Reservoir Lithology Determination from Seismic Inversion Results Using Markov Processes

Proefschrift

ter verkrijging van de graad van doctor

aan de Technische Universiteit Delft,

op gezag van de Rector Magnificus prof. Ir. K.C.A.M. Luyben,

voorzitter van het College voor Promoties,

in het openbaar te verdedigen op

vrijdag 1 september 2017 om 09:00 uur

door

Runhai Feng

Master of Science,

Nanjing University, Nanjing, P. R. China

geboren te Jiangsu, P. R. China

Dit proefschrift is goedgekeurd door promotor:

Prof. dr. Stefan M. Luthi

Copromotor:

Dr. ir. G. G. Drijkoningen

Samenstelling promotiecommissie:

Rector Magnificus	voorzitter
Prof. dr. Stefan M. Luthi	Technische Universiteit Delft
Dr. ir. G. G. Drijkoningen	Technische Universiteit Delft

Onafhankelijke leden:

Prof. dr. ir. Dries Gisolf	Technische Universiteit Delft
Prof. dr. ir. Allard W. Martinius	Technische Universiteit Delft
Prof. dr. ir. Gert Jan Weltje	Katholieke Universiteit Leuven
Dr. Philip Christie	Schlumberger Cambridge Research
Dr. Erika Angerer	OMV Exploration & Production

Reservoir Lithology Determination from Seismic Inversion Results Using Markov Processes

Runhai Feng

This thesis is dedicated to my parents

(谨以此文献给我的父母)

CONTENTS

Summary

For reservoir characterization, subsurface heterogeneity needs to be qualified in which the distribution of lithologies is an essential part since it determines the location and migration paths of hydrocarbons. Preliminary analysis of well-log data could help to identify various lithologies in a one-dimensional direction (depth), while the lateral information is missing because of the sparse locations. On the other hand, a larger areal coverage of the target reservoir could be provided by seismic data, and from the inversion thereof, inferences of lithologies could be made.

However, just like other geophysical inversions, translation of seismic inversion results to these categorical variables (lithologies) is a non-unique problem, which means that different lithologies could produce the same, or similar, property responses. In order to mitigate this problem, geological prior information should be introduced in the sense of Bayes' theorem. Thus, the main motivation for this thesis is to investigate the usage of geological prior information in the classification of reservoir lithologies from properties obtained from seismic inversion. Different methods have been tried in this process in order to fully understand their performances and to make comparisons.

A detailed synthetic model has been built as a virtual asset that has been used to assess the performance of the seismic inversion method since the inversion results are used as inputs for the determination of lithologies. Instead of using seismic attributes such as AVO attributes or acoustic impedance obtained from inversion, the full-elastic wave-equation based inversion scheme applied here can provide a high resolution properties and geometries in the subsurface. This is because of the non-linear character of the wave-equation based inversion, which takes the full elastic wave propagation into account, including internal multiple scattering as well as multiple mode-conversions.

A new Markov Chain method (A-CMC) has been developed in order to take the layers' dips into account during simulation of lithologies. Structural interpretation of the seismic data, or of the inversion results, could provide general information such as the dipping direction and angles which are essential inputs for the simulation.

In the simulation of the Markov Chain model (A-CMC), seismic interpretation is used only as a guidance for subsurface structures. The Hidden Markov Models (HMMs) is then applied to classify the lithologies with the properties from seismic inversion as inputs, while honouring the geological depositional processes contained in the Markov Models. This means that data samples in the subsurface have not been treated independently, but a relationship between adjacent pixels in the subsurface has been introduced by the use of a transition matrix.

A disadvantage of HMMs is that only the vertical coupling is considered and the lateral information is missing. The Markov Random Field (MRF) method could enforce dependencies vertically as well as horizontally. Profile Markov matrices should be designed in order to reflect the geological prior information.

Through the comparison with other statistical methods such as *k*-Means or Fuzzy Logic, we found that the approaches that make use of geological prior information outperform the methods that do not. In order to construct the prior information, be it a transition matrix or profile Markov matrices, geological understanding of the regional settings is needed and cored wells should be used.

Both in the simulation and in the prediction of lithologies, seismic data, with, or without inversion, plays a role. Because the quality of the seismic data is very critical for the accuracy and resolution of the seismic inversion, and therefore crucial for the lithology prediction process, high quality state-of-the-art seismic acquisition and processing are important for successful application of our method.

Samenvatting

Voor reservoir karakterisering moet de heterogeniteit van de ondergrond worden bepaald. In dit proces is de distributie van lithologieën een essentieel onderdeel, omdat deze de locatie en migratiepaden van de koolwaterstoffen bepaalt. Een analyse van well-log data kan helpen om verschillende te identificeren in een eendimensionale richting (diepte), terwijl laterale informatie ontbreekt als een gevolg van het feit dat de putlocaties ver uit elkaar liggen. Aan de andere kant kan seismische data zorgen voor een bedekking van de gehele oppervlakte van het doelgebied en uit de inversie van de seismische data, gevolgtrekkingen ten aanzien van de lithologieën kunnen worden gemaakt.

Echter, het vertalen van de resultaten van seismische inversie resultaten in categorische variabelen (lithologieën) is een niet uniek probleem wat betekent dat verschillende lithologische kenmerken dezelfde of vergelijkbare responsies in termen van gesteente eigenschappen kunnen geven. A-priori geologische informatie kan geïntroduceerd worden volgens de theorie van Bayes. De belangrijkste motivatie voor dit proefschrift is om het potentiele nut te onderzoeken voor het gebruik van geologische a priori informatie in de classificatie van reservoir lithologieën. Verschillende methoden worden gebruikt om dit proces beter te begrijpen en de resultaten te kunnen vergelijken.

Een gedetailleerd synthetisch model is gebouwd als een virtuele asset die wordt gebruikt om de werking van de seismische inversie methode te analyseren, omdat de inversie resultaten worden gebruikt als input voor het bepalen van lithologieën. In plaats van gebruik te maken van seismische attributen, zoals AVO en de akoestische impedantie verkregen door inversie, levert een inversie schema gebaseerd op de volledig elastische golfvergelijking een hoge resolutie van eigenschappen en geometrieën in de ondergrond op. Dit is het gevolg van het niet-lineaire karakter van deze inversie die de volledige elastische golfpropagatie, inclusief meervoudige reflecties en conversies van longitudinale naar transversale golven, beschrijft.

Een nieuwe Markov Chain methode (A-CMC) is ontwikkeld om rekening te houden met de inclinatie van de lagen tijdens de simulatie van lithologieën. Structurele interpretatie van seismische data, of van de resultaten van seismische inversie, kan informatie geven over azimut en inclinatie van de lagen, wat essentiële input is voor de simulatie.

Structurele seismische interpretatie wordt alleen gebruikt als richtlijn voor ondergrondse structuren in de simulatie van het Markov Chain model (A-CMC). De seismische inversie resultaten dienen als input voor de Hidden Markov Models (HMMs) methode om lithologieën te classificeren, waarbij rekening wordt gehouden met de geologische depositionele processen. Er wordt een transitie matrix geïntroduceerd die de afhankelijkheid van de lithologieën van nabijgelegen punten in de ondergrond beschrijft.

Een nadeel van HMMs is dat alleen de verticale koppeling wordt meegenomen en laterale informatie ontbreekt. De Markov Random Field methode kan verticale en horizontale

afhankelijkheid opleggen. Profile Markov matrices moeten ontworpen worden die deze geologische a priori informatie reflecteren.

Als we onze methode met andere statistische methode vergelijken, zoals *k*-Means of Fuzzy Logic, blijkt dat de benaderingen die geologische a priori informatie in beschouwing nemen duidelijk beter werken dan de benaderingen die dit niet doen. Geologisch begrip van de regionale omgeving is essentieel en boorkernen uit putten moeten bekeken worden om a priori geologische informatie in te kunnen bouwen in de transitie matrix of in de profile Markov matrices.

Seismische data en de inversie daarvan spelen een belangrijke rol in de simulatie en in de voorspelling van lithologieën. De kwaliteit van de seismische data is zeer belangrijk voor de nauwkeurigheid en de resolutie van de seismische inversie en daardoor voor de betrouwbaarheid van de lithologie predicties. Daarom zijn geavanceerde acquisitie en verwerking van data van groot belang voor succesvolle toepassing van onze methode.

1

Introduction

1.1. Geological Prior Information

While the industry suffers from the drop in oil price because of weak demands and rising supplies since 2015, geophysicists also suffer from the problem of non-uniqueness in inversion for a very long time. On the other hand, geological information is available and could be brought in to mitigate this problem.

In general, geological information can be "provided as a prior component of the solution i.e. information that exists before the solution is formed and which is incorporated into the solution. Such information is termed as geological prior information" (Wood and Curtis 2004). This kind of information can be derived from well logs, general knowledge of geological settings, similar oil reservoirs, previous experiences or observations of other people etc. For example, the data from well logs can provide initial or starting property models for full-waveform inversion, while general knowledge of geological settings can regulate the inversion outputs in order to exclude unrealistic results, especially during the lithology classification process, for example that a water sand cannot be on top of a gas or oil sand in a continuous reservoir. Previous experiences can be formalized as empirical rock physical models which are potentially important tools for reservoir characterization. Such models can be based on outcrop analogues when the subsurface properties cannot be characterized adequately.

1.2. Bayes' Theorem

In order to address the non-uniqueness and uncertainty which exists in geophysical inversion problems (Buland and Omre 2003), the process can be carried out by Bayes' theorem (Bayes 1764) in the probability domain. The inversion problem of non-uniqueness can be overcome by using *a priori* information on the parameters (Duijndam 1988a; b). If this is done in the form of Bayesian inference, then the prior information will be probabilistic.

The procedure of Bayesian inversion consists of combining the prior information with the information included in the data in order to derive a refined statistical distribution, i.e. a posterior model distribution (Tarantola and Valette 1982; Tarantola 1987; Duijndam 1988a; b; Scales and Snieder 1997). The solution of the inverse problem will not be limited to a single set of predicted parameters but is represented by a probability density function (PDF) in the model space (Buland and Omre 2003). The posterior PDF is the result after the data and the prior information have been assimilated. The reason for using the data is that a posterior information helps constraining the model more tightly than the prior model distribution alone (Scales and Snieder 1997).

In geophysics, usually, the prior information can come from the analysis of other datasets such as the well logs, petrophysical experiments, core data etc. In non-linear full-waveform inversion (Tarantola 1984; Lam *et al*. 2007; Gisolf and van den Berg 2010), a detailed and high-resolution geological model can provide good prior information for the inversion. One can then analyze the uncertainty of posterior results to see how credible they are (Gouveia and Scales 1998).

The simplest form of Bayes' theorem is as follows:

$$\Pr(B|A) \propto \Pr(A|B)\Pr(B) \tag{1.1}$$

where $\Pr(B)$ is the probability of event B, $\Pr(A|B)$ and $\Pr(B|A)$ are the conditional probabilities of the A, B events given that the B, A events happen respectively.

The whole procedure of Bayes' theorem in equation (1.1) can be demonstrated in Figure 1.1 in which the inversion of density and compressional velocity is performed. It can be seen that by integrating the prior information, the posterior result has been narrowed down.

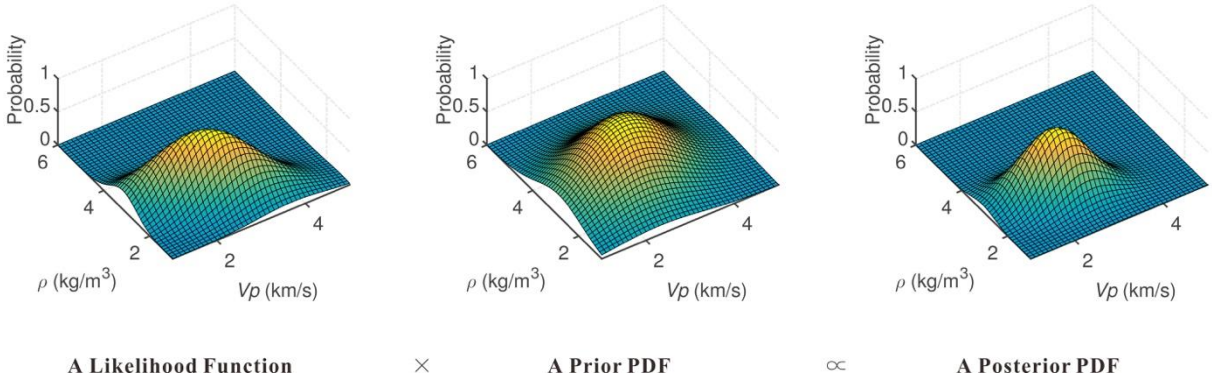


Figure 1.1: A schematic diagram of Bayes' theorem. The posterior PDF ($\Pr(B|A)$) is proportional to the product of the prior PDF ($\Pr(B)$) and the likelihood function ($\Pr(A|B)$).

1.3. Geological Modelling

As one type of the geological prior information, reservoir models can provide helps and guidance during the inversions as well as in the reservoir simulations. The theory of sequence stratigraphy is important in the construction of these models (van Wagoner 1995; 1998; Catuneanu *et al*. 2011). Because of the limited subsurface information available, one is often unable to determine a model which reflects the true stratigraphy and structure of the target. This calls for probabilistic approaches that should provide the degree of uncertainty associated with spatial reservoir characteristics (O'Byrne and Flint 1993). With the help of high-resolution sequence stratigraphy, the accuracy of a particular model can be improved, with the elements of the model being assigned probabilities close to 1.0 for geometries and spatial positions on which interpreters generally agree upon. Such a constraint provided by sequence stratigraphy and probabilistic modelling can be spelled out in the form of the Bayes inversion theorem (Ulrych *et al*. 2001), whereby one would not only indicate the result of the problem by incorporating the prior information, but also its uncertainty. In other words, sequence stratigraphy and other geological information such as well logs will be the prior knowledge in the Bayesian theorem of reservoir modelling.

In reservoir modelling, the geometrical data such as vertical sequences are used to fit the information of the subsurface, which raises concerns on the appropriateness. The sequence stratigraphic interpretation is already based on observations from the subsurface, such as well logs or cores, and may thus lead to circular reasoning. Many attributes essential to reservoir

characterization and engineering can be predicted from sequence stratigraphy, such as petrophysical properties, the degree of lithological heterogeneity, facies associations, and 3-D continuities and geometries of strata (Cross *et al*.1993). Thus, sequence stratigraphy provides also a 2-D and 3-D geometric framework. It can be implemented in the following two ways: (1) Through a high-resolution correlation process to obtain a more accurate representation of rock properties in spatial and temporal contexts; (2) Through the relationship between changes in accommodation space and the different stratal architectures and their associated facies preserved in the stratigraphic record (Cross *et al*. 1993).

Another way for the stratigraphic analysis to contribute to reservoir characterization and modelling is the transformation or conversion of information in one dimension such as well logs to 3-D predictions of stratigraphy. This means that information collected at one scale can help to establish the 3-D distributions of properties at other scales, which makes the predictions of the stratal geometries and petrophysical attributes, facies compositions and arrangements more accurate as the temporal resolution increases. The development of combined statistical and empirical methods in forward and inverse stratigraphic modeling is believed to have the potential to lead to more realistic and accurate predictions of the inter-well population of reservoir parameters (Cross *et al*.1993).

By incorporating seismic data that provide continuous 3-D images, the sparseness of well control information can be partially overcome. Constrained simulations of models using geostatistical methods can produce high-resolution 3-D realizations (Haas and Dubrule 1994).

1.4. Thesis Objectives

The main objective of this thesis is to bring geological prior information into geophysical inversions in order to produce better results. The output of the inversion is a distribution of the reservoir lithologies that are important for reservoir characterization, stratigraphic interpretations and inputs for reservoir simulations. Different methods have been introduced and compared with each other as well as in different dimensionalities (1-D vs. 2-D). As a rule in Bayes' theorem, all geological prior information had to be transformed into the probabilistic domain which is not easily to be done, since different geologists may have different interpretations and different probabilistic functions will have to be assigned. However, this problem is considered a topic aside or, at the very least, has been limited to a minimum, since the focus is on how geological prior information can help geoscientists and to what degree inversion results will be improved by it.

1.5. Thesis Outline

This thesis starts with building a detailed geological and petrophysical model of the Book Cliffs model, previously described by Hodgetts and Howell (2000) and developed by Tetyukhina *et al*. (2014). This geological model (Figure 1.2) includes eight depositional environments on which a non-linear acoustic full-waveform inversion has been applied to retrieve the media parameters.

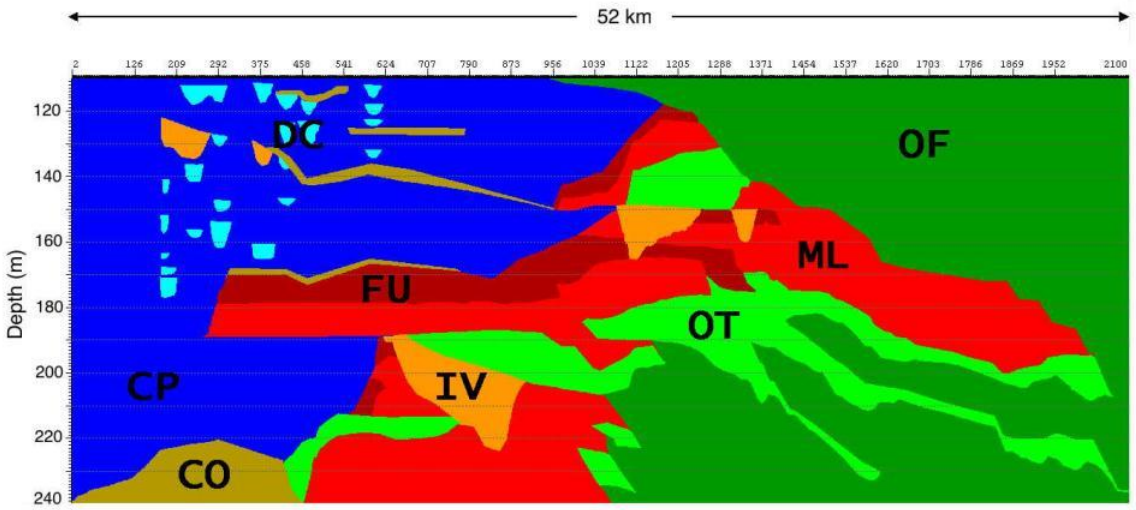


Figure 1.2: The fluvio-deltaic Book Cliffs model proposed by Tetyukhina *et al.* (2014) with the eight depositional environments termed "lithotypes". For details refer to the original publication.

However, the geological model by Tetyukhina *et al.* (2014) is relatively coarse. A more detailed model has been built for this study (Chapter 2) in order to include small reservoir units, i.e. sub-units within the eight lithotypes, that, however, may differ drastically in their reservoir properties. This model then serves as a basis for an iterative elastic wave-equation-based inversion scheme in order to retrieve the rock properties and layer geometries of the initial model.

In Chapter 3, the goal of simulating reservoir lithologies between cored wells is pursued with a new scheme of Coupled Markov Chain models. The dips of lithological layers are taken into considerations with searching angles during the Monte Carlo simulations. Seismic inversion results are used to provide information on the subsurface structures as well as the construction of horizontal matrices. They, however, have not been utilized to their fullest since they are connected to the types of lithologies intrinsically. As a kind of prior information in the simulation process, they have not been implemented in the sense of Bayes' theorem either.

Chapter 4 approaches the problem of lithological classification in which seismic inversion results are used for guidance. Since the reservoir lithologies are correlated with well-log data and seismic inversion data, lithologies are attempted to be inferred based on them. Differing from other approaches, here the seismic inversion results are proposed as inputs thereby avoiding the location bias if only wells are used. The vertical transitions between lithologies are constrained though the application of Hidden Markov Models (HMMs) during the classification process.

