistic seismic inversion based on rock-physics models, *Geophysics*, **72**(5), R87-R97.

Wyllie, M. R. J., A. R. Gregory and L. W. Gardner, 1956, Elastic wave velocities in heterogeneous and porous media, *Geophysics*, **21**(1), 41-70.

3 Petrophysical inversion of well logs

> *"I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind."*
>
> William T. Kelvin

3.1 Abstract

In this chapter we propose a computationally efficient probabilistic method to solve nonlinear inverse problems. In this method, any post-inversion (posteriori) joint probability density function (PDF) over the model parameters is represented by a weighted sum of multivariate Gaussian PDF's. A so-called mixture density neural network estimates the weights, mean vector, and covariance matrix of the Gaussians given any measured data set. In the first application, we jointly invert compressional and shear wave velocity for the joint PDF of porosity, clay content, and water saturation in a synthetic, fluid-saturated dispersed sand-shale system. Results show that if the method is applied appropriately the joint PDF estimated by the neural network is comparable to the Monte Carlo (MC) sampled a posteriori solution of the inverse problem. However, the computational cost of training and using the neural network is much lower than inversion by MC sampling (more than a factor of 10^4 in this case and potentially a much larger factor for 3D seismic inversion). To analyze the performance of the method on real exploration geophysical data, in the second application, we jointly invert P-wave impedance and Poisson's ratio logs for the joint PDF of porosity and clay content, without using any petrophysical forward function. Results at a 'blind' well show that the posterior model PDF of porosity and clay content provides a good estimate of actual porosity and clay content log values. In the third application we apply the method in conjunction with a petrophysical forward function calibrated with well logs from Akpo field

to invert P- and S-wave impedance logs for the joint PDF of porosity, clay content, and water saturation. Results show that the posterior PDF of the model parameters provides reasonable estimates of measured well logs. Errors in the posterior PDF are mainly due to errors in the petrophysical forward function. In the last application of the method we invert Backus averaged P- and S-wave impedance to obtain upscaled values of porosity, clay content, and water saturation. Results show that the upscaled values of rock and fluid properties can be biased due to errors in the petrophysical forward function. Although the results may vary from one application to another, the fast, probabilistic method of solving non-linear inverse problems developed in this chapter can be applied to invert well logs and large seismic data sets for petrophysical parameters in any application.

3.2 Introduction

Prediction of rock and fluid properties (e.g., porosity, clay content, and water saturation) from seismic data (e.g., P- and S-wave velocity) is referred to as petrophysical inversion. Petrophysical inversion of well logs is the first step in the quantitative reservoir characterization from seismic data. Results of the petrophysical inversion of well logs can be used for evaluation and quality control of the petrophysical inversion of seismic data at well locations. The main challenges in inferring properties from acoustic well logs are uncertainty in the measured values of logs (i.e., the data vector), and uncertainty in petrophysical forward relations. Therefore, any approach to the petrophysical inverse problem must address the effect of the above two sources of uncertainty on results.

The petrophysical forward relations between rock and fluid properties and seismic properties for a mixture of sand and clay can be significantly nonlinear. Rock physics theory predicts that infinitely many different values of clay content can give rise to the same velocities. A specific model proposed by Marion et al. (1992) predicts that in a dispersed mixture of sand and clay, two different values of clay content can result in the same values of P- and S-wave velocity, even when all other properties of the mixture are held constant (Appendix B, Dvorkin and Gutierrez, 2001). This means that in the petrophysical inversion, for a given pair of P- and S-wave velocity, two possible values of clay content (and porosity) should be estimated. The Monte Carlo sampling method (Mosegaard and Tarantola, 2002; Sambridge and Mosegaard, 2002) is usually used to solve this type of strongly nonlinear inverse problems. However, due to the high dimensionality of the model vector, which includes porosity, clay content, water saturation, depth (or effective pressure), bulk modulus and density of saturating fluids, and possibly bulk and shear moduli of grains, the computational cost of the Monte Carlo sampling solution of the petrophysical inverse problem can be significantly high.

In practice several simplifying assumptions have been made in the past to overcome the computational difficulty of the petrophysical inversion. Two main assumptions, which are commonly used implicitly or explicitly are: (1) applying prior knowledge to select reservoir facies and therefore decrease the nonlinearity of problem, and (2) assuming known values for effective pressure and the type of fluid and grains to decrease the dimensionality of the model space. Bosch (2004) and Bosch et al. (2007; 2009) inverted short-offset seismic data for acoustic impedance, porosity, and water saturation. They assumed that Wyllie's time-average equation (Wyllie et al., 1956; Mavko et al., 2009, p. 350) can be used to describe the relationship between total porosity and acoustic impedance. Wyllie's equation is a weakly nonlinear model in comparison to the dispersed sand-shale mixture model (Dvorkin and Gutierrez, 2001). This assumption also implicitly decreases the dimensionality of the model space to at most three parameters (P-wave impedance, porosity, and water saturation).

Bachrach (2006) inverted P- and S-wave impedance for porosity and water saturation. In that paper a second order polynomial forward function was applied to describe the relationship between bulk and shear moduli and porosity. The assumption about the polynomial type of the petrophysical relationship decreases the dimensionality of the model space to two parameters (porosity and water saturation), and in the same way as described above decreases the nonlinearity of the forward petrophysical function. In addition to that, a lithology indicator map was used to select reservoir facies before petrophysical inversion. The application of a priori knowledge about lithology also decreases the nonlinearity of the petrophysical forward function as this assumption confines the inversion to reservoir facies, for which the function between porosity and P- and S-wave impedance is injective.

Spikes et al. (2007) inverted two constant-angle stacks, for porosity, clay content, and water saturation. In that paper, they used the stiff-sand model (Dvorkin and Nur, 1996; Mavko et al., 2009, p. 260) to describe the relationship between porosity and clay content, and P- and S-wave impedance. Effective pressure, fluid and grain properties were assumed to be constant and as a result the dimensionality of the model space decreased to three (i.e., porosity, clay content, water saturation). A lithology indicator was also used to select reservoir facies and petrophysical inversion was performed only on those facies. Therefore, application of the a priori information about the type of facies decreased the nonlinearity of this problem.

Neural networks have also been used to solve nonlinear geophysical inverse problems with one-dimensional model spaces in a computationally efficient manner. Devilee et al. (1999) inverted regional surface wave dispersion velocities for crustal thickness across Eurasia, while Meier et al. (2007b) extended this to obtain a global crustal thickness map. Devilee et al. (1999) also showed how the laws of probability could be used to combine the output of multiple neural networks, each solving a 1D inverse problem, to solve a multi-dimensional inverse problem. This was used by Meier et al. (2007a) to invert for a two-parameter (average velocity and Moho depth) global crustal model that could be used, amongst other applications, to make corrections for near surface heterogeneity to remove bias in global deep mantle tomography. Meier et al. (2009) extended the data and methodology to perform petrophysical inversion for global water content and temperature in the Earth's mantle transition zone (approximately 440km – 660km depth) in an inversion that also constrained parameters in the petrophysical forward relations between temperature, water content, and seismic velocities. In all of the preceding applications, full probability density functions (PDF's) of the solution to the nonlinear inverse problems were obtained. Roth and Tarantola (1994) applied neural networks to invert synthetic common shot gathers for seismic velocity models. Saggaf et al. (2003) applied a neural network to estimate porosity values from 3D seismic data; they showed how a regularization method could be used with neural networks to improve their robustness. However, Roth and Tarantola (1994) and Saggaf et al. (2003) applied conventional neural networks to estimate just one value of model parameters as the solution of an inverse problem, and the neural networks which were used in these two applications, did not provide any information about uncertainty of the estimate.

Neural networks can also be used to classify lithofacies successions from borehole well logs. Maiti et al. (2007) and Maiti and Tiwari (2009; 2010) applied neural networks to identify lithofacies boundaries using density, neutron porosity and gamma ray logs of the German Continental Deep Drilling Project (KTB). Maiti et al. (2007) applied the super self-adapting back propagation algorithm to train the neural network, while Maiti and Tiwari (2009; 2010) applied a Hybrid Monte Carlo algorithm for training. Both of the above algorithms resulted in more robust training procedures for neural networks. However, these two papers did not address the problem of inverting data for the joint PDF of a continuous multi-dimensional model vector, as in Devilee et al. (1999) and Meier et al. (2007a,b; 2009). For other background information, Poulton (2002) provides a detailed description of mathematical theory and other geophysical applications of neural networks.

A Mixture Density Network (MDN) is a particular extension of neural networks that maps a deterministic input vector onto a PDF over uncertain output vectors (Bishop, 1994). In the original development of the MDN, it is correctly assumed that any arbitrary PDF can be modeled as a mixture (weighted sum) of Gaussian PDF's, each with an isotropic covariance matrix (i.e., one with equal diagonal elements and zero off-diagonal elements), and this form was used in the papers by Meier et al. (2007a,b; 2009) above. However, when a multi-dimensional model space is considered within a single MDN inversion the isotropic assumption causes practical difficulties, especially when the uncertainty distribution is highly variable for different parameters of the model vector. In order to solve this practical difficulty we develop an extension of MDN theory which models a PDF using a mixture of Gaussians with a covariance matrix with unequal diagonal elements. This development allows us to use the MDN to solve several nonlinear petrophysical inverse problems with multi-dimensional model and data spaces rapidly, robustly, and fully probabilistically.

In this chapter we first explain the structure of an MDN and introduce a method for designing and applying it to solve a given problem. Then we apply the MDN to solve four different petrophysical inverse problems. In each application, we first introduce the inverse problem, and then we describe the specifications of the MDN used to solve that problem. Then follows the result of applying the MDN to solve the problem. Finally, we discuss the results and conclude. In a separate conclusion section, at the end of this chapter, we give a summary of the conclusions of the four applications.

3.3 Method

3.3.1 Mixture Density Networks

A neural network is essentially a flexible function or mapping. By varying the parameters (i.e., weights) within the network, we can change the mapping. Varying the parameters to emulate a specific, desired mapping is called training the network. Networks are usually trained by fitting them to examples of the input and output values of the mapping. The set of examples used is called training data set.

One application of neural networks is therefore to estimate some given mapping from an input vector, $\mathbf{x}$, to a target vector, $\mathbf{t}$. Any uncertainty associated with the target vector in this mapping can be represented by the probability density of $\mathbf{t}$ conditioned on (or given) knowledge of the value of $\mathbf{x}$, written as $p(\mathbf{t}|\mathbf{x})$. The mixture density network (MDN) is a type of neural network that can be trained to emulate an approximation to $p(\mathbf{t}|\mathbf{x})$. Within the MDN, $p(\mathbf{t}|\mathbf{x})$ is represented by a mixture or sum of known probability densities:

$$p(\mathbf{t}\,|\,\mathbf{x}) = \sum_{i=1}^{m} \alpha_i(\mathbf{x})\, \varphi_i(\mathbf{t}\,|\,\mathbf{x}) \tag{3-1}$$

In the above equation, $\varphi_i(\mathbf{t}|\mathbf{x})$ is a known PDF and is called a kernel, m is the number of kernels, and $\alpha_i(\mathbf{x})$ is called the mixing coefficient that defines the weight of each kernel in the mixture (the sum). This representation of the probability density function is called a mixture model.

It can be shown that a mixture of densities with Gaussian kernels can approximate any PDF to any desired accuracy, given a sufficient number of kernels with appropriate parameters (McLachlan and Peel, 2000). Therefore, we assume kernels are Gaussians with diagonal covariance matrix:

$$\varphi_i(\mathbf{t}\,|\,\mathbf{x}) = \frac{1}{\prod_{k=1}^{c}\left(\sqrt{2\pi}\;\;\sigma_{ik}(\mathbf{x})\right)} \exp\left\{ -\frac{1}{2}\sum_{k=1}^{c} \frac{\left(t_k - \mu_{ik}(\mathbf{x})\right)^2}{\sigma_{ik}^2(\mathbf{x})} \right\} \tag{3-2}$$

Where c is the dimensionality of the output vector $\mathbf{t} = (t_1, \ldots, t_c)$, μ_{ik} is the k^{th}-component in the mean vector of the i^{th}-kernel, and σ_{ik} is the k^{th}-diagonal element in the covariance matrix of the i^{th}-kernel. Therefore, the mean and covariance of the i^{th} Gaussian kernel are $\boldsymbol{\mu}_i = (\mu_{i1}, \ldots, \mu_{ic})$ and $\Sigma_i = \mathrm{diag}(\sigma_{i1}, \ldots, \sigma_{ic})$, respectively. We call this kernel a diagonal-Gaussian kernel, and an MDN that uses this kind of kernel a diagonal-MDN. We chose this type of kernel because its covariance matrix has unequal diagonal elements and zero off-diagonal elements. Therefore, while we might be able to approximate multi-dimensional PDF's with a lower number of kernels than if we had used isotropic Gaussians (with equal diagonal elements as used by Meier et al. (2007a,b; 2009)) still the number of the non-zero elements in the covariance matrix of each kernel remains smaller than a kernel with a full covariance matrix so that in what follows its application requires lower computation.

Appropriate values for the parameters of the diagonal-MDN in equations (3-1) and (3-2) can be predicted by using a two-layer neural network, with one layer of hidden units from a set of data-model, $(\mathbf{t}, \mathbf{x})$, examples (i.e., training dataset) (Bishop, 1994). Since to our knowledge the use of a diagonal-MDN has not been published previously, in Appendix A we derive and present the mathematical details required for its implementation.

3.3.2 Design and implementation of the diagonal-MDN

In order to solve a particular problem with a diagonal-MDN we need to specify two parameters of the network: (1) the number of kernels in the mixture density model, and (2) the number of hidden units in the neural network.

The number of the kernels depends on the shape of the PDF that is going to be modeled. The match between the PDF and its mixture density representation improves if we increase the number of the kernels. However, a large number of kernels will result in more computations and longer training time. The appropriate number of kernels is usually selected by a trial and error procedure to give an acceptable mixture density representation of the PDF within a reasonable time.

The number of hidden units is usually determined by checking the improvement in the performance of the network as units are added in a trail and error procedure (Poulton, 2002). A simple network with a small number of hidden units can under-fit data (i.e., can not sufficiently fit the relationships embodied in the training data set) while a complex network with a large number of hidden units can over-fit data (i.e., accurately fits the training data set, but is far less accurate for data not represented within the training data set). Duda et al. (2001, pp. 310-311) give a rule of thumb to select the number of hidden units from the number of training samples by optimizing the generalization behavior of the network. This rule of thumb states that the number of weights (i.e., adaptable parameters) in the network should be less than one-tenth of the number of training samples. We always follow this rule to select the number of hidden units when the number of training samples is limited (e.g., in the second application below). When the forward function is known (e.g., in the first and third applications below) we can produce a large noisy data set that results in a small chance of over-fitting and select the number of hidden units with the trial and error procedure indicated above.

When the number of training samples is limited, in order to further mitigate against over-fitting, the so-called early stopping technique is also used. In the early stopping technique the over-fitting is controlled by monitoring the performance of the neural network on an independent set of pairs of data-model samples (the validation dataset). A description of methods for controlling over-fitting is given in Appendix A.

3.4 Applications

The diagonal-MDN described above is applied to solve four problems. The first problem is a synthetic petrophysical inverse problem with multi-dimensional model and data vectors, which shows that the solution of an inverse problem obtained using the diagonal-MDN provides a good estimate of a solution found by Monte Carlo sampling. It also shows that the MDN can be used to obtain the marginal PDF of a subset of model parameters without computing the joint PDF of all model parameters. The second application is a data-driven application of the MDN, where it is used to estimate the joint PDF of porosity and clay content from P-wave impedance and Poisson's ratio well logs. This application shows that the diagonal-MDN can be used to solve petrophysical inverse problems statistically from a limited number of real data samples, and without using a petrophysical forward function. In the third application the MDN is used to obtain an estimate of the joint PDF of porosity, clay content, and water saturation from P- and S-wave impedance logs at one of the wells in Akpo field, using a calibrated petrophysical forward function. Finally, in the last application the up-scaled P- and S-wave impedance logs are inverted to obtain the up-scaled porosity, clay content, and water saturation logs.

3.4.1 First application: Synthetic problem

In the first application we apply the diagonal-MDN inversion method to invert P- and S-wave velocity for the joint PDF of porosity, clay content, and water saturation using a synthetic data set. A petrophysical forward function, which was developed to describe the elastic response of well-dispersed mixtures of sand and clay, was applied to construct the synthetic data set. Hypothetical data uncertainty was added to the output of the petrophysical forward function and the resulting uncertain synthetic data was inverted for the joint PDF of the parameters of the model vector.

Petrophysical forward function

The petrophysical forward function in the synthetic problem is a model for a well-dispersed sand-clay mixture (Dvorkin and Gutierrez, 2001). In this model, the geometry of the sand-shale mixture is divided into two classes depending on the clay volume in the mixture. In sands and shaley sands, where clay content V_{cl} is smaller than the initial porosity of the sand matrix ϕ_s, clay minerals fill the sand pore space without disturbing the sand matrix. In shales and sandy shales, where clay content V_{cl} is larger than the initial porosity of the sand matrix ϕ_s, sand grains are suspended in the shaley matrix. Therefore, sand grains are load-bearing

Parameter	Range
c	0.0-1.0
z (m)	500-3000
ρ_s (g/cc)	2.60-2.70
K_s (GPa)	35-45
G_s (GPa)	15-50
ρ_c (g/cc)	2.50-2.60
K_c (GPa)	20-30
G_c (GPa)	3-15
s_w	0.0-1.0

Table 3-1: A priori intervals of the independent model parameters in the first application. Parameters are Uniformly distributed over the specified ranges. The upper and lower bounds are obtained from Avseth et al. (2005) and Mavko et al. (2009).

materials in sands and shaley sands, while sand and shale components are both load-bearing in sandy shales and shales. The compressional and shear wave velocity models are derived based on the above distinction between sands and shaley sands, and shales and sandy shales. The forward model is presented in Appendix B.

The compressional- and shear-wave velocity, V_p and V_s, are the parameters of the data vector in the synthetic problem, $\mathbf{x} = (V_p, V_s)$. In the Dvorkin and Gutierrez model (Appendix B), these two parameters are functions of porosity ϕ, clay content V_{cl} (i.e., volume of clay in a unit volume of rock), effective pressure p_e, depth z, density of sand particles ρ_s, bulk and shear moduli of sand particles K_s and G_s, density of clay particles ρ_c, and bulk and shear moduli of clay particles K_c and G_c. The effect of fluid on the compressional- and shear-wave velocity is modeled by the Gassmann's equation. We assume that pore space of the synthetic rock is saturated with a two-phase fluid with brine and oil components. Therefore, compressional- and shear-wave velocity are functions of bulk modulus and density of brine, K_w and ρ_w, bulk modulus and density of oil, K_{hc} and ρ_{hc}, and water saturation, s_w. In this inversion we wanted to obtain information about porosity, clay content, and water saturation, i.e. $\mathbf{t} = (\phi, V_{cl}, s_w)$; all other parameters are assumed to be so-called confounding parameters, $\mathbf{m}_{\text{conf}} = (K_s, G_s, \rho_s, K_c, G_c, \rho_c, z, p_e, K_w, \rho_w, K_{hc}, \rho_{hc})$. Variations in the confounding parameters result in additional uncertainty in estimates of the desired model parameters $\mathbf{t} = (\phi, V_{cl}, s_w)$.

A priori information about model parameters and data uncertainty

The a priori (before the inversion) intervals for independent model parameters are given in table 3-1. We assumed that model parameters were Uniformly distributed over the ranges given in that table. Effective pressure is a function of depth. The bulk modulus and density of any type of hydrocarbon (with a given value of density at standard conditions) are empirical functions of pore pressure and temperature (Batzle and Wang, 1992). The pore pressure and overburden stress and hence the effective pressure were assumed to be hydrostatic in this synthetic example. Therefore, as we explained in Appendix B bulk modulus and density of oil can be represented as functions of effective pressure. Porosity is also a function of depth and clay content. Therefore, effective pressure, porosity, bulk modulus and density of oil are dependent model parameters and are not represented explicitly in table 3-1. The density and bulk modulus of water were assumed to be constant and independent of effective pressure over the a priori effective pressure range.

We assumed that the measurement error of V_p was around 5%, and for V_s was around 7%. We also assumed that the error of V_p and V_s were uncorrelated. Therefore, the measurement error was modeled by a Gaussian distribution with zero mean. The standard deviation for V_p was 5% of its value, i.e. $\sigma_{V_p} = 0.05\ V_p$, and for V_s was 7% of its value, i.e. $\sigma_{V_s} = 0.07\ V_s$.

The problem of estimating $\mathbf{t} = (\phi, V_{cl}, s_w)$ from $\mathbf{x} = (V_p, V_s)$ is a non-unique and non-linear inverse problem. The sources of non-uniqueness (or uncertainty) in the solution are the measurement uncertainty of V_p and V_s, uncertainty in the independent confounding model parameters $\mathbf{m}_{\text{conf}} = (K_s, G_s, \rho_s, K_c, G_c, \rho_c, z)$, and the non-unique relationship between clay content and compressional- and shear-wave velocity (Appendix B). The latter source of uncertainty results in bi-modality of the solution of this inverse problem. This means that for each data vector $\mathbf{x} = (V_p, V_s)$ it is possible to have more than one value of clay content and porosity, the regions around which contain likely values of model parameters, whereas between these regions parameter values are unlikely to be correct given available data. Such inverse problems are difficult to solve without direct sampling methods. The MDN will be trained to solve this non-unique inverse problem.

MDN specifications and training data set

The training data set was constructed by systematic sampling from the a priori intervals of the independent model parameters. For the clay content, water saturation and depth, 13 equally (Uniformly) spaced samples were selected. For the bulk and shear moduli of sand,

and the bulk and shear moduli of shale, 3 equally spaced samples were selected. For density of sand and clay, 2 equally spaced samples were selected. The forward model was calculated for all $13^3 \times 3^4 \times 2^2 = 711{,}828$ samples to obtain corresponding synthetic data.

The MDN will interpolate the relationship between **t** and **x** after training. We selected samples from depth, clay content, and water saturation more densely (i.e., 13 samples from the a priori intervals) than for the other model parameters to reduce the interpolation error of the MDN on the desired model parameters, $\mathbf{t} = (\phi, c, s_w)$. Because the effect of the confounding model parameters (K_s, G_s, ρ_s, K_c, G_c, ρ_c) is integrated out by the MDN, we selected smaller number of samples from these parameters. A denser sample selection of the above parameters would improve the accuracy of the MDN, however, it would also increase the training time significantly.

To model measurement errors, two independent samples of the Gaussian noise, described in the previous section, are added to each synthetic data vector computed above. The total number of noisy training samples in this synthetic data set is therefore $2 \times 711{,}828 = 1{,}423{,}656$ (**x**, **t**) pairs.

The specifications of the diagonal-MDN for solving the petrophysical inverse problem are as follows. The input vector of the MDN is the data vector $\mathbf{x} = (V_p, V_s)$. Its outputs are the parameters of the mixture density model, i.e. α_i, μ_i, and Σ_i (equations (3-1) and (3-2)), of the model vector $\mathbf{t} = (\phi, V_{cl}, s_w)$. The number of kernels in the mixture density model, m (equation (3-1)), is determined by trial and error and is set to 15. The number of hidden units in the single hidden layer is also determined by trial and error and set to 10. In the training process, we observe that adding more kernels or hidden units do not reduces the training error significantly. The total number of weights (i.e., parameters) of the network for the selected number of kernels and hidden units in the MDN is 1185.

The over-fitting of the network in this synthetic application was controlled using the noisy training dataset. Training a neural network with noisy data decreases the chance of over-fitting (Appendix A). In addition to that, due to the large number of training samples compare to the number of weights of the network (1,423,656 to 1185), the probability of over-fitting is small (Appendix A).

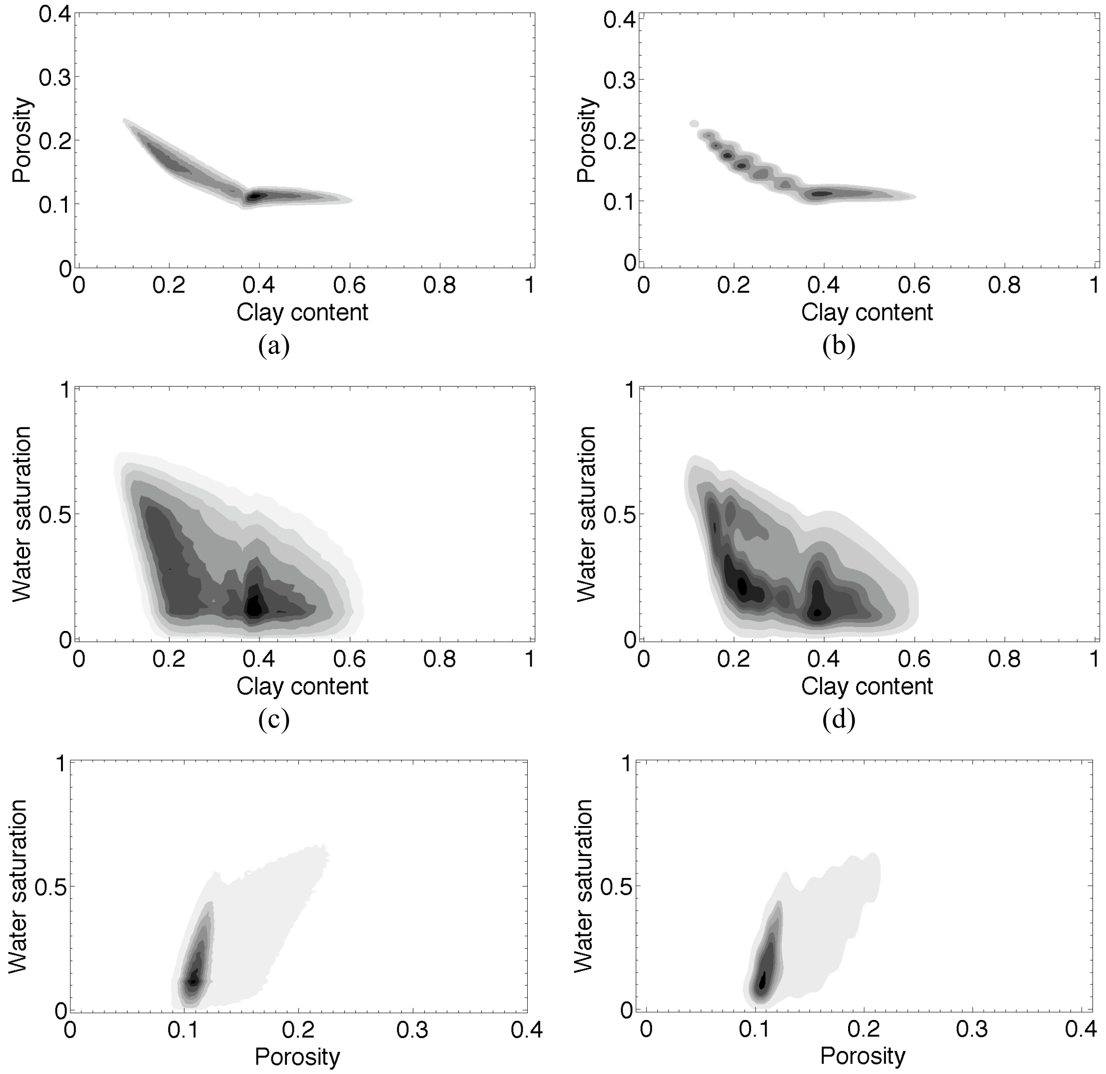
Monte Carlo sampling solution

In order to evaluate the MDN result, we use the Metropolis-Hasting algorithm (Tarantola, 2005) to obtain a comparative solution for each inverse problem. In the Metropolis-Hastings algorithm the likelihood of a given value of the measurement vector $\mathbf{x} = (V_p, V_s)$ is derived from the Gaussian probability density function representing the error. If we assume $i-1$ samples have been taken from the a posteriori solution of the inverse problem, the i-th candidate sample of the solution is constructed as follows: a sample from the a priori uniform distribution of the independent model parameters $\mathbf{m}_i = (V_{cl}, s_w, z, K_s, G_s, \rho_s, K_c, G_c, \rho_c)$ is taken. For this sample the data vector, $\mathbf{d}_i = (V_p, V_s)$, in addition to all dependent model parameters (i.e., porosity, effective pressure, bulk modulus, and density of oil) are calculated using the forward petrophysical function. The likelihood of this sample, L_i, is calculated by using the Gaussian distribution of the measurement error defined above. The sample will be accepted if $L_i/L_{i-1} \geq 1$. If $L_i/L_{i-1} < 1$, the sample will be accepted with probability $p = L_i/L_{i-1}$. In order to obtain independent samples of the a posteriori PDF, only one sample for each 10 samples selected above is accepted as a sample of the a posteriori PDF. The histogram of the selected samples can be used to infer the a posteriori joint probability density of the model parameters.

The marginal probability of the desired model parameters, $\mathbf{t} = (\phi, V_{cl}, s_w)$, is obtained by integration of the joint probability density of the model parameters over all possible values of the confounding model parameters $\mathbf{m}_{\mathrm{conf}} = (K_s, G_s, \rho_s, K_c, G_c, \rho_c, z)$.

Inversion results

Figure 3-1 shows the joint a posteriori 2D marginal PDF's of the model parameters evaluated at $(V_p, V_s) = (2818\ \mathrm{m/s}, 1675\ \mathrm{m/s})$. Figures 3-1(a), 3-1(c) and 3-1(e) show the result of the Monte Carlo sampling inversion and figures 3-1(b), 3-1(d) and 3-1(f) show the result of the diagonal-MDN inversion. The marginal PDF of porosity, clay content and water saturation, which are obtained from the diagonal-MDN inversion, are compared with the Monte Carlo sampling inversion results in figure 3-2. The comparison between the results of the Monte Carlo sampling solution and the diagonal MDN solution in figure 3-1 and figure 3-2 show that the accuracy of the diagonal-MDN solution is not perfect due to the finite number of kernels used, but may be sufficiently good for many applications, particularly given the distinct computational advantages illustrated below.



(a) (b)

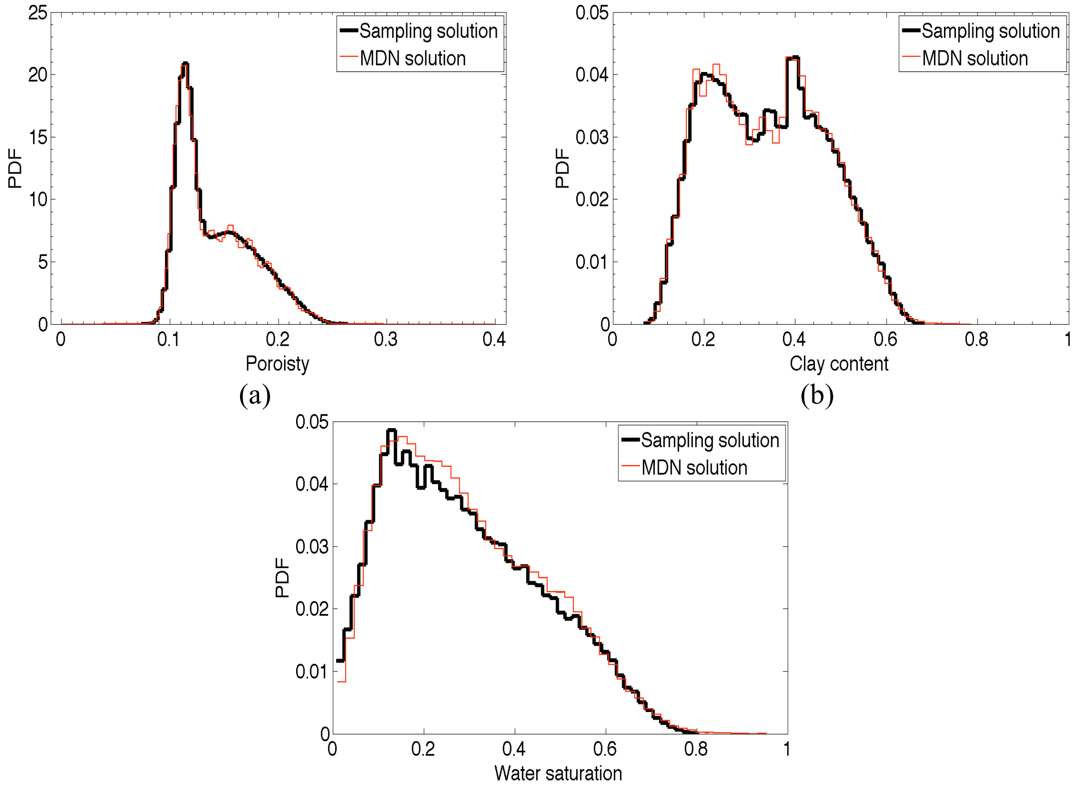
(c) (d)

(e) (f)

Figure 3-1: Inversion result for (V_p, V_s)=(2818 m/s, 1675 m/s). First row is the joint marginal PDF of porosity and clay content: (a) Monte Carlo sampling result, (b) MDN result. Second row is the joint marginal PDF of clay content and water saturation: (c) Monte Carlo sampling result, (d) MDN result. Third row is the joint marginal PDF of porosity and water saturation: (e) Monte Carlo sampling result, and (f) MDN result. Dark colors represent areas with higher probability.

Discussion

The synthetic application above shows that the diagonal-MDN solution of a non-linear inverse problem provides a good estimate of the Monte Carlo sampling solution. The main advantage of the diagonal-MDN is its speed. A single training iteration took 65.37 seconds and the total training time for this network was around 240 hours on a standard personal computer. After training, each fully nonlinear probabilistic inversion took 915 μs. Therefore, if we use this network for inversion, in 48 hours it will provide the full posterior PDF of the



(a) (b)

(c)

Figure 3-2: Inversion result for (V_p, V_s)=(2818 m/s, 1675 m/s). (a) The marginal PDF of porosity. (b) The marginal PDF of clay content. (c) The marginal PDF of water saturation. The black curve is the marginal PDF obtained from the sampling solution and the red curve is the marginal PDF obtained from the MDN solution.

model vector for 188,850,000 data points. The calculation of the Monte Carlo sampling solution by forward modeling of 500,000 samples took around 625 s and as a result, in the same total time as the MDN training and inversion (i.e., 288 hours) the MC sampling method can invert only 1700 data points. Obviously, the relative advantage of the MDN increases with the number of inverse problems to be solved, as the time required for training becomes a smaller proportion of the total time for inversion (e.g. in 1000 hours the MDN solves 2,990,000,000 problems while the grid search method solves only 5,800 problems). This is of great utility when inverting massive data sets point by point (e.g., logs from many wells, or 3D seismic cubes which typically might contain around 10^9 data points).

The second advantage of the diagonal-MDN method is its memory efficiency. Each fully probabilistic MDN inversion result can be stored by the parameters of the mixture density model. The number of such parameters depends on the number of kernels and the dimensionality of the model space, which might typically be of the order of tens or hundreds in the kind of applications discussed here. However, the Monte Carlo sampling solution of one problem typically requires saving thousands of accepted samples to represent the solution. Alternatively, histogram of MC samples could be shown but even this requires thousands of parameters for only 1D intervals in each of the three target parameters above.

The above two advantages are obtained at the expense of the accuracy of the estimated PDF. For some cases, the error in the estimate of the PDF can be large. For example, due to smoothness of Gaussians, the error in the estimation of a truly Uniform PDF (with abrupt variations in the PDF at the boundaries of the uniform interval) with Gaussian kernels can be high. The accuracy of the estimate can be improved if the number of kernels is increased, or kernels with more flexibility (e.g., Gaussians with a full instead of a diagonal covariance matrix) are used. However, both of these possible improvements require estimation of an increased number of kernel parameters, which is more time-consuming. The computational cost of the MDN training increases as the number of weights in the network increases. Therefore, solving inverse problems using neural networks with a large number of weights is computationally demanding unless large computational facilities are available.

Another drawback of using an MDN to solve inverse problems is the trial and error procedure for selecting the appropriate number of hidden units and kernels. This procedure can be slow depending on the number of training samples and size of the network. Therefore, because the performance of only a limited number of networks can be checked, the selected number of the kernels and hidden units might not be optimum.

In summary, this synthetic example shows that the diagonal-MDN can be used to estimate the a posteriori joint marginal PDF of a subset of model parameters in a nonlinear inverse problem highly efficiently if many such inversions must be performed. In the next examples we use the diagonal-MDN with real field data.

3.4.2 Second Application: Inversion of P-wave impedance and Poisson's ratio logs for porosity and clay content

The second application is a completely data-driven application of the diagonal-MDN, in which the MDN learns to invert acoustic logs for the joint PDF of rock properties. Well log data in this application is from Akpo field. The logs of P-wave velocity, S-wave velocity, bulk density, porosity, clay content, and water saturation are available for five wells in this field. The P-wave impedance and Poisson's ratio logs were derived from P-wave velocity, S-wave velocity and bulk density logs. The P-wave impedance, Poisson's ratio, porosity, clay content, and water saturation logs in addition to gamma ray, resistivity, and neutron porosity

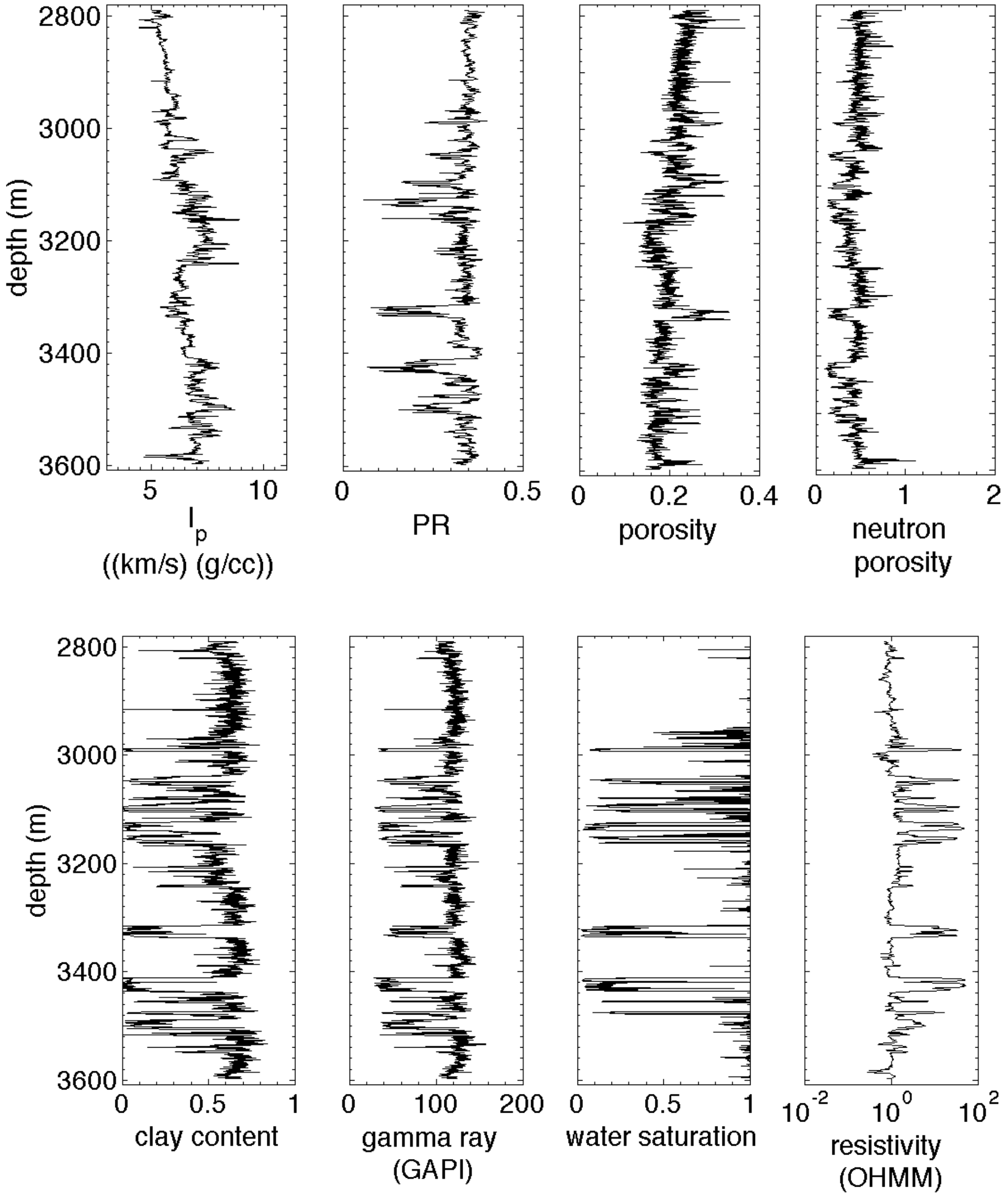
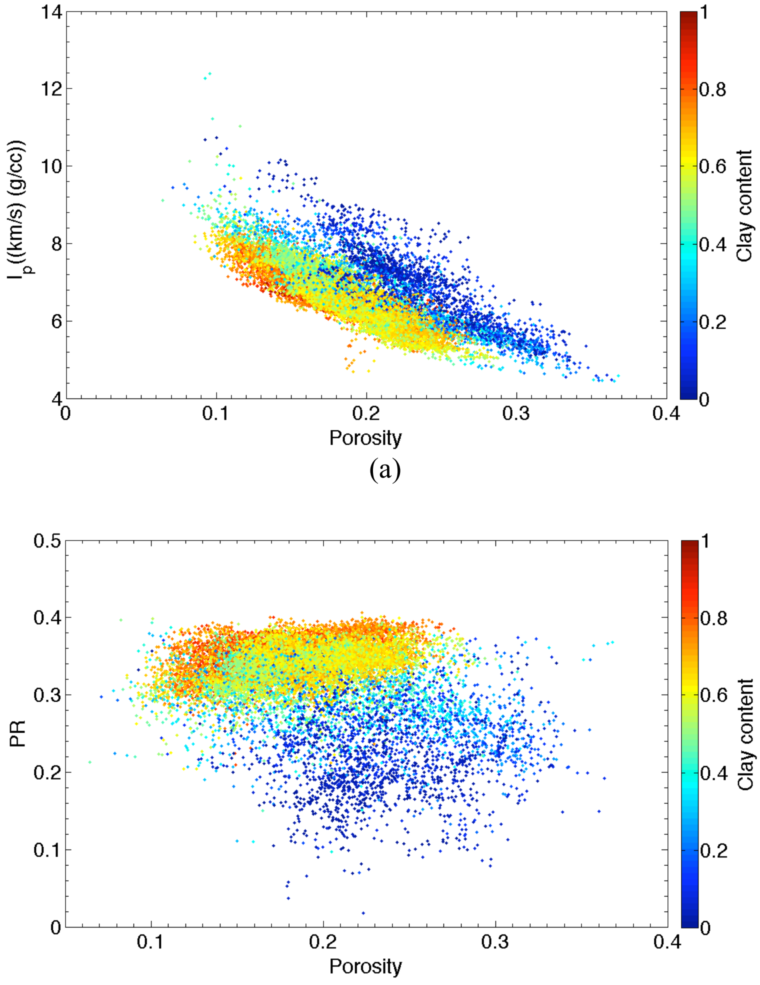


Figure 3-3: Well logs for one of the wells. Log samples of this well are used as training samples. I_p is the P-wave impedance, and *PR* is the Poisson's ratio.

logs for one of the wells are shown in figure 3-3. Conventional methods of well log analysis were used to estimate clay content from Gamma ray, water saturation from resistivity, and total porosity from neutron log (Hearst et al., 2000, pp. 368, 409, 424).