Since in Chapter 4 the horizontal coupling is missing in HMMs, the Hidden Markov Random Field (HMRF) is introduced in Chapter 5 as a 2-D lithology classifier. The Gibbs priori as well as the profile Markov matrix are used as geological prior information with the goal of guiding the vertical and horizontal continuities and the preferential transitions.

In Chapter 6, finally, the general conclusions are summarized.

References

Bayes, T. 1764. An Essay Toward Solving a Problem in the Doctrine of Chances. *Philosophical Transactions of the Royal Society of London*, **53**, 370-418.

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

Catuneanu, O., Galloway, W.E., Kendall, C.G., Miall, A.D., Posamentier, H.W., Strasser, A., Tucker, M.E. 2011. Sequence stratigraphy: methodology and nomenclature. *Newsletters on Stratigraphy*, **44**(3), 173-245.

Cross, T.A. et al. 1993. Applications of high-resolution sequence stratigraphy to reservoir analysis. *In*: Eschard, R. & Doligez, B. (eds) *Subsurface Reservoir Characterization from Outcrop Observations*, Editions Technip, Paris, 11-33.

Duijndam, A.J.W. 1988a. Bayesian Estimation in Seismic Inversion. Part Ⅰ: PRINCIPLES. *Geophysical Prospecting*, **36**(8), 878-898.

Duijndam, A.J.W. 1988b. Bayesian Estimation in Seismic Inversion. Part Ⅱ: Uncertainty ANALYSIS. *Geophysical Prospecting*, **36**(8), 899-918.

Gisolf, A. & van den Berg, P.M. 2010. *Target oriented non-linear inversion of seismic data*. P308, presented at the 72nd EAGE Annual Meeting, Barcelona, 14-17 June.

Gouveia, W.P. & Scales, J.A. 1998. Bayesian seismic waveform inversion: Parameter estimation and uncertainty analysis. *Journal of Geophysical Research: Solid Earth*, **103**(B2), 2759-2779.

Haas, A. & Dubrule, O. 1994. Geostatical inversion-a sequential method of stochastic reservoir modelling constrined by seismic data. *First break*, **12**(11), 561-569.

Hodgetts, D. & Howell, J.A. 2000. Synthetic seismic modelling of a large-scale geological cross-section from the Book Cliffs, Utah, USA. *Petroleum Geoscience*, **6**(3), 221-229.

Lam, C.H., van den Berg, P.M., and Gisolf, A. 2007. *Nonlinear inversion of focused data*. Presented at the 77th SEG Annual Meeting, San Antonio, 23-26 September.

O'Byrne, C.J. & Flint, S. 1993. High-resolution sequence stratigraphy of Cretaceous shallow marine sandstones, Book Cliffs outcrop, Utah, USA-application to reservoir modelling. *First Break*, **11**(10), 445-459.

Scales, J. A. & Snieder, R. 1997. To Bayes or not to Bayes?. *Geophysics*, **62**(4), 1045-1046.

Tarantola, A. & Valette, B. 1982. Inverse problems = quest for information. *J. geophys*, **50**(3), 150-170.

Tarantola, A. 1984. Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, **49**(8), 1259-1266.

Tarantola, A. 1987. *Inverse problem theory*. Elsevier Science Publ. Co., Inc.

Tetyukhina, D., Luthi, S.M. & Gisolf, A. 2014. Acoustic non-linear full-waveform inversion on an outcrop-based detailed geological and petrophysical model (Book Cliffs, Utah). *AAPG Bulletin*, doi: 10.1306/05251311162.

Ulrych, T.J., Sacchi, M.D. & Woodbury, A. 2001. A Bayes tour of inversion: A tutorial. *Geophysics*, **66**(1), 55-69.

Van Wagoner, J.C. 1995. Sequence stratigraphy and marine to non-marine facies architecture of foreland basin strata, Book Cliffs, Utah, USA. *In*: Van Wagoner, J.C. & Bertram, G.T. (eds) *Sequence Stratigraphy of Foreland Basin Deposits-Outcrop and Subsurface Examples from the Cretaceous of North America*. American Association of Petroleum Geologist Bulletin, Memoirs, **64**, 137-224.

Van Wagoner, J. C. 1998. Sequence stratigraphy and marine to nonmarine facies architecture of foreland basin strata, Book Cliffs, Utah, USA; reply. *AAPG Bulletin*, **82**(8), 1607-1618.

Wood, R. & Curtis, A. 2004. Geological prior information and its applications to geoscientific problems. *Geological Society, London, Special Publications*, **239**, 1-14.

2

A high-resolution geological and petrophysical model and its inversion results based on reservoir-oriented elastic wave-equation

A previous geological and petrophysical model of the fluvio-deltaic Book Cliffs outcrops contained eight lithotypes, within each of which a number of lithologies were grouped. While this model was an adequate representation of the overall depositional architecture, for reservoir-geological purposes the potential reservoir and non-reservoir lithologies needed to be separated. Here, a new and more detailed geological model is presented in which more differentiation is put on the potential reservoir lithologies. This new model contains twelve lithologies with layers down to one meter in thickness. Assuming a burial depth of three kilometers and an average clay content, representative rock physical properties are assigned to lithologies based on published data. After the model thickness has been stretched by a factor of four in order to represent a more realistic reservoir, a full-waveform forward seismic response is modelled. These data are used as inputs to an iterative elastic wave-equation based inversion scheme with the goal to retrieve the rock properties and layer geometries. The results of this conceptual study show that sandstone units in the shoreface and distributary channels, which are potential reservoirs, are successfully identified. The recovery of medium parameters has a high resolution because the nonlinear relationship between rock properties and the seismic data has been exploited.

This chapter has been published in Petroleum Geoscience.
Doi: https://doi.org/10.1144/petgeo2015-076

2.1. Introduction

In the hydrocarbon industry, it is important to maximize the seismic resolution during acquisition in order to be able to utilize these data in the following processing steps such as inversion. A high-resolution geological and petrophysical synthetic model can provide a good basis to analyze and test the limits of a reservoir-oriented wave-equation based inversion scheme.

In a previous study carried out by Tetyukhina *et al*. (2014) a relatively detailed geological model has been built based on the Book Cliffs outcrops. Eight depositional environments, termed lithotypes, were distinguished and a non-linear acoustic full-waveform inversion has been applied to the synthetic data in order to retrieve the medium parameters such as bulk rock density and compressibility.

However, the geological model presented in Tetyukhina *et al*. (2014) is considered relatively coarse from a reservoir-geological perspective. A more detailed model is needed in order to do justice to the inherent variations within the lithotypes. For example, the offshore transition zone is very heterolithic and the ratio between sandstones and mudstones is moderate (0.35-0.75; Van Wagoner 1995). This heterolithic nature is caused by the cyclic interbedding of muddy siltstones and fine-grained sandstones. From a reservoir-geological point of view, this unit contains therefore reservoir and non-reservoir lithologies in the same lithotype and lumping them into one unit is not appropriate for static reservoir modelling in which more details are needed.

For this reason, the previous model needs to be refined in order to make it more realistic with regard to its internal architecture. By analyzing the forward seismic data based on a more detailed geological and petrophysical model, interpreters can understand which geological information may be extracted, and which ones may not, with the knowledge of layer thickness and wavelength of the seismic data. Such an improved model can also serve as a benchmark for an elastic wave-equation based inversion scheme, by allowing to analyze the resolution and accuracy of the methodology. In this chapter, a high-resolution geological model is presented that serves this purpose.

Introduced by Tarantola (1984), seismic full-waveform inversion (FWI) is becoming feasible because of progress in seismic data acquisition and increased computing power. With the introduction of multicomponent digital acquisition sensors, P- and S-wave data are recorded that help to identify the subsurface anisotropies, expressed by the time shifts caused by shear-wave splitting (Stotter & Angerer 2011). Examples of successful FWI include Vigh *et al.* (2014) who used a 4C ocean-bottom cable data set with large offsets from the Gulf of Mexico to delineate the geology in terms of the elastic velocity fields. Kamath & Tsvankin (2013) performed an elastic FWI in VTI (transversely isotropic with a vertical symmetry axis) media to estimate the interval vertical P- and S-wave velocities as well as the Thomsen parameters. Butzer *et al*. (2013) presented a computationally efficient implementation of 3D elastic FWI and showed that this technique is capable of inverting the 3D scattered waves and reconstructing 3D geological structures.

Differing from the elastic FWI mentioned above, the 1.5-D non-linear inversion scheme used here is based on the integral representation of the full elastic wave equations and the number of iterations determines the order of multiples that are being used. The method takes transmission effects and multiple internal scattering into account. With PP and PS data as inputs, the elastic parameters compressibility, bulk rock density and shear compliance are inverted simultaneously instead of the first two in the acoustic inversion (Tetyukhina *et al*. 2014).

Our approach addresses topics in geology, petrophysics and geophysics. Here the new geological model is presented first, after which it is populated with petrophysical properties. Seismic forward modelling and the new elastic inversion scheme are subsequently performed in order to retrieve the geological units and their elastic properties. Analysis of the results then offers insights into the viability and accuracy of the method.

2.2. Geological Modelling

The Book Cliffs in Utah (USA) are well-studied outcrops of which several detailed high-resolution sequence-stratigraphic interpretations have been published (O'Byrne & Flint 1993; Pattison 1995; Taylor & Lovell 1995; Hodgetts & Howell 2000). They provide an excellent and, from a reservoir-geological point of view, representative dataset for testing seismic forward modelling and inversion schemes.

In the study of Tetyukhina *et al*. (2014) eight different lithotypes have been defined based on their depositional environments (Figure 1.2). Some of these lithotypes, however, have large internal variations of their reservoir properties, which makes the model less suitable if a detailed rock properties are required, specifically for those that may form potential oil or gas reservoirs. For a more detailed model, sequence stratigraphic principles are used here, because these tie changes of stratal stacking patterns to the varying accommodation space and sediment supply, providing therefore an understanding of the geometric relationship and architecture of the sedimentary strata (Catuneanu *et al*. 2011). Three different types of sequence stratigraphic units, i.e. sequences, systems tracts and parasequences, form the main framework of sequence stratigraphy (Catuneanu *et al*. 2011). They are principally defined based on their scales, stacking patterns and bounding surfaces. In the Book Cliffs, previous workers have identified five and three parasequences in the Kenilworth and Aberdeen Members, respectively, in which different systems tracts have been assigned based on their positions within the sequence and other interpretations (Balsley 1980; Taylor & Lovell 1995; Howell & Flint 2003). Ainsworth & Pattison (1994) and Pattison (1995; 2005) alternatively distinguished nine cycles or parasequences in the Kenilworth Member (Figure 2.1).

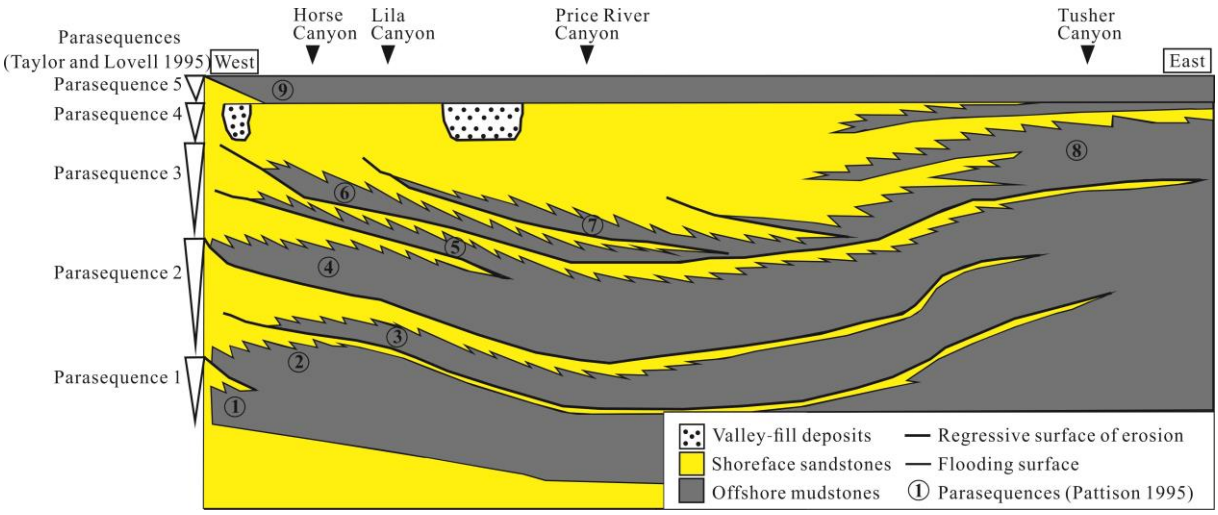


Figure 2.1: A schematic stratigraphic cross-section of the Kenilworth Member. The five-fold subdivision of Taylor and Lovell (1995) and nine-fold subdivision of Pattison (1995) are indicated (modified from Hampson (2000)).

In this chapter, the nine-fold parasequence scheme is adopted (Pattison 2005). The stratal stacking patterns of the parasequences 1 to 5 in the nine-fold scheme exhibit aggradational to progradational features of the highstand systems tract (HST), while parasequences 6 to 8 are progradational sets of the lowstand systems tract (LST) (Ainsworth & Pattison 1994; Pattison 1995; 2005). The uppermost parasequence 9 is part of the transgressive systems tract (TST), which exhibits a retrogradational stacking pattern (Figure 2.2). The three parasequences in the underlying Aberdeen Member display an aggradational stacking pattern of the highstand systems tract (Hodgetts & Howell 2000). The different parasequences are separated by coal seams in the coastal plain and by claystones in the shoreface environment (Kamola & Van Wagoner 1995) (Figure 2.2).

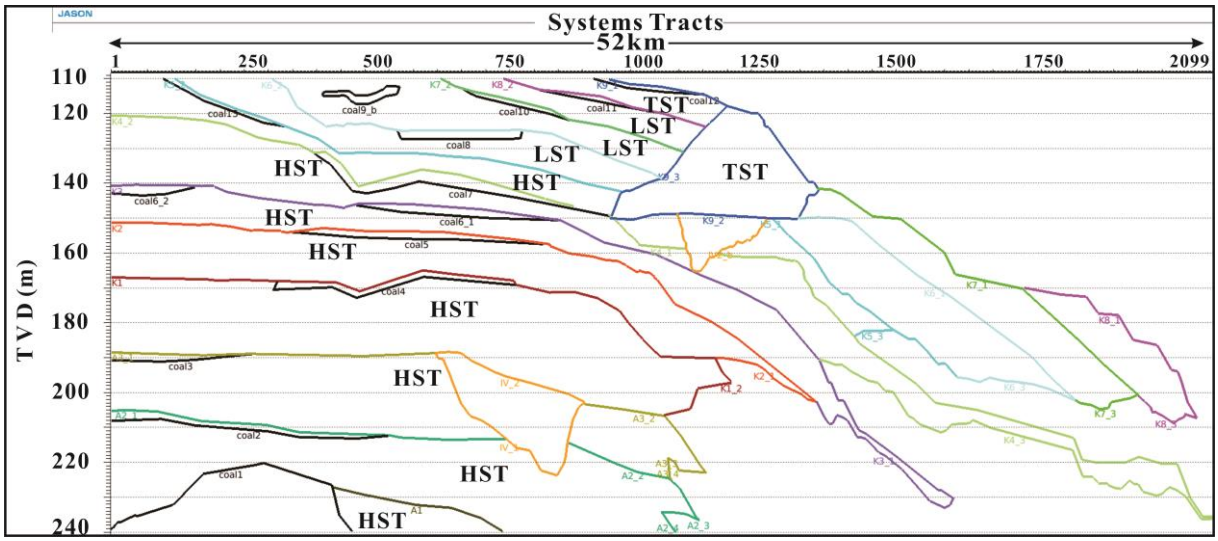


Figure 2.2: The different parasequence units and systems tracts in the Kenilworth and Aberdeen Members (HST−Highstand Systems Tract; LST−Lowstand Systems Tract; TST−Transgressive Systems Tract).

First, the previous model (Tetyukhina *et al*. 2014) has been separated into two parts, the marine and non-marine part, based on the position of the strata and their clay contents, which will influence the rock properties in the petrophysical model. Within each parasequence of the different systems tracts in the shoreface or marine part, different grain size trends occur. For example, in the HST and LST the trends are mainly coarsening-upward, while they are fining-upward in the TST (Van Wagoner *et al*. 1988). This interpretation enables us to divide the lithology of one single parasequence into smaller units or lithologies. In the parasequences of the HST and LST the sequences include from bottom to top siltstones (SS), very fine-grained sandstones (VFS), fine-grained sandstones (FS), and medium-grained sandstones (MS), while in the fining-upward parasequences of the TST the inverse occurs. Because of erosion or non-deposition not all these lithologies will be present in each parasequence. For the incised valley in the middle of the model (Figure 2.2), the lower part consists of (non-marine) fluvial sandstones, whereas the upper part is filled with interbedded marine sandstones and claystones which were probably deposited during the early base level rise (Pattison 1995).

These vertical trends, however, only apply in the marine part; in the non-marine coastal plain the trends are more complex (Coleman & Prior 1982). Here the parasequences were divided into crevasses (mainly composed of fine-grained sandstones or siltstones), lagoonal deposits (siltstones and fine-grained sandstones), and interdistributary deposits (very fine-grained sandstones and fine-grained sandstones). The distributary channels of the coastal plain often demonstrate a fining-upward trend with coarse-grained sandstones (CS) occurring at the bottom and very fine-grained sandstones at the top. This differentiates them from the incised valleys in the shoreface successions. The new refined 2-D model is shown in Figure 2.3 (Feng *et al*. 2015a).

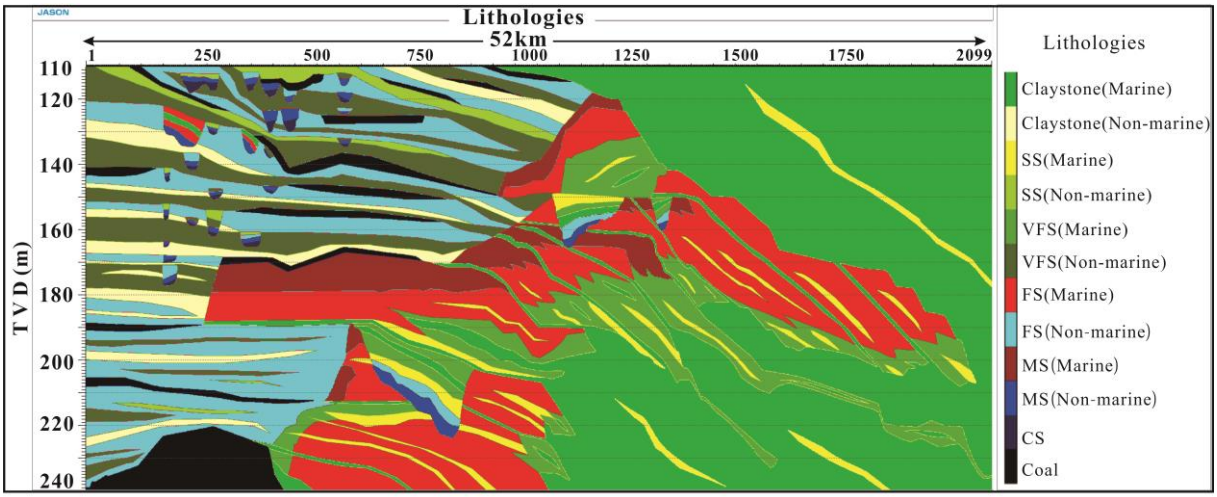


Figure 2.3: The new 2-D lithology model. Different colors represent different lithologies (CS−Coarse-grained sandstones; MS−Medium-grained sandstones; FS−Fine-grained sandstones; VFS−Very fine-grained sandstones; SS−Siltstones).

Partitioning the model into a marine and a non-marine part is done because in the marine lithologies the clay content is generally higher than that in the non-marine counterparts. With increasing grain size, the clay content decreases in the marine as well as the non-marine part.

Thus, the clay content in the fine-grained sandstones of the marine part is higher than that in the same lithology of the non-marine part, and the clay content in the fine-grained sandstones is lower than that in the very fine-grained sandstones, whether marine or non-marine. This strategy of varying the clay content affects the rock physical properties which are discussed in the following sections.

2.3. Petrophysical Modelling

The new geological model (Figure 2.3) contains numerous details in the internal architecture and sub-layers for which rock physical properties such as bulk rock density and velocity need to be assigned in order to be able to perform the seismic forward modelling. Because this model is different in character from the previous model, one cannot apply the same properties (Hodgetts & Howell 2000; Tetyukhina *et al*. 2014). The different lithologies contain at least some amounts of clay, and thus a sand-shale mixture model is used which has been developed at Stanford University over the last few decades (Han *et al*. 1986; Marion *et al*. 1992; Mavko *et al*. 1998; Nur *et al*. 1998). The clay content affects the rock properties in two ways: if its volume fraction is lower than the porosity of pure sand, the compressional velocity increases with the volume fraction of clay while the porosity decreases; if, on the contrary, the volume fraction of clay is higher than the porosity of pure sand, i.e. the rock is shaley matrix-supported, the compressional velocity decreases with the increasing volume fraction of clay while the porosity increases. This phenomenon can be explained as follows: If the volume fraction of clay is less than the porosity of pure sand ($c \leq \varphi_s$), the clay particles fill the pore space as a solid, which would otherwise be filled with a fluid, thereby stiffening the material and increasing the velocity of the mixture as the volume fraction of clay increases. If the volume fraction of clay is higher than the porosity of pure sand ($c > \varphi_s$), the sand grains will become "suspended" in the clay matrix, and, since the velocity of clay is slower than that of pure sand, an increase in the volume fraction of clay causes the velocity of the mixture to decrease (Marion *et al*. 1992). But, since clay mixtures are porous, the porosity increases as the volume fraction of clay increases.

According to Eberhart-Phillips *et al*. (1989) and Marion *et al*. (1992), the resulting density and velocities for a water-filled rock can be calculated with the following empirical equations:

For the bulk rock density, if $c \leq \varphi_s$, $\varphi = \varphi_s - c(1 - \varphi_{sh})$

$$\rho = (1 - \varphi_s)\rho_s + c(1 - \varphi_{sh})\rho_c + \varphi\rho_w \tag{2.1}$$

if $c > \varphi_s$, $\varphi = c\varphi_{sh}$

$$\rho = (1 - c)\rho_s + c(1 - \varphi_{sh})\rho_c + \varphi\rho_w \tag{2.2}$$

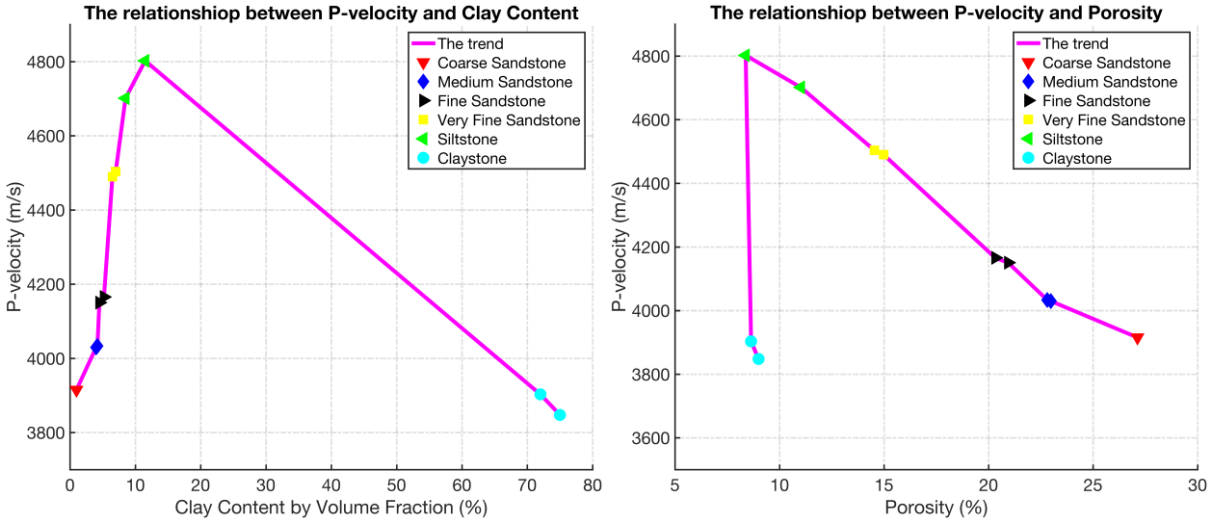
For velocities,

$$V_P = 5.77 - 6.94\varphi - 1.73\sqrt{c} + 0.446(P_e - e^{-16.7P_e}) \quad (2.3)$$

$$V_S = 3.70 - 4.94\varphi - 1.57\sqrt{c} + 0.361(P_e - e^{-16.7P_e}) \quad (2.4)$$

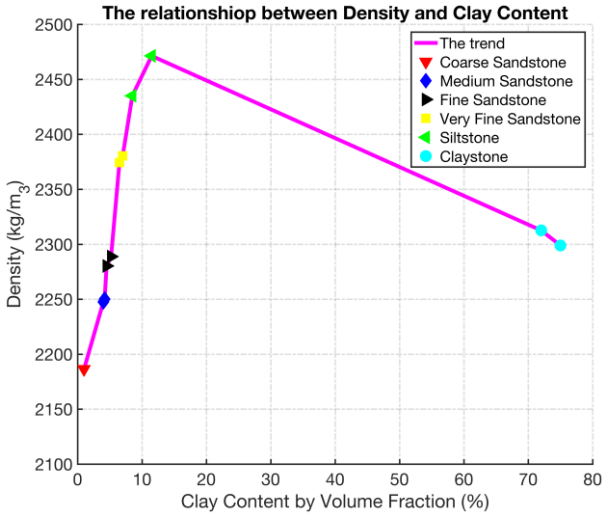
Where c is the volume fraction of clay; φ_s, φ_{sh}, φ are the porosities of pure sand, pure shale and the sand-shale mixture, respectively; ρ_s, ρ_c, ρ_w, ρ are the densities of sand (quartz), clay, water and the mixture, and P_e is the effective pressure.

After the determination of these parameters in equations (2.1) to (2.4) following Manger (1963), Hamilton (1980), Marion *et al*. (1992), Vidar *et al*. (2005), Fisher *et al*. (2007) and adjusting for P_e which accounts for the depth of the model, assuming here to have a top at 3 km depth (Eberhart-Phillips *et al*. 1989; Hofmann *et al*. 2005), the resulting relationships between bulk rock density, porosity, clay content and velocity are shown in Figure 2.4.

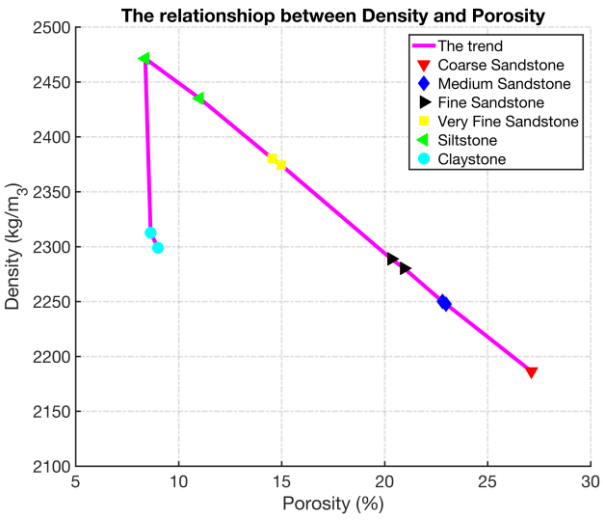


(a)

(b)



(c)



(d)

Figure 2.4: The relationships between P-velocity, clay content by volume fraction, porosity and bulk rock density of different lithologies in the sand-shale mixture model. Note that for each lithology except for the coarse-grained sandstones there are two values i.e. for the marine and non-marine parts respectively.

These graphs show the trend reversals at the critical levels of clay content and porosity. Thus, for example, in Figure 2.4c, the bulk rock density increases with decreasing grain size due to a decrease in porosity, as seen in Figure 2.4d, but decreases towards the claystone points. This means that the porosity exerts more control than the clay content on the bulk rock density because less porosity will lead to more sand even though the volume fraction of clay increases at the same time.

Finally, the entire 2-D geological model is populated with these rock properties, i.e. bulk rock density, P-velocity and S-velocity that have been defined by equations (2.1) to (2.4). The resulting 2-D petrophysical model is shown in Figure 2.5.

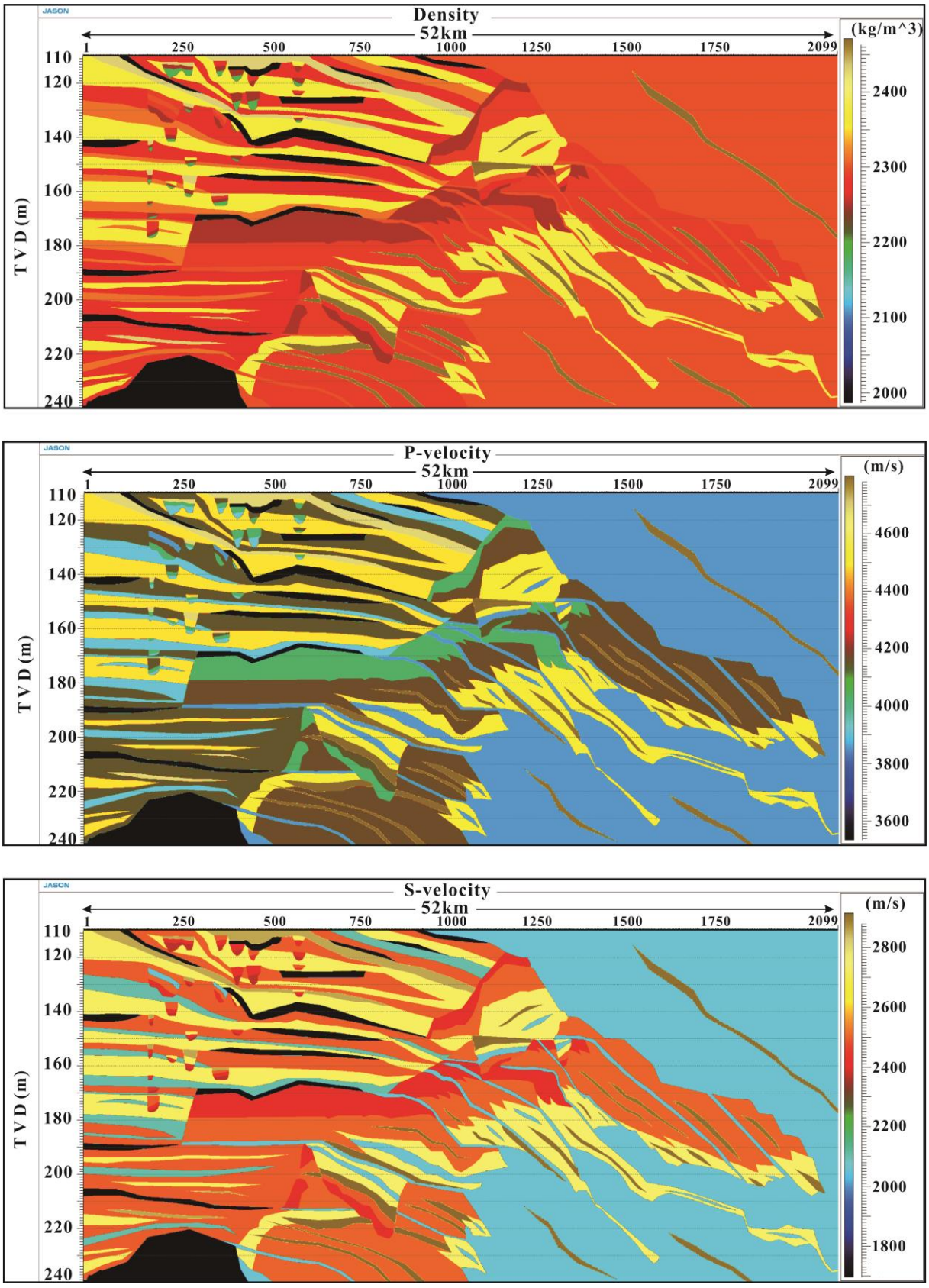


Figure 2.5: The petrophysical properties of the new geological model (Figure 2.3) in terms of bulk rock density, V_P, V_S. The horizontal width is 52 km approximately and has 2099 CMPs in total.

2.4. Seismic Modelling

The synthetic seismic data can be considered as a bridge that links the outcrop to surface seismic data from analogous subsurface formations because they can offer insights into whether certain geological features can potentially be recognized from seismic data or not. In addition, forward modelling may also help to analyze artefacts observed in real seismic data (Stafleu *et al*. 1994; Coleman *et al*. 2000).

Here the same method is used as by Tetyukhina *et al*. (2014), which is the Kennett invariant embedding method (Kennett 1983). This method is able to incorporate all the internal multiples as well as the wave-mode conversions and transmission effects. In order to simulate more realistic actual reservoir situations and increase the detectability of different parasequences, the vertical thickness of the entire model has been multiplied by a factor of four, resulting in the dimensions of 520 m in thickness and approximately 52 km in width. This model is suitable for the Kennett method because of the almost horizontal layers which satisfy the local 1-D assumption (notice that all model figures are highly exaggerated vertically; in reality, the layer dips are mostly well below 1°). The vertical discretization of the model parameters is 0.4 m and the time sampling is 4 ms in the traces. The horizontal sampling interval is 25 m and in total there are 2099 common mid-points (CMPs). The wavelet used is a zero-phase band-pass filter with trapezium corner frequencies of 6-12-60-80 Hz (Figure 2.6) (Tetyukhina *et al*. 2014). The incidence angles range from 0° to 45° at the highest compressional velocity in the entire model with a set of 10 plane waves at each CMP or trace location. White random noise has been added with a signal-to-noise ratio (SNR) of 30 dB.

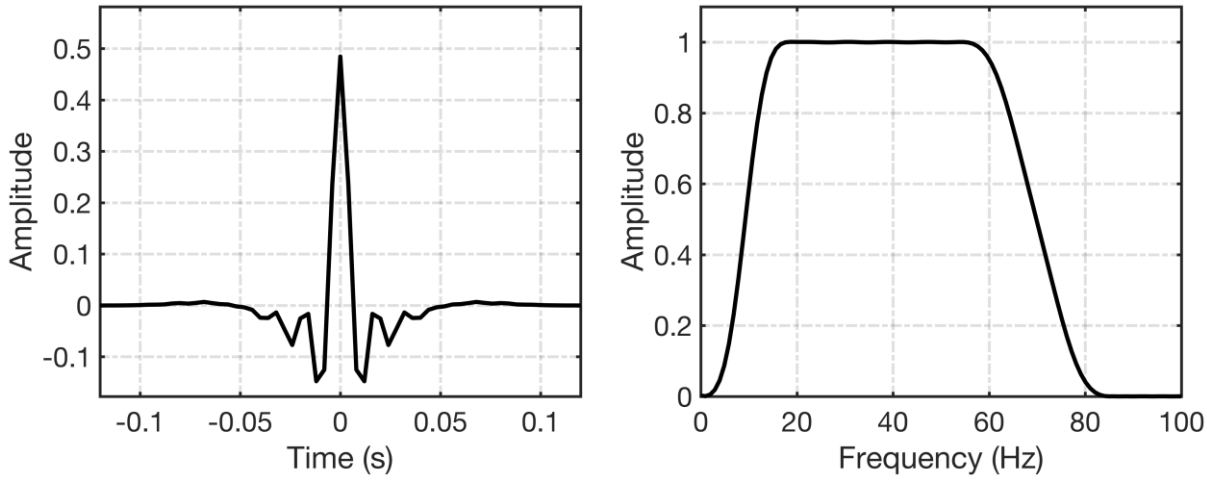


Figure 2.6: The zero-phase wavelet together with its amplitude spectrum used in the forward modelling with trapezium corner frequencies of 6-12-60-80 Hz.

The synthetic seismic data are generated directly at the top of the objective sequence. This acquisition setup is equivalent to migrating surface data, converting offsets to angles, and de-migrating the migrated data over the target interval to the target boundary marking the top of the model. We opted for the modelling the synthetic data at the top of the model in order to

set a benchmark for the resolving power of the inversion, without being side-tracked by data processing issues, substantial though they may be.

The top and bottom of the model have been tapered during the forward modelling which means the differences in properties of the top and bottom boundaries between the actual and the background media have been gradually minimized to zero in order to avoid reflections from the truncations of the properties.

The synthetic seismic data are calculated for every location separately. The outputs of the Kennett forward modelling are separate PP and PS records in the Radon domain (τ/p) which are used as inputs for the elastic wave-equation based inversion later. Figure 2.7 is showing the PP data only for normally incident plane waves. These data can be used to interpret the stacking pattern and the architecture of the parasequences. The coals, for example, are clearly imaged because of their relatively large impedance contrasts with the surrounding lithologies. The clinoform geometries of the different parasequences exhibit fairly strong amplitudes that can be traced from the offshore to the shoreface and even to the coastal plain environment. The distributary channels can also be distinguished well, which is relevant because they represent important potential reservoir units. The incised valleys are not as easily recognized as before (Tetyukhina *et al*. 2014) mainly because there are now more sub-layers and the amplitudes of the various reflectors are interfering with each other constructively or destructively (Feng *et al*. 2015a).

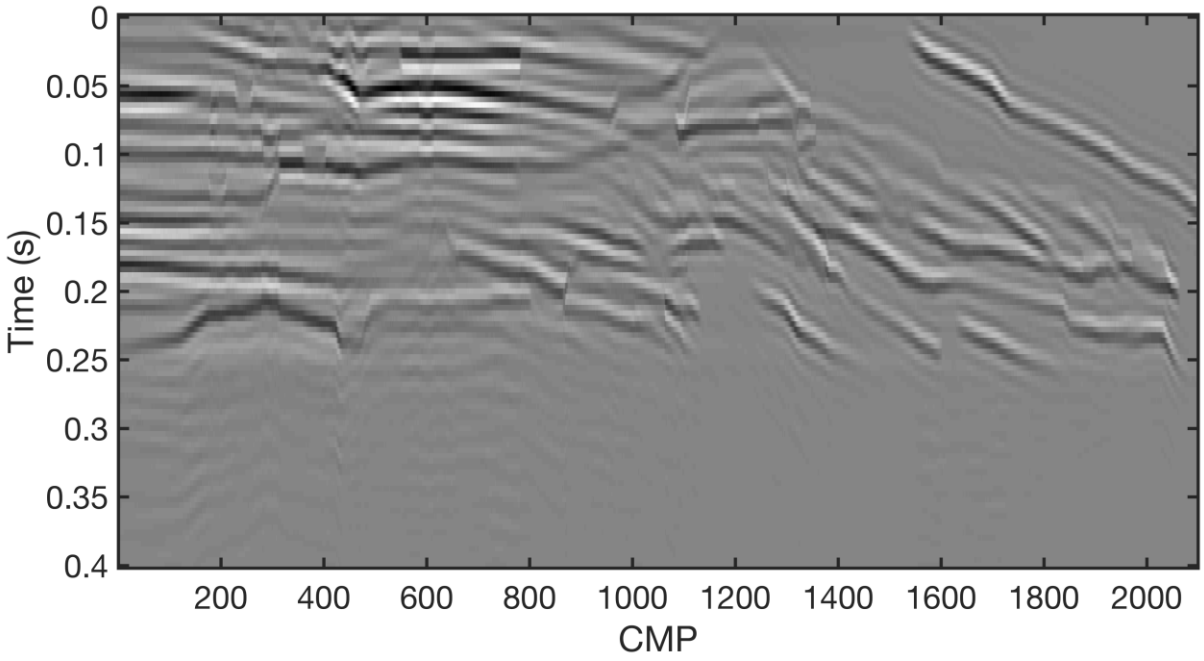


Figure 2.7: The synthetic seismic data (PP) based on the whole 2-D petrophysical model (Figure 2.5). Here data for normal incident plane waves are shown as a function of the CMP.