In conventional well logging methods it is assumed that gamma ray is linearly proportional to clay content and by using its reading in front of a 100% shale formation the corresponding value of clay content for other gamma ray readings are estimated. Resistivity log measures the electrical resistivity of a formation. Brine is a conductor of electricity, while hydrocarbon



(a)

(b)

Figure 3-4: Cross plot of log samples. (a) I_p versus Porosity. (b) PR versus Porosity. Samples in both plots are color-coded by the value of clay content.

is not; therefore, in conventional well logging methods the difference between the resistivity of brine and hydrocarbon is used to estimate the brine saturation of any formation. The reading of the neutron porosity log is proportional to the number of hydrogen atoms in a formation. Since hydrogen is found mainly in pore fluids, the neutron porosity log respond principally to porosity. In conventional well logging methods porosity is estimated from the reading of the neutron porosity log with some corrections from water saturation log.

In this example, P-wave impedance, I_p, and Poisson's ratio, PR, were jointly inverted using an MDN to obtain the joint PDF of porosity and clay content.

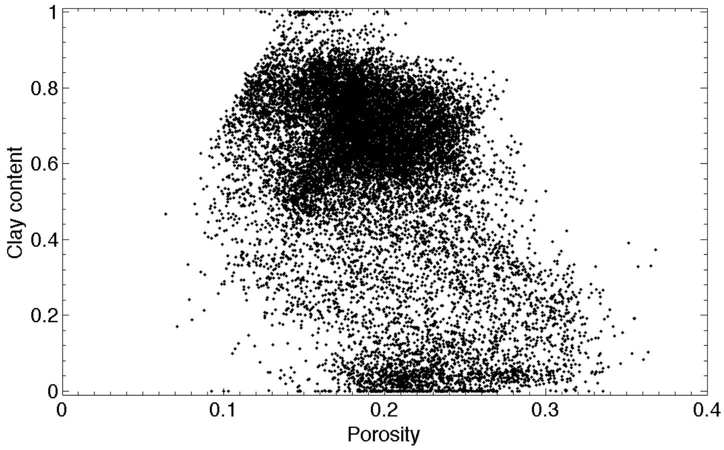


Figure 3-5: Cross plot of clay content versus porosity for well log samples used in the training.

Forward relation between model and data vector, and a priori information about model parameters

The MDN was trained with samples from four wells. The scatter plot of P-wave impedance and porosity, color-coded by clay content for samples of these four wells is shown in figure 3-4(a). The scatter of P-wave impedance for one value of porosity in this figure is due to different sources of uncertainty, including measurement and processing errors of the logs, and variations in other rock physics parameters such as clay content, water saturation, effective pressure, bulk and shear moduli of minerals, bulk modulus of fluid, and different pore-scale matrix geometries. This scatter plot shows that for a given value of porosity, I_p generally increases as the clay content value decreases.

The scatter plot of Poisson's ratio and porosity, color-coded by clay content is shown in figure 3-4(b). In this figure also, the scatter of Poisson's ratio for one value of porosity is due to lack of knowledge about different petrophysical parameters, measurement uncertainty, and processing errors. This scatter plot shows that for a given value of porosity, PR generally increases as the clay content value increases.

Figure 3-5 shows the cross plot of porosity and clay content for well log samples from the four wells used for training. Training samples are selected from these samples to represent a priori combinations of porosity and clay content; 9,885 out of 20,089 samples of the four wells are selected as the training samples.

After training I_p and PR were inverted for the joint marginal PDF of porosity and clay content at a fifth 'blind' well within the same geological context. The MDN learns to derive the joint PDF of porosity and clay content conditioned on I_p and PR, i.e. $p(\phi, V_{cl} \mid I_p, PR)$, from the training samples. All other influential parameters such as water saturation and

Number of kernels	Number of hidden units	Normalized training error	Normalized validation error
3	81	-1.15	-0.96
	54	-1.14	-0.96
	27	-1.12	-0.95
5	51	-1.16	-0.99
	34	-1.17	-1.00
	17	-1.13	-0.97
10	25	-1.19	-1.00
	17	-1.17	-0.99
	8	-1.14	-0.98

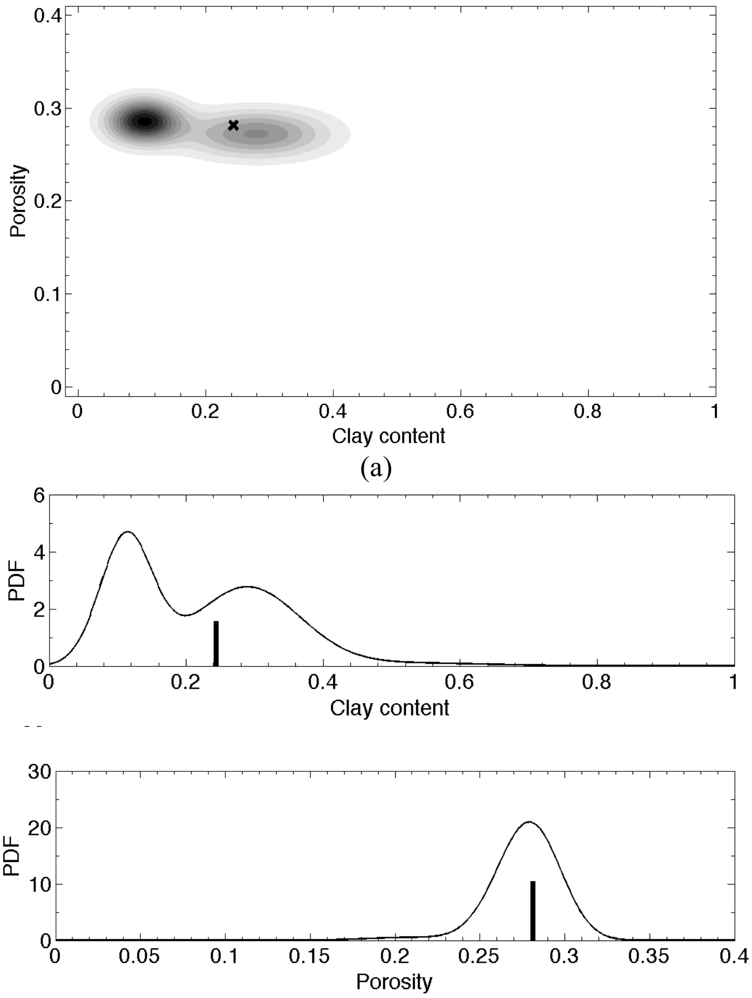
Table 3-2: Specifications of the diagonal MDN's, which are trained to select the optimum number of kernels and hidden units of the neural network in the second application.

effective pressure are considered as confounding parameters, and contribute to uncertainty of porosity and clay content as previously explained in the synthetic application above.

MDN specifications

The input vector of the MDN is P-wave impedance and Poisson's ratio, $\mathbf{x} = (I_p, PR)$, and the outputs are parameters of the mixture density model (equations (3-1) and (3-2)) of the target vector $\mathbf{t} = (\phi, V_{cl})$. The number of kernels was set to five and the number of hidden units was set to 34, using the trial and error procedure described below. Over-fitting is controlled by the early stopping technique (Appendix A). The number of samples in the validation dataset (used in the early stopping technique) is 9,885, all of which are different from the training samples.

In the trial and error procedure for selecting the optimum number of the diagonal-Gaussian kernels and hidden units, we trained 9 networks with different numbers of kernels and hidden units for 35,000 iterations, and select the simplest network that gives the minimum error on the validation dataset. Table 3-2 summarizes the specifications of different MDN's. For each of the selected number of kernels, we used the rule of thumb above to select the number of hidden units. Two other values of the number of hidden units were also selected, one smaller and another larger than the number of hidden units derived above. For example, for five kernels the rule of thumb above results in 34 hidden units; so we also test networks with 17 and 51 hidden units. Table 3-2 shows that the network with 5 kernels and 34 hidden units is the simplest network that gives the smallest error on the validation dataset.



(a)

(b)

Figure 3-6: Example of the posterior PDF of the model parameters in the second application. (a) $p(\phi, V_{cl} \mid I_p, PR)$ evaluated at I_p= 5837 (m/s)(g/cc) and $PR = 0.249$. The black cross represents the measured values from logs. Dark colors represents areas with higher probability. (b) The marginal PDF's of porosity and clay content, vertical black lines represent the measured value of logs.

Inversion results

Figure 3-6 shows the a posteriori joint PDF $p(\phi, V_{cl} \mid I_p, PR)$ evaluated at (I_p, PR) = (5837 (m/s)(g/cc), 0.249), which is a sample from the blind well. For this data point the measured values of clay content and porosity shown on the figure are 0.24 and 0.28, respectively. Figure 3-6 shows that for this sample the estimated marginal PDF of clay content is bi-modal. In this case the a priori information in the training data set were not sufficient to constrain the final estimate of the clay content PDF to a single mode. Nonetheless, the true values of ϕ and V_{cl} lie close to one of the two modes identified by the MDN.

Figure 3-7 shows the a posteriori joint PDF of porosity and clay content for another sample of the blind well with (I_p, PR) = (6146 (m/s)(g/cc), 0.337). The measured values of clay content and porosity for this sample are 0.64 and 0.22, respectively. In this case the a posteriori marginal PDF of porosity and clay content have one maximum probability value, which shows that the a priori information in the training data set have constrained the a posteriori joint PDF of these parameters to a single mode. Notice that the solution is far from isotropic in terms of lateral extent of uncertainties in the model parameters. If we had used the standard form of MDN's with isotropic covariance matrix, this solution would have had to be represented by several 'circular'-looking kernels arranged to form the observed elliptical PDF. This example shows the clear advantage and efficiency of our anisotropic MDN formulation.

Figure 3-8 shows the marginal PDF's of porosity and clay content for each data point in the blind well. The measured logs are also shown in that figure. This figure shows that in general the PDF's of porosity and clay content provide good estimates of the measured logs.

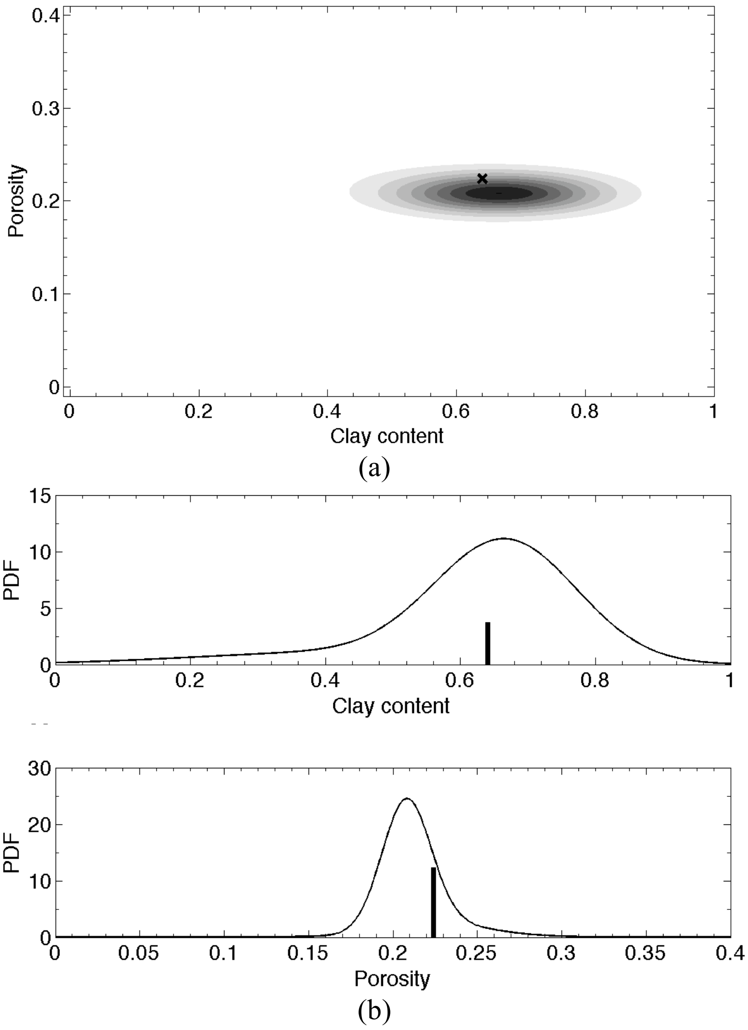
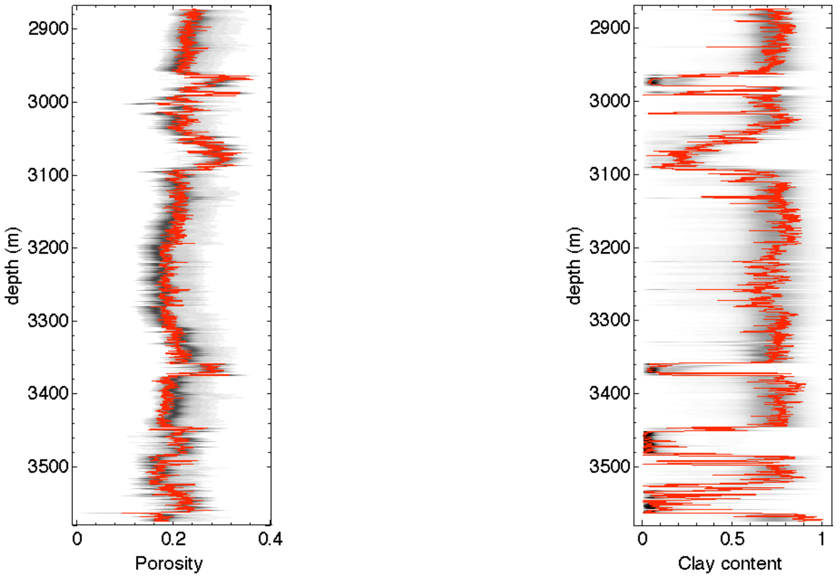


Figure 3-7: Example of the posterior PDF of the model parameters in the second application. (a) $p(\phi, V_{cl} \mid I_p, PR)$ evaluated at I_p= 6146 (m/s)(g/cc) and $PR = 0.337$. The black cross represents the measured values from logs. Dark colors represents areas with higher probability. (b) The marginal PDF's of porosity and clay content, vertical black lines represent the measured value of logs.

A potentially significant source of error in the a posteriori joint PDF of porosity and clay content is water saturation. Gassmann's equation provides the theoretical link between the P-wave impedance and water saturation. Therefore, systematic variations in the water saturation will result in systematic variations of P-wave impedance, which can affect the results of the MDN inversion. Figure 3-9(a) shows the marginal PDF's of porosity and clay content in addition to the water saturation log for the 3030 m-3100 m interval of the blind well. In this interval water saturation varies between zero and one, and the results show that its effect is generally successfully integrated out by the MDN. Figure 3-9 (b) shows the marginal PDF's of porosity and clay content for another interval with variable water saturation. Figure 3-9(a) and figure 3-9(b) show that variations of water saturation do not



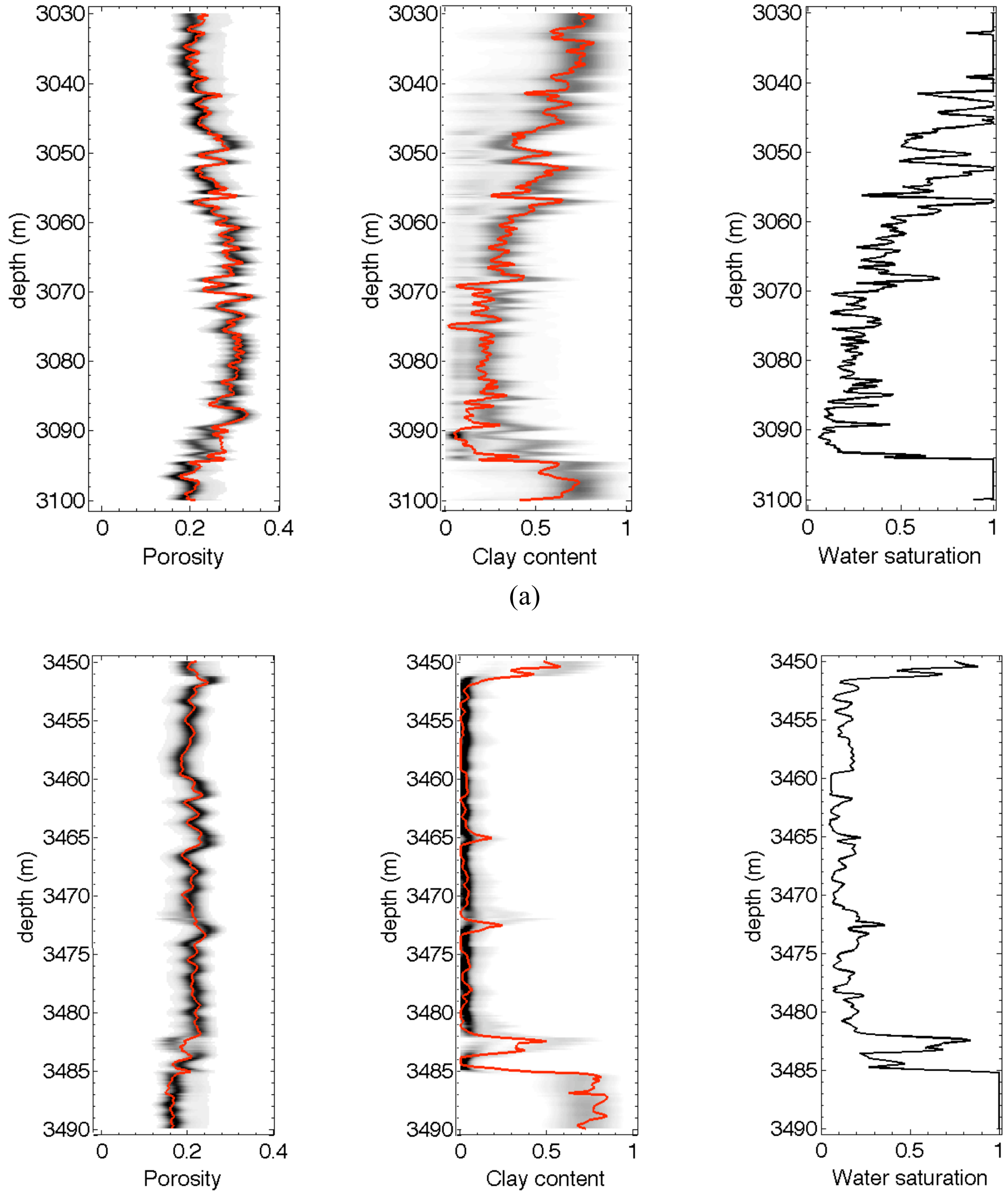
(a) (b)

Figure 3-8: The marginal a posteriori PDF's of porosity and clay content in the blind well. (a) Porosity. (b) Clay content. Darker areas represent higher probabilities. The red line is the measured log.

affect the quality of the MDN estimate of porosity and clay content logs as the effect of water saturation is generally integrated out by the MDN.

Discussion

The above field example shows that when the training samples themselves represent the petrophysical forward function, (i.e., no independent petrophysical forward function is used) the inversion result in a blind well gives a good estimate of the property logs. The estimated PDF of porosity and clay content can be multi-modal (figure 3-6) due to the behavior of the true petrophysical forward function (figure 3-4). The a posteriori uncertainty of the estimated parameters might be decreased if, for example, spatial information about the distribution of rock properties at the wellbore were used. Such information might be included in the training data set by defining model and data vectors for a group of neighboring samples (Caers and Ma, 2002) instead of for each individual sample as used above. However, due to computer memory requirements for defining such model and data vectors, we have not tested this possibility in this thesis.



(a)

(b)

Figure 3-9: The marginal a posteriori PDF's of porosity and clay content for two intervals with variable water saturation. (a) 3030 m-3100 m interval. (b) 3450 m-3490 m interval. The dark color shows the high probability area and the red line is the measured log.

In this example, we invert I_p and *PR* for porosity and clay content. Another parameter of interest in the petrophysical inversion is water saturation. In order to invert for water saturation we need to use a petrophysical forward function. In principle, application of the petrophysical forward function reduces uncertainty in the model parameters by adding theoretical information about the relation between the model and data vectors. In the next application we use a petrophysical forward function to estimate the joint PDF of porosity, clay content, and water saturation from I_p and I_s logs.

3.4.3 Third application: Inversion of P- and S-wave impedance logs for effective porosity, clay content, and water saturation

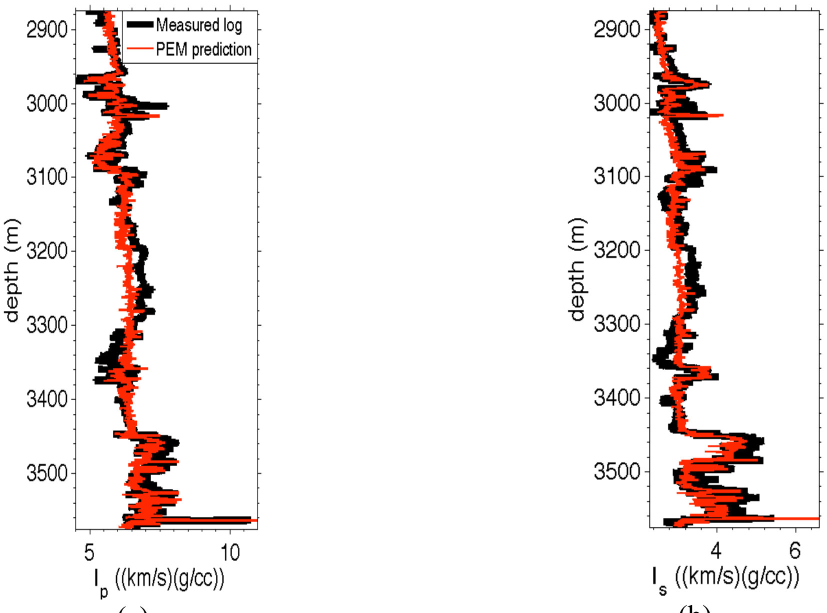
In the third application, we applied the diagonal-MDN inversion method to invert compressional and shear wave impedance logs for the joint PDF of petrophysical parameters at one of the wells in Akpo field. As in the previous application the logs of P-wave velocity, S-wave velocity, bulk density, porosity, clay content, and water saturation are available for five wells in this field. The P- and S-wave impedance logs were derived from the measured P-wave velocity, S-wave velocity and bulk density logs. The calculated P- and S-wave impedance, I_p and I_s, were jointly inverted using a diagonal-MDN to obtain the joint PDF of effective porosity, ϕ_e, clay content, V_{cl}, and water saturation, s_w.

Petrophysical forward function

The petrophysical forward relations are a combination of Gassmann's law to account for fluid substitution, a mixing law to account for the mixed lithology (sand and shale), and empirical depth trend curves to describe pressure effects on the bulk and shear moduli of reservoir rocks. The components of the output vector of this petrophysical relation, **d**, are estimated compressional- and shear-wave impedances, $\mathbf{d} = (I_p, I_s)$. The input vector is $\mathbf{m}' = (\phi_e, V_{cl}, s_w, p_e, K_{hc}, K_w, \rho_{hc}, \rho_w)$ where ϕ_e is effective porosity, V_{cl} is clay content, s_w is water saturation, p_e is effective pressure, K_{hc} and K_w are bulk modulus of hydrocarbon and brine, and ρ_{hc} and ρ_w are density of hydrocarbon and brine, respectively. Effective porosity is the void space outside of porous clay in a unit volume of rock (Dvorkin et al., 2007). A detailed description of the rock physics model is given in Appendix C.

In our inverse problem the data vector $\mathbf{d} = (I_p, I_s)$ is inverted for the marginal joint PDF of porosity, clay content, and water saturation, $\mathbf{m} = (\phi_e, V_{cl}, s_w)$. As in the first application, other input parameters of the petrophysical forward relations are treated as confounding parameters, $\mathbf{m}_{\text{conf}} = (p_e, K_{hc}, K_w, \rho_{hc}, \rho_w)$ – parameters that can increase uncertainty of the model vector **m** in the inversion process. In mathematical terms, the confounding petrophysical parameters are integrated out in the marginal a posteriori PDF of the model vector **m**:

$$\sigma_M(\mathbf{m}|\mathbf{d}) = \int_{\mathbf{m}_{\text{conf}}} \sigma_M(\mathbf{m}'|\mathbf{d})\, d\mathbf{m}_{\text{conf}} \tag{3-3}$$



(a) (b)

Figure 3-10: Comparison of measured logs (black) and the prediction of the petrophysical forward relations (red). (a) P-wave impedance (I_p). (b) S-wave impedance (I_s).

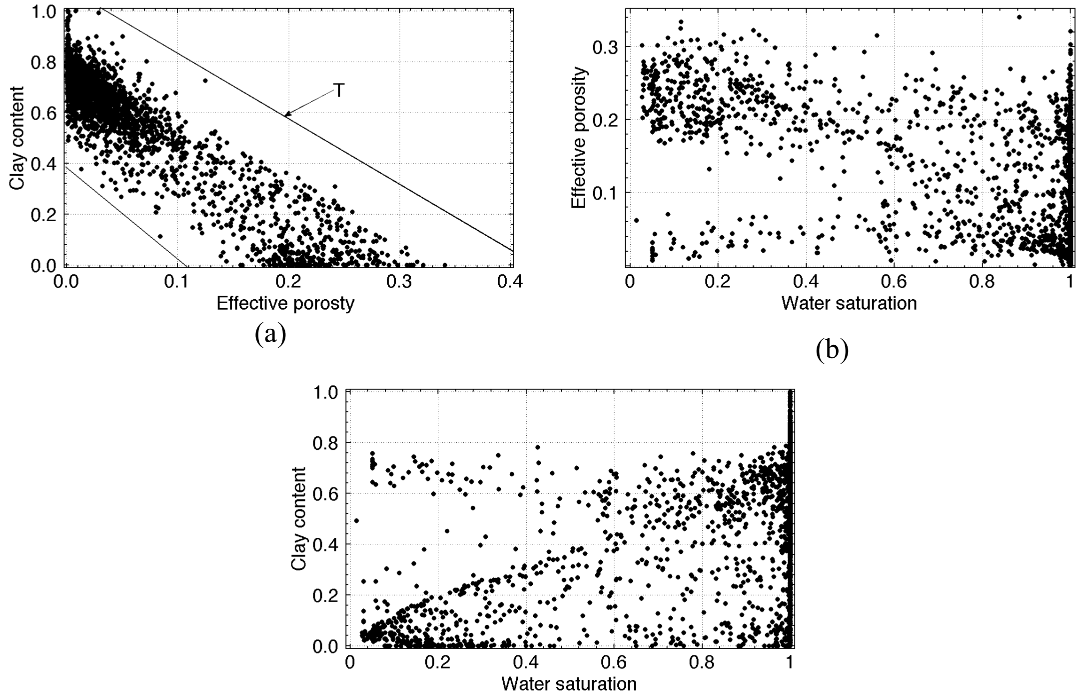
Where, $\mathbf{m}_{\text{conf}}$ and $\mathbf{m}'$ are defined above. The above equation shows that the effect of confounding model parameters is to increase uncertainty in the a posteriori model vector $\mathbf{m} = (\phi_e, V_{cl}, s_w)$.

Data uncertainty

The petrophysical forward relations in Appendix C was calibrated with data from five wells in the field under study to construct the petrophysical forward function used in the inversion process. Figure 3-10 compares the estimated values of I_p and I_s using the calibrated forward relations with the measured values of these parameters at one of the wells. The PDF's of uncertainty of data parameters were obtained from error of the calibrated forward function. Error for predicted values of I_p and I_s are equal to 4% and 6%, respectively (Chao et al., 2009). Therefore, uncertainty of I_p was modeled using a Gaussian with zero mean and standard deviation $\sigma_{Ip} = 0.04\, I_p$. Uncertainty of I_s was also modeled using a Gaussian with zero mean and standard deviation $\sigma_{Is} = 0.06\, I_s$.

A priori PDF of the input parameters of the forward petrophysical function

In the petrophysical inversion, model parameters (e.g., porosity, clay content, and water saturation) are usually assumed to be statistically independent with Uniform a priori PDF's (Bachrach, 2006; Spikes et al., 2007). In this application, the independence assumption is not



(a)

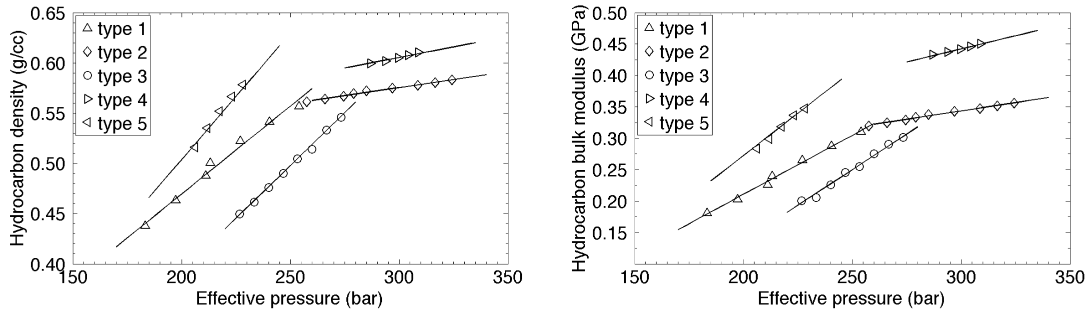
(b)

(c)

Figure 3-11: Cross plots of the parameters of the model vector for well log samples. (a) Clay content versus effective porosity. *T* represents the a priori region of the porosity-clay content plane. (b) Effective porosity versus water saturation. (c) Clay content versus water saturation.

fulfilled, because at least effective porosity and clay content are dependent variables (Appendix C). Figure 3-11(a) shows the cross plot of clay content versus effective porosity for well log samples from five wells. A correlation between these two parameters can be observed in that figure. Based on this observation the a priori PDF for effective porosity and clay content was defined as a Uniform distribution over the region $T(\phi_e, V_{cl})$, which is shown in figure 3-11(a). Figure 3-11(b) and figure 3-11(c) show effective porosity versus water saturation, and clay content versus water saturation for well log samples. These two figures show that while effective porosity and clay content values are to some extent correlated with water saturation values, there are no hard boundaries to intervals in which parameters lie as there are in figure 3-11(a). We therefore assumed that water saturation was independent of the both of effective porosity and clay content and Uniformly distributed over the [0, 1] interval. We thus assumed weak prior information about the three parameters in question.

Effective pressure changes between 173 and 332 bar in the inversion intervals we considered. There is no pervasive correlation between effective pressure and effective porosity, clay content, and water saturation. However, as we explained in the first application, the bulk modulus and density of any type of hydrocarbon (with a given value of density at standard conditions) are empirical functions of pore pressure and temperature. In



(a) (b)

Figure 3-12: Properties of the five different types of hydrocarbons in the wells. (a) Density as a function of effective pressure. (b) Bulk modulus as a function of effective pressure.

this application, the pore pressure and overburden stress are assumed to be hydrostatic and hence the effective pressure is hydrostatic, too. Therefore, the empirical relation between bulk modulus (or density) of fluid and pore pressure is transformed into a relation between bulk modulus (or density) of fluid and effective pressure, using the above assumption. Five different types of hydrocarbon were observed in the well intervals we considered. For each type, the bulk modulus as a function of effective pressure, and density as a function of effective pressure, are given at the well location and are shown in figure 3-12(a) and figure 3-12(b), respectively. The effect of temperature on the hydrocarbon bulk modulus and density was assumed negligible in this study. Brine density and bulk modulus are equal to 1.008 g/cc and 2.625 GPa, respectively, and are constant in all well intervals we considered. Based on the above observations, effective pressure was assumed Uniformly distributed between 173 and 223 bar, and bulk modulus and density were derived from effective pressure using the straight lines in figure 3-12 for each of the five different types of hydrocarbon.

The above a priori probability densities of the input parameters of the petrophysical forward function are selected to be as noninformative (conservative) as possible given the known constraints. However, they limit possible values of the parameters to physically realizable values in the field under examination.

Training data set and MDN specifications

The MDN input vector consists of the P- and S-wave-impedances $\mathbf{d} = (I_p, I_s)$, and its outputs are parameters of the mixture density model of the model vector $\mathbf{m} = (\phi_e, V_{cl}, s_w)$. The output vectors of the training data set were constructed by systematic sampling from the a priori PDF of the model parameters above. 846 equally spaced samples of porosity and clay content were selected from the T-region in $\phi_e - V_{cl}$ plane (figure 3-11(a)). 40 equally spaced

Number of kernels	Number of hidden units	Number of weights	Normalized training error	Normalized validation error
5	52	2011	-1.05	-0.94
	80	3075	-1.08	-0.94
10	27	2041	-1.10	-1.00
	40	2990	-1.14	-1.00

Table 3-3: Variation of the training and validation error due to different choices of the MDN's parameters in the third application.

samples of s_w and 6 equally spaced samples of p_e were selected from the intervals [0,1] and [173 bar , 332 bar], respectively. The values of the bulk modulus and density of each type of hydrocarbon were derived from the selected values of p_e using the straight lines in figure 3-12. The total number of model vectors constructed in this way was 642,960.

The MDN will interpolate the relationship between **d** and **m** after training. In order to reduce the interpolation error of the MDN for the desired model parameters (ϕ_e, V_{cl}, s_w) we selected samples from these parameters densely (i.e., $\Delta\phi_e = 0.009$, $\Delta V_{cl} = 0.0204$, and $\Delta s_w = 0.025$). Because the effect of the confounding model parameters (p_e, K_{hc}, ρ_{hc}, K_w, ρ_w) is integrated out by the MDN, we selected smaller number of samples from these parameters and apply the MDN to interpolate and integrate the effect of other intermediate values. A denser sample selection of the above parameters will improve the accuracy of the MDN, however, it will increase the training time significantly.

For each of 642,960 a priori model vectors selected above, one data vector $\mathbf{d} = (I_p, I_s)$ was calculated by using the forward petrophysical relations. Three samples of Gaussian noise of the forward relations (using data uncertainties defined above) were added to each of these data vectors to represent data uncertainty. The number of noisy model-data pairs in the training dataset was therefore 3×642,960 = 1,928,880.

As in the first application above, the number of required kernels (equation (3-2)), and hidden units of the MDN were selected using a trial and error procedure. Four networks with different number of parameters were trained for 100 iterations. Then error of each network over the well log samples (validation error) was measured and the network with the minimum validation error was selected to continue the training procedure (Table 3-3). According to Table 3-3, the number of kernels was set to be 10 and the number of hidden units was set to be 27. As in the first application, over-fitting of the network to the training dataset was controlled using the noisy training dataset above. The total number of the weights is 2041 and is far smaller than the number of training samples (1,928,880); therefore, the probability of over-fitting reduces significantly (Appendix A).

Results

The MDN with the specifications given above was trained using the training data set. Figure 3-13 shows the a posteriori joint PDF of the model parameters evaluated at a random data point from well logs with $(I_p, I_s) = (5351\ ((\mathrm{m/s})\ (\mathrm{g/cc})), 3166.8\ ((\mathrm{m/s})\ (\mathrm{g/cc})))$. The measured values of the model parameters for this point in the well logs are $(\phi_e, V_{cl}, s_w) = (0.27, 0.14, 0.12)$. Figures 3-13(b), 3-13(d), and 3-13(f) show the results of the MDN inversion, i.e., the marginal 2D joint PDF of porosity and clay content, clay content and water saturation, and porosity and water saturation, respectively. In order to compare these results with the MC sampling inversion results, we applied the MC sampling algorithm introduced in the first application above, to solve the same problem. Figure 3-13(a), figure 3-13 (c), and figure 3-13 (e) show the result of the MC sampling method. Figure 3-13 shows that the result of the MDN inversion method provides a reasonable estimate of the MC sampling method.

Figure 3-14 shows the marginal a posteriori PDF of porosity, clay content, and water saturation logs obtained from the inversion of each data point on the I_p and I_s logs using the diagonal-MDN. The red curves in that figure show the measured values of the model parameters along the same well profile. We observe that in general the inversion results provide reasonable estimates of the measured logs. Small biases can be seen on the estimated porosity, e.g., 3440 m- 3480 m interval. The uncertainty of clay content is larger than effective porosity, and in some intervals, e.g. 3140 m- 3360 m, we observe a bias between the measured values of the well log and the estimated PDF of the model parameters. The uncertainty of the estimated PDF of water saturation in figure 3-14(c) is much larger than porosity and clay content. The probability of all possible values of water saturation in figure 3-14(c) is large (gray area in that figure), however in general the peak of the PDF (darker areas in that figure) is near the measured values of water saturation.

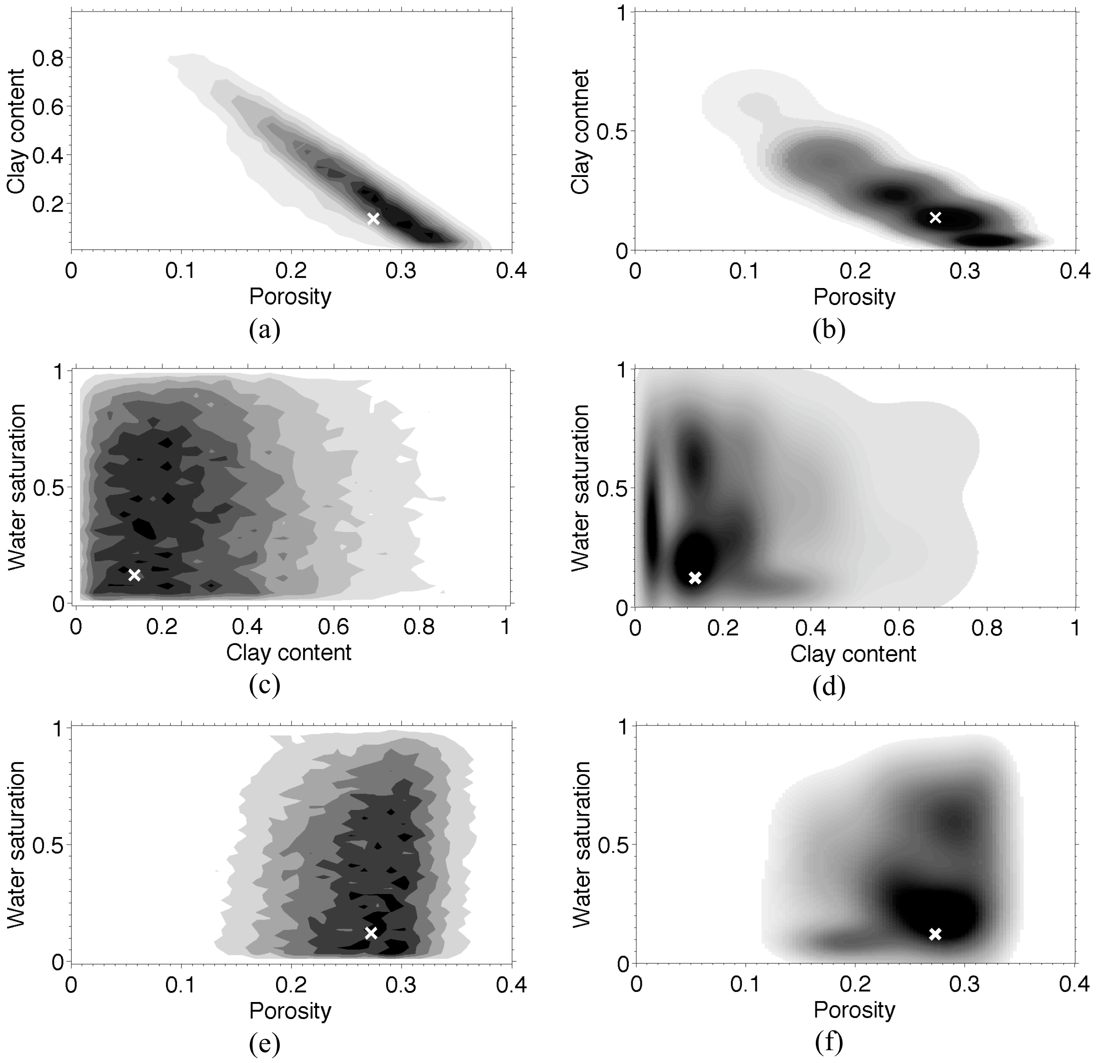


Figure 3-13: Inversion result for $(I_p, I_s) = (5351\ ((\mathrm{m/s})\ (\mathrm{g/cc})), 3166.8\ ((\mathrm{m/s})\ (\mathrm{g/cc})))$ in the third application. Model vector is $(\phi_e, V_{cl}, s_w) = (0.27, 0.14, 0.12)$. First row is the marginal PDF of porosity and clay content: (a) Monte Carlo sampling result, (b) MDN result. Second row is the marginal PDF of clay content and water saturation: (c) Monte Carlo sampling result, (d) MDN result. Third row is the marginal PDF of porosity and water saturation: (e) Monte Carlo sampling result, and (f) MDN result. Dark colors represent areas with higher probability. The white cross shows the measured values of model parameters.

Discussion

Figure 3-13 shows that the MDN solution provides a reasonable estimate of the MC sampling solution and the difference between two PDF's is due to the limited number of kernels in the MDN solution. The difference between the marginal PDF's of the model parameters and measured logs in figure 3-14 is mainly due to errors in the petrophysical forward function. Figure 3-10 compares the prediction of the petrophysical forward function

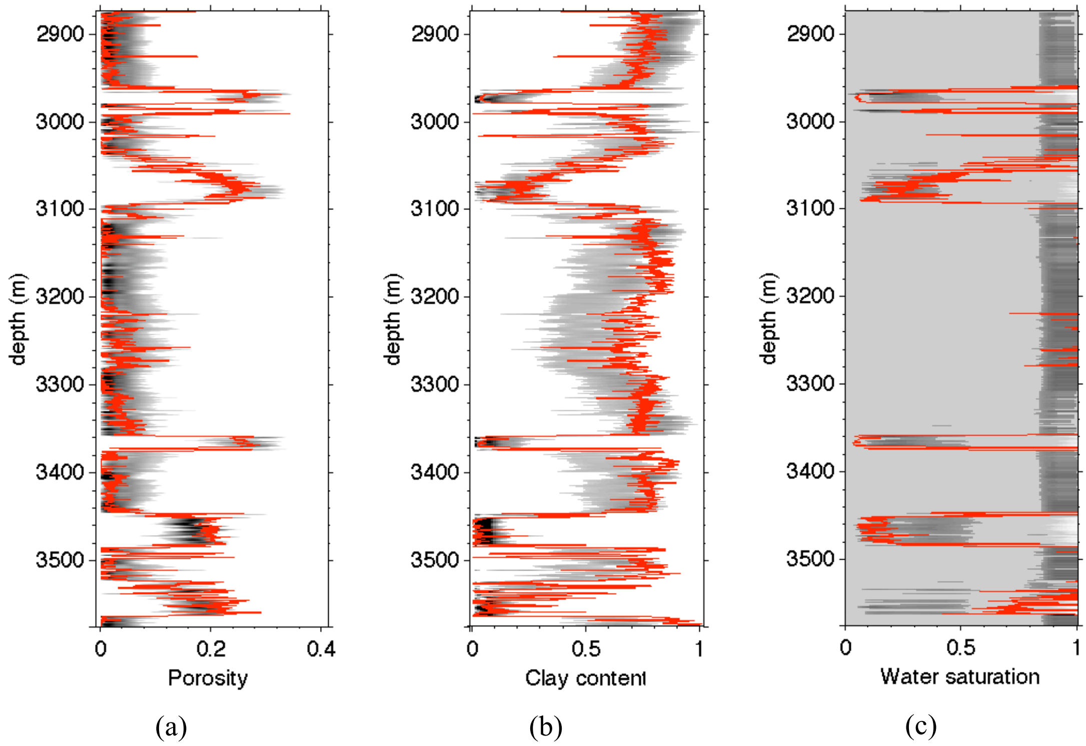


Figure 3-14: Marginal a posteriori PDF of porosity, clay content, and water saturation in the blind well. (a) Porosity. (b) Clay content. (c) Water saturation. Darker shades represent higher probabilities. The red line is the measured log.

with the measured values of I_p and I_s logs. The error in the prediction of the forward function for the intervals 3140 m- 3360 m, and 3380 m- 3450 m are large. Figure 3-14(b) shows that on the same intervals the error of the estimated PDF of clay content is also large. Any improvement in the accuracy of the petrophysical forward function can reduce the effect of this type of error on the inversion result.

Figure 3-14 shows the large uncertainty of a posteriori marginal PDF's of porosity, clay content, and water saturation. High uncertainty of the model parameters stems from uncertainty about the values of the confounding model parameters (i.e., effective pressure, bulk modulus, and density of hydrocarbon) and uncertainty in the petrophysical forward relations (figure 3-10). In particular, due to low sensitivity of the petrophysical forward function predictions to water saturation in the presence of uncertainty of the confounding model parameters, the reduction in uncertainty in the water saturation from the petrophysical inversion is small (figure 3-14(c)). Nevertheless, the inversion process does reduce the uncertainty of all model parameters. In the next chapter, we will discuss the effect of uncertainty of the confounding model parameters on the petrophysical inversion results in more detail.

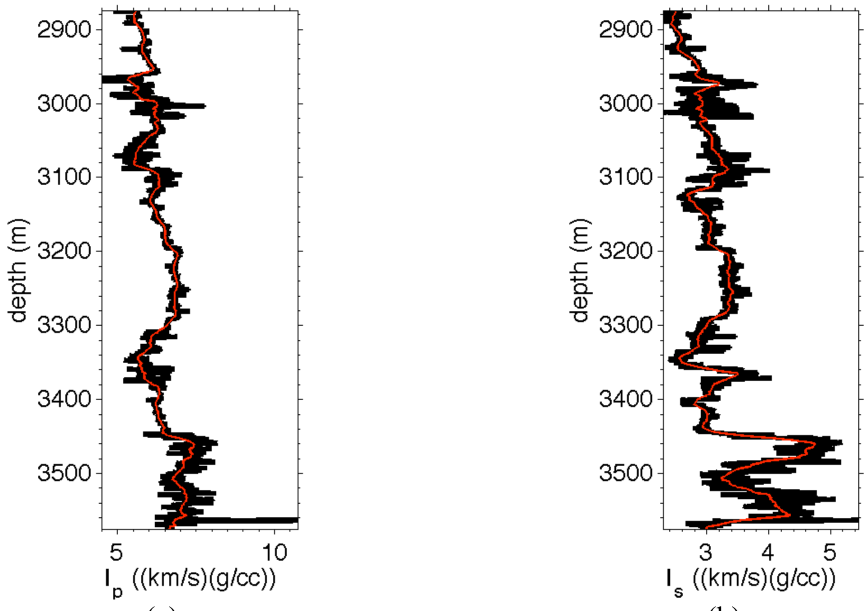
The main advantage of the diagonal-MDN in comparison to the MC sampling solution is its time and memory efficiency as described above in the first application. These advantages are obtained at the expense of the accuracy of the estimated PDF, which is not perfect but sufficiently good for most practical applications (figure 3-13 and figure 3-14). This application shows that the diagonal-MDN can be used in conjunction with a calibrated forward function to estimate the MC solution of the petrophysical inversion of acoustic well logs.

3.4.4 Fourth application: Inversion of upscaled P- and S-wave impedance logs for upscaled effective porosity, clay content, and water saturation

In the next chapter we wish to apply the MDN method to invert seismic data for petrophysical properties. Seismically-defined properties are always averaged over a spatial scale that is large compared to properties from logs. Synthetic change of the scale of well log measurements to a larger scale is called upscaling. In the fourth application, we applied the diagonal-MDN trained in the third application to obtain an estimate of the upscaled effective porosity, clay content, and water saturation, from upscaled values of P- and S-wave impedance.

Upscaled model and data vector

Seismic properties, and in particular P- and S-wave impedance depend on the frequency content of the signal, and on the support volume associated with each measurement. For a given frequency, different numerical methods can be used to change the scale of elastic properties of a medium with a known statistical description. In one limiting case, if the wavelength of the signal is much larger than the typical thickness of layers, the Backus average (Backus, 1962; Mavko et al., 2009, p. 210-215) can be used. In this application, we assumed that the conditions of the Backus average method were fulfilled. We applied this method to transform the high-resolution I_p and I_s measured logs at the well into seismic-scale logs with spatially-averaged properties similar to seismic measurements. The upscaling window used was approximately 1/4 of the seismic wavelength which is around 14 m for seismic waves with frequency of around 50 Hz and average speed of 2800 m/s. Figure 3-15 shows the high-resolution I_p and I_s measured logs and the calculated low-resolution Backus averaged logs.



(a) (b)

Figure 3-15: Comparison of the upscaled well logs and measured logs. (a) I_p, P-wave impedance. (b) I_s, S-wave impedance. Black curves are the measured logs and red curves are upscaled logs.

The value of rock and fluid properties (e.g., porosity, clay content, and water saturation) also depends on the support volume associated with each measurement. In particular, values of rock and fluid properties, which are derived from the inversion of the low-resolution I_p and I_s, can be different from well log measurements. Therefore, in order to obtain low-resolution values of rock and fluid properties at the well location, we applied the MDN inversion method to invert low-resolution (Backus averaged) I_p and I_s logs (figure 3-15), using the petrophysical forward function.

Data uncertainty, A priori model PDF, and MDN specifications

In this inversion, uncertainty of data (i.e., upscaled I_p and I_s logs) must be obtained from the difference between the upscaled logs (figure 3-15) and predictions of the petrophysical forward relations (figure 3-10). The estimated values of the standard deviation of I_p and I_s in this application were similar to those obtained in the previous application, i.e., $\sigma_{Ip} \approx 0.04\, I_p$ and $\sigma_{Is} \approx 0.06\, I_s$. A priori information about model parameters is also the same as the previous application; therefore, we applied the MDN trained in the previous application to perform inversion in this application.

Results and discussion

The upscaled I_p and I_s logs were inverted to obtain the PDF of the low-resolution porosity, clay content, and water saturation logs. The maximum a posteriori (MAP) points of the

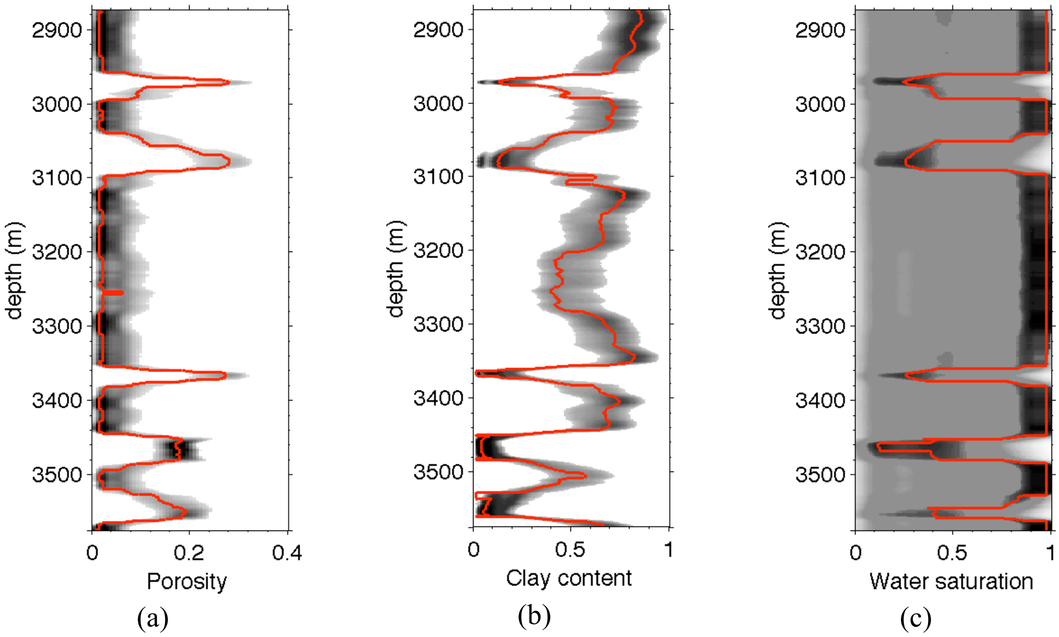


Figure 3-16: Inversion of the Backus averaged I_p and I_s. Marginal a posteriori PDF of porosity, clay content, and water saturation in one of the wells. (a) Porosity. (b) Clay content. (c) Water saturation. Darker shades represent higher probabilities. The red line is the maximum a posteriori (MAP) realization of the PDF.

resulting marginal PDF's were selected as approximations of the low-resolution porosity, clay content, and water saturation logs. The a posteriori marginal PDF's of the low-resolution porosity, clay content, and water saturation are shown in figure 3-16. The red curves in that figure show the maximum a posteriori (MAP) point of each marginal PDF.

The MAP values of the marginal PDF's of the upscaled logs (red curves) are compared with the measured logs (black curves) in figure 3-17. In general, the MAP of the inversion result varies between the upper and lower limits of the small-scale variations of log values. However, in some intervals we observe a bias in the MAP point estimates of the parameters. For example, the bias in the MAP estimates of clay content log is noticeable on the 3140 m-3320 m interval. As we explained in the previous application on the same interval we observe large errors in the predictions of the petrophysical forward function. Therefore, we conclude that biases in the MAP estimates of the upscaled porosity, clay content, and water saturation are mainly due to errors in the petrophysical forward function.

Commonly, upscaled values of rock and fluid properties are calculated by simple averaging of measured logs over the support volume of measurements (Menezes and Gosselin, 2006). In this method the upscaled values of porosity and clay content are derived using arithmetic

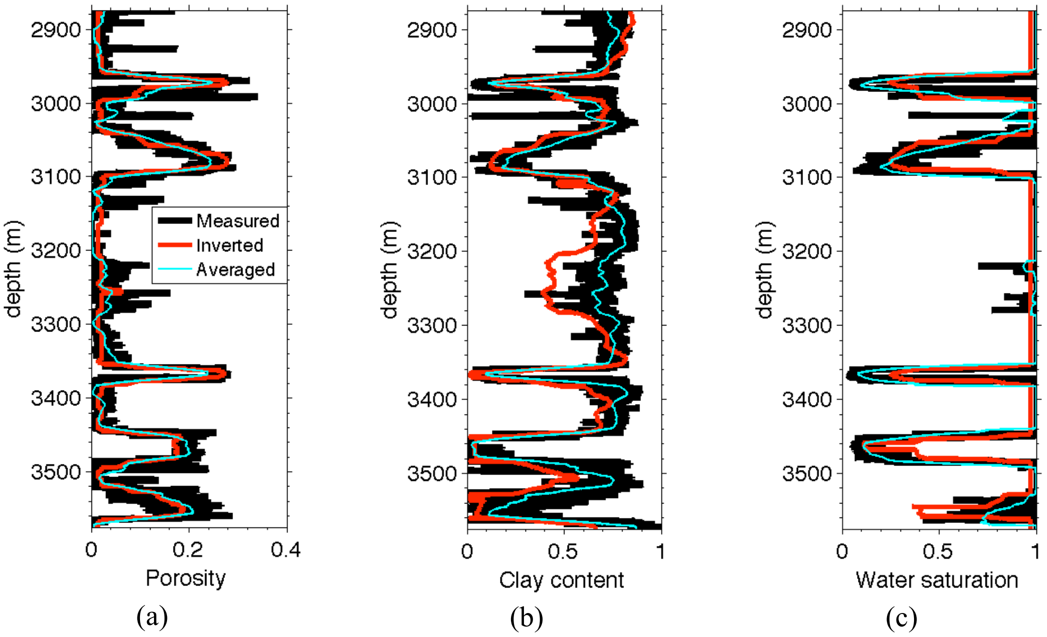


Figure 3-17: Comparison between measured well logs (black), upscaled well logs obtained from the inversion of the Backus averaged I_p and I_s logs (red) and upscaled well logs obtained by the standard averaging method (blue). (a) Porosity. (b) Clay content. (c) Water saturation.

averaging weighted by the bulk volume, while the upscaled value of water saturation is derived using arithmetic averaging weighted by the pore volume:

$$\bar{\phi} = \sum_{i=1}^{n} A_i\, h_i\, \phi_i \Bigg/ \sum_{i=1}^{n} A_i\, h_i \qquad (3\text{-}4)$$

$$\bar{V}_{cl} = \sum_{i=1}^{n} A_i\, h_i\, V_{cl_i} \Bigg/ \sum_{i=1}^{n} A_i\, h_i \qquad (3\text{-}5)$$

$$\bar{S}_w = \sum_{i=1}^{n} \phi_i\, A_i\, h_i\, S_{w_i} \Bigg/ \sum_{i=1}^{n} \phi_i\, A_i\, h_i \qquad (3\text{-}6)$$

In the above equations n is the number of measurements in the support volume, A_i and h_i are the area and height associated with each measurement in the support volume, ϕ_i, V_{cl_i}, and S_{w_i} are the small-scale measured values of porosity, clay content, and water saturation. In practice it is usually assumed that A_i and h_i are the same for all measurements and these parameters are cancelled out from the numerator and denominator of the above equations.

In order to compare the results of the petrophysical inversion method above (Figure 3-16) with the standard averaging results, we applied the above equations to compute upscaled values of porosity, clay content, and water saturation. In our application, the support volume of the measurement was assumed to be the same as the upscaling window of the Backus averaging procedure above. The results of the standard averaging upscaling procedure are shown as blue curves in figure 3-17. Because the standard averaging upscaling results are averaged values of measurements, they are not biased in comparison to the results of the petrophysical inversion method. Below we compare predictions of the petrophysical forward function using the upscaled petrophysical parameters obtained by the standard averaging upscaling method and inversion method.

We applied the petrophysical forward function to derive I_p and I_s from the upscaled rock and fluid properties in the above two cases. First, we used the results of the standard averaging upscaling method as the upscaled values of porosity, clay content, and water saturation. Then, we used the MAP points of the MDN inversion result as the upscaled values of those parameters. In both cases, upscaled values of the effective pressure were calculated using arithmetic averaging (because effective pressure is nearly a linear function of depth). Upscaled values of hydrocarbon and brine density were also calculated using arithmetic averaging. Upscaled values of hydrocarbon and brine bulk modulus were calculated using Backus averaging method. The averaging window in all above cases was the same as that used for the Backus averaging of I_p and I_s. Predictions of the petrophysical forward relations for the above two cases, in addition to Backus averaged values of I_p and I_s logs are shown in figure 3-18.

Figure 3-18 shows that the error of P- and S-wave impedance obtained using the standard averaging upscaling method is similar to those obtained using the MAP of the MDN inversion result. This similar error shows that although the standard averaging upscaled values of porosity, clay content, and water saturation are less biased (figure 3-17), their consistency with the upscaled P-wave and S-wave impedance is similar to the MDN inversion upscaled values. We note that in figure 3-18, errors in the upscaling methods are mainly due to the errors in the petrophysical forward relations and might be reduced if a more accurate forward function is used.

This application of the MDN inversion shows that upscaled values of porosity, clay content, and water saturation can be obtained from the results of the petrophysical inversion of upscaled (Backus averaged) P- and S-wave impedance. The results of the petrophysical inversion method can be biased due to errors in the petrophysical forward function.

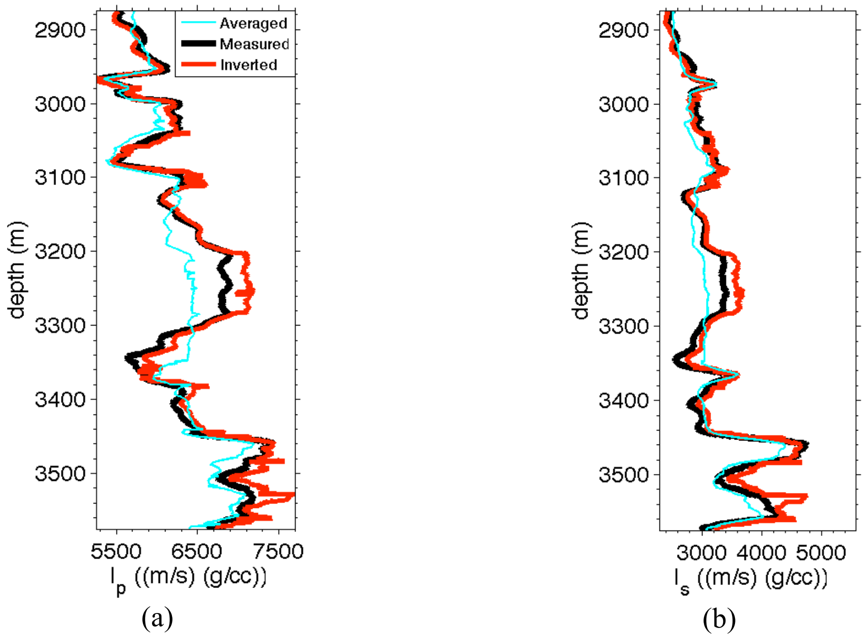


Figure 3-18: Comparison between Backus averaged measured well logs (black) and estimated well logs obtained using petrophysical forward function with upscaled rock and fluid properties from the inversion result (red) and using the standard averaging method (blue). (a) P-wave impedance. (b) S-wave impedance.

3.5 Conclusions

In this chapter, we present the diagonal-MDN inversion method as a new method to solve inverse problems. The MDN can be trained to provide a fully probabilistic solution to a nonlinear, Bayesian inverse problem. The MDN solution of a large number of similar inverse problems can be much faster to compute than the corresponding sampling-based solution, yet the MDN provides a good estimate of the sampling solution. The first and third applications above (figure 3-1 and figure 3-13) show that a posteriori PDF's of model parameters obtained using the MDN provide good approximations of the MC sampling solution. The accuracy of the MDN can be improved by using more kernels or by increasing the flexibility of the kernels. However, these improvements can significantly increase the required time for training and inversion.

The second application above shows that the MDN can be used to obtain the rock properties (i.e., porosity and clay content) from the seismic attributes logs (i.e., I_p and PR) without applying an independent petrophysical forward function. In this case with a limited number of training samples, the proper design and training procedure of the MDN is more important than the cases with a potentially unlimited number of training samples, because of the higher chance of over-fitting. The result of such completely data-driven applications of the MDN is deemed acceptable only if they are tested against data that are not used in the training procedure. The results of the inversion in the second application are acceptable in a blind well. This means that the geological setting and relations between acoustic and petrophysical properties of rocks in the blind well must be similar to those in the other wells used in the training procedure. The estimated a posteriori PDF of porosity and clay content represents the uncertainty due to uncontrolled factors, most importantly perhaps water saturation. Water saturation and its associated uncertainty can be estimated using a petrophysical forward function as we show in the third application.

The third application shows that the MDN can be used in conjunction with a petrophysical forward function to invert real P- and S-wave impedances for the joint PDF of porosity, clay content, and water saturation. The inversion results at the well location show that the marginal PDF's of the model parameters provide reasonable estimates of the measured logs. Errors in the estimated PDF are mainly due to errors in the petrophysical forward relations. The uncertainty in the a posteriori marginal PDF of water saturation is high, and this is mainly due to low sensitivity of predictions of the petrophysical forward function to water saturation, which we will investigate in more detail in chapter 4.

In the fourth application, the MDN is applied to compute the upscaled values of porosity, clay content, and water saturation from the petrophysical inversion of the Backus averaged values of P- and S-wave impedance. Due to errors in the petrophysical forward function, the results of the inversion method are biased in some intervals of the well. Any improvement in the accuracy of the forward function will result in more accurate upscaled values of porosity, clay content, and water saturation.

3.6 References

Avseth, P., T. Mukerji and G. Mavko, 2005, *Quantitative seismic interpretation: applying rock physics tools to reduce interpretation risk*. Cambridge ; New York: Cambridge University Press.

Bachrach, R., 2006, Joint estimation of porosity and saturation using stochastic rock-physics modeling, *Geophysics*, **71**, O53-O63.

Backus, G. E., 1962, Long-Wave Elastic Anisotropy Produced by Horizontal Layering, *Journal of Geophysical Research*, **67**(11), 4427-4440.

Batzle, M. and Z. Wang, 1992, Seismic Properties of Pore Fluids, *Geophysics*, **57**, 1396-1408.

Bishop, C. M., 1994, Mixture Density Networks.

Bosch, M., 2004, The optimization approach to lithological tomography: Combining seismic data and petrophysics for porosity prediction, *Geophysics*, **69**(5), 1272-1282.

Bosch, M., L. Cara, J. Rodrigues, A. Navarro and M. Diaz, 2007, A Monte Carlo approach to the joint estimation of reservoir and elastic parameters from seismic amplitudes, *Geophysics*, **72**(6), O29-O39.

Bosch, M., C. Carvajal, J. Rodrigues, A. Torres, M. Aldana and J. Sierra, 2009, Petrophysical seismic inversion conditioned to well-log data: Methods and application to a gas reservoir, *Geophysics*, **74**(2), O1-O15.

Caers, J. and X. Ma, 2002, Modeling Conditional Distributions of Facies from Seismic Using Neural Nets, *Mathematical Geology*, **34**(2), 143-167.

Chao, G., G. Lambert and H. Cumming, 2009, Analysis of Intrinsic Uncertainties of Petro-elastic Models Using Simulated Annealing, 71st EAGE Conference & Exhibition, Amsterdam.

Devilee, R. J. R., A. Curtis and K. Roy-Chowdhury, 1999, An efficient, probabilistic neural network approach to solving inverse problems: Inverting surface wave velocities for Eurasian crustal thickness, *Journal of Geophysical Research*, **104(B12)**, 28841-28856.

Duda, R. O., P. E. Hart and D. G. Stork, 2001, *Pattern Classification*: New York: Wiley.

Dvorkin, J. and M. A. Gutierrez, 2001, Grain sorting, porosity, and elasticity, from pangea.stanford.edu/~jack/Bimodal.pdf.

Dvorkin, J., G. Mavko and B. Gurevich, 2007, Fluid substitution in shaley sediment using effective porosity, *Geophysics*, **72**(3), O1-O8.

Dvorkin, J. and A. Nur, 1996, Elasticity of high-porosity sandstones: Theory for two North Sea data sets, *Geophysics*, **61**(5), 1363-1370.

Hearst, J. R., P. H. Nelson and F. L. Paillet, 2000, *Well Logging for Physical Properties: A Handbook for Geophysicists, Geologists, and Engineers*: Wiley.

Maiti, S. and R. K. Tiwari, 2009, A Hybrid Monte Carlo Method Based Artificial Neural Networks Approach for Rock Boundaries Identification: A Case Study from the KTB Bore Hole, *Pure and Applied Geophysics*, **166**(12), 2059-2090.

Maiti, S. and R. K. Tiwari, 2010, Automatic discriminations among geophysical signals via the Bayesian neural networks approach, *Geophysics*, **75**(1), E67-E78.

Maiti, S., R. K. Tiwari and H. J. Kumpel, 2007, Neural network modelling and classification of lithofacies using well log data: a case study from KTB borehole site, *Geophysical Journal International*, **169**(2), 733-746.

Marion, D., A. Nur, H. Yin and D. Han, 1992, Compressional Velocity and Porosity in Sand-clay Mixture, *Geophysics*, **57**, 554-563.

Mavko, G., T. Mukerji and J. Dvorkin, 2009, *The rock physics handbook: tools for seismic analysis of porous media*. Cambridge ; New York: Cambridge University Press.

McLachlan, G. and D. Peel, 2000, *Finite mixture models*: New York: Wiley.

Meier, U., A. Curtis and J. Trampert, 2007a, A global crustal model constrained by non-linearised inversion of fundamental mode surface waves, *Geophysical Research Letters*, **34**, L16304-L16304.

Meier, U., A. Curtis and J. Trampert, 2007b, Global crustal thickness from neural network inversion of surface wave data, *Geophysical Journal International*, **169**, 706-722.

Meier, U., J. Trampert and A. Curtis, 2009, Global variations of temperature and water content in the mantle transition zone from higher mode surface waves, *Earth Planetary Science Letters*, **282**, 91-101.

Menezes, C. and O. Gosselin, 2006, From Logs Scale to Reservoir Scale: Upscaling of the Petroelastic Model, SPE Europec/EAGE Annual Conference and Exhibition, Vienna.

Mosegaard, K. and A. Tarantola, 2002, Probabilistic approach to inverse problems, *International handbook of earthquake and engineering seismology*. W. H. K. Lee, H. Kanamori, P. C. Jennings and C. Kisslinger: Academic Press Inc.**:** 237-265.

Poulton, M. M., 2002, Neural networks as an intelligence amplification tool: A review of applications, *Geophysics*, **67**, 979-993.

Roth, G. and A. Tarantola, 1994, Neural Networks and Inversion of Seismic data, *Journal of Geophysical Research*, **99**, 6753-6768.

Saggaf, M. M., M. N. Toksoz and H. M. Mustafa, 2003, Estimation of reservoir properties from seismic data by smooth neural networks, *Geophysics*, **68**(6), 1969-1983.

Sambridge, M. and K. Mosegaard, 2002, Monte Carlo methods in geophysical inverse problems, *Reviews of Geophysics*, **40**(3), 1009.

Spikes, K., T. Mukerji, J. Dvorkin and G. Mavko, 2007, Probabilistic seismic inversion based on rock-physics models, *Geophysics*, **72**(5), R87-R97.

Tarantola, A., 2005, *Inverse Problem Theory and Methods for Model Parameter Estimation*: SIAM.

Wyllie, M. R. J., A. R. Gregory and L. W. Gardner, 1956, Elastic wave velocities in heterogeneous and porous media, *Geophysics*, **21**(1), 41-70.

4 Inversion of seismic attributes for petrophysical parameters and rock facies

> *"Everything is vague to a degree you do not realize*
> *till you have tried to make it precise."*
>
> Bertrand Russell

4.1 Abstract

In this chapter we applied the MDN inversion method of the previous chapter to predict 3D petrophysical properties from inverted pre-stack seismic data. The objective of petrophysical inversion is to estimate the joint probability density function (PDF) of model vectors consisting of porosity, clay content, and water saturation components at each point in the reservoir, from data vectors consisting of compressional- and shear-wave-impedance components at each point; obtained from the inversion of AVO seismic data. The petrophysical inverse problem is significantly non-linear, and the large number of data points in a seismic cube makes conventional methods of probabilistic inversion such as Monte Carlo sampling extremely computationally demanding to the point of being infeasible without large, parallel computing facilities.

The training dataset of the MDN is constructed by using forward petrophysical relations and includes different sources of uncertainty in the inverse problem such as variations in effective pressure, bulk modulus and density of hydrocarbon, and random noise in recorded data. Results show that the standard deviations of all model parameters are reduced after inversion, which shows that the inversion process provides information about all parameters. The reduction of uncertainty in water saturation is smaller than that for porosity or clay content; nevertheless the maximum of the a posteriori PDF of model parameters (MAP) clearly shows the boundary between brine saturated and oil saturated rocks. The MAP

estimates of different model parameters show the lateral and vertical continuity of these boundaries. Errors in the MAP estimate of different model parameters can be reduced by using more accurate petrophysical forward relations and less uncertain measured seismic attributes. We also applied the result of the petrophysical inversion to estimate the probability of non-reservoir facies, and brine- and oil-saturated reservoir facies. The accuracy of the predicted oil-saturated reservoir facies at the well location is good, but due to errors in the petrophysical inversion the predicted non-reservoir facies and brine-saturated reservoir facies are ambiguous.

4.2 Introduction

Prediction of the variation of rock and fluid properties such as porosity, clay content, and water saturation throughout reservoir volumes is essential for exploration and development of hydrocarbon reservoirs. Rock and fluid property maps obtained from such predictions can be used for optimal selection of well locations for reservoir development and production enhancement. Seismic data are usually the only source of information available throughout a field that can be used to predict the 3D distribution of properties with appropriate spatial resolution. The main challenge in inferring properties from seismic data is the ambiguous nature of geophysical information. Uncertainty enters into the problem in at least three levels: first, there is non-uniqueness in the inversion of (AVO) seismic data for the acoustic impedances of rock, second there is non-uniqueness in the petrophysical inversion of the rock impedances for rock-fluid properties given a petrophysical relationship between the acoustic properties and rock-fluid properties, and third there is ambiguity in these petrophysical relationships themselves (Doyen, 1988). Therefore, any estimate of rock and fluid property maps derived from seismic data must also represent its associated uncertainty.

Rock physics theories are used to construct petrophysical relations that provide the link between seismic data and reservoir rock properties. Theoretically, elastic moduli and density of rocks are controlled by different rock properties including porosity, clay content, fluid saturations, effective pressure, fluid densities, fluid and mineral elastic moduli, and possibly more parameters (Avseth et al., 2005; Mavko et al., 2009). Although some theories and laws appear to be approximately true in many examples of similar specific rock settings, it is difficult to address the effects of all of the micro-scale heterogeneity of rocks in any single set of relations. In practice for any particular geological basin, petrophysical relations are therefore semi-empirical, are calibrated with well logs and core data, and their theoretical uncertainty remains non-negligible (Bachrach, 2006).

Statistical rock physics has been used as a general tool to address uncertainties associated with the petrophysical relations. Mukerji et al. (2001) applied a classification method based on Monte Carlo (MC) sampling (Mosegaard and Tarantola, 2002; Sambridge and Mosegaard, 2002) to predict lithofacies (i.e., brine sand, oil sand, and shale) from attributes of seismic data. They analyzed acoustic impedance and V_p / V_s logs at wells to define the joint probability distribution function (PDF) of these parameters for each facies. The probability of each facies for data points in a 3D cube of seismic attributes (i.e., acoustic impedance and V_p / V_s) was then derived using the above joint PDF's. Geostatistical indicator simulation was applied to impose spatial correlations between facies. This study combined the statistical techniques for classification with seismic data and well log information in order to perform reservoir characterization.

Avseth et al. (2001) proposed a methodology to add information about rock physics theories to the statistical lithofacies prediction technique of Mukerji et al. (2001). They first defined brine sand, oil sand, and shale as seismic lithofacies at wells using geological observations. They applied rock physics theories and the MC sampling method to derive the joint PDF of zero-offset reflectivity and AVO gradient for different possible facies successions. Bayesian classification (Avseth et al., 2005, pp. 147-159) was then applied to predict lithofacies and pore fluid from attributes of 3D seismic data (i.e., zero-offset reflectivity and AVO gradient) using the joint PDF's estimated in the previous step. This methodology was applied for reservoir characterization in a North Sea turbidite system.

Eidsvik et al. (2004) proposed another statistical method to combine well observations and seismic AVO data to constrain the a priori distribution of lithofacies (i.e., shale, brine sand, and oil sand). Spatial coupling information about facies is included in the a priori PDF using a 2D Markov random field. The a priori PDF also models dependencies between lithofacies and elastic parameters (i.e., porosity, density, compressional- and shear-wave velocity of different facies) through nonlinear relations based on rock physics theories. The likelihood function relates zero-offset reflectivity and AVO gradient values to acoustic parameters of lithofacies (i.e., compressional- and shear-wave velocity). The MC sampling method was applied to update the above a priori PDF with the likelihood of seismic data in order to obtain samples of a posteriori PDF of facies.

Larsen et al. (2006) proposed a 1D version of Eidsvik et al. (2004) method, in which samples of a posteriori lithofacies model are obtained using a fast sampling algorithm that updates the a priori lithofacies model with the likelihood of the 1D seismic traces. In order to make sampling and representation of the posterior PDF of lithofacies tractable no spatial dependencies between lithofacies were assumed. Ulvmoen and Omre (2010) and Ulvmoen et al. (2010) generalized this methodology further to account for spatial coupling of lithofacies and applied it to obtain facies maps from 3D seismic data. Buland et al. (2008) proposed another extension of Larsen et al. (2006) method to account for possible multi-modality of seismic attributes of different lithofacies.

Gonzalez et al. (2008) proposed a technique to invert seismic data for lithofacies using multi-point geostatistics. The joint PDF of P-wave velocity and density for different lithofacies was derived from well log data. The multi-point geostatistical simulation technique was used to construct the a priori model of the lithofacies. Then a sampling algorithm was applied to update the a priori lithofacies model with seismic data using the joint PDF's estimated in the previous step. All of above lithofacies inversion techniques were tested on real data sets.

The Monte Carlo method has also been used to simulate the elastic response of earth models by exploring a priori ranges of rock and fluid properties using petrophysical forward relations. Bosch et al. (2007) demonstrated a quantitative application of the MC sampling techniques to obtain compressional impedance and porosity from short-offset seismic data. Bosch et al. (2009) extended the previous work to estimate water saturation in addition to the above two parameters and to constrain the final result by geostatistical information from well logs.