2.5. Elastic wave-equation based inversion

The inversion algorithm we use was developed by Gisolf & van den Berg (2010a; 2010b), and has been applied by Tetyukhina before but only in an acoustic mode (Tetyukhina *et al*. 2014). Here the inversion scheme has been extended to elastic, which means that wave-mode conversions will also be taken into account in the inversion process. Because the inversion scheme is based on the full elastic wave equations, it takes all internal multiple scattering, with mode conversion, into account, allowing recovery of broadband properties and providing potentially a high resolution (Gisolf *et al*. 2014).

The main goal of the elastic wave-equation based inversion is to retrieve relevant reservoir properties, and for this the overburden effects would have to be compensated for. Since the synthetic data have been generated for the target interval only, there is no need to remove the overburden in this particular case. However, in real reservoirs, extensive pre-processing will be required to achieve this (Gisolf *et al*. 2014). For example, the wavefield recorded at the surface should be back-propagated (re-datumed) to the top of the objective interval through an overburden that describes the kinematics sufficiently well, or the pre-stack migrated data are de-migrated to the target boundary. The offset should be converted to ray-parameters using a layered overburden model since the inversion is carried out in the Radon domain (τ/p). In either case, the target-oriented inversion of field data has to rely on the migration/re-datuming process to deliver suitable input data for the inversion.

The inversion is based on the integral representation of the full elastic wave equations, the so-called scattering integral (Fokkema & van den Berg 1993). The total wavefield is calculated in the full inhomogeneous medium, rather than in a constant background medium. The P and S velocities are broken down into elastic moduli, notably the compressibility $\kappa = 1/K$ (with K being the bulk modulus), and the shear compliance $M = 1/\mu$ (with μ being the shear modulus):

$$V_P = \sqrt{\frac{1}{\rho}\left(\frac{1}{\kappa} + \frac{4}{3M}\right)} \tag{2.5}$$

$$V_S = \sqrt{\frac{1}{M\rho}} \tag{2.6}$$

Instead of inverting for κ, M and ρ, the contrast functions based on the backgrounds (κ_0, M_0, ρ_0) are solved at every location:

$$\chi_\kappa(z) = \frac{\kappa(z) - \kappa_0(z)}{\kappa_0(z)} \tag{2.7}$$

$$\chi_M(z) = \frac{M(z) - M_0(z)}{M_0(z)} \tag{2.8}$$

$$\chi_\rho(z) = \frac{\rho(z) - \rho_0(z)}{\rho_0(z)} \tag{2.9}$$

The backgrounds are smooth media in which the incident field and Green's functions are calculated (Haffinger 2013) and which represent the prior knowledge before the inversion. In our synthetic exercise, we can obtain realistic backgrounds by strongly smoothing the true properties with the help of a high-cut filter (0-4 Hz). The results are shown in Figure 2.8. Although having backgrounds that describe the kinematics reasonably well is important, the non-linear inversion is able to reduce the dependency on the background models (Haffinger 2013). In real cases, the background models can be obtained from the migration velocities, well logs or even empirical rock physical considerations.

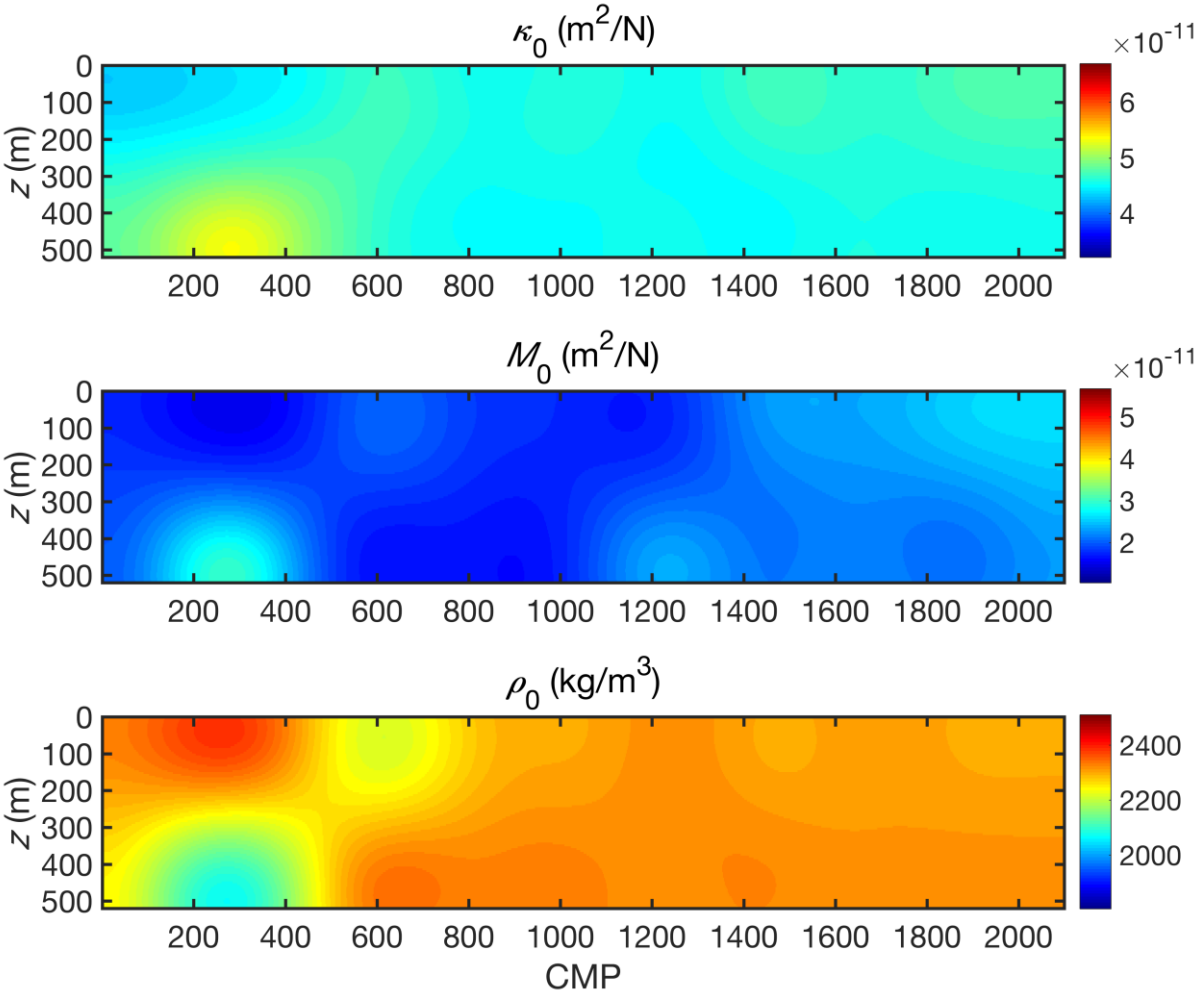


Figure 2.8: The smooth background models for compressibility, shear compliance and bulk rock density (from top to bottom).

The inversion approach is based on a local 1-D earth model with a 2-D pre-stack plane wavefield of (τ/p) PP and PS gathers (Gisolf *et al*. 2014). Since the model (Figure 2.5) is 52 km approximately in width and 520 m in thickness, meaning that the layers are almost horizontal, this approach seems suitable, both for modelling the synthetic data and for the inversion.

Because the seismic data and the inversion are based on the full wave equations, the data are non-linear in the elastic properties to be estimated. We solve this using an iterative scheme where we alternately determine the total wavefield in the object from our best knowledge of the properties, and determine the updated properties from the data and our current best knowledge of the total wavefield in the object. The first estimate of the total field is the incident field that propagates in the background models (Figure 2.8). This is a well-known linearization of the problem called the Born approximation. With this total field estimate we invert the data, leading to a first estimate of the properties in the object. Together with the wave equations these are used to calculate an update of the total wavefield that now contains first-order scattering. With this new version of the total wavefield a new inversion of the data is undertaken, leading to a better estimate of the properties, etc. The order of the multiples accounted for in the data will be determined by the number of iterations. The iteration will stop when neither the properties nor the total wavefield change significantly (Gisolf & Verschuur 2010; Tetyukhina *et al*. 2014).

Applying this scheme to all the input gathers of the pre-stack PP and PS data in the Radon (τ/p) domain, inversion results are obtained as shown in Figure 2.9 in terms of the dimensionless contrasts in compressibility, shear compliance and bulk rock density (χ_κ, χ_M, χ_ρ). The left plots are unfiltered, i.e. in the broad spatial bandwidth range generated by the inversion, while the ones on the right have been spatially filtered to the equivalence of the seismic temporal bandwidth by using a spatial band-pass filter.

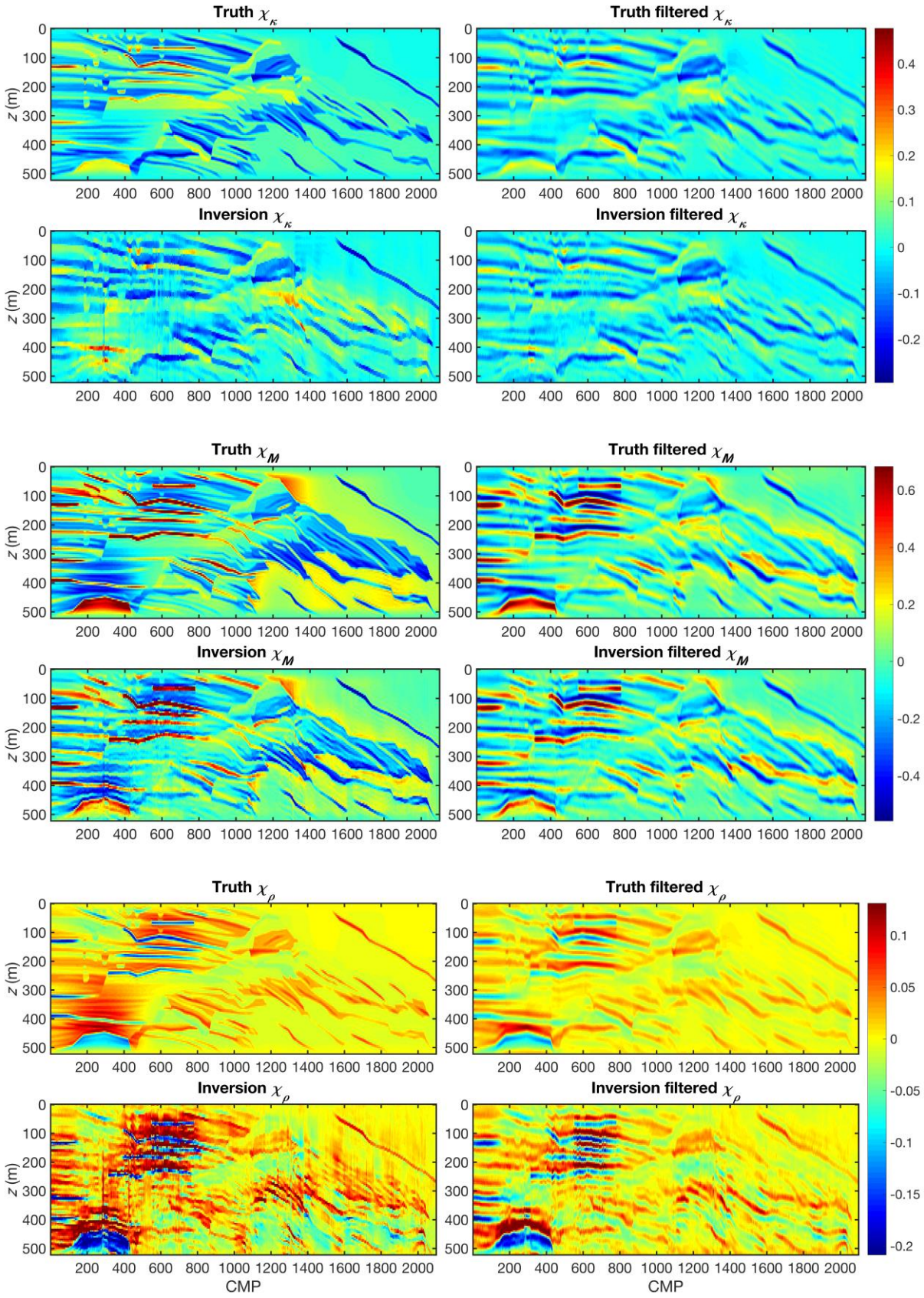


Figure 2.9: The true and inverted results for the contrasts in compressibility, shear compliance and bulk rock density (χ_κ, χ_M, χ_ρ). The left plots are unfiltered while the right ones have been filtered to the spatial equivalence of the seismic wavelet bandwidth.

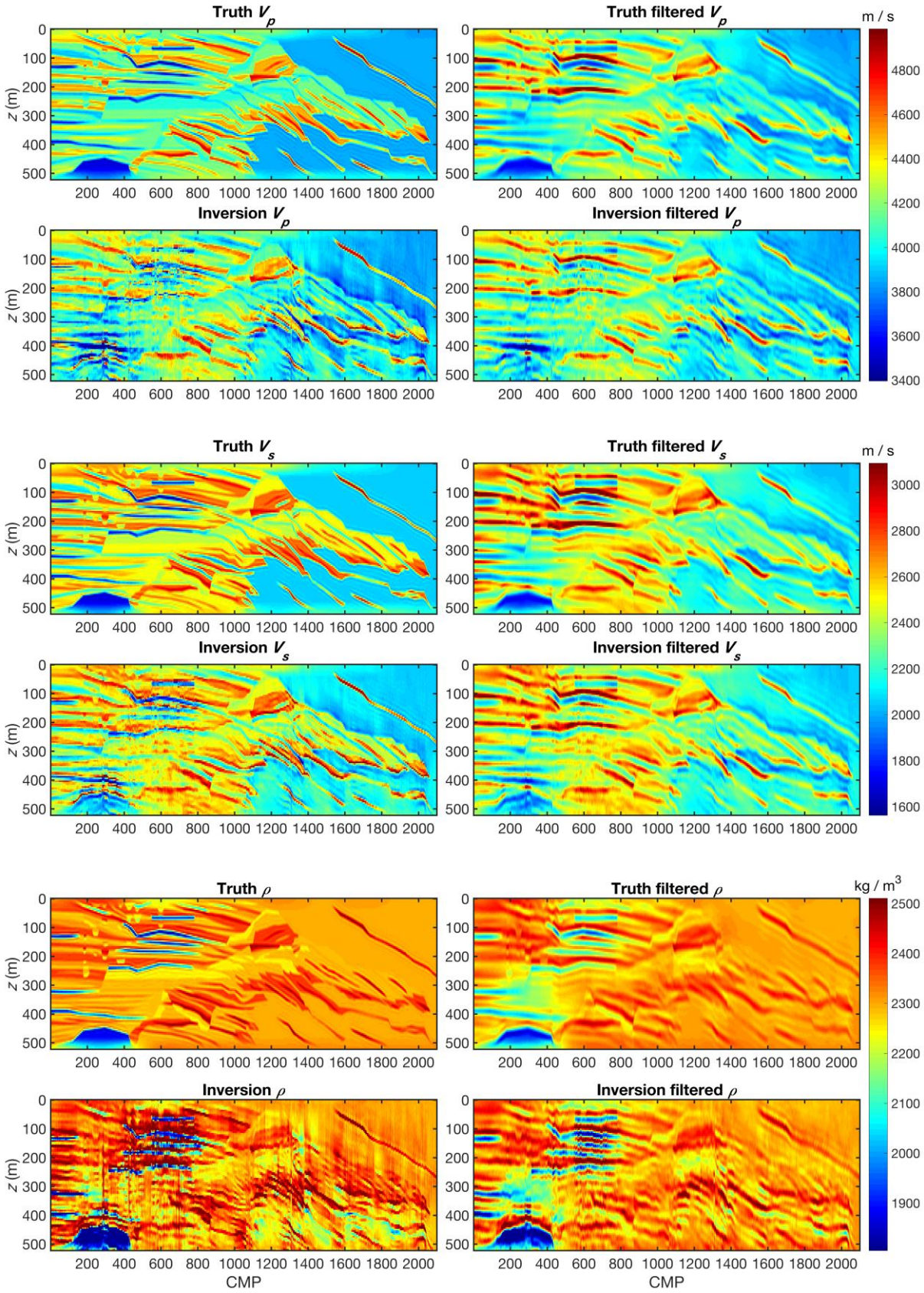


Figure 2.10: The true and inverted results for the compressional and shear velocities, as well as the bulk rock density (V_P, V_S, ρ). The left plots are unfiltered while the right ones have been filtered to the spatial equivalence of the seismic wavelet bandwidth.

In order to compare these results with the petrophysical properties, they are converted to the compressional and shear velocities (V_P, V_S) as well as the bulk rock densities (ρ) (Figure 2.10). These results show that the inversion successfully recovers the properties and geometries for the whole section. The small sandstone units in the shoreface section, which are important potential reservoir targets, are well detected on the inversion results. The incised valleys, on the other hand, are less easy to be identified, compared with the previous study (Tetyukhina *et al*. 2014), an effect caused by the interfering sub-layers within them. It is noticeable that the inversion result for the contrast of shear compliance (χ_M) is better than that for the compressibility (χ_κ), because we are using PS as well as PP data. Both elastic properties are predicted considerably better than the density (χ_ρ), mostly because the latter shows only relatively small variations that are difficult to recover. However, the coals, whose bulk rock density differs substantially from the rest, are adequately resolved in the density inversion result. The same can be seen for some of the sandstone units in the shoreface area.

2.6. Discussion and Conclusion

In this chapter, we present a new geological and petrophysical model of the Book Cliffs outcrop model, based on sequence stratigraphic and rock physical principles. Compared with the previous model by Tetyukhina *et al*. (2014), this new model has more details and includes a distinction between reservoir and non-reservoir lithologies. Therefore, it is considered to be a more realistic reservoir analogue for seismic modelling and inversion studies.

The reservoir-oriented elastic wave-equation based inversion scheme can take care of the internal multiple scattering energy as well as the wave-mode conversions, which makes the recovery of high-resolution, broadband properties possible. It can be seen that the three different medium parameters, compressibility, shear compliance and bulk rock density, have been recovered quite successfully, although with some differences in quality. Since the value range for bulk rock density for the different lithologies is much smaller than those for the two elastic medium parameters and the incidence angle is not very wide in the forward modelling, the inversion result of bulk rock density is of lesser quality than the other two. This is, for example, demonstrated by the distributary channel in the coastal plain (Figure 2.9), where the images of compressibility and shear compliance are clearer than that of bulk rock density. This is also true for the geometries of different lithologies, like the coals in the coastal plain and sandstones in the shoreface. Since the PS data are also used for the inversion, this results in a good retrieval of the shear compliance. In a real case, this kind of data will be, however, rarely available. Nevertheless, according to Haffinger *et al*. (2015), a good recovery of shear compliance can still be expected.