Spikes et al. (2007) demonstrated another application of the MC sampling to invert two constant-angle stacks of seismic data for porosity, clay content, and water saturation in an exploration setting. Bachrach (2006) also applied the MC sampling to produce porosity and water saturation maps from compressional- and shear-wave impedance as attributes of seismic data. Data in both of the above papers were from an exploration setting; therefore in order to make a priori assumptions about the model parameters as noncommittal as possible, Uniform distributions over physically realizable intervals of the model parameters were assumed as the a priori PDF. Bachrach (2006) also indicated that the a posteriori maps of porosity and water saturation were constrained by the a priori geological information that was used for obtaining compressional- and shear-wave impedance from inversion of (AVO) seismic data.

In all of the above studies, it is indicated that while in principle the MC sampling method can map uncertainty of petrophysical parameters and lithofacies, in practice applying it to invert seismic attributes for rock and fluid properties is computationally demanding. For each point in the subsurface the forward problem (in this case the petrophysical relationships and approximations of Zoeppritz' equation) must be evaluated for a large number of model samples (typically at least of the order of thousands), and this process would have to be repeated for all data points of interest in a seismic cube, which usually includes hundreds of millions of data points. In addition, storing the obtained probabilistic results (many MC samples per data point) can require large storage facilities. Consequently, MC sampling techniques can often be impossible to implement in practice other than for a small or sparse subset of the available data points.

As a solution we present the mixture density network (MDN), which is developed and tested in the previous chapter, as a new tool for probabilistic inversion of seismic attributes to obtain rock and fluid properties, and lithofacies. The MDN provides a solution to both the computational power and the memory problems associated with the MC sampling method. Previously, Devilee et al. (1999) and Meier et al. (2007a,b; 2009) show other geophysical applications of the MDN for solving tens of thousands of one-dimensional inverse problems. The latter papers invert regional and global seismological datasets for the global distribution of various rock properties and structures in the crust and upper mantle, showing that large, seismic-like data sets can be inverted efficiently and probabilistically for 3D earth models. In the previous chapter we used the MDN to invert acoustic velocity logs for porosity, clay content, and water saturation profiles down wellbores, showing that the method can be used for probabilistic inversion of well logs for relevant petrophysical parameters.

In this chapter we jointly invert industrial seismic compressional and shear impedances, I_p and I_s, for the joint probability density function (PDF) of effective porosity ϕ_e, clay content V_{cl}, and water saturation s_w, using calibrated petrophysical relations and other prior information from wells. At the end of this chapter, we applied the probabilistic result of the petrophysical inversion of seismic data to estimate the probability of lithofacies.

We first briefly summarize the MDN method of solving an inverse problem in the petrophysical inversion context, and then introduce forward petrophysical relations, a priori information about model parameters, and data uncertainties. The statistical behavior of the forward petrophysical relations due to a priori uncertainty of the model parameters, and due to noise in the data is then analyzed. Then follows the MDN inversion results. Finally we discuss the results and conclude.

4.3 Method

4.3.1 Mixture density network solution of an inverse problem

The solution to any inverse problem is a definition of the extent to which any combination of model parameter values are consistent with the data, given the data uncertainty. In mathematical terms the solution can be expressed as (Tarantola, 2005, p. 34):

$$\sigma_M(\mathbf{m}) = K\rho_M(\mathbf{m})L(\mathbf{d}|\mathbf{m}) \tag{4-1}$$

In this equation **m** is the model parameter vector, ρ_M is the a priori PDF of the model vector over the model parameter space M, and σ_M is the a posteriori PDF of the model vector, which represents the solution of the inverse problem and is normalized by a constant K. For a given value of measured data **d**, the likelihood function L measures the consistency of a model vector **m** with **d**. It represents the uncertainty of the synthetic data **f**(**m**) due to different sources such as theoretical uncertainties in the forward function **f**, and also account for measurement uncertainties in the data **d** (the vertical line represents the fact that data **d** are measured and hence have fixed values). Equation (4-1) represents a Bayesian solution because it combines information known prior to inversion, $\rho_M(\mathbf{m})$, with information from a new dataset, $L(\mathbf{f}(\mathbf{m})|\mathrm{d})$, using Bayes rule for combining probabilities.

The mixture density network (MDN) is trained to emulate $\sigma_M(\mathbf{m})$ for any measured data **d**. This is achieved using pairs of a priori samples of model and data vector pairs (**m**,**d**). The set of sample pairs is called a training data set and is constructed in the following way: samples $\mathbf{m}_i$, $i = 1,\ldots,N$, are taken according to the a priori model PDF ρ_M, and for each sample the corresponding synthetic data $\mathbf{f}(\mathbf{m}_i)$ is calculated. Several samples, $\varepsilon_{i,j}$, $j = 1,\ldots,R$, of data measurement uncertainty, as well as of the theoretical uncertainty in the forward function **f** (both of which are represented within the likelihood function L), are added to each calculated synthetic data vector $\mathbf{f}(\mathbf{m}_i)$. This results in several samples of possible synthetic data vectors for each sample of the model vector $\left\{\left(\mathbf{m}_i, \mathbf{f}(m_i) + \varepsilon_{i,j}\right) : i = 1,\ldots,N;\ \ j = 1,\ldots,R\right\}$. Using these sample pairs, which for short we denote $(\mathbf{m}_k, \mathbf{d}_k)$, $k = 1,\ldots,NR$, the network is trained to map any data vector including its uncertainty into an approximation of a posteriori PDF of model vector **m** as described in chapter 3 and Appendix A.

Point estimator and uncertainty estimation of a multi-dimensional PDF

Although the a posteriori PDF, $\sigma_M(\mathbf{m})$, contains all information about parameters of the model vector, sometimes it is useful to choose just one value (i.e. a point estimator) for the parameters of the model vector as a representative sample, plus some measure of uncertainty or deviation of the posterior distribution around this sample. Here, the maximum a posteriori (MAP) probability point of the marginal PDF for each model parameter is chosen as the point estimator. The marginal PDF of one model parameter is obtained by integrating the joint a posteriori model PDF over all other model parameters (an analytic integration since the PDF is constructed from a sum of known Gaussians). The MAP point is simply the point at which this marginal distribution takes its maximum value and hence gives the most probable model vector post inversion. This can be a good estimator when the PDF is uni-modal (has a single local maximum). However, when the a posteriori PDF of the model vector is multi-modal (has more than one local maximum) the MAP solution does not have any particular importance (Tarantola, 2005, p. 42).

The measure of uncertainty around the MAP point estimator that we use here is the standard deviation of the marginal a posteriori PDF of model parameters. For each model parameter m_i, with marginal a posteriori PDF $p(m_i)$, the standard deviation, S_{m_i}, is defined as:

$$S_{m_i} = \left[\int_{m_i} \left(m_i - \overline{m}_i\right)^2 p\left(m_i\right) \mathrm{d}m \right]^{1/2} \tag{4-2}$$

Here $\overline{m}_i$ is the mean of the m_i obtained from the marginal PDF $p(m_i)$. Note that the uncertainty measure used is not a locally linearized approximation to the standard deviation as is often used in linearized inverse problems; instead it measures the uncertainty represented by the full (perhaps multi-modal) posterior PDF of each model parameter.

4.3.2 Petrophysical forward relations, a priori model PDF, and data uncertainty

We applied the MDN inversion method to invert compressional and shear wave impedances derived from seismic data, for porosity, clay content, and water saturation in the Akpo field, which is briefly introduced in the previous chapters. The petrophysical forward relations are given in Appendix C and are the same as what we applied in the third application in the previous chapter. The a priori information about model parameters are obtained from available well logs and are also the same as what we presented in the third application in the previous chapter. Here, we describe the seismic data and its uncertainty.

Seismic data

A 3D simultaneous elastic inversion technique jointly inverted near-, mid- and far-angle sub-stacks to derive estimates of the 3-dimensional distribution of compressional- and shear-wave impedances, I_p and I_s. Five angle stacks were used as the input seismic dataset, with angle ranges 0°–13° for near, 13°–21° and 21°–29° for middle, 29°–37° for far, and 37°–45° for ultra far stacks. The sampling interval of the seismic data was 3 ms. Statistical wavelet for each angle stack was estimated at four different well locations using reflection coefficients obtained from V_p, V_s, and density logs.

Structural and stratigraphic interpretations resulted in picks of 18 different horizons in the seismic data to define the geometry of the initial model used for inversion. The a priori models for I_p and I_s in each horizon were provided by the low frequency trend of the well logs and were extrapolated over the entire model using kriging methods. The trend of a priori I_p and I_s were obtained from seismic interval velocity (provided from seismic imaging) and the distribution of the residuals was obtained from kriging of well logs (Dubrule, 2003, p. 3-39). The range of the variogram used for the kriging was defined based on a priori geological information in each of the 18 different horizons of the model geometry. The result of the kriging was then filtered with a low-pass filter (below seismic band-pass) and transferred into the model geometry as a priori I_p and I_s model. Finally, an iterative algorithm, based on the simulated annealing technique (Sen and Stoffa, 1991) was used to adjust this prior model estimates of I_p and I_s in each bin, in order to optimize the match between the input measured seismic data and the synthetic seismic response calculated by the Zoeppritz equations (Aki and Richards, 1997).

In this way, the seismic inversion technique combined geological information about the expected vertical and spatial continuity of the medium, well log data, and information from seismic data to provide an estimate of the distribution of I_p and I_s. The result for one section, which includes one of the wells, is shown in figure 4-1.

Uncertainty of the estimated value of I_p and I_s can be estimated either directly from the inversion algorithm (Buland et al., 2003; Buland and Omre, 2003) or by statistical comparison of upscaled I_p and I_s values derived from well logs with seismic-inversion results (Bachrach, 2006). In this example, the latter method was applied and uncertainty of I_p was

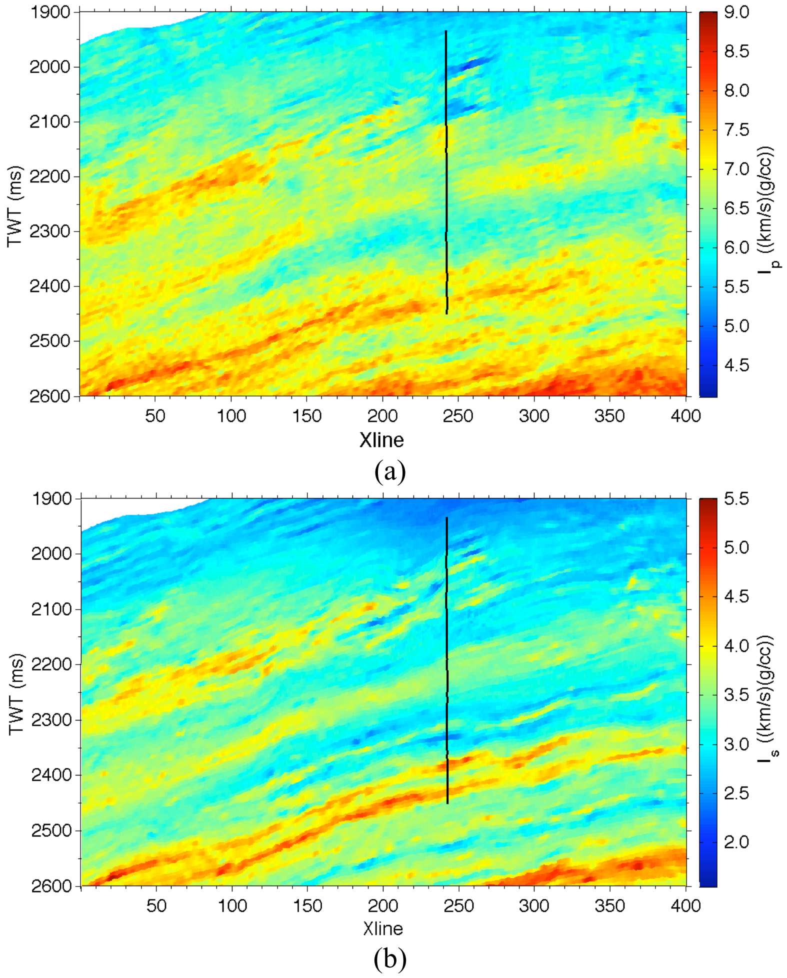
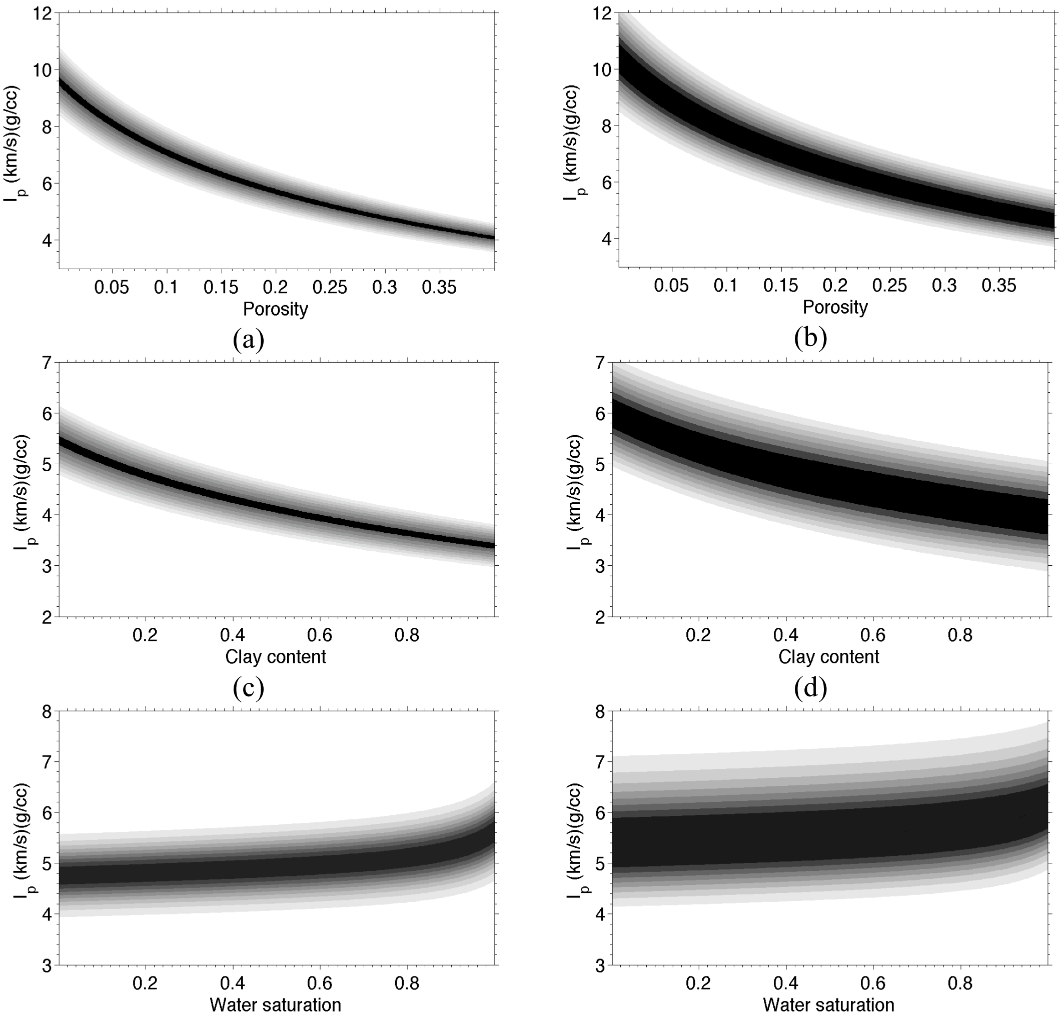


Figure 4-1: The P- and S-wave-impedance for one cross section of the 3D seismic dataset. (a) P-wave impedance. (b) S-wave impedance. The scale of the TWT and Xline axes of the plots are from real data but the numbers are fictitious.

modeled with a Gaussian distribution with zero mean and standard deviation $\sigma_{I_p} = 0.06\ I_p$. Uncertainty of I_s was modeled with another Gaussian distribution with zero mean and standard deviation $\sigma_{I_s} = 0.08\ I_s$. It is important to note that by using this method we assumed that I_p and I_s estimates derived from seismic inversion were unbiased. This assumption may be violated where the low-frequency background model used in the seismic inversion is biased.

4.3.3 Forward modeling of I_p and I_s

In this petrophysical inversion the parameters of the model vector are effective porosity, clay content, and water saturation, $\mathbf{m} = (\phi_e, V_{cl}, s_w)$. The parameters of the data vector are compressional- and shear-wave impedance, $\mathbf{d} = (I_p, I_s)$. Other inputs of the petrophysical



(a) (b) (c) (d) (e) (f)

Figure 4-2: Uncertainty in the predictions of the petrophysical forward relations for I_p. (a) I_p versus porosity when other input parameters of the petrophysical forward relations are held constant. Ambiguity (gray area) is due to the overall effect of theoretical and measurement uncertainty. (b) I_p versus porosity when confounding parameters of the forward rock physics model are varied. The thicker black and dark gray area shows additional uncertainty due to variations of the confounding parameters. (c,d) Corresponding plots for I_p versus clay content. (e,f) Corresponding plots for I_p versus water saturation.

forward function, which are effective pressure, bulk modulus of brine and hydrocarbon, and density of brine and hydrocarbon, are considered as confounding model parameters, $\mathbf{m}_{\text{conf}} = (p_e, K_{hc}, K_w, \rho_{hc}, \rho_w)$.

The behavior of the petrophysical forward relations over the a priori model space, in the presence of uncertainty in data, in the forward relations, and in the confounding model parameters, is studied in this section. Figure 4-2(a) shows I_p as a function of porosity, when all other input parameters of the petrophysical forward relations are held constant at possible realistic values: V_{cl} =0.2, s_w =0.1, p_e =200 bar, K_{hc} =0.22 GPa, K_w =2.62 GPa, ρ_{hc} =0.47 g/cc, and ρ_w =1.008 g/cc. In the petrophysical inversion process, we apply the predictions of the petrophysical forward relations to infer information about model parameters from inverted seismic data. Uncertainty in the petrophysical forward relations (i.e., theoretical uncertainty) is assumed to be Gaussian and is given in Appendix C. Uncertainty in the inverted seismic data (i.e., measurement uncertainty) is also assumed to be Gaussian and is given by σ_{I_p} and σ_{I_s} as derived above. Therefore, the total measurement and theoretical uncertainty, which is represented by the likelihood function in equation (4-1), is Gaussian with a covariance matrix equal to the sum of the covariance matrices of measurement and theoretical uncertainty (Tarantola, 2005, p. 35). The gray area in figure 4-2(a) shows the total uncertainty of the petrophysical forward function prediction for different values of porosity when other input parameters are known and held fixed. Notice that these plots are differ from plots that represent the theoretical uncertainty alone, as might be found in other papers (e.g., Tarantola and Valette, 1982). This is because in the MDN inversion methodology, as we described in chapter 3, data uncertainties are also added to the theoretical forward function predictions.

Figure 4-2(b) shows I_p as a function of porosity for the constant values of clay content and water saturation given above, but when the confounding parameters (i.e., effective pressure, bulk modulus and density of hydrocarbon) are varied according to their a priori distributions. The thicker black and dark gray area in figure 4-2(b) in comparison with figure 4-2(a) shows that the uncertainty of the prediction of the petrophysical forward function increases due to variations in the confounding parameters.

Figure 4-2(c) shows I_p as a function of clay content, when porosity is equal to 0.3 and the values of water saturation, effective pressure, bulk modulus and density of hydrocarbon, and bulk modulus and density of water are held constant at the values given above. The light gray area in that figure shows the total uncertainty (sum of the measurement and theoretical uncertainty) of the prediction of the petrophysical forward function when other input parameters are known. In figure 4-2(d) I_p is shown as a function of clay content when porosity and water saturation are constant, and effective pressure as well as all other confounding parameters vary according to their a priori distributions. Again the difference between the thickness of the black and dark gray area in figure 4-2(c) and figure 4-2(d) shows the effect of variations of the confounding parameters on the uncertainty of the prediction of the petrophysical forward function.

In the same manner as figure 4-2(a) and figure 4-2(c) for porosity and clay content, figure 4-2(e) shows the prediction of the petrophysical forward relations for I_p as a function of water saturation with its associated uncertainty, when all other input parameters of the forward petrophysical relations are held constant at the values given above. Figure 4-2(f) shows the effect of the variations of the confounding parameters on uncertainty of the predictions of the petrophysical forward relations as a function of water saturation and for the constant values of porosity and clay content given above.

Figure 4-3(a), figure 4-3(c), and figure 4-3(e) show the total effect of the theoretical and measurement uncertainty on the predictions of the petrophysical forward relations for I_s. Figure 4-3(b), figure 4-3(d), and figure 4-3(f) show the additional uncertainty of the predictions of the forward petrophysical relations for I_s due to variations in the confounding input parameters – similarly to figure 4-2 for I_p.

Figure 4-2 and figure 4-3 show that although I_p and I_s are strongly varying functions of porosity, and clay content, and weakly varying functions of water saturation, inference of the petrophysical parameter values from I_p and I_s estimates may well be obscured by the high uncertainty of I_p and I_s due to uncertainty in the seismic processing, and due to confounding parameters of the forward petrophysical relations $\mathbf{m}_{\mathrm{conf}} = (p_e, K_{hc}, K_w, \rho_{hc}, \rho_w)$. Therefore, with only two data (I_p and I_s) for these three unknowns (porosity, clay content, and water saturation) we clearly do not expect a unique solution for the petrophysical inversion. Hence, herein we aim principally to assess how much information the seismic data provides to reduce uncertainty in these three parameters. We could also do the same for any of the other parameters in the petrophysical forward function.

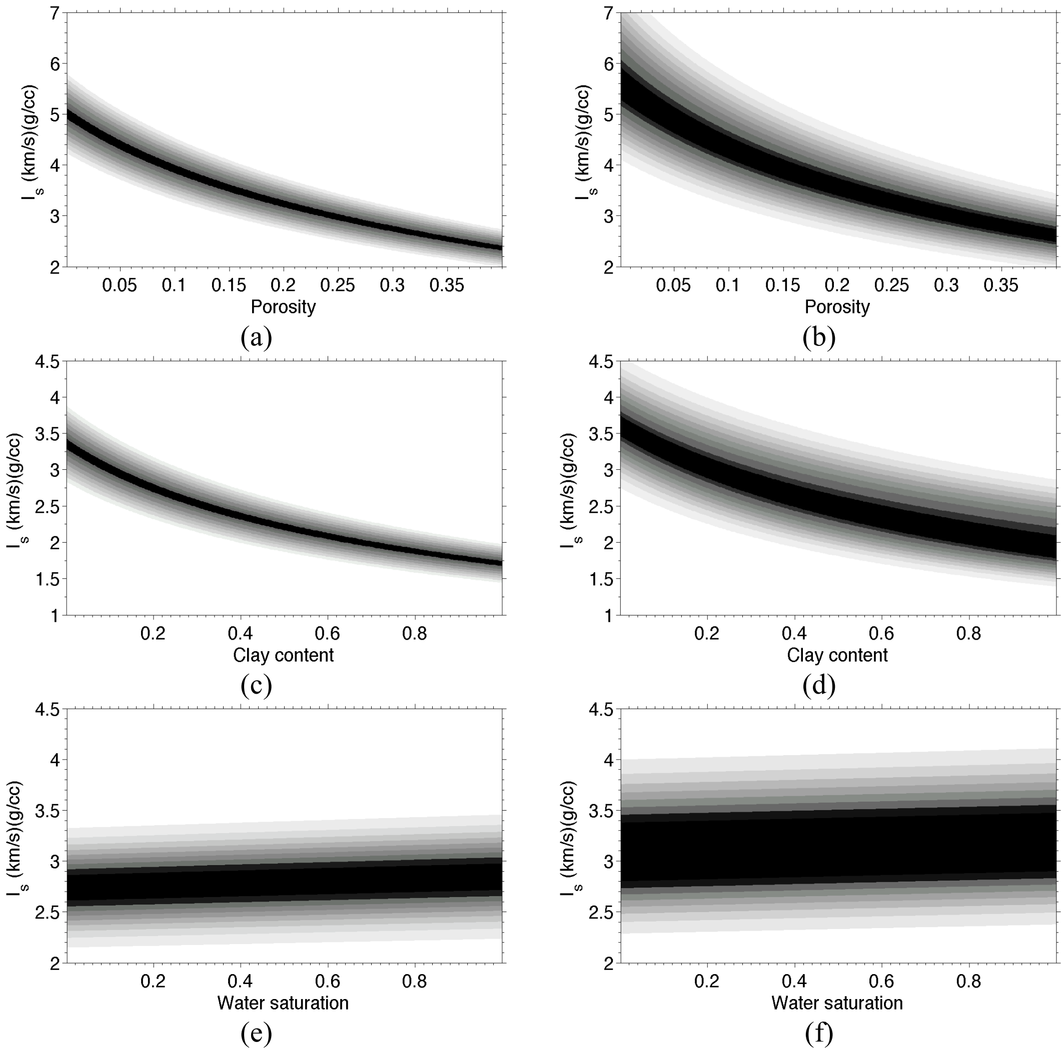


Figure 4-3: Uncertainty in the predictions of the petrophysical forward relations for I_s. (a) I_s versus porosity when other input parameters of the petrophysical forward relations are held constant. Ambiguity (gray area) is due to the overall effect of theoretical and measurement uncertainty. (b) I_s versus porosity when confounding parameters of the forward rock physics model are varied. The thicker black and dark gray area shows additional uncertainty due to variations of the confounding parameters. (c,d) Corresponding plots for I_s versus clay content. (e,f) Corresponding plots for I_s versus water saturation.

4.3.4 Training data set and MDN specifications

The MDN input vector consists of the P- and S-wave-impedances $\mathbf{d} = (I_p, I_s)$, and its outputs are parameters of the mixture density model of the posterior PDF of the model vector $\mathbf{m} = (\phi_e, V_{cl}, s_w)$. The output vectors of the training data set of the MDN were constructed by systematic sampling from the a priori PDF of model parameters. As in the third application of the previous chapter, 846 equally spaced samples of porosity and clay content were selected from the *T*-region in the $\phi_e - V_{cl}$ plane (figure 3-11(a)). 40 equally spaced samples of s_w and 6 equally spaced samples of p_e were then selected from the intervals [0,1] and [173 bar , 332 bar], respectively. The values of the bulk modulus and density of each type of hydrocarbon were derived from the selected values of p_e using the straight lines in figure 3-12. The total number of model vectors constructed in this way was 642,960.

As we explained in the previous chapter the above number of samples of the desired model parameters, i.e., (ϕ_e, V_{cl}, s_w), will reduce the interpolation error of the MDN inversion result. We selected smaller number of samples from the confounding model parameters, i.e., (p_e, K_{hc}, K_w, ρ_{hc}, ρ_w), because the effect of these parameters are integrated out by the MDN. A denser sample selection of the above parameters will improve the accuracy of the MDN, however, it will increase the training time significantly.

For each of the 642,960 a priori model vectors selected above, one data vector (I_p, I_s) was calculated by using the forward petrophysical relations. Three samples of Gaussian noise with the covariance matrix equal to the sum of the measurement and theoretical covariance matrices were added to each of these data vectors to represent data uncertainty. The number of noisy model-data pairs derived above was therefore $3\times642{,}960 = 1{,}928{,}880$ which were used as the training data set.

The specifications of the MDN are the same as what we used in the third application of the previous chapter. The number of required kernels was 10 and the number of hidden units was 27. Training with noisy dataset and the large number of training samples control over-fitting of the network.

4.3.5 Change of scale of the original well log data and seismic impedances

As explained in the fourth application of Chapter 3, the measured value of P- and S-wave impedance, porosity, clay content, and water saturation depend on the frequency content of the measuring signal, and on the support volume associated with each measurement. As described in Chapter 3, we applied Backus averaging method to transform the high-resolution I_p and I_s logs measured at well locations to seismic-scale logs with spatially-averaged properties similar to seismic measurements.

Brown and Seifer (1997) mentioned that velocity measurements in different directions using different frequencies requires considerations of (1) elastic scattering, (2) intrinsic attenuation, (3) volume effects, and (4) path effects. Separation of these effects is important in reservoir characterization using multi-frequency signals. They showed by an example that due to effect of the above parameters the velocity profile at a well obtained using upscaled sonic log and seismic measurements can be different. In the same way the above parameters can affect the petrophysical parameters obtained by petrophysical inversion of multi-frequency velocity data. In this study we do not consider the effect of the above parameters separately, and assume that measurement uncertainty, which is introduced above, can represent the effect of these parameters. A more detailed study of data from Akpo field is required to consider the effect of the above parameters on the final petrophysical inversion result.

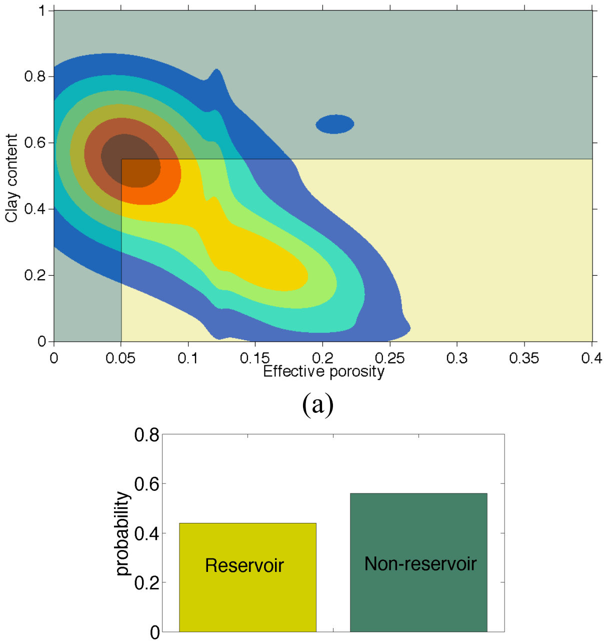
We wish to use the well logs as a "ground-truthing" check on the petrophysical inversion results of seismic I_p and I_s data. However, for the reasons stated above, the spatially-averaged rock and fluid properties (porosity, clay content, water saturation), which are derived from the inversion of the seismic scale I_p and I_s, are different from well log measurements. The values of porosity, clay content, and water saturation used for ground-truthing must be consistent with the scale of the seismic I_p and I_s measurement. We used the MAP of the marginal PDF of porosity, clay content, and water saturation (figure 3-16) obtained from the petrophysical inversion of the Backus averaged I_p and I_s logs in the fourth application of the previous chapter (section 3.4.4) as ground-truthing check.

4.3.6 Facies probabilities from the petrophysical inversion results

In the past, prediction of facies maps from 3D seismic data has been one of the main applications of the petrophysical inversion (Avseth et al., 2001; Mukerji et al., 2001; Eidsvik et al., 2004; Larsen et al., 2006; Buland et al., 2008; Gonzalez et al., 2008; Grana and Della Rossa, 2010; Ulvmoen and Omre, 2010; Ulvmoen et al., 2010). As we discussed in Chapter 2 and Appendix C the petrophysical forward function is defined for different lithofacies, which are identified by the value of rock and fluid properties. In Chapter 2 we also discussed how to obtain the probability of each facies $\Im_i$, conditioned on the value of the rock and fluid properties $\mathbf{m}$, $p(\Im_i \mid \mathbf{m})$. The probability of each facies after the petrophysical inversion is derived from the a posteriori joint PDF of the model parameters, σ_M as:

$$p(\Im_i) = \int_{\mathbf{m}} p(\Im_i \mid \mathbf{m})\, \sigma_M(\mathbf{m})\, d\mathbf{m} \tag{4-3}$$

Based on the information provided in Appendix C three different lithofacies are defined for this application: 1) non-reservoir facies, with effective porosity smaller than 5% or clay content larger than 55%, 2) oil-saturated reservoir facies, with effective porosity larger than



(a)

(b)

Figure 4-4: Facies probabilities from the petrophysical inversion result. (a) Contour plot is the joint a posteriori PDF of effective porosity and clay content obtained using the MDN inversion. Yellow area represents the reservoir facies; Probability of this facies is uniform over the yellow area and zero elsewhere. Green area represents non-reservoir facies; Probability of this facies is uniform over the green area and zero elsewhere. (b) The result of the integration of the MDN inversion result over the reservoir (yellow bar) and non-reservoir facies (green bar).

5%, clay content smaller than 55%, and water saturation smaller than 50%, and 3) brine-saturated reservoir facies, with effective porosity larger than 5%, clay content smaller than 55%, and water saturation larger than 50%.

Figure 4-4 shows an example of calculation of the probability of the reservoir and non-reservoir facies from the a posteriori joint PDF of porosity and clay content obtained from the inversion of well logs. The probability of reservoir (non-reservoir) facies over the model space, $p(\text{res} \mid \phi, V_{cl})$ ($p(\text{non-res} \mid \phi, V_{cl})$), is uniform over the yellow (green) area in that figure. The result of the integration in the equation (4-3), which represents the probability of reservoir and non-reservoir facies, is represented in figure 4-4(b). Note that in this application, facies probabilities are obtained efficiently from the analytical integration of a mixture of Gaussian distributions over rectangle areas in the model space. Below, we apply this method to compute the probability of three facies above from the result of the petrophysical inversion of seismic data.

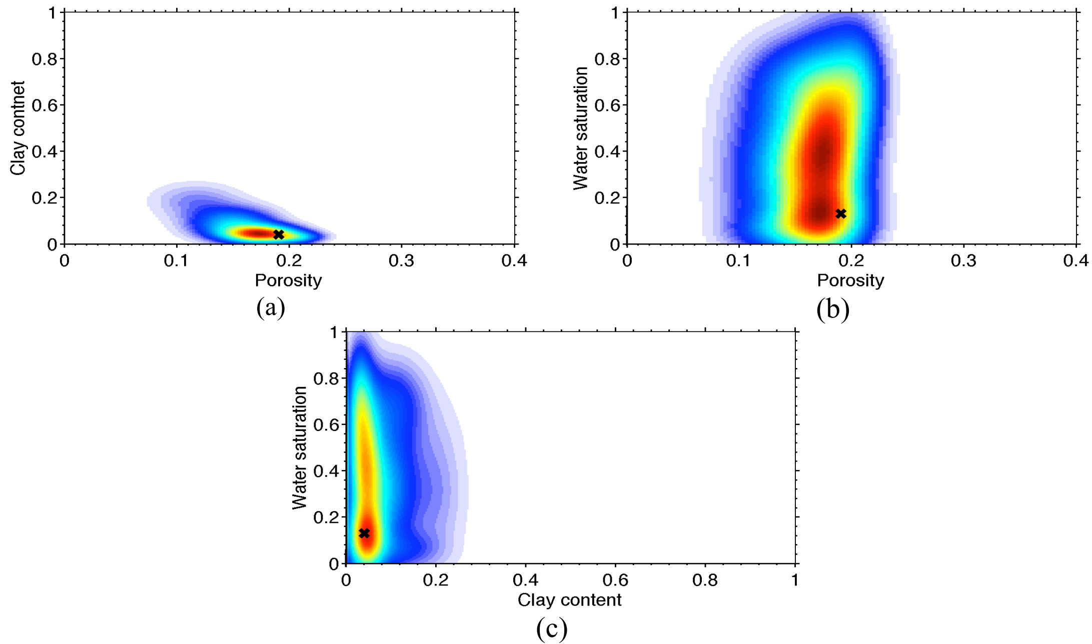


Figure 4-5: Marginal a posteriori joint PDF of the model parameters for I_p= 7345 kg/(m^{2}s) and I_s= 4658 kg/(m^{2}s). (a) Marginal joint PDF of effective porosity and clay content. (b) Marginal joint PDF of effective porosity and water saturation. (c) Marginal joint PDF of clay content and water saturation. Hot colors show high probability zones. The black cross is the value of low-resolution logs for this data point.

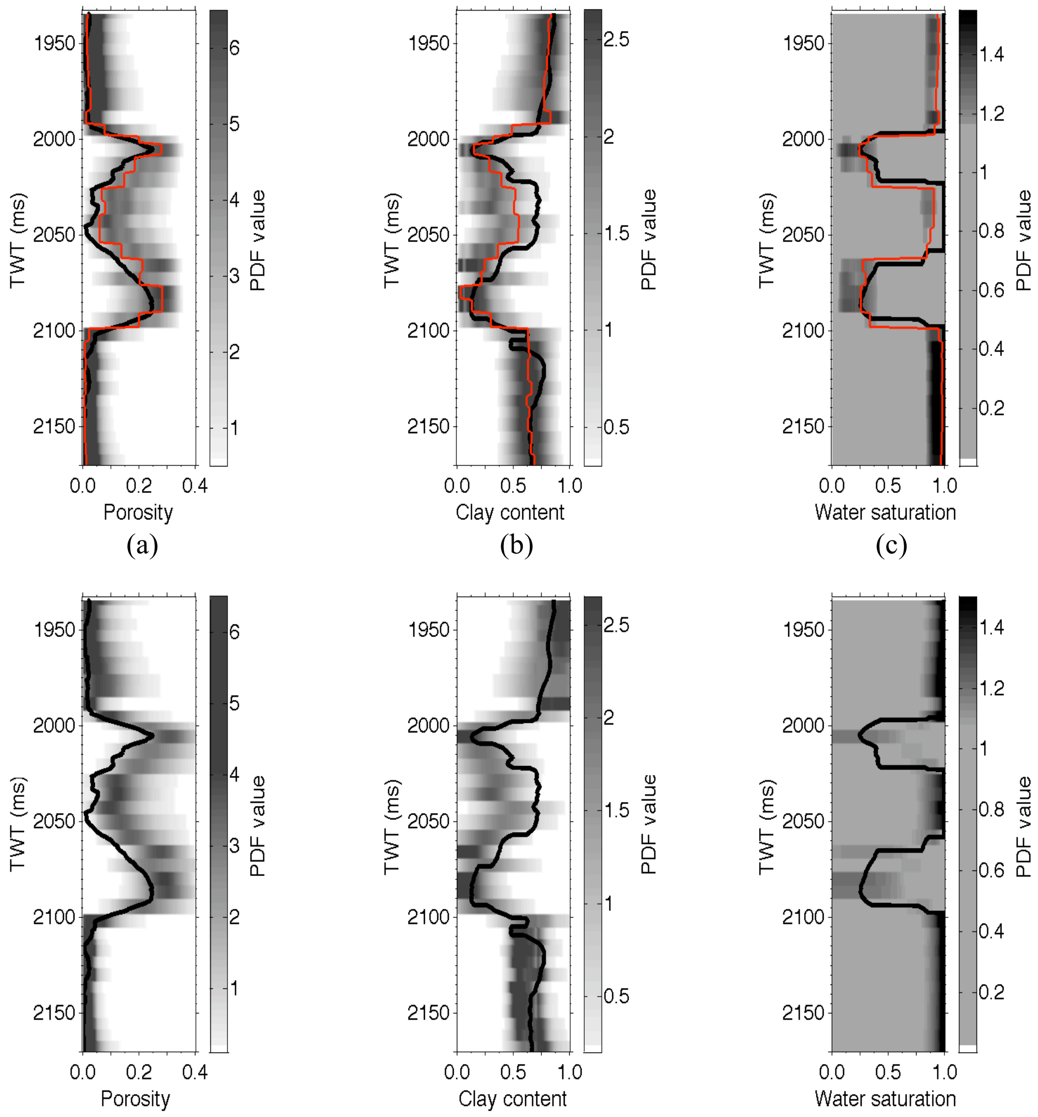
4.4 Results

The MDN with the specifications given above was trained using the training data set to estimate the joint PDF of model vector $\mathbf{m} = (\phi_e, V_{cl}, s_w)$ from data vector $\mathbf{d} = (I_p, I_s)$ using the methods of the Chapter 3. In this section we present results of the inversion.

Figure 4-5 shows the joint marginal PDF's of the model parameters evaluated at a randomly chosen data point from one well in the field with I_p = 7345 kg/(m^{2}s) and I_s = 4658 kg/(m^{2}s). The values of the low-resolution log of the model parameters for this data point are ϕ_e =0.19, V_{cl} =0.04, and, s_w =0.13, which are shown by crosses in figure 4-5. The point estimates of porosity, clay content, and water saturation are obtained as the MAP points of the 1-dimensional PDF's of those parameters and are equal to 0.17 for porosity, 0.04 for clay content, and 0.12 for water saturation.

In order to assess the reduction of uncertainty of the model parameters due to the inversion process (and hence due to the seismic data), we compare the posterior and the prior standard deviation of the model parameters. The standard deviation for the posterior (prior) porosity PDF is 0.02 (0.09), for the clay content PDF is 0.05 (0.26), and for the water saturation PDF is 0.17 (0.30). The relative reduction in all three posterior standard deviations shows that the inversion process decreases the uncertainty of the entire model vector. Remaining uncertainty in the model parameters is mainly due to uncertainty in effective pressure, other confounding parameters, theoretical uncertainty of the petrophysical forward relations, and measurement noise in I_p and I_s estimates obtained from AVO seismic data.

Figure 4-6 and figure 4-7 show the marginal PDF's of porosity, clay content, and water saturation, which are derived from the inversion of seismically derived I_p and I_s values lying in two intervals along one well profile (performance at other wells is similar). The value of water saturation in these two intervals varies between 10% and 100%. The first row in figure 4-6 and figure 4-7 show results of the MDN inversion of seismic data. In order to assess the accuracy of the MDN solution, the Monte Carlo sampling method of the previous chapter was also used to invert I_p and I_s values derived from seismic data for the joint PDF of model vector $\mathbf{m}$. The a priori information about model parameters, and the forward petrophysical relations used in the Monte Carlo sampling inversion were the same as the MDN inversion. The marginal PDF's of the model parameters, which are obtained from the Monte Carlo sampling, are shown in the second row of figure 4-6 and figure 4-7. Comparison between the results of the MDN inversion and Monte Carlo sampling inversion shows that the MDN solution provides a reasonable approximation of the Monte Carlo sampling solution.



(a) (b) (c)

(d) (e) (f)

Figure 4-6: Marginal a posteriori PDF's of the model parameters from the inversion of seismically derived I_p and I_s for the depth interval 1935 ms–2165 ms, in one of the wells. First row is the MDN solution: (a) Porosity. (b) Clay content. (c) Water saturation. The solid red line is the maximum a posteriori (MAP) of the marginal PDF's derived from the petrophysical inversion of seismic data. Second row is the Monte Carlo sampling solution: (d) Porosity. (e) Clay content. (f) Water saturation. Solid black line is the low-resolution log in the both rows. The scale of the TWT axis of the plots is from real data but the numbers are fictitious.

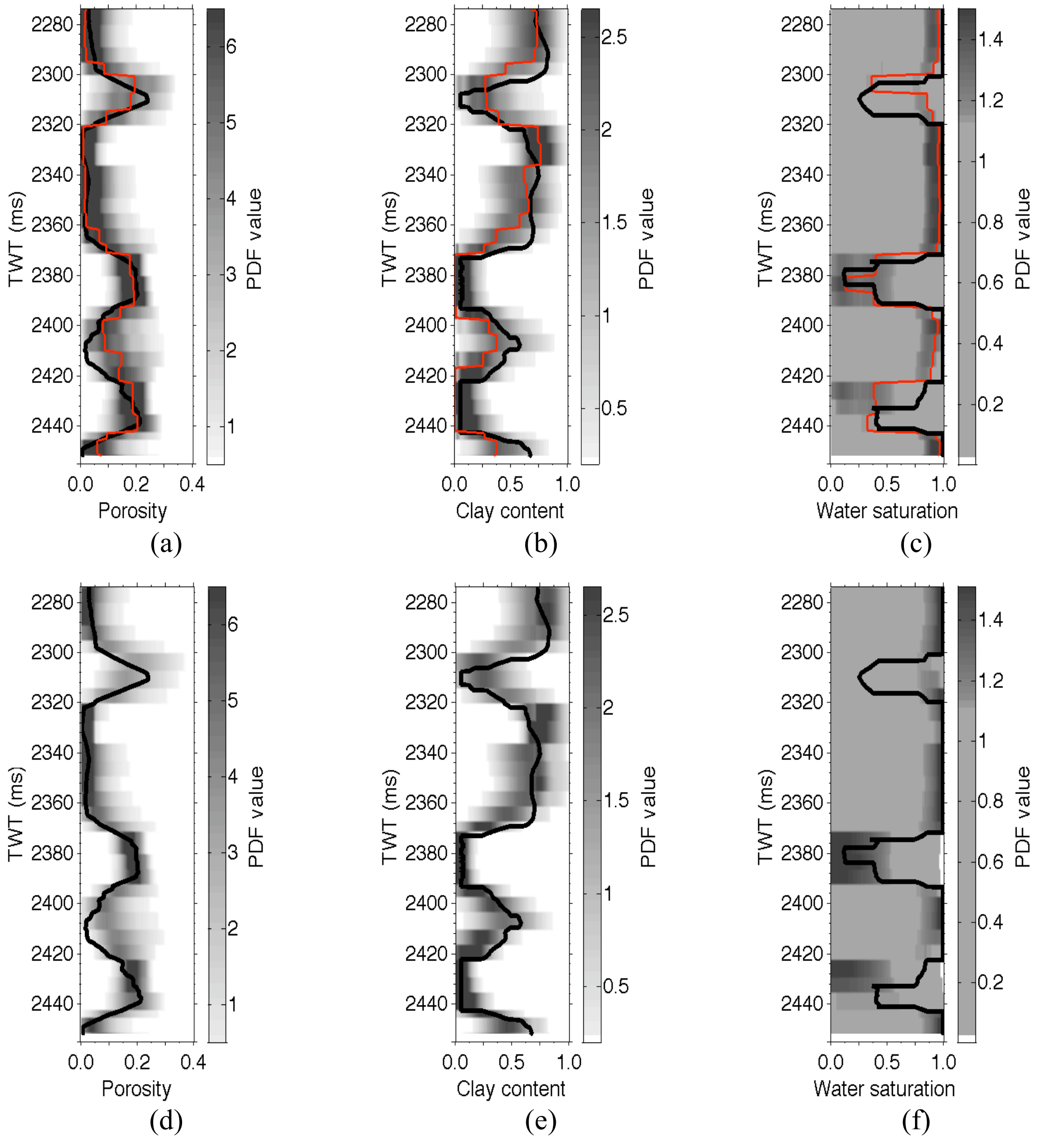


Figure 4-7: Marginal a posteriori PDF's of the model parameters from the inversion of seismically derived I_p and I_s for the depth interval 2275 ms–2455 ms in the same well as in figure 4-8. First row is the MDN solution: (a) Porosity. (b) Clay content. (c) Water saturation. The solid red line is the maximum a posteriori (MAP) of the marginal PDF's derived from the petrophysical inversion of seismic data. Second row is the Monte Carlo sampling solution: (d) Porosity. (e) Clay content. (f) Water saturation. Solid black line is the low-resolution log in the both rows. The scale of the TWT axis of the plots is from real data but the numbers are fictitious.

Figure 4-8(a) shows the MAP of the marginal PDF of porosity obtained from inverting the I_p and I_s cross-sections in figure 4-1. Figure 4-8(b) shows the standard deviation of the marginal distribution of porosity estimates in this cross-section. The highest value of the color bar corresponds to the a priori standard deviation of porosity, and throughout the cross-section the a posteriori standard deviation is smaller than this value (the color of the a priori

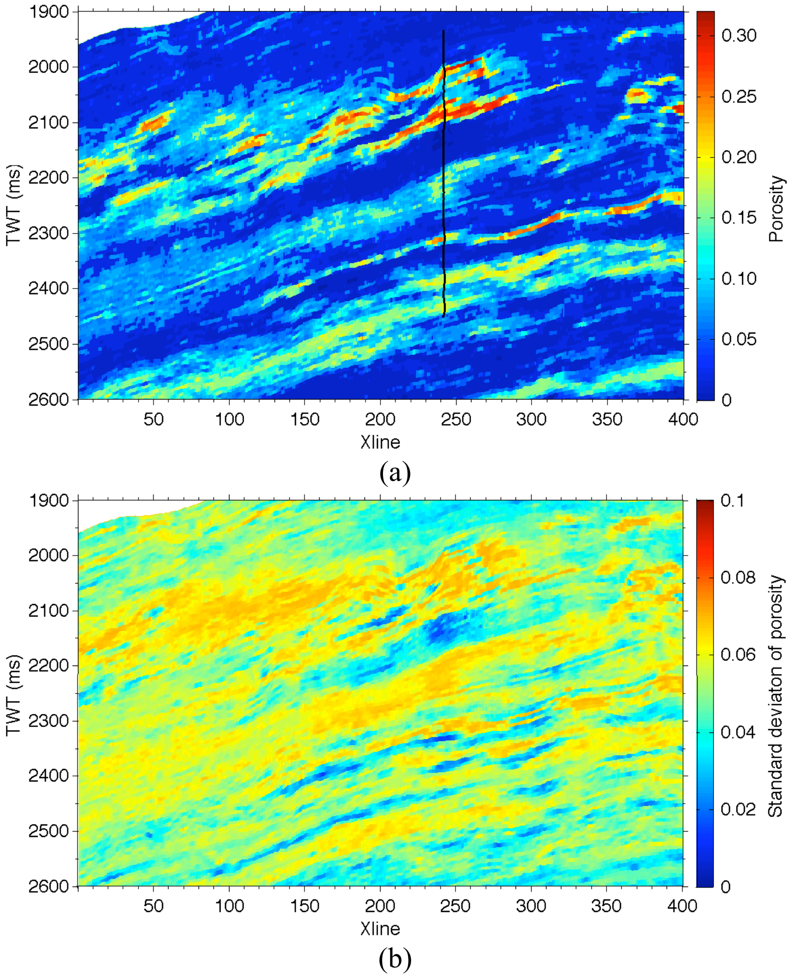
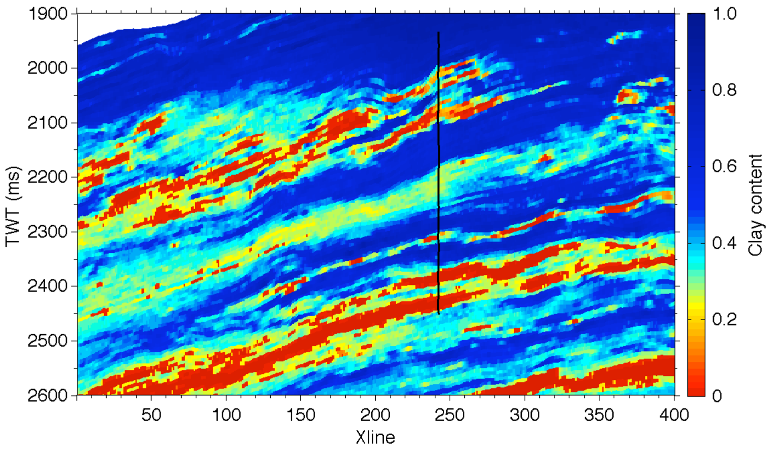


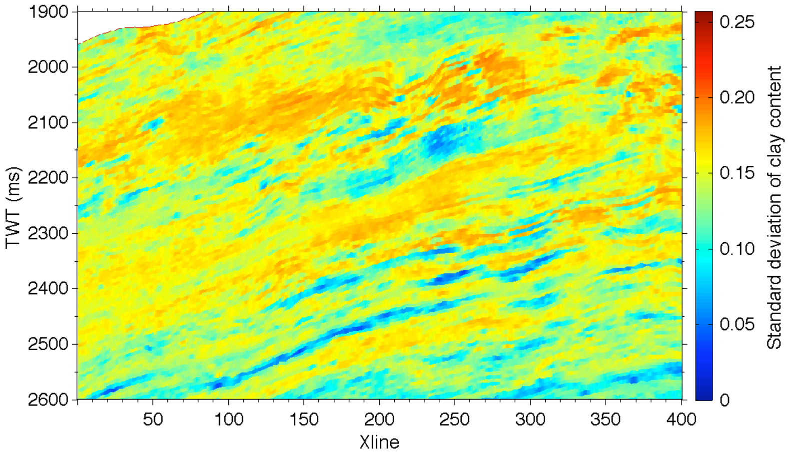
Figure 4-8: Estimated porosity and its associated uncertainty for one cross-section from the seismic cube, which includes one of the wells. (a) Maximum a posteriori of the marginal porosity PDF. (b) Standard deviation of porosity. The scale of the TWT and Xline axes of the plots are from real data but the numbers are fictitious.

standard deviation is dark red, and the hottest color in figure 4-8(b) is light orange). Figure 4-9(a) and figure 4-10(a) show the MAP of the marginal PDF's of clay content and water saturation, respectively, obtained from inverting I_p and I_s sections in figure 4-1. Figure 4-9(b) and figure 4-10(b) show the standard deviations of clay content and water saturation in this section. Again, the highest values of the color bars correspond to the a priori standard deviation values and it is evident that for all data points the a posteriori standard deviation is smaller than this value.

Figure 4-11 shows the MAP of the model parameters for 30 neighboring traces of the well in the in line section perpendicular to the previous cross line section. Figure 4-8, figure 4-9 and figure 4-10 in addition to figure 4-11 show this method gives a 3D estimate of the underground rock and fluid properties. As indicated previously, in addition to the MAP and



(a)

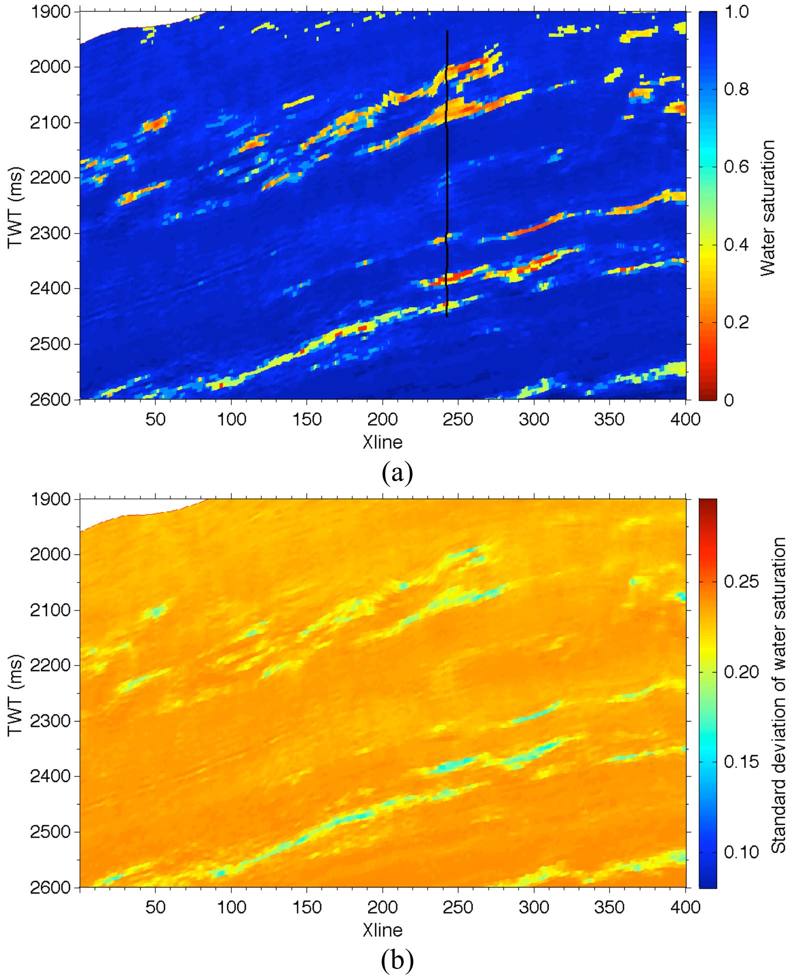


(b)

Figure 4-9: Estimated clay content and its associated uncertainty for one cross-section from the seismic cube, which includes one of the wells. (a) Maximum a posteriori of the marginal clay content PDF. (b) Standard deviation of clay content. The scale of the TWT and Xline axes of the plots are from real data but the numbers are fictitious.

standard deviation of the model parameters, any other statistical properties of the model parameters can be calculated from the estimated a posteriori joint PDF $p(\phi_e, c, s_w \mid I_p, I_s)$ for any location spanned by the 3D seismic data set.

Figure 4-12(a) shows the probability of the non-reservoir and, oil- and brine-saturated reservoir facies obtained from the integration of the estimated joint PDF of the model parameters (equation (4-3)). Figure 4-12(b) shows the most probable facies and figure 4-12(c, d) show the observed facies and water saturation logs at the well location. Comparison between figure 4-12(b) and figure 4-12(c, d) shows that the facies map obtained from seismic data is comparable with well observations. However, in many cases the facies map obtained from petrophysical inversion predicts brine-saturated reservoir facies instead of shale facies.



(a)

(b)

Figure 4-10: Estimated water saturation and its associated uncertainty for one cross-section of the seismic cube, which includes one of the wells. (a) Maximum a posteriori of the marginal water saturation PDF. (b) Standard deviation of water saturation. The scale of the TWT and Xline axes of the plots are from real data but the numbers are fictitious.

Figure 4-13 shows the probability of facies obtained from the petrophysical inversion of I_p and I_s sections in figure 4-1. As we expected figure 4-12(a) and figure 4-13 show that the uncertainty of fluid discrimination from seismic data is high. Figure 4-14 shows the facies probability for 30 neighboring traces of the well in the inline section perpendicular to the cross line section in figure 4-13. Again figure 4-13 and figure 4-14 show this method can give a detailed 3D description of rock facies probabilities within seismic resolution.

4.5 Discussions

Figure 4-6 and figure 4-7 show that the MAP of the marginal a posteriori PDF of the model parameters is a reasonable estimate of the values of the low-resolution logs. The difference between this estimate and the low-resolution logs can be related to the accuracy of inverted

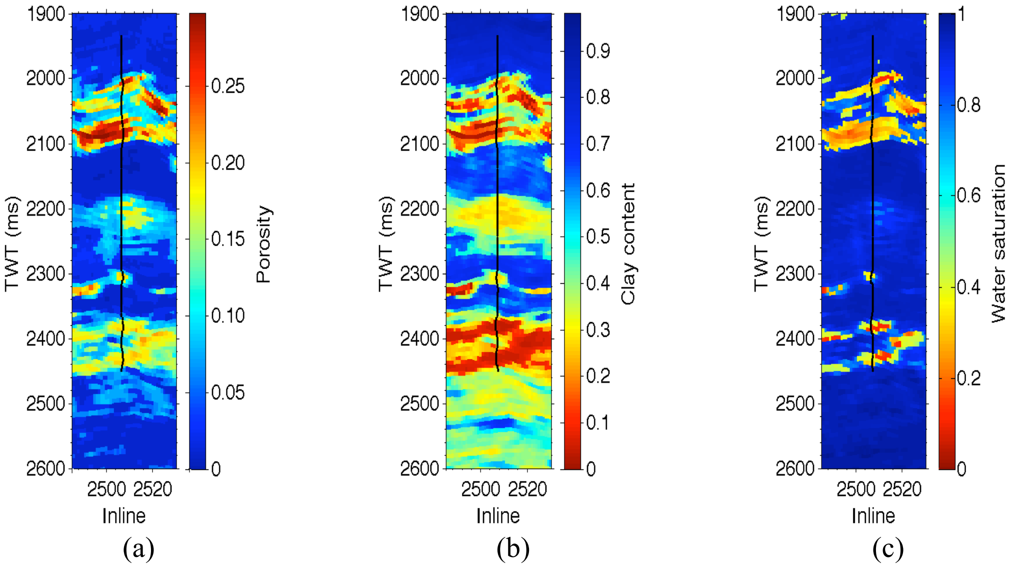
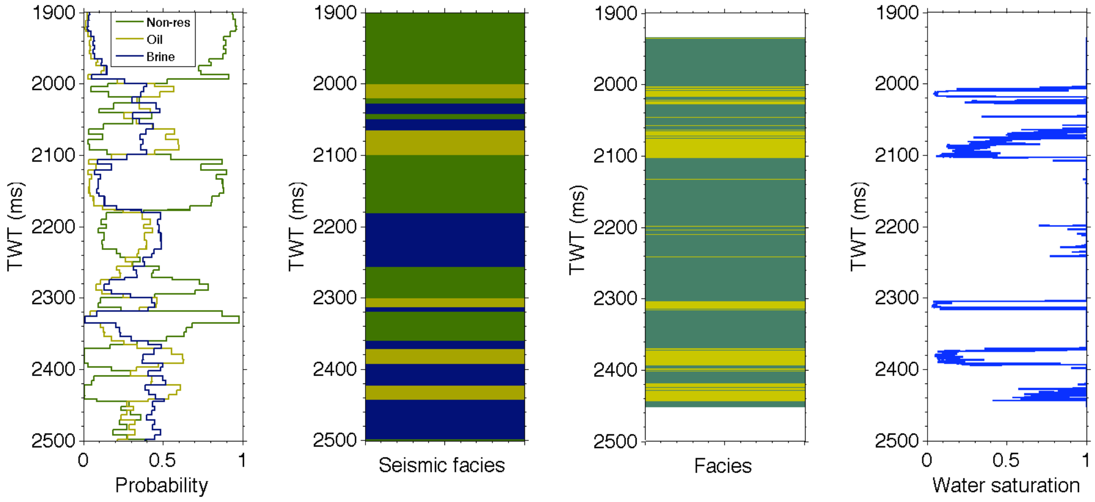


Figure 4-11: Estimated model parameters around the well in the in line section. (a) Porosity. (b) Clay content. (c) Water saturation. The scales of the TWT and inline axes of the plots are from real data but the numbers are fictitious.

seismic I_p and I_s, and the accuracy of the forward petrophysical relations. In the fourth application in Chapter 3 (section 3.4.4), we discussed the effect of the accuracy of the petrophysical forward relations on the petrophysical inversion result of seismic scale logs. The same effect is existed in the inversion of the seismic attributes because we apply the same petrophysical forward relations in the both cases. Any improvement in the accuracy of the petrophysical forward relations can reduce the effect of this type of error.

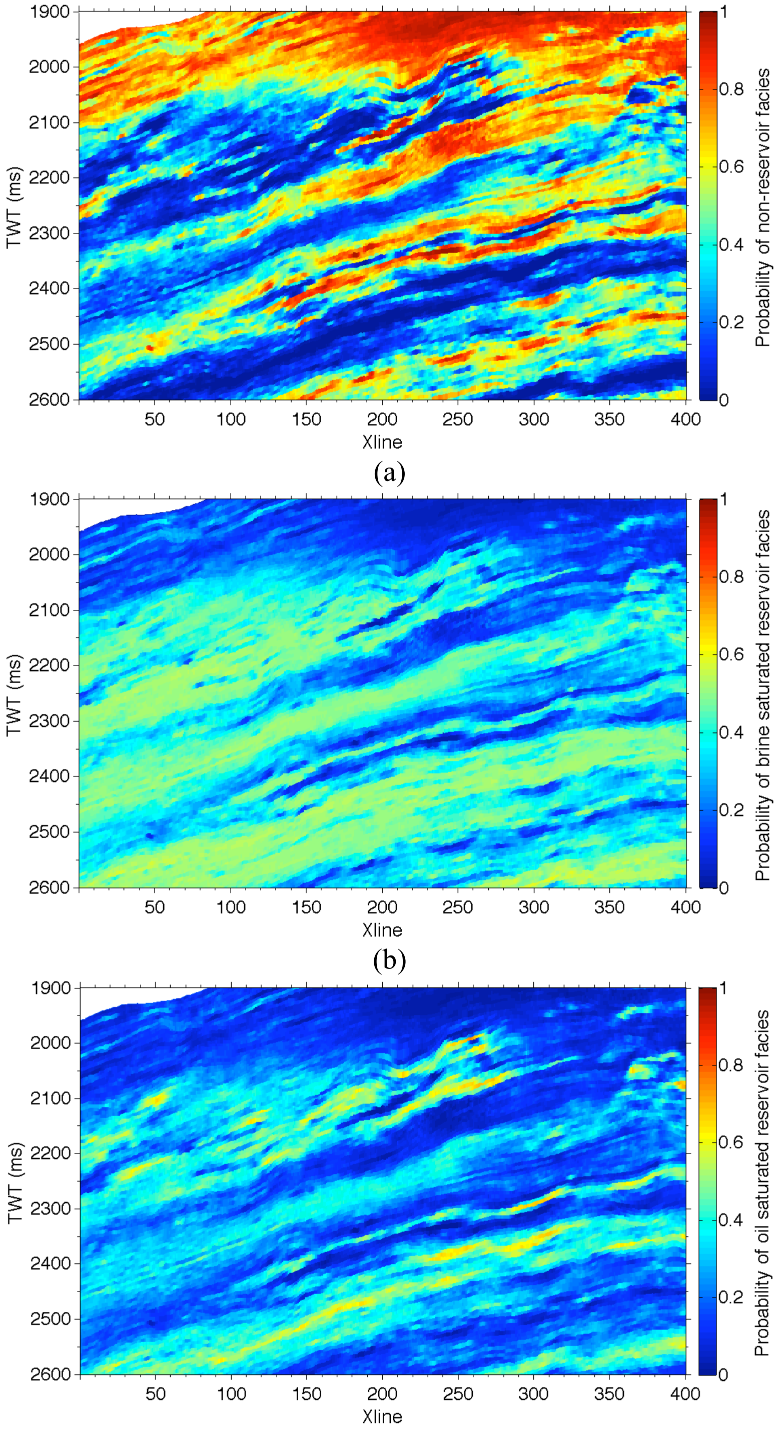
Figure 4-15 shows I_p and I_s obtained from seismic data and compare it with upscaled measured logs for the two intervals of figure 4-6 and figure 4-7. This figure shows that seismic inversion for I_s in the interval 2025 ms–2070 ms has large biases (in some cases near 25%; around two times larger than the 14% error estimated as the sum of the theoretical and measurement errors above). Therefore, we observe large errors in the inverted values of porosity and clay content in that interval in figure 4-6(a) and (b). We also see large errors in the inverted values of I_p in the shale interval 2100 ms–2150 ms, which result in errors in the inverted values of clay content in the same interval in figure 4-6(b). The errors in the above intervals show the effect of the accuracy of the processed seismic data on the petrophysical inversion result. Any improvement in the accuracy of the inverted seismic data can reduce the effect of this type of error.



(a) (b) (c) (d)

Figure 4-12: Facies inversion results at the well location. (a) Probability of different facies. Green curve represents non-reservoir facies. Blue represents brine saturated reservoir facies. Yellow represents oil saturated reservoir facies. (b) Maximum probability seismic facies at the well location. (c) Facies logs at the well location shales represented by green, sandstone, shaley sandstone, and sandy shales are shown by yellow. (d) Water saturation well log. The scales of the TWT axes of the plots are from real data but the numbers are fictitious.

Figure 4-5, figure 4-6 and figure 4-7 show that the uncertainty of porosity, clay content, and water saturation is large. The high uncertainty of the model parameters stems from uncertainty about the confounding model parameters (i.e., effective pressure, bulk modulus, and density of hydrocarbon), uncertainty of the petrophysical forward relations, and measurement uncertainty in seismic I_p and I_s. In addition of course, we invert only two data for three parameters, and hence simply by appealing to dimensionality arguments a unique solution is impossible without strong a priori information. Figure 4-2 and figure 4-3 show the effect of theoretical and measurement uncertainty in addition to the effect of the confounding model parameters on the uncertainty of I_p and I_s. In particular for water saturation, in figure 4-2(f), we observe that uncertainty in the confounding model parameters results in large uncertainty in I_p values. Figure 4-2(f) shows that even when porosity and clay content values are known, for a given value of I_p, the uncertainty of water saturation is large. This means that the reduction in uncertainty in the water saturation from the petrophysical inversion should be small (Figure 4-5, figure 4-6 and figure 4-7). Nevertheless the inversion process does reduce the uncertainty of all model parameters as can be seen in figure 4-8(b), figure 4-9(b) and figure 4-10(b).



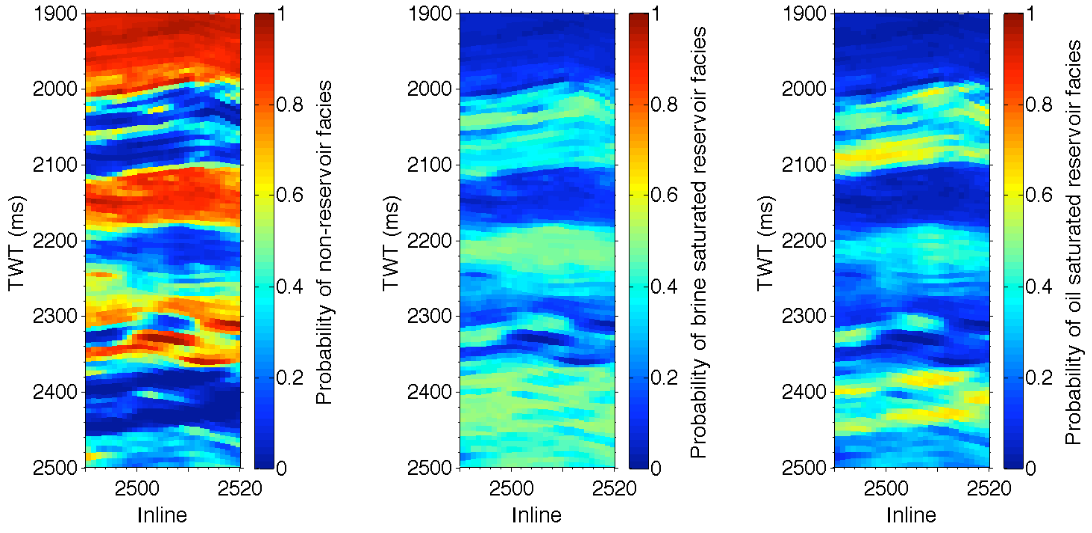
(a)

(b)

(c)

Figure 4-13: Probability of rock facies obtained from the petrophysical inversion. (a) Probability of non-reservoir facies. (b) Probability of the brine-saturated reservoir facies. (c) Probability of the oil-saturated reservoir facies. The scales of the TWT and Xline axes of the plots are from real data but the numbers are fictitious.

Figure 4-8(a), figure 4-9(a) and figure 4-10(a) in conjunction with figure 4-11 show that if this method is applied to invert 3D seismic I_p and I_s sections, the estimates of porosity, clay

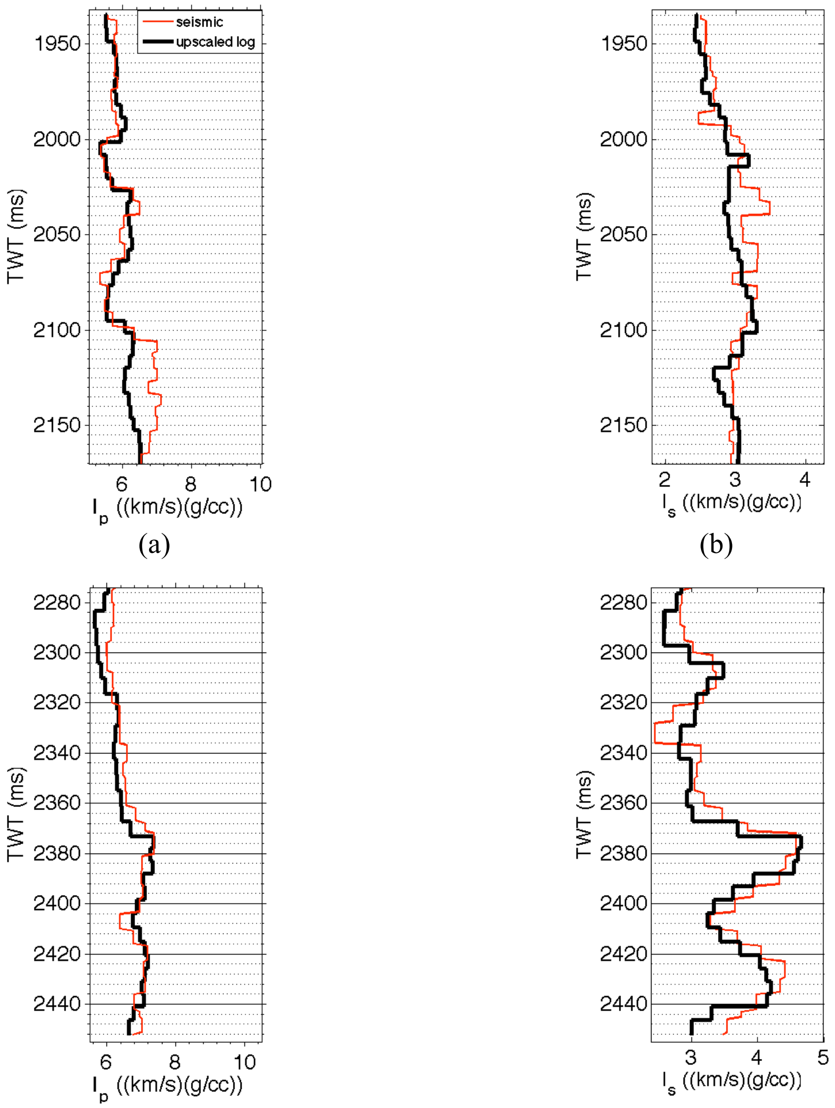


(a) (b) (c)

Figure 4-14: Probability of rock facies obtained from the petrophysical inversion around the well in the in line section. (a) Probability of non-reservoir facies. (b) Probability of the brine-saturated reservoir facies. (c) Probability of the oil-saturated reservoir facies. The scales of the TWT and inline axes of the plots are from real data but the numbers are fictitious.

content, and water saturation can provide a detailed 3D description of rock properties in a reservoir. This is of great value and importance for exploration and reservoir development plans as it can help to locate possible sources of hydrocarbon inside a reservoir. The accuracy and resolution of this description depends on the accuracy and resolution of 3D seismic impedance cubes, and on the accuracy of the petrophysical relationships used.

The estimated uncertainties in figure 4-8(b), figure 4-9(b) and figure 4-10(b) are mainly due to the effect of confounding model parameters, theoretical uncertainty of the petrophysical forward function, and measurement uncertainty of the inverted seismic P- and S-impedance. As we explained in the seismic data section, the measurement uncertainty is obtained from the difference between upscaled well logs and inverted seismic traces at the well locations. This approach to uncertainty measurement implies that the inverted seismic data is unbiased comparing to the upscaled well log. However, as figure 4-15 shows this assumption can be violated in the Akpo field. Several factors can contribute to the bias observed in figure 4-15, for example as mentioned above the intrinsic dispersion can cause such biases, another possible source of bias can be existence of low effective pressure zones which are missed by the back ground trend used in seismic inversion. Effects of such parameters were assumed to be nonsystematic error and modeled by a Gaussian PDF with zero mean. Therefore, the biases observed in figure 4-15 result in biased estimates of the petrophysical parameters. A more coherent approach to uncertainty estimation of the AVO inversion (Buland and Omre,



(a) (b)

(c) (d)

Figure 4-15: Comparison between seismically derived I_p and I_s (red) and upscaled measured logs (black). The first row is for the depth interval 1933 ms– 2170 ms: (a) I_p. (b) I_s. The second row is for the depth interval 2274 ms– 2455 ms: (c) I_p. (d) I_s. The scale of the TWT axis of the plots is from real data but the numbers are fictitious.

2003; Buland et al., 2003) result can reduce the biases and errors introduced due to the above parameters.

In the same manner as the measurement uncertainty, the theoretical uncertainty is also assumed to be unbiased and is modeled using a Gaussian PDF with zero mean. In chapter 2 we discussed that due to application of Reuss average in several stages of the petrophysical forward modeling, the estimated P- and S-wave velocity obtained from the petrophysical forward function can be biased. In addition, the petrophysical forward model applied in this chapter, and introduced in Appendix C, is a single facies model. Therefore, it cannot model the facies uncertainty properly. The above sources of error result in biased predictions of the petrophysical forward function in figure 3-10 that consequently result in biased estimate of the petrophysical parameters in figure 4-6 and figure 4-7. More geological a priori information can reduce biases in the petrophysical forward function predictions and consequently decrease errors and biases in the petrophysical inversion result due to the above parameters.

Comparison between figure 4-12(b) and figure 4-12(c, d) shows that the method can detect the oil-saturated reservoir facies with good accuracy. Figure 4-12(a) show that the probabilities of oil- and brine-saturated reservoir facies are very similar. The cause of this ambiguity in the type of the fluid is the high uncertainty in the a posteriori estimate of water saturation, which we discussed above. The ambiguity of the fluid type can result in unrealistic realization of facies sequence at the well location. For example in figure 4-12(b), in the interval 2050 ms–2070 ms, we observe brine-saturated reservoir facies occurs above oil-saturated reservoir facies, which is physically impossible. However, figure 4-12(a) shows that at this interval the probability of oil-saturated reservoir facies is slightly smaller than the brine-saturated reservoir facies and therefore, the uncertainty of this prediction is high. The unrealistic realization of facies sequences can also be due to the resolution of seismic data. For example in figure 4-12(b), in the interval 2360 ms–2380 ms we also observe brine-saturated reservoir facies on the top of the oil-saturated reservoir facies. Facies log and water saturation log in figure 4-12(c, d) shows that a very thin layer of shale separates brine- and oil-saturated reservoirs at that interval.

The transform from petrophysical inversion results to facies probabilities causes ambiguities between brine-saturated reservoir facies and non-reservoir facies. Comparison between observed facies log in figure 4-12(c, d) and estimated seismic facies log in figure 4-12(b) shows these ambiguities in the intervals 2030 ms–2060 ms, 2180 ms–2260 ms, and 2390 ms–2420 ms. The main cause of this error is the error in the petrophysical inversion as can be seen in figure 4-6 and figure 4-7 for the first and third intervals above. In particular error in the interval 2180 ms–2260 ms is due to petrophysical forward function error on the corresponding depth interval 3180 m–3280 m in figure 3-10. Errors in the intervals 2030 ms–2060 ms and 2390 ms–2420 ms are due to errors in the seismic estimates of I_s on the same intervals in figure 4-15(a) and figure 4-15(b), respectively. All of the above errors could be reduced if more accurate petrophysical forward function and seismic data were used.

Due to large uncertainty associated with data, forward petrophysical relations, and confounding model parameters the a posteriori PDF of the model parameters and especially water saturation will always have large uncertainty. Therefore, in order to address this uncertainty appropriately, any petrophysical inversion method must be probabilistic. The MDN method is a time and memory efficient method for probabilistic nonlinear inversion. The nonlinear inversion of each cross line section, which included 170,322 data points and resulted in full joint posterior PDF's for ϕ_e, c, and s_w, took 340 seconds on a regular desktop computer. The number of cross line sections is 1461 so inverting the whole 3D seismic cube with 248,840,442 data points takes only around 138 hours on the same desktop computer. What is more, the mixture density neural network encoded the full joint PDF of all model parameters for all data points into only 2041 scalar values (i.e., the number of MDN weights), which requires 20.5 KBytes memory for storage. Note that there is a trade-off between time and memory efficiency of the MDN inversion technique, and the accuracy of the a posteriori PDF of the model parameters estimated by this technique (Appendix A). An MDN with a larger number of kernels usually results in a smaller training and validation error (Table 3-2), however, the time and memory required to train such a network increase exponentially. Another drawback of using MDN is that, the selection of the number of kernels and hidden units is a trial and error procedure, which can take a long time. However, this time is usually much smaller than the time required for petrophysical inversion using MC sampling method.

The inversion of 248,840,442 data points in the seismic cube using the Monte Carlo sampling method we employed in the Results section will take around 27,500 hours (almost 3 years) on the same desktop computer. The storage of the full joint PDF of ϕ_e, c, and s_w for each spatial data point when Monte Carlo sampling is used requires around 3.2 MBytes which means that for all data points around 2.4 TBytes (2.4×10^{12} Bytes) of memory is required. Using conventional inversion methods such as MC sampling to invert 3D seismic cubes is thus impractical unless large parallel computing resources are available. To our knowledge this application is the first fully probabilistic, nonlinear, petrophysical inversion method that is applicable to large seismic cubes on a standard desktop computer.