Other inversion procedures convert the seismic data to compressional and shear velocities as well as bulk rock density. However, if the bulk rock density result is inadequate, good images of compressional and shear velocities are difficult to achieve in this method. Therefore, we choose to invert for the compressibility and shear compliance, since they are better reservoir indicators. Especially in 4D or time-lapse inversion, compressibility and shear compliance can

help to identify property changes in the reservoir because of their complementary property behavior (Gisolf & van den Berg 2010b; Feng *et al*. 2015b).

This inversion scheme is target (reservoir) oriented, which means the data have been brought down to the target level by migration, or re-datuming. In this conceptual study, this problem has been avoided by generating the synthetic seismic data directly at the top boundary of the target model. In real data projects, re-datuming can be achieved by local demigration of the migrated input data to a target horizon.

In migration/demigration processes, all internal scattering over the target interval is preserved. Multiples generated in the overburden that leak into the primary target time window should be removed. However, in this chapter the emphasis is on the extraction of quantitative geological information from seismic data and not so much on the details of preparing the seismic data for the inversion step. For a detailed discussion of all the challenges in the pre-processing steps for providing suitable data sets to inversion, we refer to the work by Gisolf *et al*. (2014) and Haffinger (2013; 2015).

We have used a 1.5-D assumption in the inversion approach, which means the layers are locally horizontal to sub-horizontal (dips are well below 0.5 degrees in the coastal plain and below 1 degree in the shoreface units). In this synthetic case, this approach seems suitable because of the very large horizontal width and small vertical thickness. Lateral heterogeneity of the properties is nonetheless well captured with this locally 1.5-D approach. In real cases, it is considered unlikely that within a 500 m thick sedimentary sequence there would be strong changes in local dip over short vertical distances, unless there are structural deformations involved. Laterally, locations 25 m apart are here treated completely independently. This approach may therefore be mostly suitable on passive continental margins with relatively young sediments.

As an extension of this research, the relationship between the rock properties (bulk rock density, compressibility, shear compliance) and the reservoir parameters (porosity, clay content, hydrocarbon saturation) will be investigated with the goal to make the method more directly applicable to the reservoir characterization (following chapters). Thus this model may become a virtual asset that can be used to assess the advantages and limitations of various inversion methods.

References

Ainsworth, R.B. & Pattison, S.A.J. 1994. Where have all the lowstand gone? Evidence for attached lowstand systems tracts in the Western Interior of North America. *Geology*, **22**, 415-418.

Balsley, J.K. 1980. Cretaceous wave-dominated delta systems, Book Cliffs, east-central Utah. *American Association of Petroleum Geologists*, Field Guide.

Butzer, S., Kurzmann, A. & Bohlen, T. 2013. 3D elastic full-waveform inversion of small-scale heterogeneities in transmission geometry. *Geophysical Prospecting*, **61**, 1238-1251.

Catuneanu, O., Galloway, W.E., Kendall, C.G., Miall, A.D., Posamentier, H.W., Strasser, A. & Tucker, M.E. 2011. Sequence stratigraphy: methodology and nomenclature. *Newsletters on Stratigraphy*, **44**, 173-245.

Coleman, J.M. & Prior, D.B. 1982. Deltaic environments of deposition. *In*: Scholle, P.A. & Spearing, D. (ed.) *Sandstone Depositional Environments*. American Association of Petroleum Geologist Bulletin, Memoirs, **31**, 139-178.

Coleman, J.L.Jr., Sheppard, III.F.C. & Jones, T.K. 2000. Seismic resolution of submarine channel architecture as indicated by outcrop analogs. *In*: Bouma, A.H. & Stone, C.G. (eds) *Fine-grained turbidite systems.* American Association of Petroleum Geologist Bulletin, Memoirs, **72**, 119–126.

Eberhart-Phillips, D., Han, D.H. & Zoback, M.D. 1989. Empirical relationships among seismic velocity, effective pressure, porosity, and clay content in sandstone. *Geophysics*, **54**, 82-89.

Feng, R., Sharma, S., Luthi, S.M. & Gisolf, D. 2015a. *An outcrop-based detailed geological model to test automated interpretation of seismic inversion results*. Th N102 15, presented at the 77th EAGE Annual Meeting, Madrid, 1-4 June.

Feng, R.H., Luthi, S.M., Gisolf, D. & Sharma, S. 2015b. *Non-linear full-waveform inversion (FWI-res) of time-lapse seismic data on a higher-resolution geological and petrophysical model, Book Cliffs (Utah, USA)*. Presented at the 85th SEG Annual Meeting, New Orleans, 18-23 October.

Fisher, Q.J., Harris, S.D., Casey, M. & Knipe, R.J. 2007. Influence of grain size and geothermal gradient on the ductile-to-brittle transition in arenaceous sedimentary rocks: implications for fault structure and fluid flow. *In*: Lewis, H. & Couples, G.D. (eds) *The Relationship between Damage and Localization*. Geological Society, London, Special Publications, **289**, 105-121.

Fokkema, J.T. & van den Berg, P.M. 1993. *Seismic application of acoustic reciprocity*. Elsevier.

Gisolf, A. & Verschuur, D.J. 2010. *The principles of quantitative acoustical imaging*. EAGE Publications bv.

Gisolf, A. & van den Berg, P.M. 2010a. *Target oriented non-linear inversion of seismic data*. P308, presented at the 72nd EAGE Annual Meeting, Barcelona, 14-17 June.

Gisolf, A. & van den Berg, P.M. 2010b. *Target-oriented non-linear inversion of time-lapse seismic data*. Presented at the 80th SEG Annual Meeting, Denver, 17-22 October.

Gisolf, A., Huis in't Veld, R., Haffinger, P., Hanitzsch, C., Doulgeris, P. & Veeken, P.C.H. 2014. *Non-linear full wavefield inversion applied to carboniferous reservoirs in the Wingate Gas Field (SNS, Offshore UK)*. Th E106 08, presented at the 76th EAGE Annual Meeting, Amsterdam, 16-19 June.

Haffinger, P. 2013. Seismic broadband full waveform inversion by shot/receiver refocusing. PhD thesis, Delft University of Technology.

Haffinger, P., von Wussow, P., Doulgeris, P., Henke, C. & Gisolf, A. 2015. *Reservoir delineation by applying a nonlinear AVO technique – A case study in the Nile Delta*. Tu P9 07, presented at the 77th EAGE Annual Meeting, Madrid, 1-4 June.

Hamilton, E.L. 1980. Geoacoustic modelling of the sea floor. *Journal of the Acoustical Society of America*, **68**, 1313-1340.

Hampson, G.J. 2000. Discontinuity surface, clinoforms, and facies architecture in wave-dominated, shoreface-shelf parasequence. *Journal of Sedimentary Research*, **70**, 325-340.

Han, D.H., Nur, A. & Morgan, D. 1986. Effect of porosity and clay content on wave velocity in sandstones. *Geophysics*, **51**, 2093-2107.

Hofmann, R., Xu, X., Batzle, M., Prasad, M., Furre, A.K. & Pillitteri, A. 2005. Effective pressure or what is the effect of pressure?. *The Leading Edge*, **24**, 1256-1260.

Howell, J.A. & Flint, S.S. 2003. Sequences and systems tracts in the Book Cliffs. *In*: Coe, A. (ed.) *The Sedimentary Record of Sea-Level Change*. Cambridge University Press, Cambridge.

Hodgetts, D. & Howell, J.A. 2000. Synthetic seismic modelling of a large-scale geological cross-section from the Book Cliffs, Utah, USA. *Petroleum Geoscience*, **6**, 221-229.

Kamath, N. & Tsvankin, I. 2013. Full-waveform inversion of multicomponent data for horizontally layered VTI media. *Geophysics*, **78**, WC113-WC121.

Kamola, D.L. & Van Wagoner, J.C. 1995. Stratigraphy and facies architecture of parasequences with examples from the Spring Canyon Member, Blackhawk Formation, Utah. *In*: Van Wagoner, J.C. & Bertram, G.T. (eds) *Sequence Stratigraphy of Foreland*

Basin Deposits-Outcrop and Subsurface Examples from the Cretaceous of North America. American Association of Petroleum Geologist Bulletin, Memoirs, **64**,27-54.

Kennett, B.L.N. 1983. *Seismic wave propagation in stratified media*. Cambridge University Press, Cambridge.

Manger, G.E. 1963. *Porosity and bulk density of sedimentary rocks*. Geological Survey Bulletin, 1-60.

Marion, D., Nur, A., Yin, H. & Han, D. 1992. Compressional velocity and porosity in sand-clay mixtures. *Geophysics*, **57**, 554-563.

Mavko, G., Mukerji, T. & Dvorkin, J. 1998. *The Rock Physics Handbook*. Cambridge University Press, Cambridge.

Nur, A., Mavko, G., Dvorkin, J. & Galmundi, D. 1998. Critical porosity: a key to relating physical properties to porosity in rocks. *The Leading Edge*, **17**, 357–362.

O'Byrne, C.J. & Flint, S.S. 1993. High-resolution sequence stratigraphy of Cretaceous shallow marine sandstones, Book Cliffs outcrop, Utah, USA-application to reservoir modelling. *First Break*, **11**, 445-459.

Pattison, S.A.J. 1995. Sequence stratigraphic significance of sharp-based lowstand shoreface deposits, Kenilworth Member, Book Cliffs, Utah. *American Association of Petroleum Geologist Bulletin*, **79**, 444-462.

Pattison, S.A.J. 2005. Storm-influenced prodelta turbidite complex in the lower Kenilworth Member at Hatch Mesa, Book Cliffs, Utah, USA: Implication for shallow marine facies models. *Journal of Sedimentary Research*, **75**, 420-439.

Stafleu, J., Everts, A.J.W. & Kenter, J.A.M. 1994. Seismic models of a prograding carbonate platform: Vercors, southeast France. *Marine and Petroleum Geology*, **11**, 514–527.

Stotter, C. & Angerer, E. 2011. Evaluation of 3C microelectromechanical system data on a 2D line: Direct comparison with conventional vertical-component geophone arrays and PS-wave analysis. *Geophysics*, **76**, B79-B87.

Taylor, D.R. & Lovell, R.W.W. 1995. High-frequency sequence stratigraphy and paleogeography of the Kenilworth Member, Blackhawk Formation, Book Cliffs, Utah, USA. *In*: Van Wagoner, J.C. & Bertram, G.T. (eds) *Sequence Stratigraphy of Foreland Basin Deposits-Outcrop and Subsurface Examples from the Cretaceous of North America*. American Association of Petroleum Geologist Bulletin, Memoirs, **64**, 257-275.

Tarantola, A. 1984. Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, **49**, 1259-1266.

Tetyukhina, D., Luthi, S.M. & Gisolf, D. 2014. Acoustic non-linear full-waveform inversion on an outcrop-based detailed geological and petrophysical model (Book Cliffs, Utah). *American Association of Petroleum Geologist Bulletin*, **98**, 119-134.

Van Wagoner, J.C., Posamentier, H.W., Mitchum, R.M. et al. 1988. An overview of the fundamentals of sequence stratigraphy and key definitions. *In*: Wilgus, C.K., Posamentier, H.W., Ross, C.K. & Kendall, C.G.St.C. (eds) *Sea Level Changes: An Integrated Approach*. Society for Sedimentary Geology, Special Publication, **42**, 39-45.

Van Wagoner, J.C. 1995. Sequence stratigraphy and marine to non-marine facies architecture of foreland basin strata, Book Cliffs, Utah, USA. *In*: Van Wagoner, J.C. & Bertram, G.T. (eds) *Sequence Stratigraphy of Foreland Basin Deposits-Outcrop and Subsurface Examples from the Cretaceous of North America*. American Association of Petroleum Geologist Bulletin, Memoirs, **64**, 137-224.

Vidar, S., Knut, B. & Nazmul, H.M. 2005. Velocity-depth trends in Mesozoic and Cenozoic sediments from the Norwegian Shelf. *American Association of Petroleum Geologist Bulletin*, **89**, 359-381.

Vigh, D., Jiao, K., Watts, D. & Sun, D. 2014. Elastic full-waveform inversion application using multicomponent measurements of seismic data collection. *Geophysics*, **79**, R63-R77.

3

Simulating Reservoir Lithology by an Actively Conditioned Markov Chain Model

Coupled Markov chain models can be used to simulate reservoir lithologies between wells, by conditioning them on the observed data in the cored wells. However, with this method, only the state at the same depth as the current cell is going to be used for conditioning, which may be a problem if the geological layers are dipping. This will cause the simulated lithological layers to be broken or to become discontinuous across the reservoir. In order to address this problem, an actively conditioning process is proposed here, in which a tolerance angle is predefined. The states contained in the region constrained by the tolerance angle will be employed for conditioning in the horizontal chain first, after which a coupling concept with the vertical chain is implemented. In order to use the same horizontal transition matrix for different future states, the tolerance angle has to be small. This allows the method to work in reservoirs without complex structures caused by depositional processes or tectonic deformations. Directional artefacts in the modelling process are avoided through a careful choice of the simulation path. The tolerance angle and dipping direction of the strata can be obtained from a correlation between wells, or from seismic data, which are available in most hydrocarbon reservoirs, either by interpretation or by inversion that can also assist the construction of horizontal probability matrix.

3.1. Introduction

Characterization of subsurface heterogeneity is important for exploration and development of hydrocarbon reservoirs, because the distribution of lithologies determines the location and migration paths of the hydrocarbons. While borehole logs provide sufficient information of subsurface lithologies in a 1-D direction (depth), they provide little information on their lateral distribution. Additionally, in most cases, the density of borehole logs is relatively sparse compared with the total volume of the reservoir.

A variety of methods are available that can be used to solve this problem (Carle and Fogg 1996; Haldoren and Damsleth 1990). Among them, the Markov chain is a promising tool to describe the spatial structure (Carle *et al*. 1998; Parks *et al*. 2000; Davis 2002; Sartore *et al*. 2016; Weissmann et *al*. 1999). Carle and Fogg (1997) modelled the spatial variability for categorical variables with a multidimensional Markov Chain in which the transition rate matrices of 1-D continuous-lag Markov Chain models are interpolated laterally in order to obtain a 3-D Markov Chain. Lin and Harbaugh (1984) modelled the lithological or sedimentological structures with 2-D and 3-D Markov chains.

Other types of methods such as multiple-point statistics (Strebelle 2002; Daly and Caers 2010; Tahmasebi *et al*. 2012) and Markov Random Fields (Ulvmoen and Omre 2009; Rimstad *et al*. 2012) or its sub-class — Markov mesh models (Kjønsberg and Kolbjørnsen 2008; Stien and Kolbjørnsen 2011) could produce good results when the requirement of training images or randomly sampled data points are satisfied which may not be the case in some of the reservoirs since there are only several cored wells. On the other hand, even though the training images may exist, there is no inbuilt mechanism to ensure that they are in any way consistent with well data (Daly and Caers 2010).

In order to perform simulations conditioned by the observed data such as cored wells only, an extension of the Coupled Markov Chain model (CMC) developed by Elfeki and Dekking (2001) makes it possible to perform such conditional simulations on any number of wells in defining the future states. In fact, any number of wells can be included in the conditional simulation which makes the method practical.

In CMC, the information of the top or bottom boundary needs to be known beforehand in order to initiate the simulation. Hence, this scheme becomes unsuitable for reservoirs at depth of which the boundary properties are not known. Another drawback is that future states are not used for conditioning fully in the horizontal chain of CMC. For example, only the future state at the same row as the current simulating cell is employed. In real reservoirs, layers often dip to some degree due to depositional or tectonic processes. After transforming the geological cross-section into the grid cell domain needed for the Monte Carlo simulation, future states may not be at the same row as the simulating cell. This can make the simulation discontinuous in the middle part of the profile. In this chapter, a new scheme is proposed to use all future states for actively conditioning, through a tolerance angle. This allows the dipping of geological layers to be taken into account and more continuous simulation results can be obtained.

As the start of the simulation, the top or bottom boundary is simulated first using the horizontal Markov Chain model, in which the concept of actively conditioning is embedded.

Firstly, a short review of the two-dimensional Coupled Markov Chain theory will be presented, then the new Markov Chain scheme is proposed, with applications to a very simple synthetic example and parts of the complex Book Cliffs model (Chapter 2), as case studies. The results will be evaluated in the Discussion section, followed by the Conclusion.

3.2. Two-dimensional Coupled Markov Chain (CMC)

Elfeki and Dekking (2001) presented a two-dimensional coupled Markov chain (CMC) model in which the conditional probability of the future states is considered using the following numerical expression:

$$\Pr\left(Z_{i,j} = S_k \,\middle|\, Z_{i,j-1} = S_l, Z_{i-1,j} = S_m,\ Z_{i,Nx} = S_q\right)$$

$$= \frac{p_{lk}^{h} p_{kq}^{h(Nx-j)} p_{mk}^{v}}{\sum_{f=1}^{n} p_{lf}^{h} p_{fq}^{h(Nx-j)} p_{mf}^{v}} \qquad k = 1, \cdots, N \tag{3.1}$$

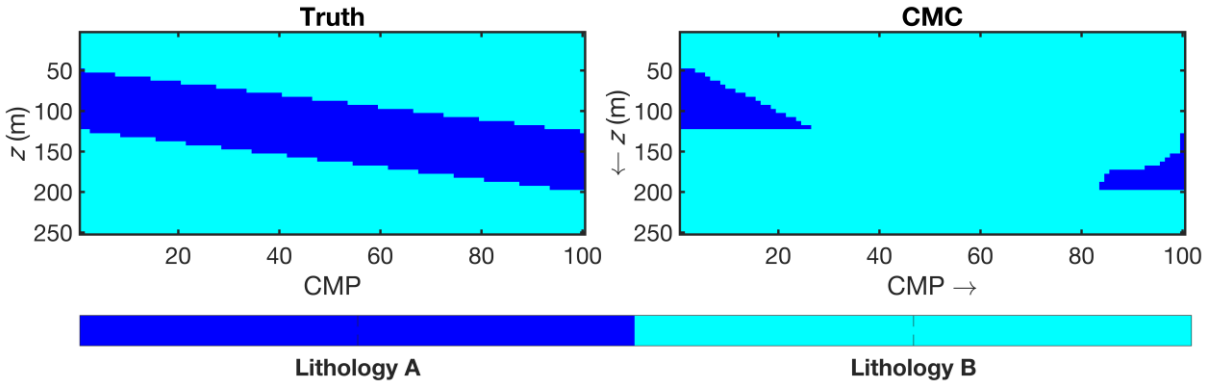
Here $Z_{i,j}$ is a random variable taking a value in the state space $\{S_1, S_2, \cdots, S_N\}$ with N states in the system; i, j are row and column numbers in the cell domain of size $(Nz \times Nx)$; the superscripts h and v represent the chains in the horizontal and vertical directions; $p_{kq}^{h(Nx-j)}$ is the $(Nx-j)^{\text{th}}$ step horizontal transition probability; p_{mk}^{v} denotes the vertical probability of transition from state S_m to state S_k.