4.6 Conclusions

The MDN inversion method of the previous chapter has been applied to jointly invert compressional- and shear-wave impedances for the joint probability density function (PDF) of porosity, clay content, and water saturation in a 3D seismic cube. The resulting property

PDF’s are obtained from the integration of geophysical information, well logs, and rock physics information in an exploration setting. In order to make the a priori assumptions about the model parameters as noncommittal as possible, following the approach of Bachrach (2006) and Spikes et al. (2007) we assume a Uniform a priori PDF for the model parameters. We also indicate that the information about vertical and spatial geological continuity has been incorporated in the process of the seismic inversion for I_p and I_s, and hence indirectly constrains the final property maps.

The inversion results at a well location show that the maximum a posteriori (MAP) point of the marginal PDF’s of the model parameters provide reasonable estimates of the corresponding low-resolution log values where these are available. Residual errors correspond to uncertainty in the forward petrophysical relations, uncertainty of the seismic inversion process to obtain I_p and I_s, and differences between frequency content and support volume of log measurements and seismic measurements. The uncertainty in the a posteriori PDF of water saturation is high, and this can be explained by relatively low sensitivity of the petrophysical forward function to variations of water saturation in the presence of seismic inversion uncertainty, and uncertainty in the confounding model parameters. Nevertheless the inversion process reduces the a priori uncertainty of all model parameters.

The result of the petrophysical inversion gives a detailed description (to within the seismic resolution) of rock and fluid properties in the reservoir that can be used for exploration and development planning. In particular, it can be used to find areas with high effective porosity and low clay content (pay zones), and also areas with possible sources of hydrocarbon based on the estimated water saturation.

The probability of facies can also be estimated from the petrophysical inversion result. Lithofacies in this chapter are defined based on the value of petrophysical parameters (i.e., effective porosity, clay content, and water saturation)–in the same way as they are defined in the construction of the petrophysical forward function (Chapter 2; Appendix C). Facies probability maps obtained in this way integrates geophysical information, well logs, well observations, and rock physics information.

Predicted facies sequences at the well location show that the method can predict oil-saturated reservoir facies with a good accuracy. Due to errors in the petrophysical inversion, in some areas we observe ambiguities between non-reservoir and brine-saturated reservoir facies. These errors could be reduced if a more accurate petrophysical forward function and seismic data were used.

The advantages of the MDN inversion method over other probabilistic inversion methods are its memory and computational efficiency. Due to the large size of 3D seismic cubes these two properties are essential for any nonlinear probabilistic petrophysical inversion. As the inversion method is independent of the particular seismic attributes chosen in this case study (I_p and I_s), it can be used to invert any set of pertinent seismic attributes such as compressional- and shear-wave velocity and bulk density, or compressional impedance and Poisson's ratio. To our knowledge this application is the first fully probabilistic, nonlinear petrophysical inversion method that is applicable to large 3D seismic cubes on a standard desktop computer.

4.7 References

Aki, K. and P. G. Richards, 1997, *Quantitative Seismology: Theory and Methods*: University Science Books,U.S.

Avseth, P., T. Mukerji, A. Jorstad, G. Mavko and T. Veggeland, 2001, Seismic Reservoir Mapping from 3-D AVO in a North Sea Turbidite System, *Geophysics*, **66**, 1157-1176.

Avseth, P., T. Mukerji and G. Mavko, 2005, *Quantitative seismic interpretation: applying rock physics tools to reduce interpretation risk*. Cambridge ; New York: Cambridge University Press.

Bachrach, R., 2006, Joint estimation of porosity and saturation using stochastic rock-physics modeling, *Geophysics*, **71**, O53-O63.

Bosch, M., L. Cara, J. Rodrigues, A. Navarro and M. Diaz, 2007, A Monte Carlo approach to the joint estimation of reservoir and elastic parameters from seismic amplitudes, *Geophysics*, **72**(6), O29-O39.

Bosch, M., C. Carvajal, J. Rodrigues, A. Torres, M. Aldana and J. Sierra, 2009, Petrophysical seismic inversion conditioned to well-log data: Methods and application to a gas reservoir, *Geophysics*, **74**(2), O1-O15.

Brown, R. L. and D. Seifer, 1997, Velocity dispersion: A tool for characterizing reservoir rocks, *Geophysics*, **62**(2), 477-486.

Buland, A., O. Kolbjornsen, R. Hauge, O. Skjaeveland and K. Duffaut, 2008, Bayesian lithology and fluid prediction from seismic prestack data, *Geophysics*, **73**(3), C13-C21.

Buland, A., O. Kolbjornsen and H. Omre, 2003, Rapid spatially coupled AVO inversion in the Fourier domain, *Geophysics*, **68**(3), 824-836.

Buland, A. and H. Omre, 2003, Bayesian linearized AVO inversion, *Geophysics*, **68**(1), 185-198.

Devilee, R. J. R., A. Curtis and K. Roy-Chowdhury, 1999, An efficient, probabilistic neural network approach to solving inverse problems: Inverting surface wave velocities for Eurasian crustal thickness, *Journal of Geophysical Research*, **104(B12)**, 28841-28856.

Doyen, P. M., 1988, Porosity from seismic data: A geostatistical approach, *Geophysics*, **53**, 1263-1275.

Dubrule, O., 2003, *Geostatistics for seismic data integration in earth models*: Society of Exploration Geophysicists.

Eidsvik, J., P. Avseth, H. Omre, T. Mukerji and G. Mavko, 2004, Stochastic Reservoir Characterization using Prestack Seismic Data, *Geophysics*, **69**, 780-993.

Gonzalez, E. F., T. Mukerji and G. Mavko, 2008, Seismic inversion combining rock physics and multiple-point geostatistics, *Geophysics*, **73**(1), R11-R21.

Grana, D. and E. Della Rossa, 2010, Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion, *Geophysics*, **75**(3), O21-O37.

Larsen, A. L., M. Ulvmoen, H. Omre and A. Buland, 2006, Bayesian lithology/fluid prediction and simulation on the basis of a Markov-chain prior model, *Geophysics*, **71**(5), R69-R78.

Mavko, G., T. Mukerji and J. Dvorkin, 2009, *The rock physics handbook: tools for seismic analysis of porous media*. Cambridge ; New York: Cambridge University Press.

Meier, U., A. Curtis and J. Trampert, 2007a, A global crustal model constrained by non-linearised inversion of fundamental mode surface waves, *Geophysical Research Letters*, **34**, L16304-L16304.

Meier, U., A. Curtis and J. Trampert, 2007b, Global crustal thickness from neural network inversion of surface wave data, *Geophysical Journal International*, **169**, 706-722.

Meier, U., J. Trampert and A. Curtis, 2009, Global variations of temperature and water content in the mantle transition zone from higher mode surface waves, *Earth Planetary Science Letters*, **282**, 91-101.

Mosegaard, K. and A. Tarantola, 2002, Probabilistic approach to inverse problems, *International handbook of earthquake and engineering seismology*. W. H. K. Lee, H. Kanamori, P. C. Jennings and C. Kisslinger: Academic Press Inc.**:** 237-265.

Mukerji, T., A. Jorstad, P. Avseth, G. Mavko and J. R. Granli, 2001, Mapping lithofacies and pore-fluid probabilities in a North Sea reservoir: Seismic inversions and statistical rock physics, *Geophysics*, **66**, 988-1001.

Sambridge, M. and K. Mosegaard, 2002, Monte Carlo methods in geophysical inverse problems, *Reviews of Geophysics*, **40**(3), 1009.

Sen, M. K. and P. L. Stoffa, 1991, Nonlinear one-dimensional seismic waveform inversion using simulated annealing, *Geophysics*, **56**(10), 1624-1638.

Spikes, K., T. Mukerji, J. Dvorkin and G. Mavko, 2007, Probabilistic seismic inversion based on rock-physics models, *Geophysics*, **72**(5), R87-R97.

Tarantola, A., 2005, *Inverse Problem Theory and Methods for Model Parameter Estimation*: SIAM.

Tarantola, A. and B. Valette, 1982, Inverse Problems = Quest for Information, *Journal of Geophysics*, **50**, 159-170.

Ulvmoen, M. and H. Omre, 2010, Improved resolution in Bayesian lithology/fluid inversion from prestack seismic data and well observations: Part 1 --- Methodology, *Geophysics*, **75**(2), R21-R35.

Ulvmoen, M., H. Omre and A. Buland, 2010, Improved resolution in Bayesian lithology/fluid inversion from prestack seismic data and well observations: Part 2 --- Real case study, *Geophysics*, **75**(2), B73-B82.

5 Discussion

> *"It is not unreasonable that we grapple with problems.*
> […] *Our responsibility is to do what we can, learn what*
> *we can, improve the solutions, and pass them on."*
>
> Richard Feynman

In this thesis we propose an efficient method to solve petrophysical inverse problems. As we observed in chapter 4, the accuracy of the solution of the petrophysical inversion depends on the accuracy of the petrophysical forward function, the seismic data, and uncertainty in the confounding model parameters of the forward function. Uncertainty in the confounding parameters depends on the a priori information about these parameters. For example, in previous applications of the petrophysical inversion of seismic data (Bachrach, 2006; Spikes et al., 2007; Bosch et al., 2009; Grana and Della Rossa, 2010) a constant effective pressure is assumed and the inversion is performed over a subset of seismic data (e.g., a short depth window). This assumption reduces uncertainty in the a posteriori PDF of solution of the petrophysical inversion. However, selection of such intervals of relatively constant pressure requires strong a priori information about the distribution of effective pressure in the field and in many locations such information will not be available.

As we discussed in chapter 4, the main effect of uncertainty in confounding model parameters is on the uncertainty of water saturation. Due to the low sensitivity of compressional- and shear-wave impedance to water saturation, in the presence of uncertainty in confounding parameters, inference about water saturation in the petrophysical inversion of seismic data is highly ambiguous. High uncertainty in the a posteriori PDF of water saturation has also been observed in the previous applications of the petrophysical inversion of seismic data (Bachrach, 2006; Spikes et al., 2007; Grana and Della Rossa, 2010). Bosch et al. (2009) show that application of geostatistical methods might decrease water saturation uncertainty.

As we discussed in chapter 4, uncertainties in seismic attributes (e.g., compressional- and shear-wave impedance) also increase uncertainty in the a posteriori information about the petrophysical parameters in a petrophysical inverse problem. In particular, we observed that biases in the inverted values of the seismic attributes result in biases in the inverted values of the petrophysical parameters at the well location. Any improvement in the accuracy of the inverted seismic attributes might decrease the aforementioned biases in the inverted petrophysical parameters. Hence, one possible approach to reduce such uncertainty is to perform petrophysical inversion directly from (AVO) seismic data. This inversion can be a one step inversion where AVO seismic data are inverted for the petrophysical parameters (Spikes et al., 2007). In this approach the AVO seismic data can be used to reduce biases in the seismic attributes and consequently in the petrophysical parameters at the well locations.

As we observed in chapters 3 and 4, in the petrophysical inversion of seismic data uncertainty in the petrophysical forward function is another source of uncertainty in the a posteriori petrophysical parameters. Biases in the petrophysical forward function predictions result in biases in the inverted values of the petrophysical parameters at the well location.

As we discussed in chapter 2, simplifying assumptions about facies uncertainty can result in large biases in the predictions of the petrophysical forward function. Results of chapter 2 show that counter-intuitively, introducing uncertainty in the facies reduces uncertainty in the petrophysical forward function. This is because without this additional variations in facies, an incorrect petrophysical model is fit to well intervals; increasing estimated uncertainty is the only way the predictions can be made compatible with the well data. We also indicate that geological a priori information and core observations are critical information in the construction of accurate petrophysical forward relations. Uncertainty in the petrophysical forward function can be reduced further if more prior geological information about facies is applied. Also, the presented method can be used with well logs from several other wells in Akpo field to construct a general petrophysical forward function with smaller uncertainty than the forward function applied in the chapters 3 and 4.

The prior geological information can improve the accuracy of the probabilistic definition of facies in the petrophysical forward function. In the previous examples of facies inversion of seismic data (Avseth et al., 2001; Mukerji et al., 2001; Larsen et al., 2006; Gonzalez et al., 2008; Ulvmoen and Omre, 2010; Ulvmoen et al., 2010), seismic attributes (e.g., compressional and shear-wave impedance) are directly inverted for facies. In our approach to facies inversion in chapter 4, we explicitly add information in rock physics theories to constrain the final facies maps obtained from seismic data. However, in chapter 4 we observe that errors in the petrophysical forward function can result in errors in the final facies maps.

The applied inversion method in chapter 4 integrates geophysical data with petrophysical theories and well observations to estimate petrophysical parameters with their associated uncertainties. In chapter 3 (Section 3.4.2) we apply the MDN inversion method to estimate petrophysical parameters statistically, without applying any petrophysical forward function. As we discussed in chapter 3, the MDN trained with such a statistical method is applicable to invert seismic attributes, only if its results are confirmed at blind wells. This statistical inversion method can be further developed to include spatial variations of the petrophysical parameters to further reduce uncertainty in the a posteriori PDF of the petrophysical parameters (Caers and Ma, 2002; Saggaf et al., 2003).

The comparison between the petrophysical inversion results using the petrophysical forward function (Section 3.4.3) and without using the petrophysical forward function (Section 3.4.2), show that although in principle applying the petrophysical forward function must reduce uncertainty in the petrophysical inversion result, in practice due to errors in the petrophysical forward function we might not see any perceivable reduction in the a posteriori uncertainty in the petrophysical parameters. However, inverting for some petrophysical parameters, and in particular fluid saturations, requires applying theoretical information in a petrophysical forward function.

The MDN inversion is a computationally efficient method for solving nonlinear, multidimensional, and repeated inverse problems. As we discussed in chapter 3, the error in the MDN estimate of a sharp-edged PDF can be large. It might be possible to decrease this error by applying other types (other than spherical or diagonal Gaussian) of the kernels in the MDN structure. For example, it might be possible to extend the work of Williams (1996) to more than one Gaussian kernel with a full covariance matrix. However, we must note that application of more complicated kernels in the MDN can significantly increase the training time and memory usage of the network. In chapters 3 and 4, we observe that another source of error in the MDN inversion result is the multi-modality of the Gaussian mixture model of a PDF. In Appendix A, we show that applying a committee of networks can reduce the aforementioned error. Another basic difficulty in applying the MDN inversion method is the selection of the required number of kernels and hidden units. Trial and error procedure to select the appropriate number of the above parameters of the MDN can be time consuming, however, this is compensated with the efficiency of the inversion step. In addition, as we indicate in Appendix A, trained networks in the trial and error procedure might be used in a committee of networks to decrease the error of the MDN inversion result.

In this thesis, we have not considered the possibility of integration of the probabilistic petrophysical inversion result with geostatistical a priori models. The probabilistic result of the petrophysical inversion, presented in this thesis, is independent of any purely geostatistical data (i.e., geostatistical data which are independent of seismic data, for example kriging models of the petrophysical parameters obtained from well logs). Therefore, probabilistic integration of geostatistical data with the petrophysical inversion result might decrease the a posteriori uncertainty of the petrophysical parameters.

MDN's have already been successful in regional and global seismology (Meier et al., 2007a,b; 2009). Hence, while this thesis contributes to extend the domain of seismological application even further, it also suggests that MDN inversion is a fairly generally applicable method to inverse problems that can be decomposed into very many smaller sub-problems. Hence, MDN could be applicable to rock boundaries identification using well log data (Maiti and Tiwari, 2009) or 4D seismic inversion for estimation of changes in fluid saturation and effective pressure (Bachrach and Dutta, 2004).

The MDN inversion method can be used to invert any other set of seismic attributes for which a suitable training dataset can be measured or synthesized for example compressional-wave velocity or Poisson's ratio. It can also be used to invert seismic data for any other petrophysical parameter such as effective pressure. What is more, as shown in chapters 3 and 4, since it is fully Bayesian and probabilistic it can be used to invert (impose constraints on) for more unknowns (e.g., 3 petrophysical parameters) than there are data available (previously, 2 independent data at each point), without creating an ill-posed inverse problem. Although, the results may vary from one application to another, the fast, probabilistic, and robust method of solving non-linear inverse problems in conjunction with the probabilistic method of constructing a petrophysical forward function developed in this thesis can be applied fairly generally to invert well logs or large seismic datasets for the petrophysical parameters in any geophysical application.

5.1 References

Avseth, P., T. Mukerji, A. Jorstad, G. Mavko and T. Veggeland, 2001, Seismic Reservoir Mapping from 3-D AVO in a North Sea Turbidite System, *Geophysics*, **66**, 1157-1176.

Bachrach, R., 2006, Joint estimation of porosity and saturation using stochastic rock-physics modeling, *Geophysics*, **71**, O53-O63.

Bachrach, R. and N. Dutta, 2004, Joint estimation of porosity and saturation and of effective stress and saturation for 3D and 4D seismic reservoir characterization using stochastic rock physics modeling and Bayesian inversion, *SEG Technical Program Expanded Abstracts*, **23**(1), 1515-1518.

Bosch, M., C. Carvajal, J. Rodrigues, A. Torres, M. Aldana and J. Sierra, 2009, Petrophysical seismic inversion conditioned to well-log data: Methods and application to a gas reservoir, *Geophysics*, **74**(2), O1-O15.

Caers, J. and X. Ma, 2002, Modeling Conditional Distributions of Facies from Seismic Using Neural Nets, *Mathematical Geology*, **34**(2), 143-167.

Gonzalez, E. F., T. Mukerji and G. Mavko, 2008, Seismic inversion combining rock physics and multiple-point geostatistics, *Geophysics*, **73**(1), R11-R21.

Grana, D. and E. Della Rossa, 2010, Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion, *Geophysics*, **75**(3), O21-O37.

Larsen, A. L., M. Ulvmoen, H. Omre and A. Buland, 2006, Bayesian lithology/fluid prediction and simulation on the basis of a Markov-chain prior model, *Geophysics*, **71**(5), R69-R78.

Maiti, S. and R. K. Tiwari, 2009, A Hybrid Monte Carlo Method Based Artificial Neural Networks Approach for Rock Boundaries Identification: A Case Study from the KTB Bore Hole, *Pure and Applied Geophysics*, **166**(12), 2059-2090.

Meier, U., A. Curtis and J. Trampert, 2007a, A global crustal model constrained by non-linearised inversion of fundamental mode surface waves, *Geophysical Research Letters*, **34**, L16304-L16304.

Meier, U., A. Curtis and J. Trampert, 2007b, Global crustal thickness from neural network inversion of surface wave data, *Geophysical Journal International*, **169**, 706-722.

Meier, U., J. Trampert and A. Curtis, 2009, Global variations of temperature and water content in the mantle transition zone from higher mode surface waves, *Earth Planetary Science Letters*, **282**, 91-101.

Mukerji, T., A. Jorstad, P. Avseth, G. Mavko and J. R. Granli, 2001, Mapping lithofacies and pore-fluid probabilities in a North Sea reservoir: Seismic inversions and statistical rock physics, *Geophysics*, **66**, 988-1001.

Saggaf, M. M., M. N. Toksoz and H. M. Mustafa, 2003, Estimation of reservoir properties from seismic data by smooth neural networks, *Geophysics*, **68**(6), 1969-1983.

Spikes, K., T. Mukerji, J. Dvorkin and G. Mavko, 2007, Probabilistic seismic inversion based on rock-physics models, *Geophysics*, **72**(5), R87-R97.

Ulvmoen, M. and H. Omre, 2010, Improved resolution in Bayesian lithology/fluid inversion from prestack seismic data and well observations: Part 1 --- Methodology, *Geophysics*, **75**(2), R21-R35.

Ulvmoen, M., H. Omre and A. Buland, 2010, Improved resolution in Bayesian lithology/fluid inversion from prestack seismic data and well observations: Part 2 --- Real case study, *Geophysics*, **75**(2), B73-B82.

Williams, P. M., 1996, Using Neural Networks to Model Conditional Multivariate Densities, *Neural Computation*, **8**(4), 843-854.

6 Conclusions

In this thesis, we present a methodology to obtain petrophysical parameters (e.g., porosity, clay content, and water saturation) and facies probabilities from seismic attributes (e.g., compressional- and shear wave impedance). In this method, we first apply prior geological information and core observations to develop a general petrophysical forward function for a field. Facies uncertainty is also included in the petrophysical forward function. In the second step, we apply the Mixture Density Network (MDN) inversion method to invert seismic attribute well logs for the petrophysical parameters at well locations using the petrophysical forward function derived in the previous step. In the third step, comparison between the inversion results at wells and measured well logs of the petrophysical parameters are used to assess applicability of the petrophysical forward function for the inversion. In the fourth step, the MDN is trained with a new dataset that simulates uncertainty in the seismic data and in the petrophysical forward function, and then is applied to invert 3D cubes of seismic attributes for the probability density function (PDF) of the petrophysical parameters. Finally, the estimated PDF of the petrophysical parameters is applied in conjunction with the a priori geological information to obtain the probability of different facies defined in the petrophysical forward function.

Results of applying above method with well log, seismic data, and a given petrophysical forward function show that the inversion result can provide an acceptable estimate of the MC sampling solution. Uncertainty and error in the final MDN inversion result depend on uncertainty and error in the seismic attributes and in the petrophysical forward function. Any improvement in the above sources of uncertainty may decrease error in the MDN petrophysical inversion result. The main advantage of the MDN inversion method is its computational efficiency, which is of great utility in the inversion of 3D seismic cubes with up to hundreds of millions of data points.

The methodology and results presented in this thesis can be extended and improved in the following ways: 1) the petrophysical inversion results provided in this thesis might be improved if a more accurate petrophysical forward function, and a less uncertain seismic dataset is applied, 2) the a posteriori uncertainty of the petrophysical parameters might be reduced if the MDN inversion results provided in this thesis are integrated with the result of a geostatistical model of the petrophysical parameters, 3) the MDN estimates of the Monte Carlo sampled solution of the a posteriori PDF of model parameters in an inverse problem might be more accurate if more flexible kernels are applied within the MDN structure, and 4) the methodology presented in this thesis is applied with seismic data from a mainly siliciclastic reservoir; an extension of this work can be application of the method with data from a carbonate reservoir.

A Neural Networks

> *"Civilization advances by extending the number of important operations which we can perform without thinking about them."*
>
> Alfred N. Whitehead

A.1 Structure of neural networks

A neural network is a computational structure that is composed of many simple processing units called neurons. Figure A-1 shows a single neuron, labeled with index *j*. The input of the neuron *j*, $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_d)$, is a vector of real numbers with dimension *d*. In the neuron, the input vector $\tilde{\mathbf{x}}$ is linearly transformed to a scalar a_j after multiplying with a weight vector $\mathbf{w}_j = (w_{j0}, \ldots, w_{jd})$:

$$a_j = \sum_{i=0}^{d} w_{ji}\, \tilde{x}_i \tag{A-1}$$

In the above representation $\tilde{x}_0$ is always equal to one, and $w_{j0} = b_j$ represents a constant bias value. A function g_j, which is referred to as activation function, maps the scalar value a_j into the output of the neuron $\tilde{z}_j$. The weight vector is the variable parameter of the neuron and is estimated from a set of samples; it memorizes the relationship between the input and output of the neuron and therefore has the role of a memory in the neuron.

The activation function can impose any required non-linearity on the output of the neuron. The form of this function for a particular neuron depends on the role of the neuron in the network. The most popular forms of the activation function are:

Linear: $$g(a) = a \tag{A-2}$$

Hyperbolic tangent: $$g(a) = \tanh(a) \tag{A-3}$$

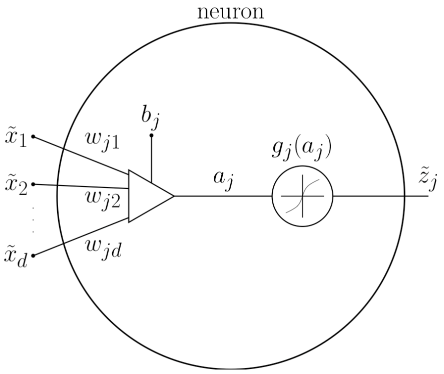


Figure A-1: A single artificial neuron.

The connection between neurons in a neural network can assume any topology. In this thesis we only consider two layer feed-forward neural networks with two layers of interconnected neurons: hidden layer and output layer (Figure A-2). The activation function of the neurons in the hidden layer, g_1, and those for the neurons in the output layer, g_2, are hyperbolic tangent and linear, respectively. The hidden units impose the required non-linearity on the input vector and the output units rescale the outputs of the hidden units to fit the target vectors of the network.

In a two layer feed-forward neural network, with input parameters x_r, the output of a given hidden unit z_j, with weight values w_{jr} is represented as:

$$z_j = g_1\left(\sum_{r=0}^{d} w_{jr} x_r\right) \tag{A-4}$$

The output of each unit in the hidden layer, z_j, is fed to all units in the output layer. The output of a given output unit y_k, with weight values w_{kj} is represented as:

$$y_k = g_2\left(\sum_{j=0}^{M} w_{kj} z_j\right) \tag{A-5}$$

In the above equation M is the number of hidden units in the hidden layer of the neural network. In what follows, we refer to hidden units by using the index j, and refer to output units by using index k.

The two layer feed-forward neural network can approximate any continuous function from one finite-dimensional space to another arbitrarily well, provided the number of hidden units M is sufficiently high. What is more, if the activation function of the units in the hidden layer is hyperbolic tangent then the network can simultaneously approximate both a function and its derivative and the error of the approximation decreases as the number of hidden units increases (Bishop, 1995, pp. 130-131).

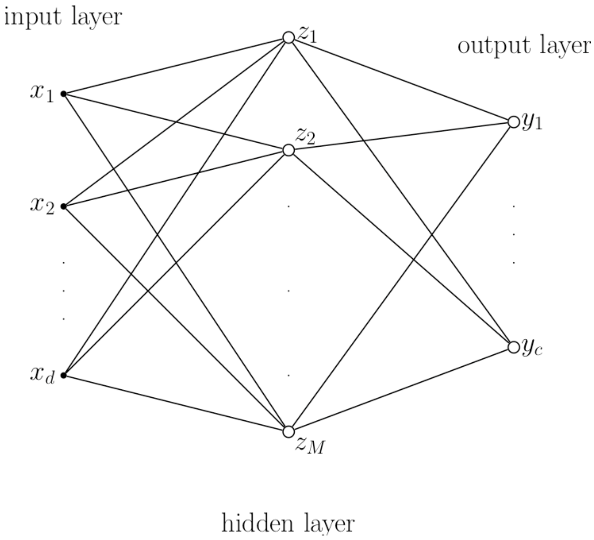


Figure A-2: A two layer feed-forward neural network.

A.2 Training of neural networks and back propagation algorithm

Varying the weights of neurons within a neural network to emulate a specific, desired mapping between its input and output parameters is called training the network. Networks are usually trained by fitting them to examples of the input and output (target) values of the mapping. The set of examples used is called the training dataset. Training is based on the definition of an error function that measures the difference between outputs of the network and target values in the training dataset. In the training process, the error function is minimized with respect to the weights of neurons in the network.

Here we introduce the error back-propagation algorithm (Rumelhart et al., 1986; Bishop, 1995, pp. 140-148) for training a two layer feed-forward neural network. We assume that a training dataset ($\mathbf{x}_n$, $\mathbf{t}_n$) and an error function E are given. The aim of training is to find an optimum set of weights of the neural network, $\mathbf{w}^*$, that minimizes the error E between the output vector $\mathbf{y}(\mathbf{x}_n; \mathbf{w}^*)$ and the corresponding target vector $\mathbf{t}_n$, for a given input vector $\mathbf{x}_n$. We assume that error function $E = E(\mathbf{y}) = E(y_1,\ldots,y_c)$ is a differentiable function of the weights of the output layer. We can thus find the derivatives of the error function with respect to the weights and these derivatives can be used to find the optimum set of weights $\mathbf{w}^*$. Back propagation algorithm evaluates derivatives of the error function by propagating errors backwards through the network in a computationally efficient procedure.

The error function E is usually assumed to be an additive function. Therefore, without loss of generality, we demonstrate the back propagation for the error function of a single training sample E_n and remove the index n of the training sample in all equations in the derivation of the back propagation algorithm below.

Assume that we have supplied the network with a sample pair ($\mathbf{x}$, $\mathbf{t}$) and calculated the activation and the output of a given neuron l, with input vector $\tilde{\mathbf{x}}$ and output $\tilde{z}_l$, using equations (A-1), and (A-4) (for hidden units) or (A-5) (for output units), respectively:

$$a_l = \sum_i w_{li}\tilde{x}_i \tag{A-6}$$

$$z_l = g(a_l) \tag{A-7}$$

The error function E depends on the weights of the neuron l, w_{li}, only through the activation a_l. We use the chain rule to calculate the derivative of E with respect to w_{li}:

$$\frac{\partial E}{\partial w_{li}} = \frac{\partial E}{\partial a_l}\underbrace{\frac{\partial a_l}{\partial w_{li}}}_{=\tilde{x}_i} \tag{A-8}$$

The second derivative of the right hand side of the above equation is obtained using equation (A-6).

The first derivative of the right hand side of the above equation is called the error δ_l. The form of the error δ_l is different for the hidden and the output layers. For a neuron in the output layer (i.e., $y_k = g_2(a_k)$), with activation function g_2, the error is obtained using the chain rule:

$$\delta_k = \frac{\partial E}{\partial a_k} = g_2'(a_k)\frac{\partial E}{\partial y_k} \tag{A-9}$$

Here $g_2'(a)$ is the derivative of the activation function $g_2(a)$. The second term in the right hand side depends on the definition of the error function E.

For a unit in the hidden layer (i.e., $z_j = g_1(a_j)$), the error function depends on the a_j via all activations of the output units of the network and, using the chain rule, we obtain the error δ_j as:

$$\delta_j = \frac{\partial E}{\partial a_j} = \sum_k \frac{\partial E}{\partial a_k}\frac{\partial a_k}{\partial a_j} \tag{A-10}$$

In the above equation the summation is over all output units. The first term on the right hand side of the above equation is the definition of δ_k (equation (A-9)). From equation (A-6) we obtain the following form for the second term:

$$\frac{\partial a_k}{\partial a_j} = w_{kj} g_1'(a_j) \tag{A-11}$$

The error δ_j is therefore obtained using (A-10) and (A-11) as:

$$\delta_j = g_1'(a_j) \sum_k w_{kj} \delta_k \tag{A-12}$$

Using equations (A-6) to (A-12), the back propagation algorithm to obtain derivatives of the error function for one sample, E_n, with respect to the network weights, w_{li}, can be represented as:

1. Apply the input vector $\mathbf{x}_n$ to the network and calculate activations a_l (equation (A-6)) and outputs z_l (equation (A-7)) for each neuron (e.g., neuron l).
2. Evaluate the error δ_k for all output units from equation (A-9).
3. Back propagate the errors δ_k using equation (A-12) to obtain the error δ_j for all hidden units.
4. Calculate necessary derivatives for the weights of the network using equation (A-8), and using the results of the previous steps.

The main importance of the back propagation algorithm in the training of neural networks is its computational efficiency (Bishop, 1995, pp. 146-147).

The derivative of the total error E over all training samples is equal to the sum of errors of individual training samples E_n and is obtained as:

$$\frac{\partial E}{\partial w_{li}} = \sum_n \frac{\partial E_n}{\partial w_{li}} \tag{A-13}$$

In the training process, the evaluated derivatives of the error function $E(\mathbf{w})$ are applied to find the minimum of this function in the weight space $\mathbf{W}$. A simple example of a minimization algorithm is the fixed-step gradient decent technique (Bishop, 1995, p. 95). This algorithm starts by an initial guess about the optimum weight vector, $\mathbf{w}$, and then updates the weight vector by moving a small distance in the direction of the fastest decrease of the error function E, i.e. in the direction of the gradient of the error function $\nabla_{\mathbf{w}} E$. By iterating this process we generate a sequence of weight vectors, $\mathbf{w}^{(\tau)}$, whose components are calculated using:

$$w_{li}^{(\tau+1)} = w_{li}^{(\tau)} - \eta \left. \frac{\partial E}{\partial w_{li}} \right|_{\mathbf{w}^{(\tau)}} \tag{A-14}$$

In the above equation η is a small positive number, which is called the learning rate and controls the rate of convergence of the algorithm. Under suitable conditions the above sequence of weights will converge to a point at which E is minimized. Bishop (1995, ch. 7) presented several iterative optimization algorithms that can be used in the network training.

The training samples can be presented to the network in two different ways: 1) batch training, and 2) on-line training (Bishop, 1995, p. 146). In the on-line training weights are updated after presenting each training sample ($\mathbf{x}_n$, $\mathbf{t}_n$) to the network. In the batch training weights are updated after presenting all training samples to the network. In the batch training second order derivatives of the error function with respect to weights can be used to enhance the training procedure, while second order derivatives cannot be easily incorporated into the on-line training (Duda et al., 2001, p. 316). However, the on-line training is more appropriate for applications with large training datasets.

A.3 Error functions

In the network training we try to model the underlying generator of data, so that the best possible predictions for output vector $\mathbf{t}$ can be made when the trained network is presented with a new value of the input vector $\mathbf{x}$. The most complete description of the generator of the data is the joint probability density function (PDF) of the input-target pairs, $p(\mathbf{x}, \mathbf{t})$, which can be represented as the product of conditional density of the target data, conditioned on the input data $p(\mathbf{t}|\,\mathbf{x})$, and the unconditional density of the input data $p(\mathbf{x})$:

$$p(\mathbf{x},\mathbf{t}) = p(\mathbf{t}|\mathbf{x})\,p(\mathbf{x}) \tag{A-15}$$

For the purpose of making predictions of $\mathbf{t}$ for new values of $\mathbf{x}$, we only need to model the conditional density $p(\mathbf{t}|\,\mathbf{x})$. Therefore, the neural network can be considered as a framework for modeling the conditional PDF $p(\mathbf{t}|\,\mathbf{x})$.

In the training of a neural network with a given dataset ($\mathbf{x}_n$, $\mathbf{t}_n$), $n = 1,\ldots,N$, the weights of the network are obtained by maximizing the likelihood of the training samples. If we assume that each sample is drawn independently from the joint PDF $p(\mathbf{x}, \mathbf{t})$, the likelihood L is obtained as:

$$L=\prod_n p(\mathbf{x}_n,\mathbf{t}_n) \qquad \text{(A-16)}$$

$$=\prod_n p(\mathbf{t}_n|\mathbf{x}_n)p(\mathbf{x}_n)$$

Error function E is defined as the negative logarithm of the likelihood:

$$E=-\ln L=-\sum_n \ln p(\mathbf{t}_n|\mathbf{x}_n)-\sum_n \ln p(\mathbf{x}_n) \qquad \text{(A-17)}$$

The second term of the right hand side of the above equation does not depend on the weights of the neural network and can be dropped in the minimization process of the error function. Therefore, the error function of the neural network can be represented as:

$$E=-\sum_n \ln p(\mathbf{t}_n|\mathbf{x}_n) \qquad \text{(A-18)}$$

Different choices of the error function arise from different assumptions about the conditional distribution $p(\mathbf{t}|\ \mathbf{x})$. We can assume that the conditional distribution, $p(\mathbf{t}|\ \mathbf{x})$, is a Gaussian with spherical covariance matrix (i.e., the identity matrix multiplied by a constant) represented as:

$$p(\mathbf{t}|\mathbf{x})=\frac{1}{(2\pi)^{c/2}\sigma^c}\exp\left\{-\frac{\sum_{s=1}^{c}\left(t_s-y_s(\mathbf{x};\mathbf{w})\right)^2}{2\sigma^2}\right\} \qquad \text{(A-19)}$$

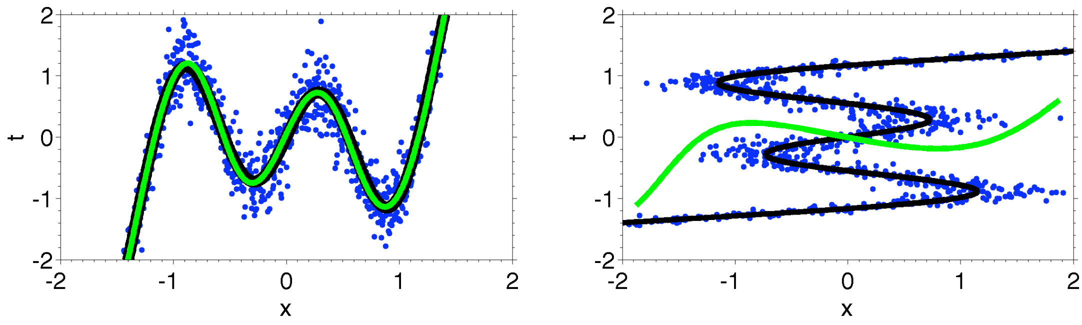
Where $\mathbf{t}=(t_1,\ldots,t_c)$ is the target vector, and $y_s(\mathbf{x};\mathbf{w})$ is the s-th output of the network, which is a function of both the weight vector $\mathbf{w}$ and the input vector $\mathbf{x}$. Therefore, using equation (A-18), the error function in this case can be represented as:

$$E=\frac{1}{2\sigma^2}\sum_n\sum_{s=1}^{c}\left(t_s-y_s(\mathbf{x}_n;\mathbf{w})\right)^2+Nc\ln\sigma+\frac{Nc}{2}\ln(2\pi) \qquad \text{(A-20)}$$

The second and third terms in the right hand side of the above equation in addition to the overall factor $1/\sigma^2$ are independent of the weight vector $\mathbf{w}$ and can be omitted in the optimization process. Therefore the simplified representation of the error function, by assuming a Gaussian distribution with a spherical covariance matrix for the conditional PDF $p(\mathbf{t}|\ \mathbf{x})$, can be represented as:

$$E=\frac{1}{2}\sum_n\sum_{s=1}^{c}\left(t_s-y_s(\mathbf{x};\mathbf{w})\right)^2 \qquad \text{(A-21)}$$

The above representation is the conventional sum-of-square error function.



(a) (b)

Figure A-3: Predictions of a neural network with the sum-of-square error function. (a) Gaussian data generator. For each value of the input *x* target value *t* is a random sample from a Gaussian PDF with a known variance. (b) Non-Gaussian data generator (inverse of the Gaussian data generator in (a)). For each value of the input *x,* target values *t* are samples of a multi-modal and non-Gaussian PDF. Black curves are the data generator function. Blue dots are noisy training samples of the data generator. Green curves are the predictions of the neural network with the sum-of-square error function.

Application of the sum-of-square error function does not require the target data to have a Gaussian distribution. However, if the PDF of the target data is strongly non-Gaussian the sum-of-square error function can result in large errors in the predictions of the data generator. Figure A-3 shows the result of the application of a neural network with the sum-of-square error function to model two different data generators. The first data generator (the black curve in Figure A-3(a)) has a Gaussian PDF and is correctly estimated by the neural network (green curve in that figure). The second data generator (the black curve in Figure A-3(b)) has a non-Gaussian PDF and therefore, the neural network estimate of it (green curve in that figure) is strongly erroneous. This example shows the limitation of the sum-of-square error for modeling data generators with a non-Gaussian PDF.

A.3.1 Mixture density network with diagonal Gaussian kernels

A general framework to model Gaussian or non-Gaussian conditional PDF's is the mixture model. Mixture models represent any given PDF as a linear combination of adaptive kernel PDF's. The mixture model of a conditional PDF $p(\mathbf{t}|\ \mathbf{x})$ is represented as:

$$p(\mathbf{t}|\mathbf{x}) = \sum_{i=1}^{m} \alpha_i(\mathbf{x}) \varphi_i(\mathbf{t}|\mathbf{x}) \tag{A-22}$$

In the above equation m is the number of kernels in the mixture. $\alpha_i(\mathbf{x})$ is called mixing coefficient and represents the probability (conditioned on $\mathbf{x}$) of the target vector $\mathbf{t}$ having generated from the *i*-th kernel. The function $\varphi_i(\mathbf{t}|\mathbf{x})$ is the i-th kernel, and represents the conditional PDF of the target vector $\mathbf{t}$ for the *i*-th kernel.