In this method, the top or bottom boundary is needed in order to start the simulation. However, this kind of information is often missing in subsurface reservoirs. This problem can be addressed simply by using the horizontal probability matrix first in order to obtain the top or bottom states:

$$\Pr\left(Z_{i,j} = S_k \,\middle|\, Z_{i,j-1} = S_l, Z_{i,Nx} = S_q\right) = \frac{p_{lk} p_{kq}^{h(Nx-j)}}{p_{lq}^{h(Nx-j+1)}} \tag{3.2}$$

$$i = 1 \text{ or } Nz;\ k = 1, \cdots, N$$

However, equation (3.2) only considers the future state of the well at the top or bottom of the interval. This equation will later be modified slightly in order to account for the dip in the layers (equation (3.16) in Appendix A). Similarly, equation (3.1) only takes the future state at the same row as the current cell. After generating stochastic realizations of a simple example with two wells and two dipping lithologies, the geological layer (lithology A) turns out to be discontinuous as shown in Figure 3.1.



(a) (b)

Figure 3.1: Truth (**a**) and the stochastic simulation result (**b**). 100 realizations have been generated and the indicator function is used (Elfeki and Dekking 2001). The vertical and horizontal sampling intervals are 5 m and 25 m respectively.

In this example, two cored wells are located at the beginning and the end of the model (CMP=1 and 100, respectively; CMP is the common midpoint in seismic acquisition). The vertical probability matrix is estimated from the cored wells and the horizontal probability matrix is calculated from the geological profile or the truth. The standard for the sampling interval is to take the value of the minimum length and thickness of the geological units (Qi *et al*. 2016). The horizontal probability matrix is used to generate states of the top or bottom boundary first, which is missing in this synthetic case and the following examples. The arrows (Figure 3.1b) indicate the directions of the simulation path, i.e. towards the right while downward.

In the following examples this basic configuration will be kept the same or similar, only cored wells are available in the cross-section, or together with seismic inversions. Different arrows represent different paths during the simulation. The vertical and horizontal matrices are calculated from wells and geological profiles or by the application of Walther's Law (Middleton 1973), with the horizontal matrix being used to generate the boundary states first. After realizations or simulations, the final output image shows the lithology which appears most frequently at every cell as the indicator function used by Elfeki and Dekking (2001). In order to have a stable result, the number of simulations should be large enough which will increase the computation time on the other hand. After testing, it will be kept to be 100 to balance the time and stability.

3.3. Actively Conditioned CMC (A-CMC)

In the new scheme of actively conditioned CMC, in the following termed A-CMC, future states in the cored wells are taken into account over a predefined tolerance angle range. A sketch of this methodology is shown in Figure 3.2.

In the beginning, the states in the interval between the predefined tolerance angle (α) of the cored well at the right boundary are employed for conditioning (Figure 3.2a). As the current

cell is moving towards the right, the number of cells to be searched is getting smaller which means less future states will be used (Figure 3.2b). As it gets closer to the right boundary, the number of cells to be searched as future states continuously decreases until finally only one cell with a future state at the same row with the current cell is being found (Figure 3.2c). This means that A-CMC becomes CMC at this stage. Since the tolerance angle is small, the probability matrices for different future states to be used for conditioning are nearly identical to the horizontal one (see Appendix B).

Here the vertical and horizontal sampling intervals are 5 m and 25 m respectively, so the tolerance angle should be smaller than 11° ($\tan(11°) \approx \frac{5}{25}$) in order to have only one future state for conditioning at the end of each row in the simulation process. For most reservoirs, the horizontal extent of the lithologies is typically much larger than the vertical thickness, which allows a coarser horizontal sampling interval to be used. Hence, the tolerance angle has to be small enough so that only one future state is used for conditioning at the end of each row, and also for the validation in the usage of the same horizontal probability matrix (see Appendix B). At the same time, the dipping geometry of the geological layers should be considered as well in the selection of a tolerance angle.

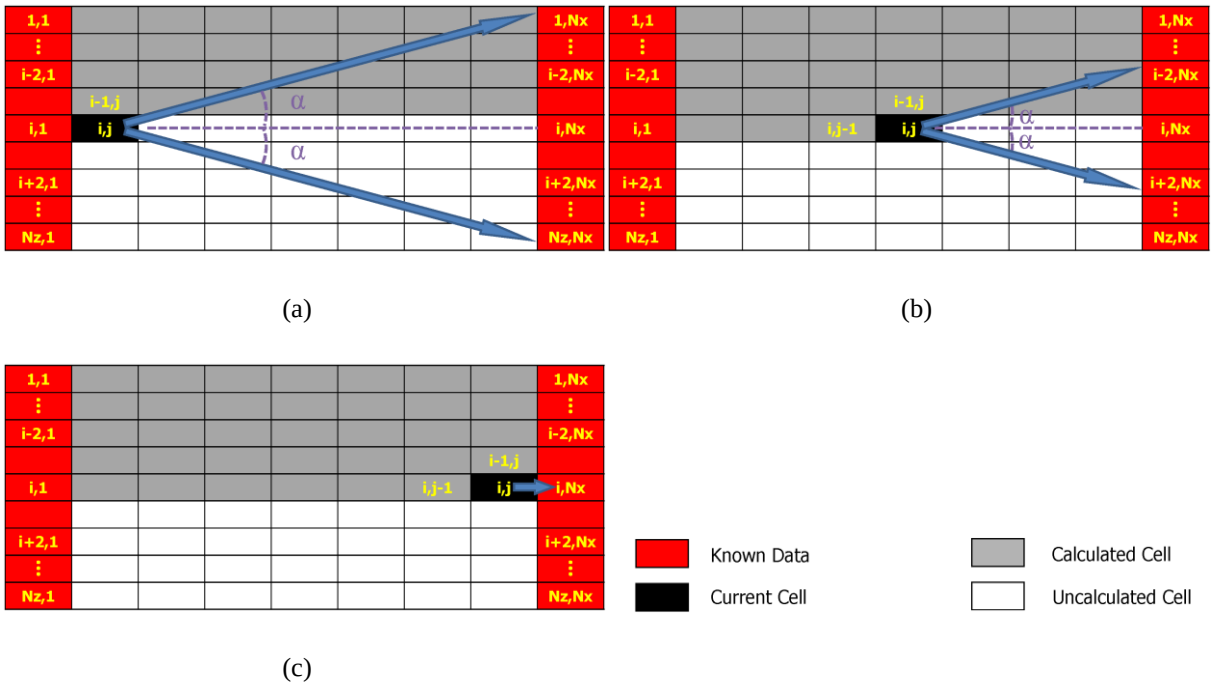


Figure 3.2: A schematic view of the A-CMC. See text for explanations.

Before testing this new scheme by applying equation (3.15) and equation (3.16) in Appendix A, a metric needs to be defined to assess the quality of the simulation results, in addition to comparing them with the geological profile by a simple subjective observation. The Confusion Matrix (Kohavi and Provost 1998) and Matthews Correlation Coefficient (*MCC*) (Matthews 1975; Jurman *et al*. 2012) are proposed as a visual inspection and numerical index, respectively.

Confusion Matrix shows the true and predicted classes in which each column represents the prediction and each row the truth. It summarizes the result in terms of the number of samples in which the diagonal values of the matrix are the numbers of correct predictions.

Matthews' Correlation Coefficient (*MCC*) is a numerical quantifier of the Confusion Matrix that can be used, because most of the time it is not easy to carry out a visual inspection of the table of a Confusion Matrix. The mathematical expression of *MCC* is as follows:

$$MCC = \frac{\sum_{k,l,m=1}^{N}(C_{kk}C_{ml} - C_{lk}C_{km})}{\sqrt{\sum_{k=1}^{N}\left[\left(\sum_{l=1}^{N} C_{lk}\right)\left(\sum_{f,g=1,f\neq k}^{N} C_{gf}\right)\right]}\sqrt{\sum_{k=1}^{N}\left[\left(\sum_{l=1}^{N} C_{kl}\right)\left(\sum_{f,g=1,f\neq k}^{N} C_{fg}\right)\right]}} \tag{3.3}$$

in which C_{km} is the element from the k^{th} row and m^{th}column of the Confusion Matrix *C*.

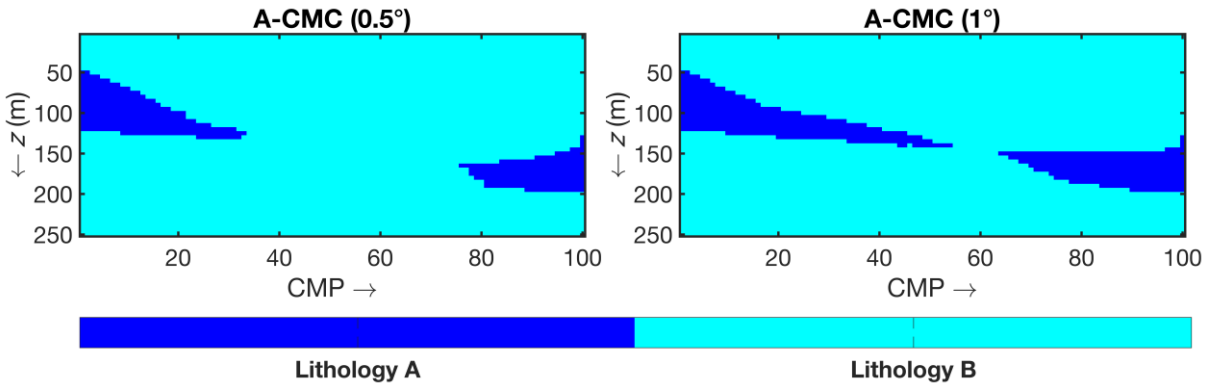
MCC values fall inside the interval [-1,1], where 1 represents a perfect classification and -1 means all values in Confusion Matrix are zeros except for two symmetrical elements. When all elements are equal or zeros except for one column, *MCC* equals to 0. The following examples illustrate this:

$$C = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix} \quad C = \begin{pmatrix} 0 & 0 & 9 \\ 0 & 0 & 0 \\ 9 & 0 & 0 \end{pmatrix} \quad C = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} \quad C = \begin{pmatrix} 6 & 0 & 0 \\ 6 & 0 & 0 \\ 6 & 0 & 0 \end{pmatrix}$$

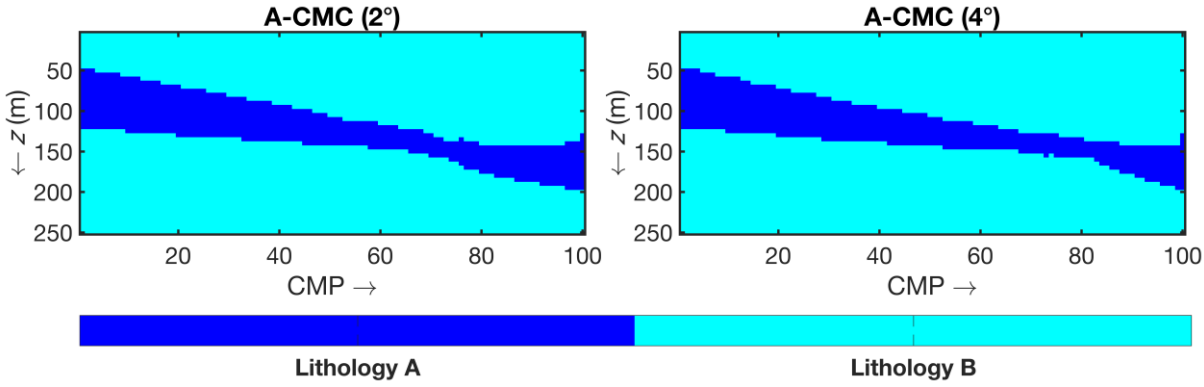
$$MCC = 1 \qquad MCC = -1 \qquad MCC = 0 \qquad MCC = 0$$

3.4. Simple Synthetic Example

In order to test our method, the simple synthetic model in Figure 3.1a is used again, with one dipping layer and the tolerance angle being set to different values. Figures 3.3a to 3.3d show the simulated results after using equation (3.15) and equation (3.16) of the A-CMC as shown in Appendix A.



(a) (b)

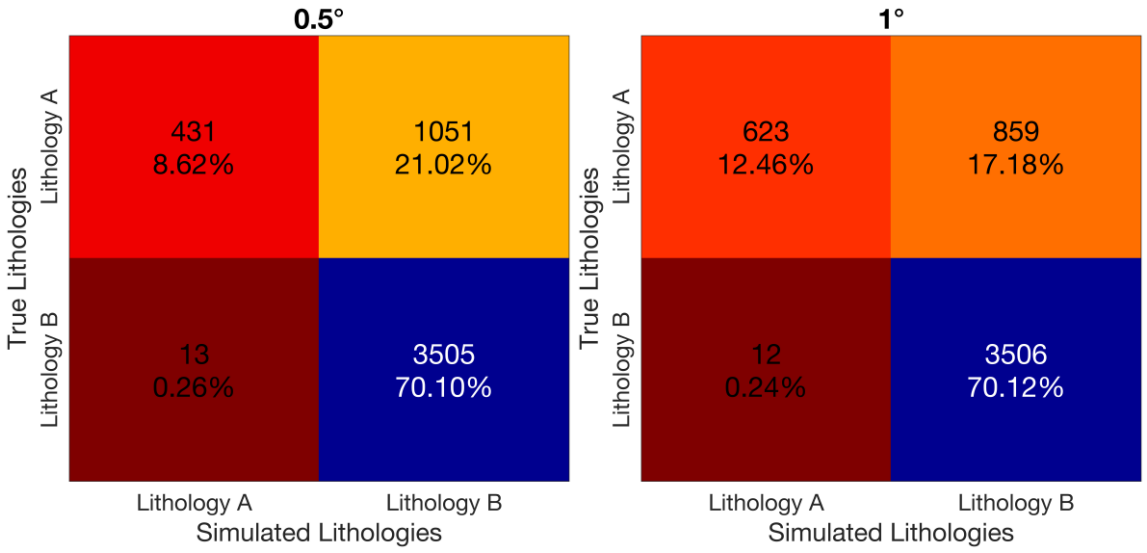


(c) (d)

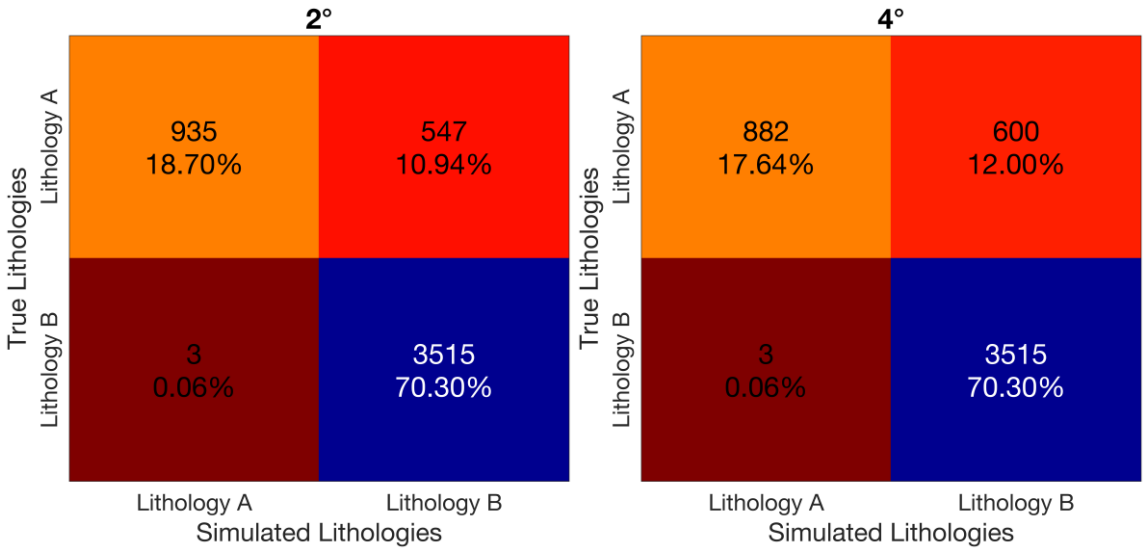
Figure 3.3: Simulated results of A-CMC under different tolerance angles: (**a**) 0.5°; (**b**) 1°; (**c**) 2°; (**d**) 4°. Two cored wells are located at CMP=1 and CMP=100. The truth is shown in Figure 1a. The simulation path is towards the right while downward as indicated by arrows.

It shows that with increasing tolerance angles more information of the right-hand side well is used for conditioning. In Figures 3.3a and 3.3b, the lithology A is not connected because the tolerance angle is too small. As the tolerance angle keeps increasing, lithology A is becoming more continuous (Figures 3.3c and 3.3d). However, there is a thinning effect at the right-hand side that is caused by the gradual move-away from the left well and directional artefacts from the previous top and left states.

In order to quantify these results, the Confusion Matrix and the *MCC* are calculated and displayed in Figure 3.4.



(a) (b)



(c) (d)

Figure 3.4: Confusion matrices for different tolerance angles: (**a**) 0.5° (*MCC*=0.4610); (**b**) 1° (*MCC*=0.5719); (**c**) 2° (*MCC*=0.7370); (**d**) 4° (*MCC*=0.7111). The percentages under the data number are for all samples.

The *MCC* is increasing with increased tolerance angles up to 2° while it decreases slightly for 4°. The reason for this is that when the tolerance angle is too large, more future information is used for conditioning which will include some unwanted information.

3.5. Complex Synthetic Examples (Book Cliffs)

The synthetic model presented before is very simple with only one dipping layer. In this section, the geological model created by Feng *et al*. (2017) (Chapter 2) is used as a more complex example to test the approach proposed here. The Book Cliffs outcrops in Utah (USA) are a well-studied reservoir analogue, and detailed high-resolution sequence stratigraphic interpretations have been proposed (O'Byrne *et al*. 1991, 1993; Taylor *et al*. 1991). Tetyukhina *et al*. (2014) developed a 2-D reservoir and petrophysical model for a part of the Book Cliffs section. Based on sequence stratigraphic principles, Feng *et al*. (2017) added more details, which makes the new model suitable for reservoir-geological purposes such as seismic inversion and reservoir characterization. The model is synthetic but of considerable realistic complexity, simulating a fluvio-deltaic sequence. Here only parts of this model are used, focusing on those where there are dipping layers.

3.5.1. Book Cliffs Model I

Figure 3.5 shows part of an incised valley cutting into floodplain deposits in the lower-middle part of the Book Cliffs model (Figure 2.3). The cross-section is 2.5 km in length and 100 m in thickness.

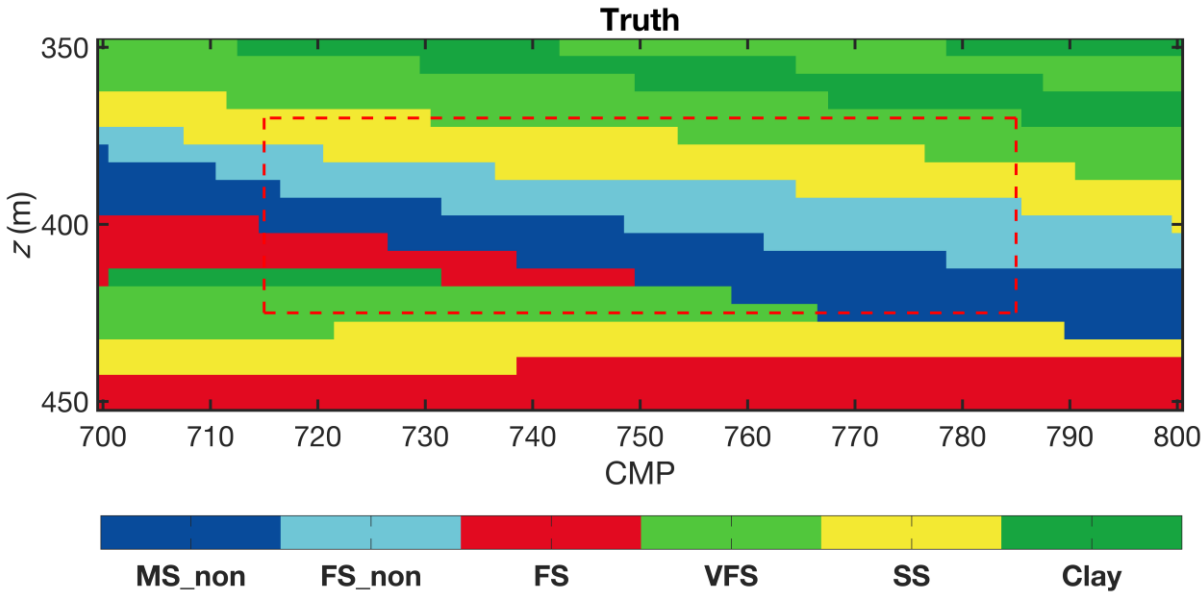


Figure 3.5: Book Cliffs Model I. MS_non−Medium-grained sandstone (non-marine); FS_non−Fine-grained sandstone (non-marine); FS−Fine-grained sandstone; VFS−Very fine-grained sandstone; SS−Siltstone. The first two are distinguished from marine lithologies in other parts of the overall model.