The mixture model of the conditional PDF (equation (A-22)) can be used to define another error function (equation (A-18)) for a neural network. This neural network is called the mixture density network (MDN). In an MDN the mixing coefficients and the parameters of the kernels are functions of the input vector **x** and are represented by the outputs.

Various choices for the kernel functions are possible, for example, Bishop (1994) and Nabney (2004, pp. 167-184) applied spherical Gaussian PDF's. It can be shown that a mixture model with (spherical) Gaussian kernels can approximate any given PDF arbitrarily well, provided the mixing coefficients and parameters of the Gaussian kernels (i.e., mean vector and covariance matrix) are correctly chosen (Bishop, 1995, p. 214; McLachlan and Peel, 2000).

In this thesis, we apply the diagonal Gaussian PDF's (i.e., the covariance matrix of the Gaussian kernel is diagonal) as the kernels of the mixture density model. Since to our knowledge the mathematical details required to implement an MDN with the diagonal Gaussian kernels has not been published previously, here we present these details. The diagonal Gaussian kernel is given as:

$$\varphi_i(\mathbf{t}\,|\,\mathbf{x}) = \frac{1}{\prod_{k=1}^{c}\left(\sqrt{2\pi}\;\;\sigma_{ik}(\mathbf{x})\right)} \exp\left\{-\frac{1}{2}\sum_{k=1}^{c}\frac{\left(t_k-\mu_{ik}(\mathbf{x})\right)^2}{\sigma_{ik}^2(\mathbf{x})}\right\} \qquad \text{(A-23)}$$

In the above equation c is the dimensionality of the output vector $\mathbf{t} = (t_1,\ldots, t_c)$, μ_{ik} is the k$^{\text{th}}$-component in the mean vector of the i$^{\text{th}}$-kernel, and σ_{ik} is the k$^{\text{th}}$-diagonal element in the covariance matrix of the i$^{\text{th}}$-kernel. Therefore, the mean and covariance of the i$^{\text{th}}$ Gaussian kernel are $\boldsymbol{\mu}_i = (\mu_1,\ldots, \mu_c)$ and $\Sigma_i = \mathrm{diag}(\sigma_{i1},\ldots,\sigma_{ic})$, respectively.

First, we discuss the link between the parameters of the Gaussian mixture model (i.e., α_i, μ_{ik}, σ_{ik}) and the outputs of the neural network. In order to have a valid representation of the conditional probability density function in equation (A-22), the mixing coefficients must be positive and their sum must be equal to one, i.e. $\sum_{i=1}^{m}\alpha_i(\mathbf{x}) = 1$. Standard deviations must also be positive. To satisfy these conditions we define the mixture model parameters, α_i, μ_{ik}, and σ_{ik}, to be related to the corresponding neural network outputs, z_i^{α}, z_{ik}^{μ}, and z_{ik}^{σ}, by:

$$\alpha_i = \frac{\exp(z_i^{\alpha})}{\sum_{i=1}^{m}\exp(z_i^{\alpha})} \qquad i = 1,\ldots,m \qquad \text{(A-24)}$$

$$\sigma_{ik} = \exp(z_{ik}^{\sigma}) \qquad i = 1,\dots,m \quad k = 1,\dots,c \tag{A-25}$$

$$\mu_{ik} = z_{ik}^{\mu} \qquad i = 1,\dots,m \quad k = 1,\dots,c \tag{A-26}$$

Note that the total number of output units in an MDN with m diagonal Gaussian kernels is $(2c + 1) \times m$ where c is the dimensionality of the target vector $\mathbf{t}$.

In the back propagation algorithm, in order to obtain the error δ_k (equation (A-9)) for an output neuron of the diagonal-MDN, derivatives of the error function E (equation (A-18)) with respect to the network outputs are required. For a given training sample ($\mathbf{x}_j$, $\mathbf{t}_j$), by substituting the diagonal Gaussian kernel (equation (A-23)) in the mixture density model of the conditional PDF in equation (A-22) and then by substituting the mixture density model into the error function E_j (equation (A-18)) we obtain the following derivatives:

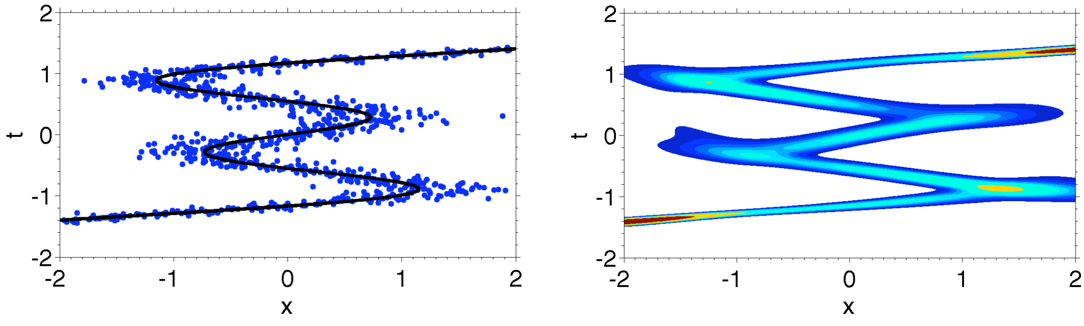
$$\frac{\partial E_j}{\partial z_i^{\alpha}} = \alpha_i - \frac{\alpha_i \varphi_i}{\sum_{i=1}^{m} \alpha_i \varphi_i} \tag{A-27}$$

$$\frac{\partial E_j}{\partial z_{ik}^{\sigma}} = -\left(\frac{\alpha_i \varphi_i}{\sum_{i=1}^{m} \alpha_i \varphi_i}\right)\left(\frac{\left(t_k - \mu_{ik}(\mathbf{x}_j)\right)^2}{\sigma_{ik}^2(\mathbf{x}_j)} - 1\right) \tag{A-28}$$

$$\frac{\partial E_j}{\partial z_{ik}^{\mu}} = -\left(\frac{\alpha_i \varphi_i}{\sum_{i=1}^{m} \alpha_i \varphi_i}\right)\left(\frac{\left(t_k - \mu_{ik}(\mathbf{x}_j)\right)}{\sigma_{ik}^2(\mathbf{x}_j)}\right) \tag{A-29}$$

In the above equations values of α_i, μ_{ik}, and σ_{ik} are computed at the sample point ($\mathbf{x}_j$, $\mathbf{t}_j$). In this work we wrote the necessary codes to implement and train a diagonal-MDN using the back propagation algorithm with the above derivatives. These were used for all of the methods and results in this thesis.

As an example, we apply the MDN to find the data generator of the above example in Figure A-3(b). The result is shown in Figure A-4(b). Figure A-4(b) shows that the MDN can model a data generator with a strongly non-Gaussian PDF (i.e., multi-modal PDF as several target values t is possible for a given input vector x). We investigate the ability of the MDN to model arbitrary conditional PDF's in more detail below (Section A.6).



(a) (b)

Figure A-4: Prediction of the MDN with the mixture model error function. (a) Noisy samples of the non-Gaussian data generator (black curve). (b) MDN estimate of the conditional PDF $p(\mathbf{t}|\mathbf{x})$. Warm colors represent high probability areas.

A.4 Pre-processing of training dataset and initialization of weights

A.4.1 Pre-processing of training dataset

In the neural networks, the influence of the parameters of the input vector $\mathbf{x}$ on the weight update (equations (A-8) and (A-14)) depends on typical values of these parameters. Typical values of the input parameters can differ significantly due to measurement units. Furthermore, the relative size of the input parameters may not reflect the relative importance of these parameters in determining the outputs of the neural network. In order to eliminate the undesirable effect of the relative size of the input parameters, we can apply a linear transformation to rescale these parameters into relatively similar values. This procedure is called the pre-processing of the input vector.

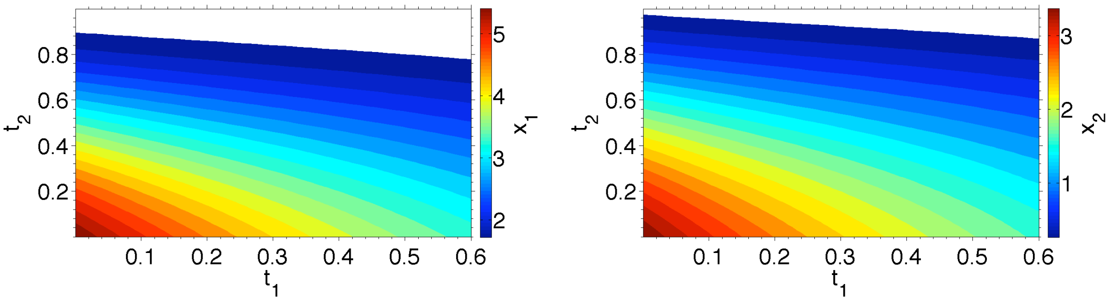
In this study, we apply the whitening algorithm to transform the input vectors of the training dataset, $\mathbf{x} = (x_1,\ldots,x_d)^T$, to a new dataset with zero mean and unit covariance matrix. Consider a training dataset $\{\mathbf{x}_j, \mathbf{t}_j\}$, with n statistically independent samples. The mean vector, $\bar{\mathbf{x}}$, and covariance matrix, $\boldsymbol{\Sigma}$, of the sample input vectors $\mathbf{x}_j$ are given by:

$$\bar{\mathbf{x}} = \frac{1}{n}\sum_{j=1}^{n}\mathbf{x}_j \tag{A-30}$$

$$\Sigma = \frac{1}{n-1}\sum_{j=1}^{n-1}\left(\mathbf{x}_j - \bar{\mathbf{x}}\right)\left(\mathbf{x}_j - \bar{\mathbf{x}}\right)^T \tag{A-31}$$

The whitened input vector, $\hat{\mathbf{x}}_j$, is then given by:

$$\hat{\mathbf{x}}_j = -\Lambda^{-1/2}\mathbf{U}^T\left(\mathbf{x}_j - \bar{\mathbf{x}}\right) \tag{A-32}$$



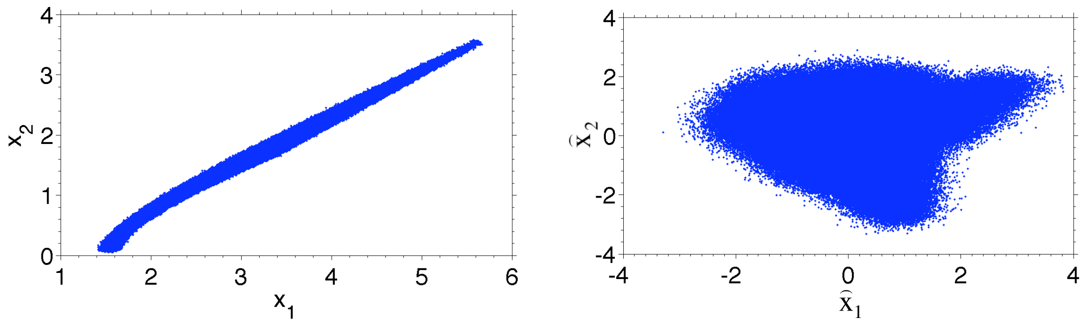
(a) (b)

Figure A-5: The forward function **f** between **t** = (t_1, t_2) and **x** = (x_1, x_2). (a) x_1 as function of **t**. (b) x_2 as function of **t**.

Here Λ and **U** are the eigenvalue and the eigenvector matrix of the covariance matrix **Σ**, respectively. It can be shown that the transformed input data $\hat{\mathbf{x}}_j$ has zero mean and unit covariance matrix (Fukunga, 1990; Bishop, 1995, p. 299). In the applications in this thesis the whitening algorithm is used to pre-process training datasets.

We explain the effect of the pre-processing by an example. Consider the function **t** = **f**(**x**) between **x** = (x_1, x_2) and **t** = (t_1, t_2), which is shown in Figure A-5. Noisy samples of the input vector **x** (Figure A-6) are used to train an MDN, with two different datasets. No pre-processing algorithm is applied on the first dataset (Figure A-6(a)), while the whitening algorithm is applied to pre-process the input vector **x** (Figure A-6(b)) of the second dataset. We applied the MDN in both cases to find the conditional PDF $p(\mathbf{t}|\mathbf{x})$ for a random input vector **x** = (2.070, 0.827). The estimate of the MDN in two cases, in addition to the Monte Carlo sampling solution of the problem is shown in Figure A-7. Figure A-7 shows that the MDN provides a better approximation of the MC sampling solution (Figure A-7(c)) when the pre-processing algorithm is used (Figure A-7(a)). In addition more kernels of the mixture density model is applied when the pre-processing algorithm is used than when it is not used (compare Figure A-7(a) with Figure A-7(b)).

In order to explain the effect of the pre-processing further, we draw 10,000 random samples of **t** from the three conditional PDF's $p(\mathbf{t}|\ \mathbf{x})$ shown in Figure A-7 and forward-modeled them using the forward function **f** to obtain the corresponding synthetic data **x**. The joint PDF's of data parameters are shown in Figure A-8. The high probability zones of these PDF's are comparable. However, the uncertainties of the synthetic data **x** obtained using the forward modeling of the MDN result without pre-processing (Figure A-8(b)) are higher than the other two cases (Figure A-8(a, c)). Also, a correlation between the uncertainty of x_1 and x_2 can be seen when the pre-processing algorithm is not used. This correlation does not exist in the other two cases.



(a) (b)

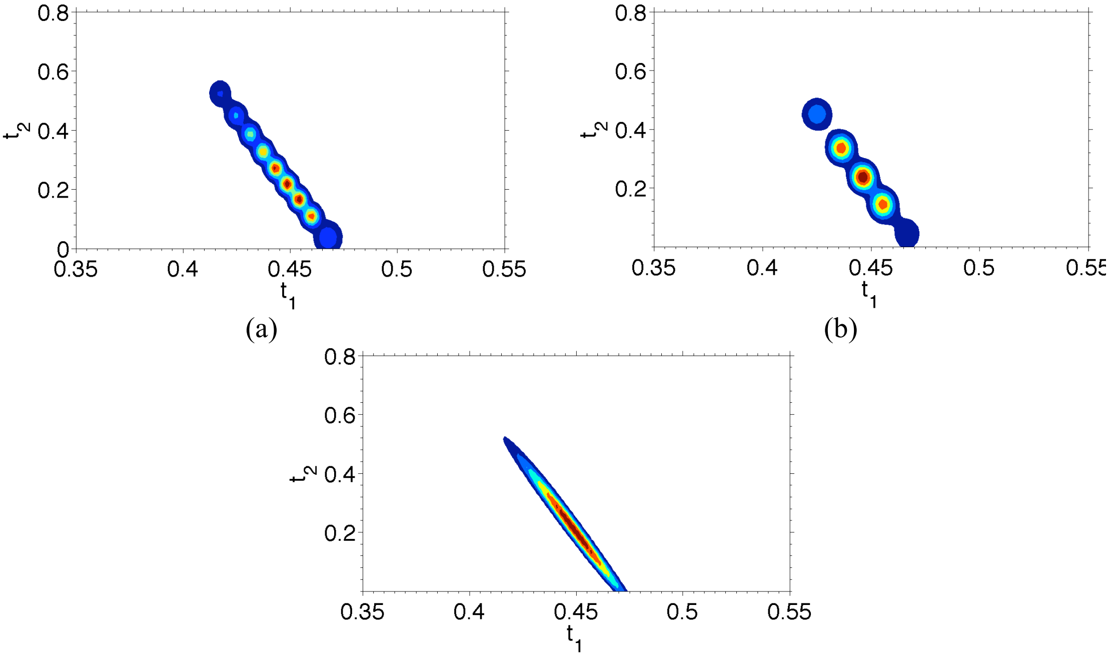
Figure A-6: Noisy samples of the input vector of the neural network. (a) $\mathbf{x} = (x_1, x_2)$: input vector of training samples without pre-processing. (b) $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2)$: input vector of training samples with pre-processing.

The improvement in the result of the MDN with preprocessing stems from two facts: first, the linear transform of the input parameters enhances the selection of the random values of initial weights of the network, which facilitates the back-propagation search for the optimum weight values. Therefore, with the same number of iterations, the MDN with pre-processing usually finds a smaller minimum of the error function than the MDN without pre-processing. Second, the preprocessing algorithm above removes the spurious correlations between the parameters of the input vector of the training dataset. In the example above the forward modeling result of the MC sampling solution (Figure A-8(c)) shows almost no correlation between x_1 and x_2. Forward modeling of the MDN solution with pre-processing (Figure A-8(a)) also shows nearly no correlation between x_1 and x_2. However, forward modeling of the MDN solution without pre-processing (Figure A-8(b)) shows a correlation between x_1 and x_2.

Figure A-6(a) shows a strong correlation between x_1 and x_2 in the training dataset. When the network is trained without preprocessing, it detects this global correlation between x_1 and x_2, and erroneously shows it in the joint PDF of the parameters for a single isolated data sample (Figure A-8(b)). However, with pre-processing this global correlation is removed and the result of the MDN is improved. Note that the pre-processing algorithm only improves the uncertainty estimate for single data points and x_1 and x_2 remain globally correlated.

A.4.2 Initialization of the weights of a neural network

As we explained in section A.2, at the beginning of the minimization algorithm, for updating the weights of a network, some initial values must be assigned to the weights. Initial values are usually set randomly in order to avoid problems due to symmetries in the network (Bishop, 1995, p. 261). The weights are also chosen to be small. Large values of the weights



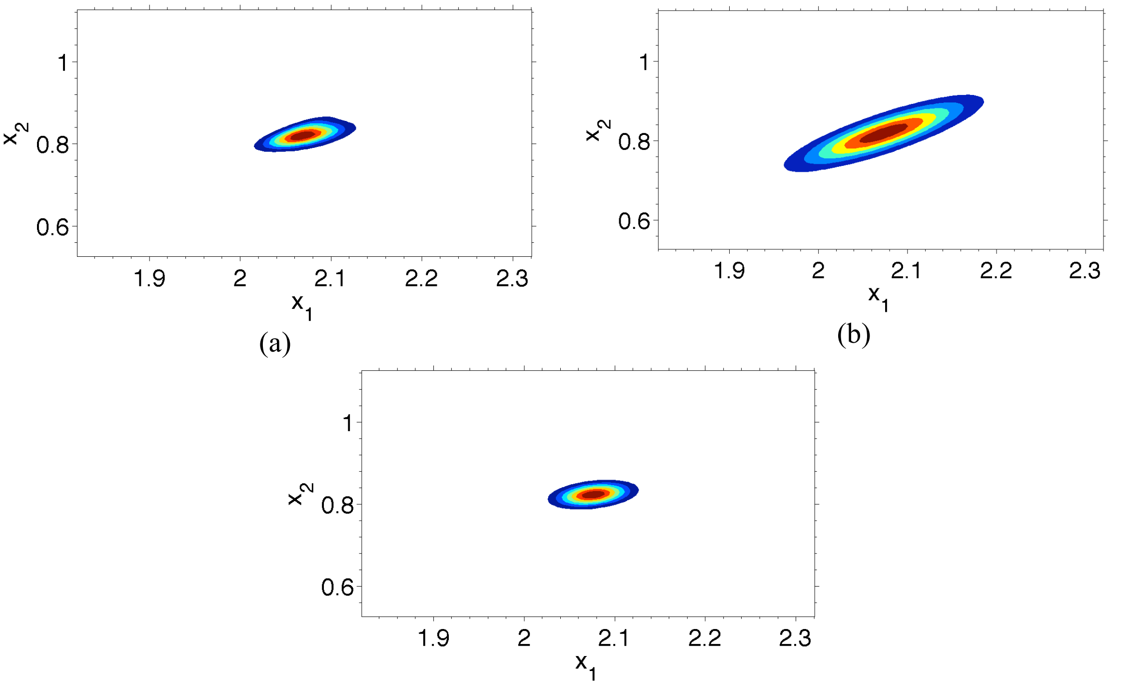
(a) (b)

(c)

Figure A-7: Conditional PDF $p(\mathbf{t}|\ \mathbf{x})$ for $\mathbf{x} = (2.070, 0.827)$. (a) Estimated PDF using the MDN with pre-processing. (b) Estimated PDF using the MDN without pre-processing. (c) Estimated PDF using the MC sampling solution. Warm colors represent high probability areas.

derive the hyperbolic tangent activation functions into the saturation region, where the derivatives are near zero. Near zero values of the derivatives of the activation function in equation (A-12) result in near zero values of the derivatives of the error function E in equation (A-14), and hence will result in a slow weight update procedure. In the other hand, if the weights are too small, all the hyperbolic tangent functions will be approximately linear, which can again lead to slow training. This suggests that in order to have an acceptable training rate, the input of the hyperbolic tangent activation functions should be of order unity after initialization. Therefore, the chosen random initial weights must satisfy this condition.

Here we obtain the statistical properties of the initial weights, required to satisfy the above condition. We assume that the whitening algorithm in the previous section is applied to transform the input vectors of the training dataset to a new dataset with zero mean, i.e. $\langle x_i \rangle = 0$, and unit variance, i.e. $\langle x_i^2 \rangle = 1$. The notation $\langle \cdot \rangle$ is used to denote the average value of a random variable. The weights are usually generated from a simple Gaussian PDF with a spherical covariance matrix and a zero mean. The mean is assumed to be zero because there is no reason to prefer any other specific point in the weight space. Below, we describe a procedure to determine the variance of the Gaussian PDF. For a neuron in the hidden layer of the network the average value of the input of the activation function (equation (A-1)) is obtained as:



(a)

(b)

(c)

Figure A-8: The PDF of the input vector $p(\mathbf{x})$ from forward modeling of the conditional PDF $p(\mathbf{t}|\ \mathbf{x})$. (a) Forward modeling of the MDN result with pre-processing. (b) Forward modeling of the MDN result without pre-processing. (c) Forward modeling of the MC sampling result. Warm colors represent high probability areas.

$$\langle a\rangle = \sum_{i=0}^{d}\langle w_i x_i\rangle = \sum_{i=0}^{d}\langle w_i\rangle\langle x_i\rangle = 0 \tag{A-33}$$

In the above equation we use the fact that the weights of the network, w_i, and the parameters of the input vector, x_i, are uncorrelated. Note that the average of the whitened parameters of the input vector is zero, i.e. $\langle x_i\rangle = 0$. The variance of a is obtained as:

$$\langle a^2\rangle = \left\langle\left(\sum_{i=0}^{d} w_i x_i\right)\left(\sum_{i=0}^{d} w_i x_i\right)\right\rangle = \sum_{i=0}^{d}\langle w_i^2\rangle\langle x_i^2\rangle = \sigma^2(d+1) \tag{A-34}$$

In the above equation σ^2 is the variance of the PDF of the weights, and we use the fact that weights are uncorrelated, i.e. $\langle w_i w_j\rangle=0$ for $i\neq j$, together with that the variance of the whitened parameters of the input vector is equal to unity, $\langle x_i^2\rangle = 1$.

As we explained above, if the activation function of the neurons in the first hidden layer is tangent hyperbolic, we would like a to be of order unity. Therefore, equation (A-34) suggests that the variance of the PDF of the initial weights should scale like $\sigma \propto (d+1)^{-1/2}$. The same argument can be applied to obtain the PDF of the initial weights of other neurons in the network, if we assume that their input parameters are appropriately distributed (Bishop, 1995, pp. 261-262).

A.5 Learning and generalization

As we mentioned in Section A.3 the aim of the network training is to build a statistical model of the process that generates the training dataset. It means that the network, as the model of the data generator, not only should be able to regenerates the training dataset, but also should have a good generalization behavior, that is, to make good predictions for a new input vector. In order to optimize the generalization behavior of the network we need to control its effective complexity. In neural networks changing the number of the adaptive parameters in the network (i.e., the number of neurons in the hidden layer) can vary the complexity. A simple network with a small number of hidden units can under-fit data (i.e., can not sufficiently fit the relationships embodied in the training data set) while a complex network with a large number of hidden units can over-fit data (i.e., accurately fits the training dataset, but is far less accurate for data not represented within the training data set). A common way to find the optimum number of the hidden units in the network is a trial and error procedure, in which we compare the behavior of a range of networks with different number of hidden units. In this section we introduce other ways to optimize the generalization behavior of neural networks.

Bishop (1995, pp. 377-380) gives a detailed discussion about the relationship between the number of training samples and generalization behavior of neural networks. One corollary of that discussion about networks with specific activation and output functions is that the minimum number of training patterns needed to predict the target value correctly for a fraction $1-\varepsilon$ of new input examples is:

$$N \cong W/\varepsilon \qquad \text{(A-35)}$$

Here W is the number of the weights in the network. For example, in order to have 90% accuracy in the network predictions, we need around ten times as many training samples as there are weights in the network (Duda et al., 2001, pp. 310-311). We use this result to obtain an upper limit on the number of the weights for a given training dataset.

A.5.1 Regularization, training with noisy dataset, and early stopping

Regularization is a common method to control the complexity of neural networks. In this method a penalty term is added to the standard error function of the network:

$$E_{reg} \cong E + \nu E_w \tag{A-36}$$

In the above equation E is the standard error function of the network given by equation (A-18), E_w is the penalty term, which is called regularizer, and ν is called regularization parameter and is a coefficient that controls the extent of influence of the penalty term on the total error. During the training procedure, the standard error function decreases constantly and the regularizer increases constantly. At the start of the training the decreasing effect of the standard error is dominant. But as the training progresses, eventually, the effect of the regularizer becomes dominant and at that point the generalization behavior of the network is optimum.

One of the simplest forms of a regularizer is the weight decay, which consists of the sum of squares of the weights in the network:

$$E_w = \frac{1}{2}\sum_{i=1}^{W} w_i^2 \tag{A-37}$$

Here W is the number of the weights in the network. It has been shown that as the magnitude of the weights increases, the tendency of a network to over-fit fine details of the training dataset increases (Bishop, 1995, p. 339; MacKay, 2003, p. 479; Meier et al., 2007b). Therefore, the above regularization term, which favors small values of weights, decreases the chance of over-fitting.

Over-fitted neural networks are generally fit fine details of the training dataset, and hence they usually have high curvature. The regularizer can be selected to explicitly penalize the curvature or equivalently the second derivatives of the network output function $\mathbf{y} = (y_1,\ldots,y_c)$ with respect to the input parameters $\mathbf{x} = (x_1,\ldots, x_d)$:

$$E_w = \frac{1}{2}\sum_{n}\sum_{i=1}^{d}\sum_{k=1}^{c}\left(\frac{\partial^2 y_k}{\partial x_i^2}\right)_n^2 \tag{A-38}$$

In the above equation n is the number of the samples in the training dataset. The derivatives of the regularized error function (equation (A-36)) with the above regularizer are required in the back propagation algorithm for training. However, evaluation of these derivatives with respect to the weights of the network is computationally demanding. Webb (1994) showed that the effect of adding small noise on the input vectors of the training dataset is equivalent to applying a regularizer of the form given in equation (A-38). Therefore, instead of applying the above regularizer, we can add small amounts of noise to the training dataset (Bishop, 1995, pp. 346-349).

Another way to control the effective complexity of a network is the procedure of early stopping. In this procedure a validation dataset, which is independent of the training dataset, is applied to control the complexity of the network. As training progresses, the standard error function (equation (A-18)) computed on the training dataset constantly decreases. The error on the validation dataset also shows an initial decrease, however, as the network starts to over-fit fine details of the training dataset, this error starts to increase. At this point the network has the best generalization performance and training must be stopped. It can be shown that early stopping procedure is equivalent to regularization with weight decay regularizer (Bishop, 1995, pp. 345, 380-381). The example below shows the effect of the regularization and early stopping on the effective generalization of neural networks.

We applied both the weight decay regularizer and the early stopping method to control the complexity of the MDN in the second application of chapter 3 (section 3.4.2). First, we applied the weight decay regularizer with a network with 34 hidden units and 5 kernels. Eleven different values of the regularization parameter ν were used to train the network. Equations (A-36) and (A-37) suggest that as both E_w and ν are positive numbers, the regularized error E_{reg} should be an increasing function of the regularization value, when the standard error is computed on the training dataset. The blue curve in Figure A-9(a) shows the evolution of the regularized error E_{reg} as a function of the regularization parameter for the training dataset. This figure confirms that E_{reg}, calculated on the training dataset, is an increasing function of ν.

In order to study the generalization performance of the network, we evaluate the error function on an independent validation dataset. The red curve in Figure A-9(a) shows the regularized error E_{reg} on the validation dataset. Equation (A-36) suggests that due to the contribution of the penalty term, a large value of ν results in a large value of the regularized error on the validation dataset. On the other hand, for small values of ν, the over-fitting of the network to the training dataset increases the standard error on the validation dataset; this also results in a large value of the regularized error on the validation dataset. Therefore, at the optimum value of the regularization parameter the value of the error on the validation dataset should assume its minimum. The red curve in Figure A-9(a) shows that for the above example the validation error reaches to its minimum at $\nu \cong 3$ ($\ln \nu \cong 1.1$) and this value can be used as the optimum value of the regularization parameter.

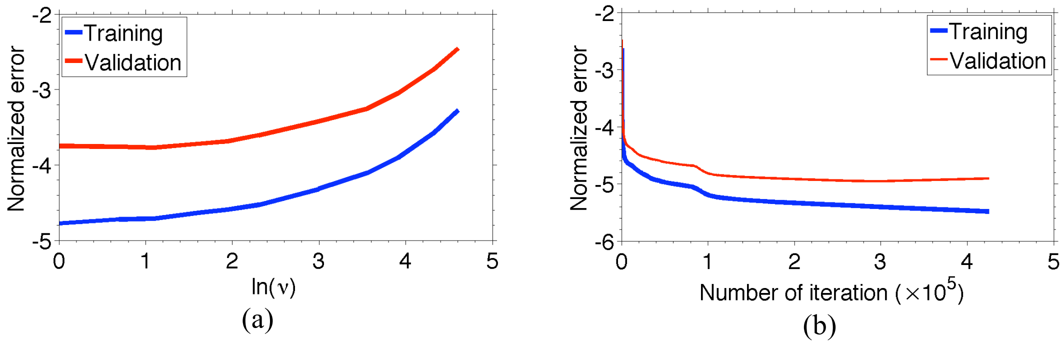


Figure A-9: Evolution of error function on the training and validation dataset for different techniques used to select the effective complexity of the network. (a) Normalized error as a function of regularization parameter in the regularization technique. (b) Normalized error as a function of the number of iterations of the optimization algorithm in the early stopping technique. The blue curve is the error on the training dataset and the red curve is the error on the validation dataset.

For the results presented in the section 3.4.2, we applied the early stopping technique to control the effective complexity of the network. The number of the hidden units and kernels was 34 and 5, respectively. The standard error function (equation (A-18)) was used in the training procedure. Initial weights of the network were selected from a Gaussian PDF with a zero mean and a variance equal to 3 (see section A.4, below equation (A-34)). Error as a function of the number of iterations of the error minimization algorithm for both the training and the validation datasets are shown in Figure A-9(b). As we can see both the training error and the validation error initially decrease, but after 3×10^5 iterations the error on the validation dataset starts to increase slightly while the error on the training dataset continue to decrease.

A.5.2 Committee of networks

In this thesis we explain that in order to select a neural network with the optimum effective complexity, several neural networks are usually trained and then the best network is selected, on the basis of the performance on an independent validation dataset. This approach has two disadvantages: first, the effort in the training of the non-optimum networks is wasted, and second, although the selected network has the best generalization performance on the validation dataset, its performance might not be the best on a new test data. The committee of networks is an idea that can be used to overcome the above drawbacks.

A committee of networks is a set of trained networks with different specifications; for example, they can have different number of weights (e.g., number of hidden units, or number of kernels in an MDN), or trained to different local minima of the error function. The output of a committee of L trained networks with outputs, $\mathbf{y}_i$, $i = 1,\ldots,L$, for an input vector $\mathbf{x}$ can be written as:

$$\mathbf{y}_{\mathrm{COM}}(\mathbf{x}) = \frac{1}{L}\sum_{i=1}^{L}\mathbf{y}_i(\mathbf{x}) \qquad \text{(A-39)}$$

Error function of the committee E_{COM} is defined using the committee outputs in the above equation. It can be shown that (Bishop, 1995, pp. 364-369) the error function of the committee is always smaller than the average of the error function of individual networks in the committee:

$$E_{\mathrm{COM}} \le \frac{1}{L}\sum_{i=1}^{L}E_i \qquad \text{(A-40)}$$

The above result holds for any convex error function. In order to improve the performance of the committee, the simple average in equation (A-39) can be replaced by a weighted average to give greater weights to networks with smaller error.

The concept of the committee of networks arises naturally in the Gaussian paradigm of the Bayesian approach to the network training (MacKay, 1992; Bishop, 1995, pp. 387-425). In this paradigm, the posterior PDF of the weights after presenting training dataset D to the network is assumed to have a Gaussian distribution. Such an approach does not allow for the presence of multiple, non-equivalent minima of the error function in the weight space. In a committee of networks, a network is trained several times from different initial weights to discover several non-equivalent minima of the error function. The posterior PDF of the weights of the committee can be approximated using a set of Gaussians, one centered on each local minimum m_i:

$$\begin{aligned} p(\mathbf{w} \mid D) &\approx \sum_i p(\mathbf{w}, m_i \mid D) \\ &= \sum_i p(\mathbf{w} \mid m_i, D)\; P(m_i \mid D) \end{aligned} \qquad \text{(A-41)}$$

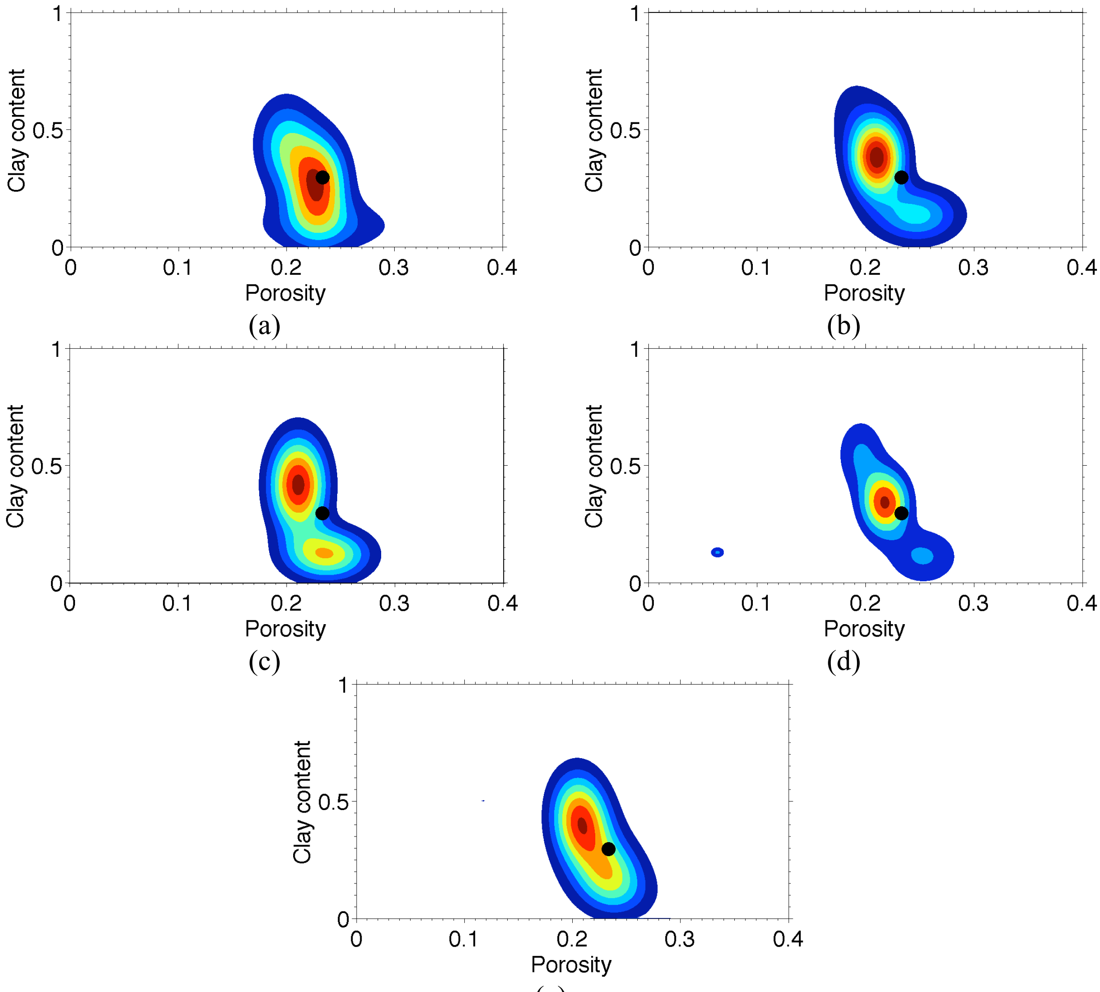
In the above equation $P(m_i \mid D)$ denotes the likelihood of the local minimum m_i given the training dataset D. The above equation can be used to determine other quantities by integration over the whole of weight space. For example, the mean output predicted by the committee is represented as:

$$\begin{aligned}\bar{\mathbf{y}} &= \int \mathbf{y}(\mathbf{x};\mathbf{w})\,p(\mathbf{w} \mid D)\,d\mathbf{w} \\ &= \sum_i P(m_i \mid D)\left[\int_{\Gamma_i} \mathbf{y}(\mathbf{x};\mathbf{w})\,p(\mathbf{w} \mid m_i, D)\,d\mathbf{w}\right] \\ &= \sum_i P(m_i \mid D)\,\bar{\mathbf{y}}_i\end{aligned} \tag{A-42}$$

Where Γ_i denotes the region in the weight space surrounding the i-th local minimum, and $\bar{\mathbf{y}}_i$ is the corresponding network prediction averaged over this region. The above equation shows that the predicted output by the committee is a linear combination of the predictions made by the members of the committee. The above result can be extended to different networks with different numbers of hidden or output units (Bishop, 1995, pp. 422-424). In practice, only a limited number of networks are trained and the weighted sum in the equation (A-42) is replaced by the simple average similar to equation (A-39). The example below shows one application of the committee of networks.

We trained 132 different MDN's with the same dataset used in well log inversion in section 3.4.2. The number of the kernels was assumed to be 3, 5, 7, or 10. For each value of the number of kernels, we selected three different numbers of hidden units. The maximum number of hidden units obtained using the rule of thumb given in the paragraph below equation (A-35). The other two values of the number of hidden units were equal to one-third and two-third of the maximum number of the hidden units. For example, for the MDN with five kernels we used 34, 22,and 11 hidden units in three different networks. Each of the above networks was trained with 11 different sets of initial weights. The training stopped after 35000 iterations of the error minimization algorithm and we accepted the network if the error function calculated on the training dataset was smaller than a selected threshold.

Figure A-10 shows the inversion result for (I_p, PR) = (6664.6 (g/cc)×(m/s), 0.286) obtained using four MDN's and the average result of all 132 networks. Note that one obvious consequence of using committee of MDN's is the removal of the multi-modality of the MDN's result. In addition, due to averaging over the whole weight space, the selection of the optimum number of kernels, hidden units, or number of iterations is not necessary. Possible outliers, or over-fitting effects are removed due to the averaging process. For example, Figure A-10(d) shows that the network with 10 kernels tends to over-fit training dataset as the estimated variance is small, and some irrelevant areas of the porosity-clay content plane, around porosity equal to 0.06 and clay content equal to 0.15, has non-zero probability; but as Figure A-10(e) shows the committee removes these effect.



(a) (b) (c) (d) (e)

Figure A-10: Committee of networks generalization. Joint PDF of porosity and clay content obtained using different networks and a committee of networks. The input vector for inversion is (I_p, PR) = (6664.6 (g/cc)×(m/s), 0.286). (a) Inversion result of the MDN with 3 kernels. (b) Inversion result of the MDN with 5 kernels. (c) Inversion result of the MDN with 7 kernels. (d) Inversion result of the MDN with 10 kernels. (e) Inversion result of the committee of 132 different MDN's. Warm colors represent high probability areas. The black dot is the measured value of porosity and clay content (ϕ, V_{cl}) = (0.233, 0.296).

A.6 Gaussian mixture model: Effect of the type of the kernels

Error in the result of an MDN depends on the type of the kernel. In this section, we compare the result of a spherical MDN (i.e., the covariance matrix of the Gaussian kernel is a scalar multiple of the identity matrix) with the result of a diagonal MDN (i.e., the covariance matrix of the Gaussian kernel is diagonal, with unequal diagonal elements).

Bishop (1994; 1995, pp. 212-222) developed the theory of the MDN using spherical Gaussian PDF's and Nabney (2004, pp. 167-184) developed the necessary computer code to

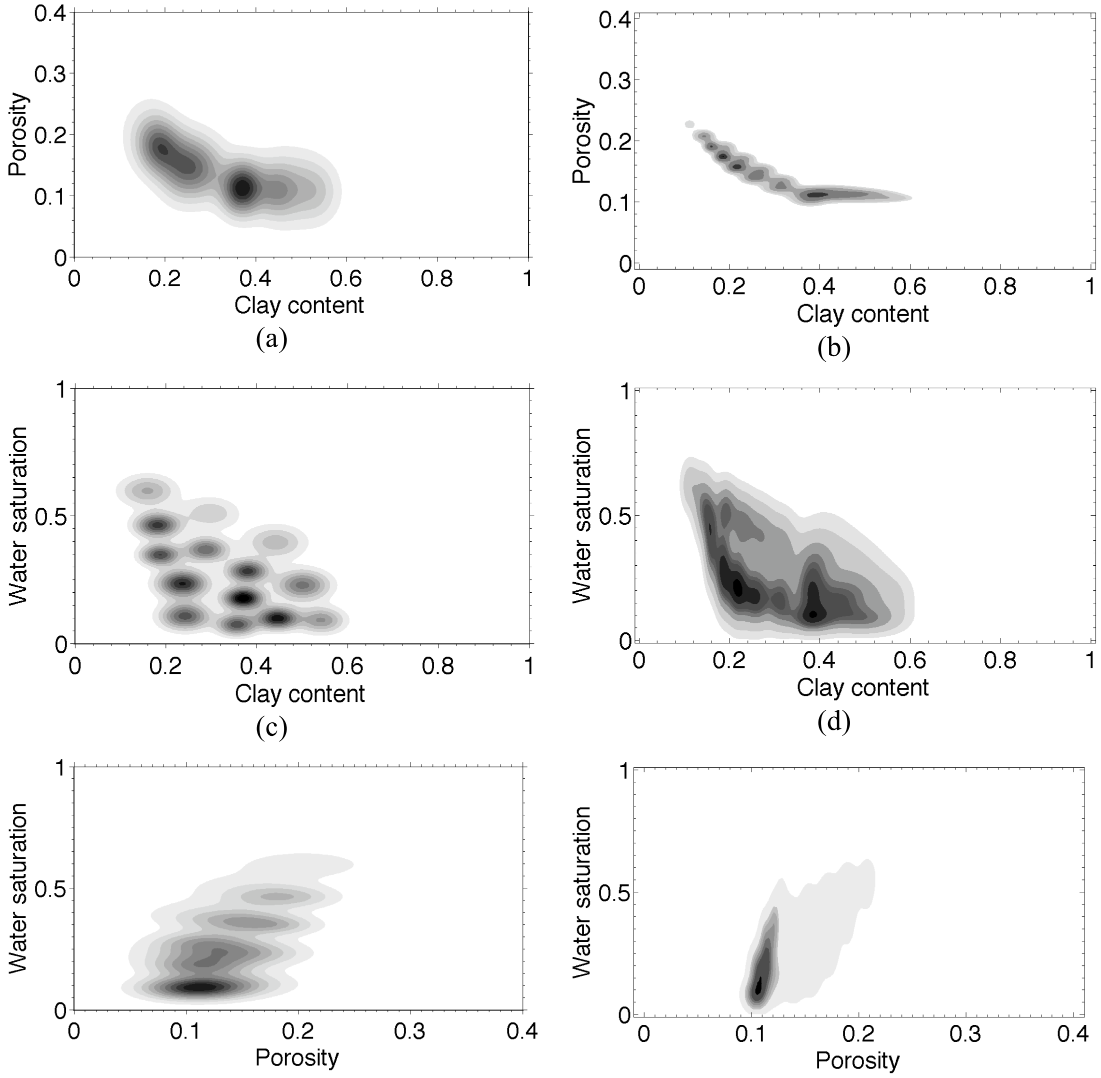
Network	Number of Gaussian kernels	Number of weights of the network	Normalized error
Diagonal	15	1185	−1
Spherical 1	15	855	− 0.74
Spherical 2	18	1020	− 0.76
Spherical 3	21	1185	− 0.77
Spherical 4	23	1350	−0.79

Table A-1: Specifications of the diagonal and spherical MDN's used to solve the synthetic problem.

implement that theory. While the spherical Gaussian kernel is a universal approximator, (i.e., it can model any PDF arbitrarily well provided enough number of kernels) in real data applications with a multi-dimensional target space the spherical assumption can cause practical difficulties. In particular, when the uncertainty distribution is highly variable in different dimensions of target space, the number of required spherical kernels can be large. As we explained above, in order to improve the result of the MDN for a multi-dimensional target space, we extend the theory of MDN's to apply Gaussians with diagonal covariance matrix. The diagonal Gaussian kernels are more flexible than spherical kernels; therefore, we can approximate a multi-dimensional PDF with a lower number of kernels.

We solved the synthetic problem in chapter 3 (section 3.4.1) with four spherical MDN's with different number of kernels. Table A-1 summarizes the specifications of the different MDN's. The number of hidden units for all of the spherical networks was the same as in the diagonal network and was equal to 10. The training was stopped when the change in the error function after a new iteration of the optimization algorithm was less than 0.5%. As Table A-1 shows, the best training error achieved by the spherical-MDN's is 20% larger than that for the diagonal-MDN. For the spherical-MDN with 21 kernels, the number of weights and biases is equal to the number of weights and biases of the diagonal-MDN. As Table A-1 shows, the training error for this spherical-MDN is around 23% larger than the error for the diagonal-MDN.

Figure A-11 shows the result of the diagonal-MDN in addition to the result of the spherical MDN with 21 kernels. This figure clearly shows that the inversion result obtained by the spherical-MDN is less accurate than the result of the comparable diagonal-MDN. Figure A-11 show that the uncertainty of the joint PDF of porosity and clay content, clay content and water saturation, and porosity and water saturation, obtained using the spherical MDN are larger than these obtained using the diagonal-MDN, the latter of which provides a

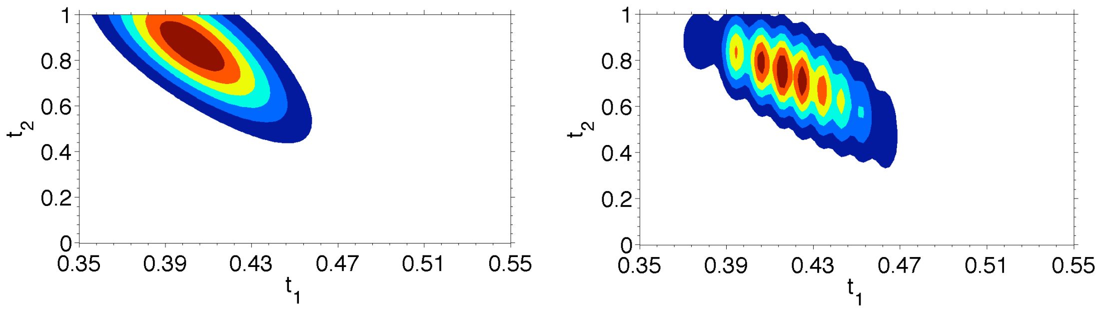


(a) (b) (c) (d) (e) (f)

Figure A-11: Comparison between the result of the MDN inversion using the spherical Gaussian kernels and the diagonal Gaussian kernels. Inversion result for (V_p, V_s)=(2818 m/s, 1675 m/s). First row is the joint marginal PDF of porosity and clay content: (a) The spherical MDN result. (b) The diagonal MDN result. Second row is the joint marginal PDF of clay content and water saturation: (c) The spherical MDN result. (d) The diagonal MDN result. Third row is the joint marginal PDF of porosity and water saturation: (e) The spherical MDN result. (f) The diagonal MDN result. Dark colors represent areas with higher probability.

better estimate of the MC sampling solution (figure 3-1). The above example confirms that applying the diagonal-MDN results in better approximations of multi-dimensional PDF's than applying a spherical-MDN with the same number of parameters.

The diagonal Gaussian kernels as presented above are, however, less flexible than Gaussian kernels with a full covariance matrix. Applying full covariance matrices with an MDN is computationally more complex than the diagonal MDN, because we need to derive an



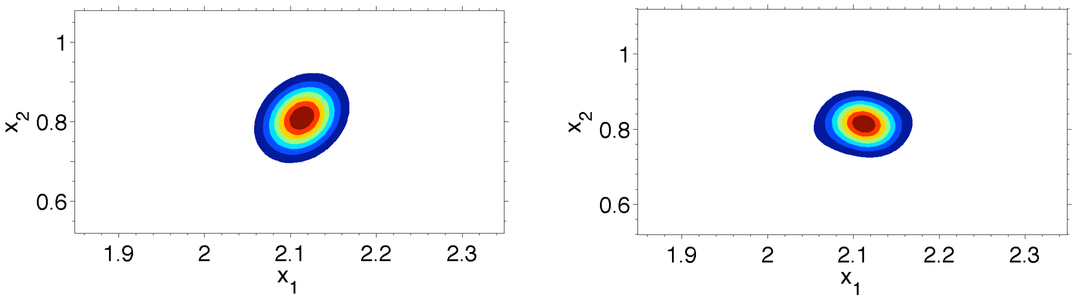
(a) (b)

Figure A-12: Error in the estimate of the MDN at the boundaries of the model space. Conditional PDF $p(\mathbf{t}|\ \mathbf{x})$ at $\mathbf{x} = (1.770, 0.782)$. (a) Estimated PDF using the MC sampling solution. (b) Estimated PDF using the MDN. Warm colors represent high probability areas.

analytical representation of the derivatives of the error function (equation (A-18)) with respect to the elements of the inverse of the covariance matrix– similar to those that we have already derived for the diagonal covariance matrix (equations (A-27) to (A-29)). What is more, as we explained above in order to estimate valid values of the diagonal elements of the covariance matrix, we need to represent these parameters as positive functions of the network outputs (equation (A-25)). In the case of a full covariance matrix, this problem will be more complicated because the covariance matrix and its inverse should be symmetric and positive definite. Therefore, we need to parameterize the positive definite matrices in such a way that: (1) the parameters can freely assume any real values (because the outputs of a neural network can assume any real values), (2) the determinant is a simple expression of the parameters, and (3) the correspondence is bijective. Williams (1996) developed such a parameterization and applied it to build an MDN with a single Gaussian kernel with a full covariance matrix. However, due to the increase in the number of the outputs of the neural network, extension of that method to cases with more than one kernel is computationally expensive and can result in instability of the network during training.

The error of the Gaussian mixture model of a PDF with abrupt cut-offs (such as uniform distribution at the boundaries, or exponential distribution at the origin) can be large. Gaussian distribution cannot be fitted to sharp cut-offs because of its shape and smoothness.

In our applications such cases can happen when inverting for model parameters near the boundaries of the model space. Figure A-12(a), shows an example where the maximum of the a posteriori PDF of the model parameters are near the boundary of t_2. In this example, the forward problem is the same as what we used in section A.4.1 above. Figure A-12(b) shows the result of the MDN inversion for this problem. This figure shows that the error of the MDN is large.



(a) (b)

Figure A-13: Error in the estimate of the MDN at the boundaries of the model space. The PDF of the input vector $p(\mathbf{x})$ from forward modeling of the conditional PDF $p(\mathbf{t}|\,\mathbf{x})$. (a) Forward modeling of the MC sampling result. (b) Forward modeling of the MDN result. Warm colors represent high probability areas.

The results of the forward modeling from the samples of the two PDF's in Figure A-12 are shown in Figure A-13. This figure shows that although the high probability areas of the joint PDF of (x_1, x_2) obtained from the MDN inversion result provides a good estimate of that obtained from the MC sampling inversion result, the boundary effect error in the MDN inversion result can cause relatively large errors in the estimate of data uncertainty.

One possible solution to the above problem is to use more kernels in the mixture density model. However, increasing the number of kernels will increase the computational cost of the solution. Another possible solution is to apply a mathematical transformation to parameters with sharp cut-off boundaries in their a priori PDF, and span them on a wider interval, which might smoothen sharp cut-off boundaries. In the petrophysical applications a logarithmic transform (Bosch et al., 2009) of porosity or water saturation can be used for that purpose. However, such approximations will result in sparse distribution of the limited training samples over the transformed model space, and consequently can increase the interpolation error of the MDN. It might also be possible to use other types of kernel such as exponential PDF's in combination with Gaussian PDF to model sharp cut-off boundaries.

A.7 References

Bishop, C. M., 1994, Mixture Density Networks.

Bishop, C. M., 1995, *Neural Networks for Pattern Recognition*: Oxford University Press.

Bosch, M., C. Carvajal, J. Rodrigues, A. Torres, M. Aldana and J. Sierra, 2009, Petrophysical seismic inversion conditioned to well-log data: Methods and application to a gas reservoir, *Geophysics*, 74(2), O1-O15.

Duda, R. O., P. E. Hart and D. G. Stork, 2001, *Pattern Classification*: New York: Wiley.

Fukunga, K., 1990, *Introduction to Statistical Pattern Recognition*: San Diego: Academic Press.

MacKay, D. J. C., 1992, Bayesian Interpolation, *Neural Computation*, 4(3), 415-447.

MacKay, D. J. C., 2003, *Information theory, inference, and learning algorithms*: Cambridge University Press.

McLachlan, G. and D. Peel, 2000, *Finite mixture models*: New York: Wiley.

Meier, U., A. Curtis and J. Trampert, 2007, Global crustal thickness from neural network inversion of surface wave data, *Geophysical Journal International*, 169, 706-722.

Nabney, I. T., 2004, *Netlab: Algorithms for Pattern Recognition*: London: Springer Verlag.

Rumelhart, D. E., G. E. Hinton and R. J. Williams, 1986, Learning representations by back-propagating errors, *Nature*, 323(6088), 533-536.

Webb, A. R., 1994, Functional approximation by feed-forward networks: a least-squares approach to generalization, *IEEE Transactions on Neural Networks*, 5(3), 363-371.

Williams, P. M., 1996, Using Neural Networks to Model Conditional Multivariate Densities, *Neural Computation*, 8(4), 843-854.

B Dvorkin-Gutierrez petrophysical forward model

Dvorkin and Gutierrez (2001) proposed the following forward rock physics model for a dispersed sand-clay mixture. The forward model is defined for two classes of facies: (1) sands and shaley sands, and (2) shales and sandy shales. For sands and shaley sands, a sand matrix with porosity, ϕ_s, is assumed; clay particles are dispersed in pore space between sand grains and cause a decrease in porosity. Therefore, as the clay content increases the pore space fills with stiffer material and the bulk and shear moduli increase. When the amount of clay exceeds the pure-sand porosity ϕ_s, the sand matrix starts to collapse and clay particles fill the contact between sand grains, which results in a softer rock. Therefore, for clay content values larger than ϕ_s, the bulk and shear moduli decrease as clay content increases. This behavior is modeled mathematically as follows:

$$V_{cl} < \phi_s : \qquad \phi = \phi_s - V_{cl}(1-\phi_c) \tag{B-1}$$

$$V_{cl} \geq \phi_s : \qquad \phi = V_{cl}\phi_c \tag{B-2}$$

In the above equations, ϕ is the porosity of the mixture, ϕ_c is the porosity of pure shale, ϕ_s is the porosity of pure sand, and V_{cl} is the clay content.

In shales and sandy shale mixture, the softer element, which is clay, envelops the stiffer element, which is sand. Such a topology, where softer element envelops the stiffer element, is a realization of the Hashin-Shtrikman lower bound. Therefore the bulk modulus, K_{mix}, and shear modulus, G_{mix}, of shales and sandy shale mixture can be modeled as (Hashin and Shtrikman, 1963; Mavko et al., 2009):

$$V_{cl} \geq \phi_s :$$

$$K_{\text{mix}} = \left[\frac{V_{cl}}{K_2 + 4/3G_2} + \frac{1 - V_{cl}}{K_s + 4/3G_2} \right]^{-1} - \frac{4}{3} G_2 \tag{B-3}$$

$$G_{\text{mix}} = \left[\frac{V_{cl}}{G_2 + Z_2} + \frac{1 - V_{cl}}{G_s + Z_2} \right]^{-1} - Z_2, \qquad Z_2 = \frac{G_2}{6} \frac{9K_2 + 8G_2}{K_2 + 2G_2} \tag{B-4}$$

In the above equations K_2 and G_2 are the bulk and shear moduli of fluid saturated pure shale matrix, respectively, and K_s and G_s are the bulk and shear moduli of sand particles. K_2 and G_2 are derived as explained below.

In sands and shaley sands clay particles fall in the pore space of the sand framework and do not significantly affect its stiffness. Dvorkin and Gutierrez (2001) argued that the bulk and shear moduli of sands and shaley sands can be described by the modified Hashin-Shtrikman lower bound where the soft end member is pure sand matrix and the stiff end member is the mixture of sand and clay at the critical clay content $V_{cl} = \phi_s$:

$$V_{cl} < \phi_s :$$

$$K_{\text{mix}} = \left[\frac{1 - V_{cl}/\phi_s}{K_1 + 4/3G_1} + \frac{V_{cl}/\phi_s}{K_{cc} + 4/3G_1} \right]^{-1} - \frac{4}{3} G_1 \tag{B-5}$$

$$G_{\text{mix}} = \left[\frac{1 - V_{cl}/\phi_s}{G_1 + Z_1} + \frac{V_{cl}/\phi_s}{G_{cc} + Z_1} \right]^{-1} - Z_1, \qquad Z_1 = \frac{G_1}{6} \frac{9K_1 + 8G_1}{K_1 + 2G_1} \tag{B-6}$$

In the above equations K_1 and G_1 are the bulk and shear moduli of pure sand matrix, respectively, and K_{cc} and G_{cc} are given by K_{mix} and G_{mix} from equations (B-3) and (B-4) at $V_{cl} = \phi_s$. K_1 and G_1 are derived as explained below.

The elastic moduli of the pure dry sand matrix (K_1 and G_1 in equations (B-5) and (B-6)) and pure dry shale matrix (K_2 and G_2 in equations (B-3) and (B-4)) can be estimated from intervals with these rock types at wellbores, provided such intervals with the texture of pure matrices can be found in wells under examination. Another way to estimate these elastic moduli is by using calibrated rock physics models. In chapter 2 we present an example of such models (Bachrach and Avseth, 2008). Here we present another theory, which is called the Hertz-Mindlin theory (Mindlin, 1949; Mavko et al., 2009). We apply the Hertz-Mindlin theory in the synthetic problem in chapter 3. In the Hertz-Mindlin theory the bulk and shear moduli of the pure dry sand matrix and pure dry shale matrix are represented as:

$$K_{i\_dry} = \left[\frac{n_i^2 (1-\phi_i)^2 \mu_i}{18\pi^2 (1-\nu_i)^2} P \right]^{1/3} \tag{B-7}$$

$$\mu_{i\_dry} = \frac{5-4\nu_i}{5(2-\nu_i)} \left[\frac{3 n_i^2 (1-\phi_i)^2 \mu_i}{2\pi^2 (1-\nu_i)^2} P \right]^{1/3} \tag{B-8}$$

In the above equations i is an index, which can be 's' for pure sand matrix and 'c' for pure shale matrix. P is the effective pressure, μ_i and ν_i are the shear modulus and Poisson's ratio of the grain materials (i.e., sand particles for sands and clay particles for shale), and n_i is the coordination number (i.e., the average number of contacts per grain). The coordination number can be a function of both porosity and effective pressure (Bachrach and Avseth, 2008). In our synthetic application in chapter 3 we use Murphy approximation (Avseth et al., 2005) to represent coordination number as a function of porosity as:

$$n_i = 20 - 34\phi_i + 14\phi_i^2 \tag{B-9}$$

The effective pressure is defined as the difference between overburden stress and pore pressure, which are assumed to be hydrostatic. Therefore, effective pressure is represented as a function of depth as:

$$P = g\int_0^z (\rho_b - \rho_f)\, dz \tag{B-10}$$

where ρ_b is bulk density, ρ_f is fluid density, z is depth, and g is the gravitational constant.

Porosity of the pure-sand and clay matrix is a decreasing function of depth due to the compaction effect (Avseth et al., 2005). The relationship between porosity and depth is usually approximated by an exponential function:

$$\phi_i = \phi_{i0} \exp(-\gamma_i z) \tag{B-11}$$

Where ϕ_{i0} is the depositional porosity (or critical porosity) of sand or shale, and γ_i is a constant with different values for sand and shale deposits.

The density of the fluid saturated rock, ρ_{mix}, is given as:

$$V_{cl} < \phi_s: \qquad \rho_{\mathrm{mix}} = (1-\phi_s)\rho_s + V_{cl}(1-\phi_c)\rho_c + \phi\rho_f \tag{B-12}$$

$$V_{cl} \geq \phi_s : \qquad \rho_{\text{mix}} = (1 - V_{cl})\rho_s + V_{cl}(1 - \phi_c)\rho_c + \phi\rho_f \tag{B-13}$$

where ρ_s, ρ_c, and ρ_f are the density of sand particles, clay particles, and fluid, respectively.

The bulk and shear modulus of the fluid saturated pure sand and fluid saturated pure shale matrices, are given by Gassman's law as:

$$K_j = K_i \frac{K_{i\_dry}/(K_i - K_{i\_dry}) + K_f/(\phi_i(K_i - K_f))}{1 + K_{i\_dry}/(K_i - K_{i\_dry}) + K_f/(\phi_i(K_i - K_f))} \tag{B-14}$$

$$G_j = G_{i\_dry} \tag{B-15}$$

where j is equal to 1 for fluid saturated pure sand matrix, with i equal to 's'. For fluid saturated pure shale matrix j is equal to 2 and i is equal to 'c'. K_{i_dry} and G_{i_dry} are the bulk and shear moduli of the dry frame matrices of sand and shale and are derived from equations (B-7) and (B-8).

The fluid bulk modulus and density, K_f and ρ_f, are functions of fluid saturation. For a mixture of brine and oil if we assume the pore fluid is uniformly distributed in the pores, these parameters are given as:

$$K_f = \left[\frac{s_w}{K_w} + \frac{1 - s_w}{K_{hc}} \right]^{-1} \tag{B-16}$$

$$\rho_f = s_w\rho_w + (1 - s_w)\rho_{hc} \tag{B-17}$$

where s_w is water saturation, K_w and ρ_w are the bulk modulus and density of brine, K_{hc} and ρ_{hc} are the bulk modulus and density of oil. The bulk moduli and densities of brine and oil are functions of pore pressure and temperature (Batzle and Wang, 1992).

Using equations (B-3) and (B-4) for the bulk and shear moduli of sandy shales and shales, (B-5) and (B-6) for the bulk and shear moduli of shaley sands and sand, and (B-12) and (B-13) for the density of shaley sands and sandy shales, respectively, the compressional and shear wave velocity of the dispersed sand and shale mixture are given by:

$$V_p = \sqrt{\left(K_{\text{mix}} + 4/3G_{\text{mix}}\right)/\rho_{\text{mix}}} \tag{B-18}$$

$$V_s = \sqrt{G_{\text{mix}}/\rho_{\text{mix}}} \tag{B-19}$$

Equations (B-18) and (B-19) imply that the mixture of sand and shale is isotropic and elastic.

In the synthetic application in chapter 3, we assumed that the depositional porosity of sand, ϕ_{s0}, was equal to 0.45, the depositional porosity of shale, ϕ_{c0}, was equal to 0.60, the compaction factor of sand, γ_s, was equal to 0.127 km^{-1}, the compaction factor of shale, γ_c, was equal to, .45 km^{-1}, the bulk modulus of brine, K_w, was equal to 2.80 GPa and the density of brine was equal to 1.09 g/cc. We assumed that both the pore pressure and overburden stress are hydrostatic. As a result effective pressure, which is a function of the difference between above two pressures, is also hydrostatic and can be represented as $P=(\rho_{\mathrm{mix}}-\rho_f)gz$. The density, gravity and gas-oil-ratio of oil, which were required to estimate the density and bulk modulus of oil as a function of pressure were assumed to be equal to 0.78 g/cc, 32 API, and 64 $\mathrm{Sm}^3/\mathrm{Sm}^3$, respectively, at standard conditions. All of the above data were obtained from Avseth et al. (2005). Shale is not a granular composite such as sand. Therefore, the validity of applying equations (B-7), (B-8), and (B-14) to pure shale is questionable. However, there is evidence that these equations provide reasonable elastic property estimates for shale (Dvorkin and Gutierrez, 2001; Avseth et al., 2005, p. 98). We do not promote applying the above equations for pure shale and we only applied them in the synthetic case in chapter 3 to show that the MDN inversion method can solve inverse problems with high dimensional model space.

The above petrophysical model can predict the same values of V_p and V_s for sandy shales and shaley sands. Figure B-1 shows the predictions of the above equations where $K_1 = 37$ GPa, $G_1 = 44$ GPa, $K_2 = 25$ GPa, $G_2 = 9$ GPa, $z = 0.5$ km, $s_w = 1$, $\phi_{s0} = 0.45$, $\phi_{c0} = 0.60$, $\gamma_s = 0.127\ \mathrm{km}^{-1}$, $\gamma_c = 0.45\ \mathrm{km}^{-1}$, $\rho_s = 2.65$ g/cc, $\rho_c = 2.55$ g/cc, $K_w = 2.95$ GPa, and $\rho_w = 1.044$ g/cc. Equations (B-1) and (B-2) define two separate intervals of the clay content axis for sand and shaley sand, and shale and sandy shale, which can be seen in Figure B-1(a). At the transition between shaley sand and sandy shale, where the pore space is completely filled with clay particles, i.e., at $V_{cl} = \phi_s$, the porosity reaches to its minimum. At the same value of clay content the compressional and shear wave velocities reach to their maximum, as clay particles, which are stiffer than pore fluid completely fill the pore space. After transition between shaley sand and sandy shale, the clay particles fill the contact between sand grains and soften the sand matrix, and thus compressional and shear velocities decrease. The described behavior of compressional and shear-wave velocity as a function of clay content is shown in Figure B-1(b) and Figure B-1(d).

In the same manner, the behavior of compressional and shear-wave velocity as a function of porosity are shown in Figure B-1(c) and Figure B-1(e). These two figures show that as we indicated above the maximum of compressional and shear velocity occurs at the minimum of porosity. The separation between V_p-porosity relation for sandy shale and shaley sand

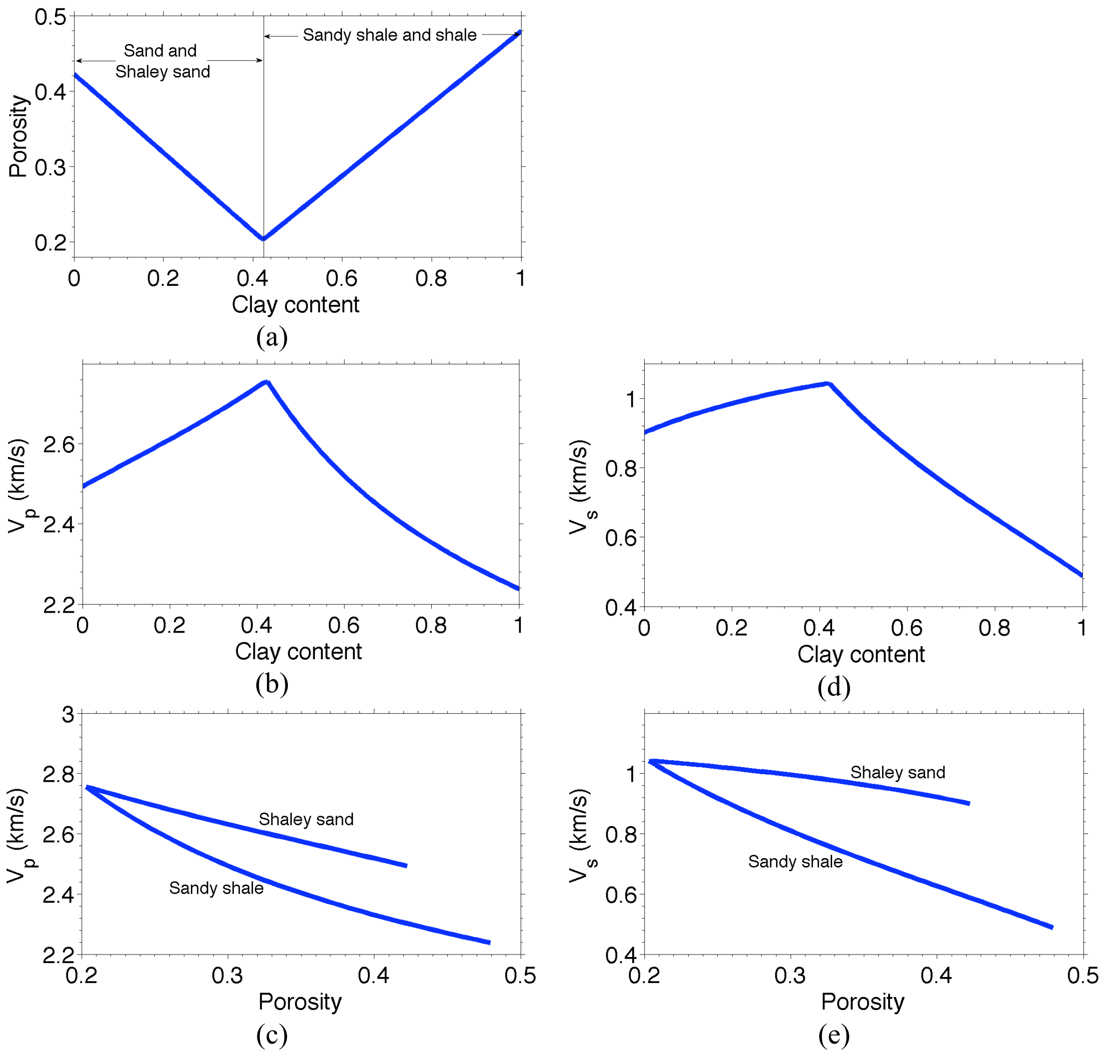


Figure B-1: Predictions of Dvorkin-Gutierrez petrophysical forward model. (a) Porosity as a function of clay content. (b) Compressional velocity as a function of clay content. (c) Compressional velocity as a function of porosity. (d) Shear velocity as a function of clay content. (e) Shear velocity as a function of porosity. Note the maximum of V_p and minimum of porosity at the transition from shaley sand to sandy shale, where $V_{cl} = \phi_s$.

depends on the contrast between elastic moduli of sand and shale matrices. Large differences between the elastic moduli of the pure sand matrix (i.e., K_1 and K_2) and pure shale matrix (i.e., G_1 and G_2) result in a large separation between V_p-porosity relation for sandy shale and shaley sand.

Figure B-1 shows that two different values of clay content (or porosity) can result in the same value of V_p. It means that inversion of a given value of V_p can result in two different values of clay content (or porosity). As a result, as we discussed in chapter 3 and chapter 4, the petrophysical inverse problem is required to be solved probabilistically to address the uncertainty due to the above strong non-linearity in the petrophysical forward function.

B.1 References

Avseth, P., T. Mukerji and G. Mavko, 2005, *Quantitative seismic interpretation: applying rock physics tools to reduce interpretation risk*. Cambridge ; New York: Cambridge University Press.

Bachrach, R. and P. Avseth, 2008, Rock physics modeling of unconsolidated sands: Accounting for nonuniform contacts and heterogeneous stress fields in the effective media approximation with applications to hydrocarbon exploration, *Geophysics*, 73(6), E197-E209.

Batzle, M. and Z. Wang, 1992, Seismic Properties of Pore Fluids, *Geophysics*, 57, 1396-1408.

Dvorkin, J. and M. A. Gutierrez, 2001, Grain sorting, porosity, and elasticity, from pangea.stanford.edu/~jack/Bimodal.pdf.

Hashin, Z. and S. Shtrikman, 1963, A variational approach to the theory of the elastic behaviour of multiphase materials, *Journal of the Mechanics and Physics of Solids*, 11(2), 127-140.

Mavko, G., T. Mukerji and J. Dvorkin, 2009, *The rock physics handbook: tools for seismic analysis of porous media*. Cambridge ; New York: Cambridge University Press.

Mindlin, R. D., 1949, Compliance of bodies in contact, *Journal of Applied Mechanics*, 16, 259-268.

C Petrophysical forward relations with effective porosity

The petrophysical forward relations used in chapter 4 is similar to the model proposed by Dvorkin et al. (2007), in the sense that both models assume that compressional and shear-wave velocity are functions of effective porosity instead of total porosity. In the petrophysical forward relationships the porous wet shale is treated as a part of solid grain and the porosity within this shale is excluded from the total porosity. The intrinsic porosity of clay ϕ_{clay} is obtained from samples of shale in the field (Dvorkin et al., 2007). The effective porosity, ϕ_e, is defined as the void space outside the porous clay (in a unit volume of rock):

$$\phi_e = \phi_t - V_{cl}\phi_{\text{clay}} \qquad \text{(C-1)}$$

Where ϕ_t is the total porosity, and V_{cl} is the volume occupied by porous clay (clay content).

The bulk and shear moduli of wet shale, $K_{P\text{clay}}$ and $\mu_{P\text{clay}}$, depend on the intrinsic porosity of clay ϕ_{clay} as well as on the effective pressure, the texture and mineralogy of clay, and the bulk modulus of saturating fluid which is brine. These moduli can be estimated by using a rock physics model calibrated with data from wells. Here, by assuming a constant intrinsic porosity, the bulk and shear moduli of wet shale are modeled as functions of effective pressure using hyperbolic tangent fits (Chao et al., 2009).

The bulk and shear moduli of the dry frame of non-shale part of the rock, K_d and μ_d, depend on effective pressure and effective porosity. One way of estimating these moduli is by using those of a 100% brine saturated interval, provided that such an interval with the same texture and properties can be found in the well under examination (Dvorkin et al., 2007). Another way is by using a rock physics model that is calibrated with well log data (Avseth et al., 2005, ch. 2). Here, the bulk and shear moduli of dry frame are modeled as functions of effective pressure and effective porosity using hyperbolic tangent fits (Chao et al., 2009).

The Gassmann theory is used to perform fluid substitution to obtain the bulk and shear moduli of the fluid saturated non-shale part of the rock, K_{QS} and μ_{QS}:

$$K_{QS} = K_d + \left[1 - \frac{K_d}{K_0}\right]^2 \Bigg/ \left[\frac{\phi_e}{K_f} + \frac{(1-\phi_e)}{K_0} + \frac{K_d}{K_0^2}\right] \tag{C-2}$$

$$\mu_{QS} = \mu_d \tag{C-3}$$

Here K_d is the bulk modulus of the dry non-shale frame, K_0 is the bulk modulus of the mineral, which is quartz for saturated sand, K_f is the bulk modulus of saturating fluid, and μ_d is the shear modulus of the dry non-shale frame.

The pore fluid bulk modulus K_f is the bulk modulus of the mixture of brine and hydrocarbon inside the effective pore space. By assuming a homogeneous mixture of brine and hydrocarbon K_f is obtained as:

$$K_f = \left[\frac{S_W}{K_W} + \frac{1-S_W}{K_H}\right]^{-1} \tag{C-4}$$

In the above equation, S_W is brine saturation, K_W is the bulk modulus of brine, and K_H is the bulk modulus of hydrocarbon.

The petrophysical forward function is defined for two facies: (1) reservoir facies with clay content value smaller than 55% ($V_{cl} < 0.55$) and effective porosity value larger than 5% ($\phi_e > 0.05$), (2) non-reservoir facies with clay content value larger than 55% ($V_{cl} > 0.55$) or effective porosity value smaller than 5% ($\phi_e < 0.05$).

In the non-reservoir facies wet shale envelops sand grains. The Reuss average is used to model the bulk K and shear G moduli of the mixture:

$$K = \left(\sum_{i=1}^{m} f_{S_i} K_{S_i}^{-1}\right)^{-1} \tag{C-5}$$

$$\mu = \left(\sum_{i=1}^{m} f_{S_i} \mu_{S_i}^{-1}\right)^{-1} \tag{C-6}$$

here, $m = 2, f_{S1} = V_{cl}$ is the volume fraction of wet shale, $f_{S2} = 1 - f_{S1}$ is the volume fraction of sand grains, $K_{S1} = K_{P\text{clay}}$ and $\mu_{S1} = \mu_{P\text{clay}}$ are the bulk and shear moduli of wet shale, and $K_{S2} = K_Q$ and $\mu_{S2} = \mu_Q$ are the bulk and shear moduli of sand grains.

The bulk and shear moduli of the reservoir facies are also modeled using Reuss average (equations (C-5) and (C-6)). In this case $m = 2, f_{S1} = V_{cl}$ is the volume fraction of wet shale, $f_{S_2} = 1 - f_{S_1}$ is the volume fraction of sand frame, $K_{S1} = K_{P\text{clay}}$ and $\mu_{S1} = \mu_{P\text{clay}}$ are the bulk and shear moduli of wet shale, and $K_{S2} = K_{QS}$ and $\mu_{S2} = \mu_{QS}$ are the bulk and shear moduli of fluid saturated sand matrix (equations (C-2) and (C-3)). Additional terms can be added to Reuss average to account for other mixture components like quartz particles in the wet shale, but we do not discuss these details in this appendix.

Porosity in general and porosity of wet shale in particular are decreasing functions of depth. Therefore, density of wet shale is an increasing function of depth, d, and modeled by an empirical relation as:

$$\rho_{sh}(d) = ad^b \qquad \text{(C-7)}$$

Parameters a and b in the above equation are empirical constants which are estimated in the calibration process.

Density of the reservoir and non-reservoir facies obtained from the density of wet shale ρ_{sh}, sand grains ρ_q, and pore fluid ρ_f as:

$$\rho = V_{cl}\,\rho_{sh} + (1 - V_{cl} - \phi_e)\rho_q + \phi_e\,\rho_f \qquad \text{(C-8)}$$

The pore fluid density ρ_f is the density of the mixture of brine and hydrocarbon inside the effective pore space. By assuming a homogeneous mixture of brine and hydrocarbon, it is obtained as $\rho_f = S_W\,\rho_W + (1 - S_W)\rho_{HC}$. Where, S_W is brine saturation, ρ_W is the density of brine, and ρ_H is the density of hydrocarbon.

Assuming that the earth is isotropic and linearly elastic, the seismic behavior of sediments can be completely characterized by three parameters: bulk modulus, shear modulus, and bulk density. The P- and S-wave impedance for reservoir and non-reservoir facies are derived from the bulk and shear moduli of saturated rock (equations (C-5) and (C-6)), and bulk density (equation (C-8)):

$$I_p = \sqrt{\rho\left(K + 4/3\mu\right)} \qquad \text{(C-9)}$$

$$I_p = \sqrt{\rho\,\mu} \qquad \text{(C-10)}$$

The above petrophysical forward relations are calibrated with data from five wells in the field under examination in this thesis. The calibration error I_p and I_s are equal to 4% and 6%, respectively (Chao et al., 2009). The calibrated petrophysical forward function and its uncertainty are used in the petrophysical inversion in chapters 3 and 4.

C.1 References

Avseth, P., T. Mukerji and G. Mavko, 2005, *Quantitative seismic interpretation: applying rock physics tools to reduce interpretation risk*. Cambridge ; New York: Cambridge University Press.

Chao, G., G. Lambert and H. Cumming, 2009, Analysis of Intrinsic Uncertainties of Petro-elastic Models Using Simulated Annealing, 71st EAGE Conference & Exhibition, Amsterdam.

Dvorkin, J., G. Mavko and B. Gurevich, 2007, Fluid substitution in shaley sediment using effective porosity, *Geophysics*, 72(3), O1-O8.