The horizontal and vertical transitional matrices are given in Table 3.1 which can be obtained by scanning of cored wells and normalizing afterwards, together with the two sampling intervals. Different lithologies contained in the model are coded with the numbers 1 to 6 with decreasing grain size (1 for MS_non, 6 for Clay). In order to eliminate the abrupt changes during the simulation processes since all the transitions are possible in reality even though have not been revealed in the wells, all zeros in the transition probabilities are set to a very small value (0.0001) following Elfeki and Dekking (2005), after which the entries in the diagonals are modified by $p_{ii} = p_{ii} - M \cdot 0.0001$ (M is number of zeros in each row).

Horizontal sampling interval = 25 m							Vertical sampling interval = 5 m						
Horizontal transition probability matrix							Vertical transition probability matrix						
State	1	2	3	4	5	6	State	1	2	3	4	5	6
1	0.9768	0.0001	0.0132	0.0066	0.0032	0.0001	1	0.7496	0.2500	0.0001	0.0001	0.0001	0.0001
2	0.0262	0.9734	0.0001	0.0001	0.0001	0.0001	2	0.0001	0.3333	0.0001	0.0001	0.6663	0.0001
3	0.0001	0.0001	0.9941	0.0001	0.0028	0.0028	3	0.1111	0.0001	0.6664	0.0001	0.2222	0.0001
4	0.0001	0.0001	0.0001	0.9841	0.0098	0.0058	4	0.0001	0.0001	0.1000	0.6997	0.0001	0.2000
5	0.0001	0.0132	0.0001	0.0022	0.9843	0.0001	5	0.1250	0.0001	0.0001	0.3750	0.4997	0.0001
6	0.0001	0.0001	0.0049	0.0294	0.0001	0.9654	6	0.0001	0.0001	0.0001	0.4996	0.0001	0.5000

Table 3.1: Input dataset for the Book Cliffs Model I.

Comparing the lithologies in two wells at CMP = 700 and 800 (Figure 3.6) gives an approximate small tolerance angle of 0.8°.

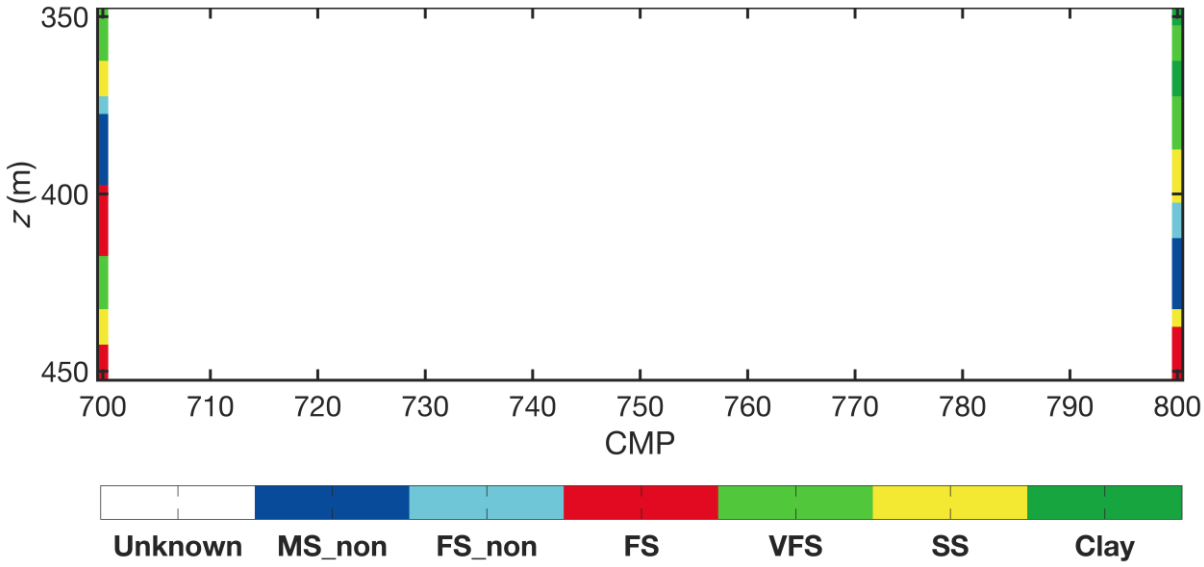
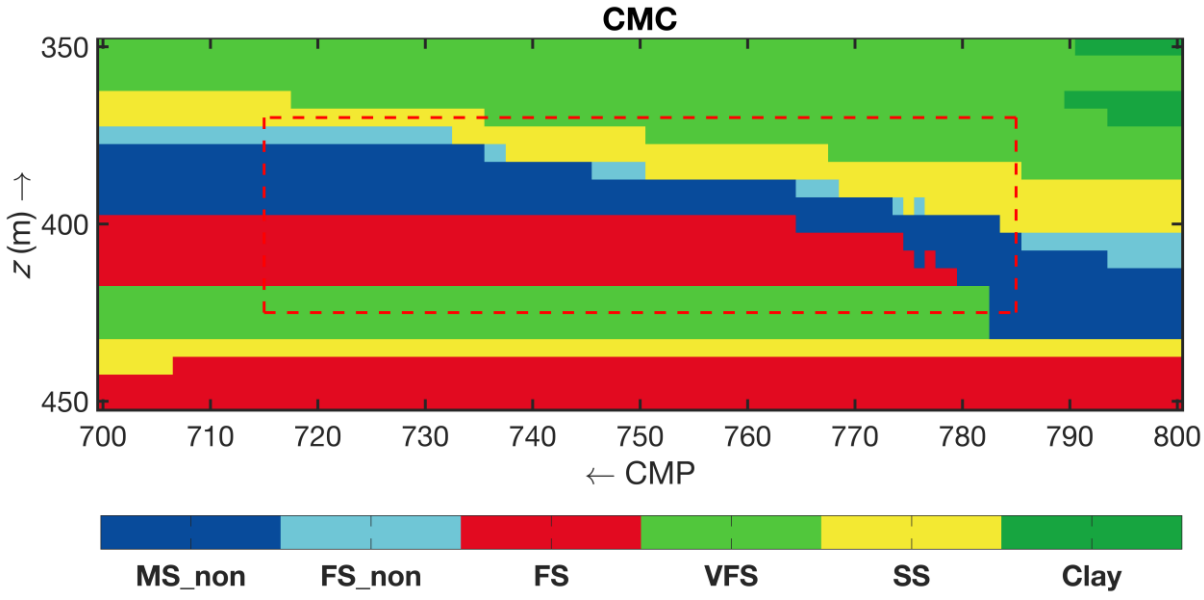
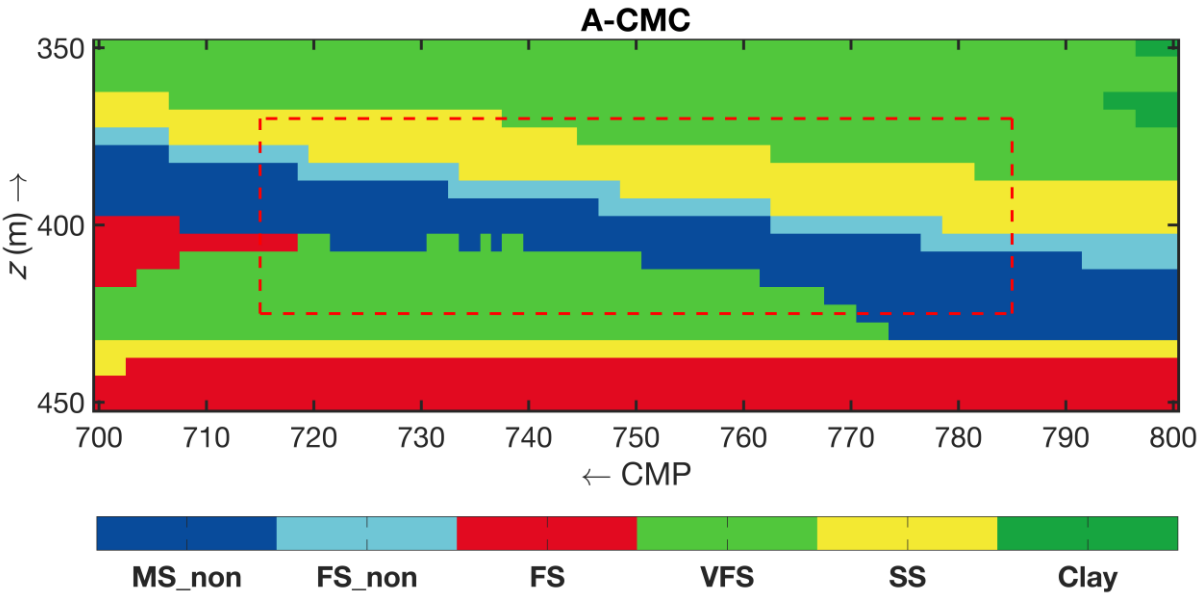


Figure 3.6: Starting model for simulations with two pseudo wells drilled at CMPs 700 and 800.

The simulation result after 100 realizations using A-CMC is shown in Figure 3.7, together with the result of CMC. The simulation path has been chosen to be towards the left and upward to compensate for directional artefacts.



(a)

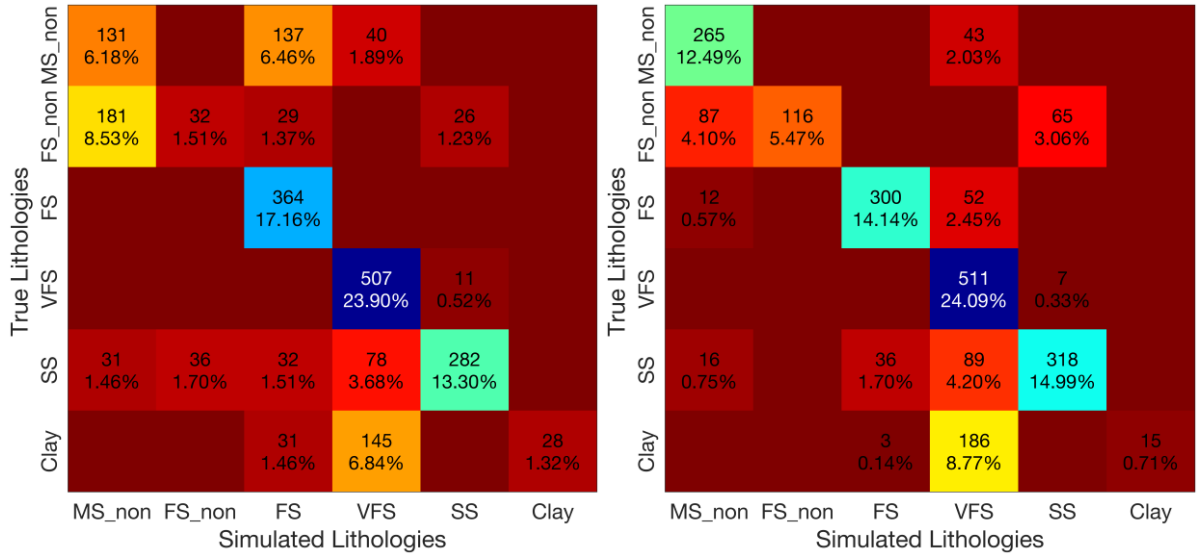


(b)

Figure 3.7: Results of CMC (**a**) and A-CMC (**b**). The simulation paths are indicated by arrows. The area inside the red dashed rectangle shows the major difference between CMC and A-CMC.

Figure 3.7 shows that with the help of a tolerance angle, FS_non is becoming continuous, which is stepwise discontinuous in CMC as indicated by the red dashed rectangle. SS and MS_non are continuous in both methods, which is caused by the larger values in the diagonal of the vertical transition probability matrix. The slope of MS_non has also been reproduced better than in CMC (Figure 3.7b).

For the clay, there is no conditioning information in the left well (CMP = 700) and therefore it cannot be simulated continuously at the top, nor does it appear at depths between 410 m and 420 m in both results.



(a) (b)

Figure 3.8: Confusion matrices of the two schemes: (**a**) CMC (*MCC*=0.5600); (**b**) A-CMC (*MCC*=0.6674).

The Confusion Matrix indicates that the new scheme (A-CMC) is numerically better than the old one (CMC) (Figure 3.8). The data samples are more compact and most of the values in the diagonal show higher values, especially for FS_non. However, FS has not been simulated better, based on the evaluation of confusion matrix, since in CMC (Figure 3.8a) all of FS has been simulated correctly while it is more distributed in A-CMC (Figure 3.8b). After comparing the simulated result (Figure 3.7a) with the truth (Figure 3.5), it can be realized that the fact that FS has been simulated correctly in terms of data samples in CMC is due to the compensation at different depths.

3.5.2. Book Cliffs Model II

The second example is also from the improved Book Cliffs model (Figure 2.3) and features different lithologies at depths between 30 m and 160 m (Figure 3.9). The lithologies are coded by numbers from 1 to 4 according to decreases in grain size (1 for FS_non, 4 for Clay_non), with coals being lithology 5.

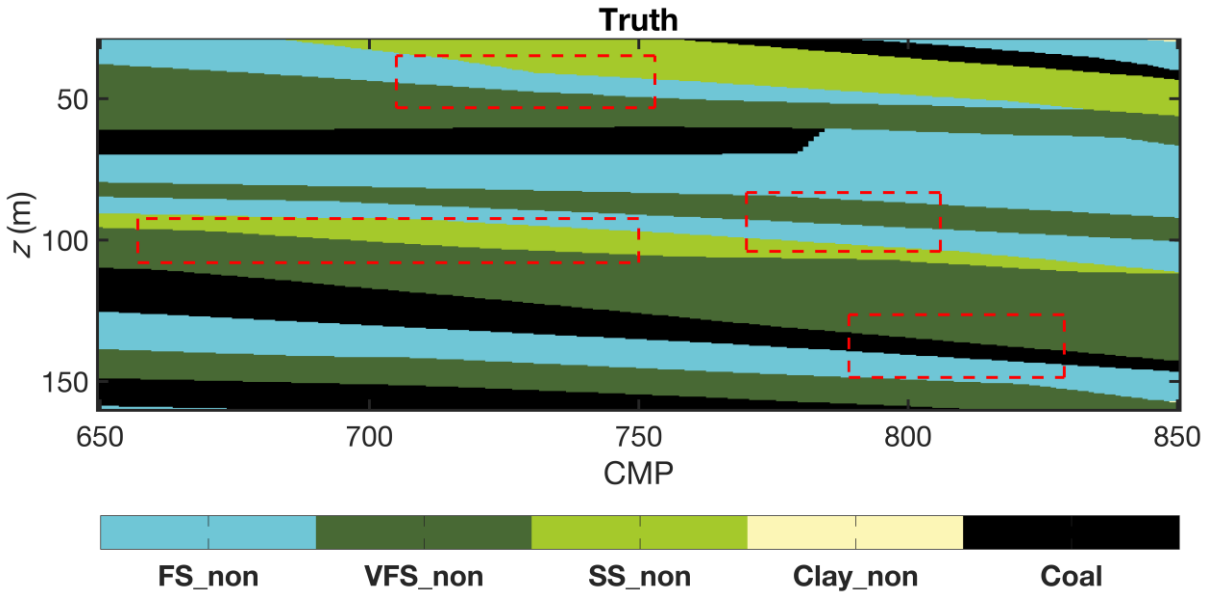


Figure 3.9: Book Cliffs Model II. FS_non−Fine-grained sandstone (non-marine); VFS_non−Very fine-grained sandstone (non-marine); SS_non−Siltstone (non-marine); Clay_non−Claystone (non-marine). Note that only very small proportions of Clay_non are in the upper- and bottom-right corners.

Instead of having two pseudo wells, one extra well has been assumed to be drilled in the middle of the cross-section (Figure 3.10). In this case, the simulation will be firstly performed between CMPs 650 and 750, then 750 and 850 afterwards.

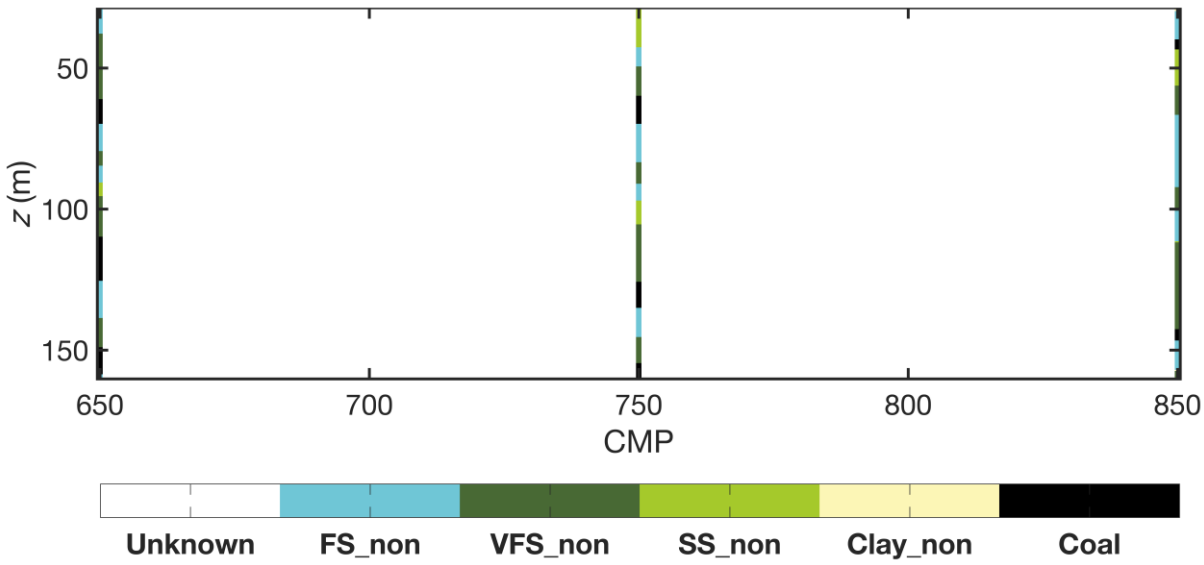
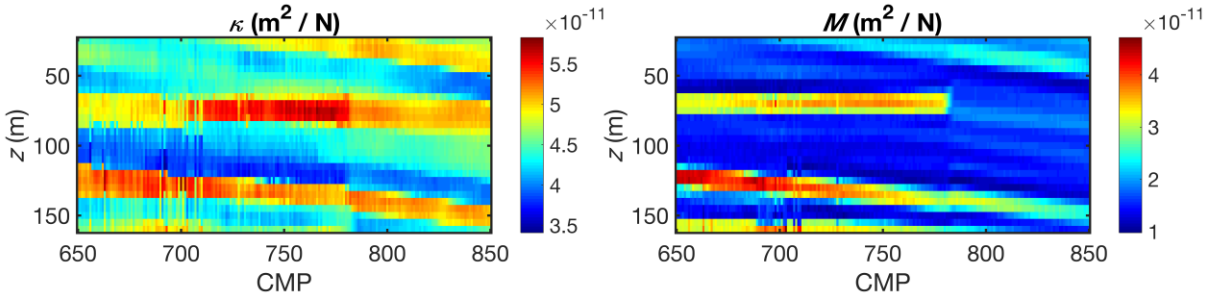


Figure 3.10: Starting model for simulations with three pseudo wells located at CMPs 650, 750 and 850.

In the above examples, the vertical matrix is estimated from the cored wells which are usually available in most of the reservoirs. The horizontal matrix is obtained from the geological profiles or the truth which may be missing (even though it is known here). On the other hand, geological understandings of the depositional environments as well as the sedimentary structures could provide some hints for the construction of horizontal matrices. Walther's Law (Middleton 1973) states the lithologies that are observed in the vertical depositional sequence must also be deposited in adjacent transects at another scale (Parks *et al*. 2000; Elfeki and Dekking 2005). With an application of this law, the lateral variability can be inferred.

Another prior knowledge to implement the new mode of CMC is the predefined tolerance angle which is estimated by a comparison of the cored wells before. This procedure could encounter some problems when same lithologies appear at different depths in the wells which will make people confused about how to correlate them. Seismic data or seismic inversion results provide images of the subsurface structures in terms of reflection amplitudes or property values which can help to estimate the tolerance angle as well as the inference of horizontal matrices.

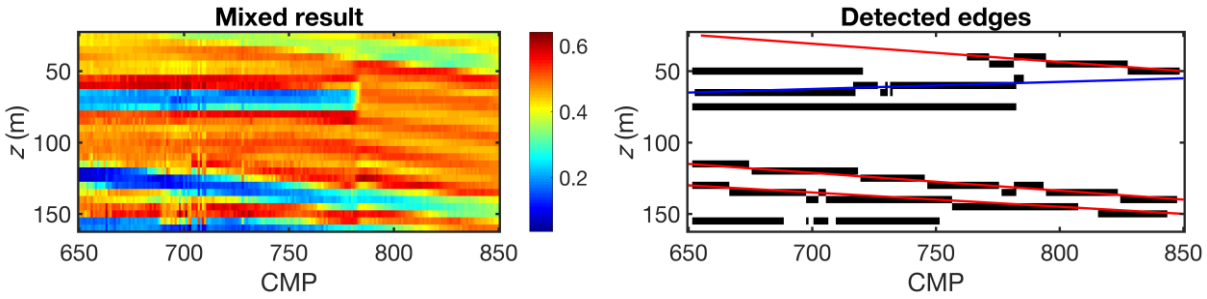
In this example, an effort for the estimation of tolerance angle and construction of horizontal matrix will be pursued based on the results of full-waveform inversion, or FWI. In this inversion scheme, all internal transmission effects and internal multiple scattering/mode-conversion are considered, allowing a recovery of broadband properties and providing a high resolution (Gisolf and van den Berg, 2010a; 2010b). Figure 3.11 shows the inversion results in terms of compressibility and shear compliance. Considering the small interval (130 m) of the cross-section and the highest resolution provided by inversion, the results are pretty good since some thin layers have been detected (Figure 2.10).



(a) (b)

Figure 3.11: Inversion results in terms of compressibility (a) ($\kappa = 1/K$, with K being the bulk modulus) and shear compliance (b) ($M = 1/\mu$, with μ being the shear modulus).

Then in order to estimate the tolerance angle based on the inversion results, edges or boundaries of lithological layers are going to be detected. Either of the inverted properties (κ and M) (Figure 3.11) could be the input for the edge detection. However, in order to mitigate the output difference caused by different inputs as well as to utilize the inversion results to the fullest and to highlight the property behaviors of different lithologies (Feng *et al*. 2015; 2017), the parameters of κ and M need to be mixed and the procedure is $(\kappa - M)/(\kappa + M)$. The mixed result is shown in Figure 3.12, together with the detected edges. A canny edge detector is used which is adaptable and can be tailored to recognition of edges under various environments (Canny 1986).



(a) (b)

Figure 3.12: Mixed result (a) as well as the detected edges of lithological layers (b).

The detected edges in Figure 3.12b shows that the layers are mainly dipping to the right as indicated by the red lines which is also reflected by the property values (Figure 3.11 and Figure 3.12a) and the dipping angle is 0.2° approximately which is going to be the tolerance angle in the simulation process. However, some layers show different dipping trends demonstrated by the blue line which will be ignored since only the main dipping direction should be taken into account.

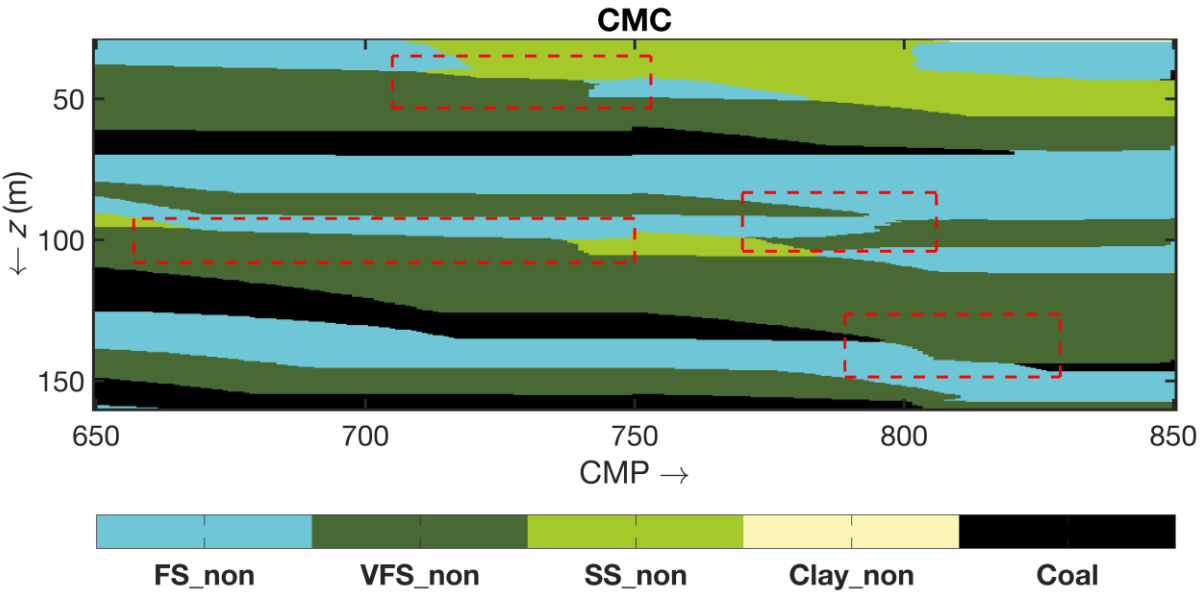
Based on the estimated small dipping angle (0.2°) and the application of Walther's Law (Middleton 1973), the horizontal matrix and the vertical matrix that is obtained from the cored wells are shown in Table 3.2.

Horizontal sampling interval = 25 m					
Horizontal transition probability matrix					
State	1	2	3	4	5
1	0.9900	0.0025	0.0025	0.0025	0.0025
2	0.0025	0.9900	0.0025	0.0025	0.0025
3	0.0025	0.0025	0.9900	0.0025	0.0025
4	0.0025	0.0025	0.0025	0.9900	0.0025
5	0.0025	0.0025	0.0025	0.0025	0.9900
Vertical sampling interval = 0.4 m					
Vertical transition probability matrix					
State	1	2	3	4	5
1	0.9622	0.0236	0.0094	0.0024	0.0024
2	0.0100	0.9698	0.0001	0.0001	0.0200
3	0.0127	0.0318	0.9553	0.0001	0.0001
4	0.2500	0.2500	0.0001	0.4998	0.0001
5	0.0354	0.0001	0.0044	0.0001	0.9600

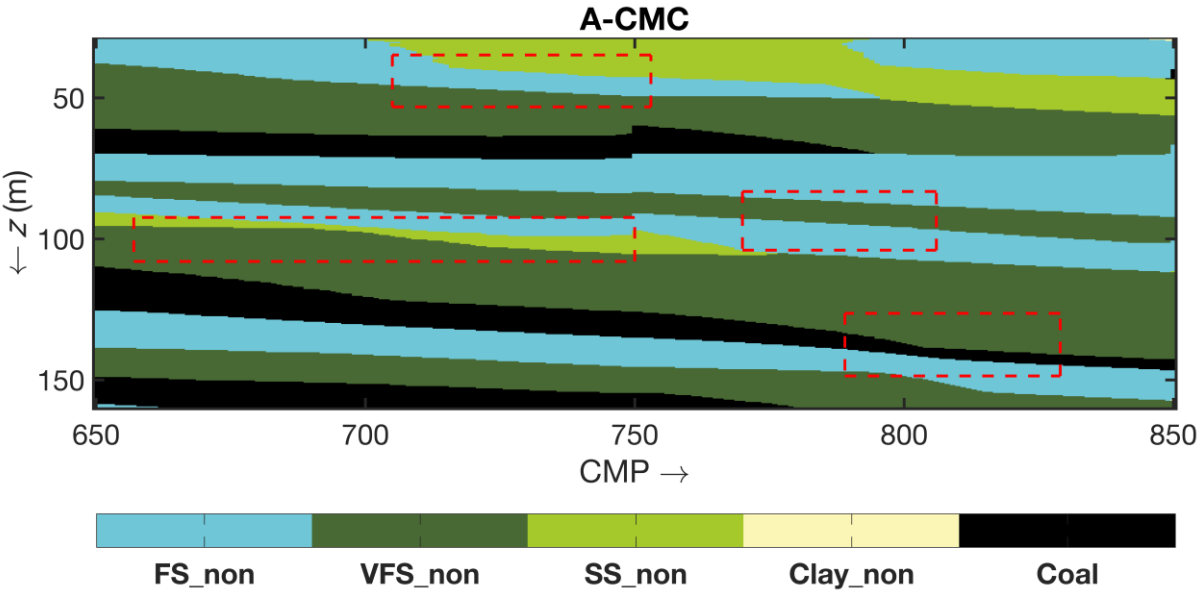
Table 3.2: Input Dataset for Book Cliffs model II.

Diagonal elements of the horizontal matrix are chosen to be large (0.99) since the dipping is very small (0.2°) and the transition from one lithology to itself is dominant. For the changes between different lithologies, there is no information and has been made all equiprobable according to $p_{ij}^{h} = (1 - p_{ii}^{h})/(n - 1)$ where n is the number of lithologies (Elfeki and Dekking 2005). Zeros in the vertical matrix has been set to 0.0001 in order to eliminate abrupt changes.

After simulations, the results of CMC and A-CMC are shown in Figure 3.13. Compared with the result of CMC (Figure 3.13a), the layers are continuous by using A-CMC (Figure 3.13b) as indicated by red dashed rectangles which should be the case in reality (Figure 3.9).



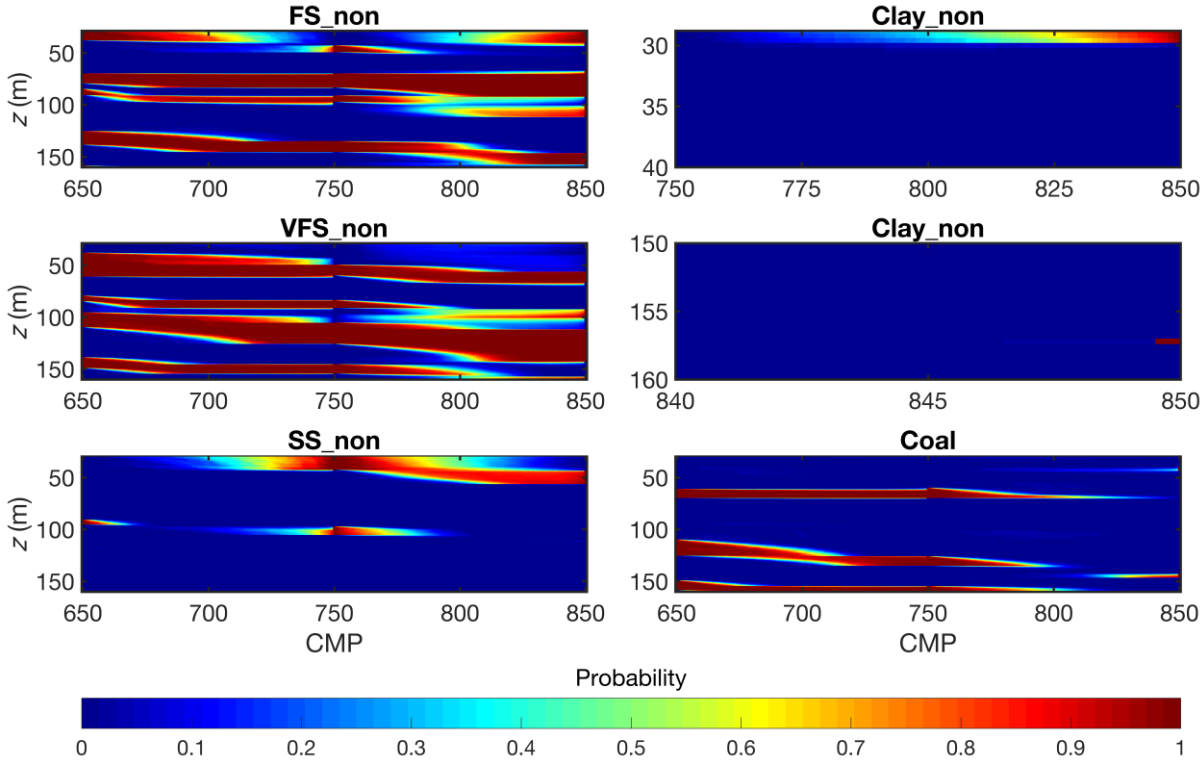
(a)



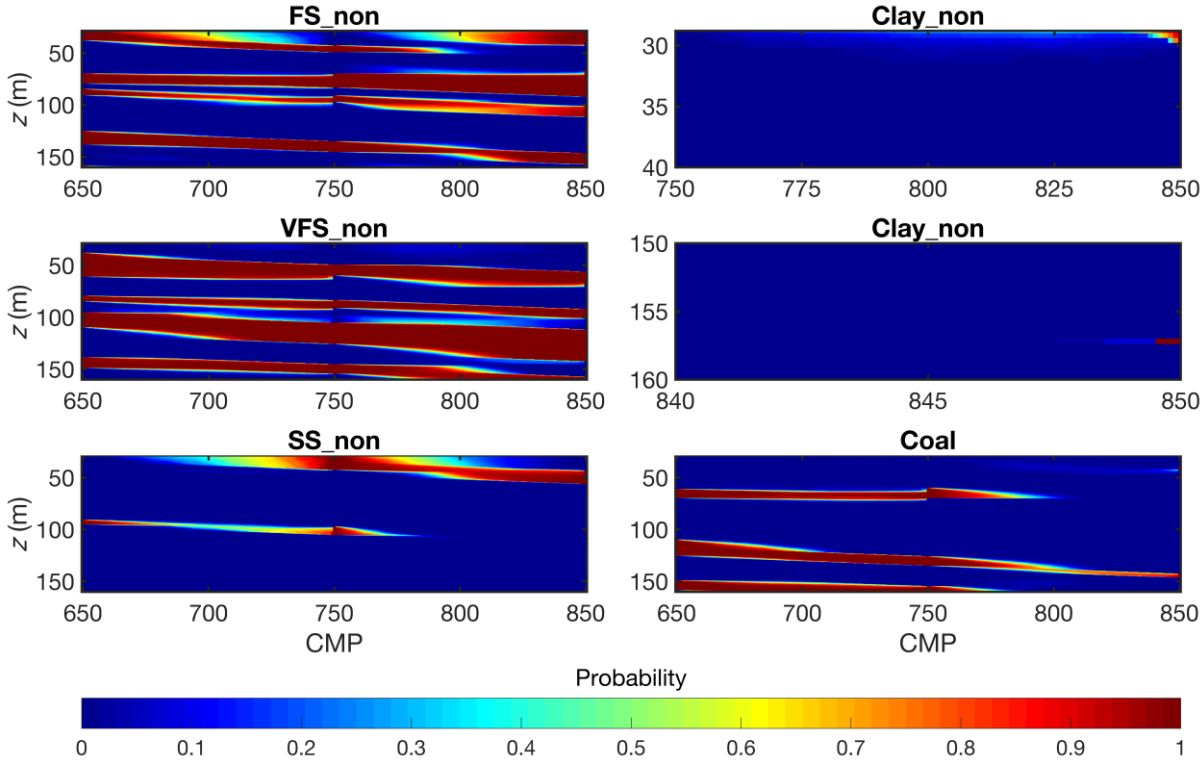
(b)

Figure 3.13: Simulated results of CMC (a) and A-CMC (b).

An indicator function (Elfeki and Dekking 2001) is selected to display the simulated results (Figures 3.1b, 3.3, 3.7, 3.13) in which the most occurred lithologies of the 100 simulations at every cell are chosen. The probability maps for different lithologies of the cross-section are shown in Figure 3.14 in which the point-wise uncertainty can be assessed. For Clay_non, small areas are shown since it only appears at the upper- and bottom-right parts.



(a)



(b)

Figure 3.14: Probability maps in terms of lithologies by CMC (a) and A-CMC (b).

With an incorporation of a tolerance angle, the point-wise uncertainty of lithologies has been decreased which leads to more continuous layers (Figure 3.13b).

3.6. Discussions

In the old 2-D Markov Chain model (CMC), a prior known future state may not be at the same vertical height in the cell domain, which may result in discontinuous layers in the simulation. However, in the new method (A-CMC) presented here, a tolerance angle is defined that assists in the active search of future states to be used for conditioning. The horizontal chain is conditioned based on different future states, and the vertical chain is coupled afterwards (see Appendix A). After numerous realizations, a more continuous distribution of the states (predicted lithologies) can be expected (Figure 3.14).

Another advantage of this new scheme is that slopes in the layers are simulated better than before, such as those of MS_non in Model I (Figure 3.7b) and Coal in Model II (Figure 3.13b), which appear horizontal towards the conditioning well in CMC. The same effect can also be seen in the examples discussed by Elfeki and Dekking (2001), where more wells are provided (Figure 4 in Elfeki and Dekking, 2001).

The tolerance angle is assumed to be small in order to use the same horizontal probability matrix for different future states to be used for conditioning. For reservoirs in sedimentary environments without strong tectonic activities, the dipping angle of the stratal layers is usually not more than a few degrees.

In the examples presented here, the tolerance angles are estimated by comparing the lithologies in the two wells at the leftmost and rightmost and an interpretation of the seismic inversion results as additional soft data could provide indications of subsurface structures (Artun *et al*. 2005). Alternately, these inversion results can be included within the simulations as hard constraints since they are closely related to the types of lithologies which will be discussed in the next chapters of this thesis.

The vertical and horizontal sampling intervals also influence the estimation of tolerance angle because there should be one future state to be used for conditioning at the end of each row in the simulation. For example, if the ratio between the vertical interval (dz) and horizontal interval (dx) is very small, the maximum oblique angle (see Appendix B) could be smaller than the dipping angle of the geological layers, then the tolerance angle will be the maximum oblique angle calculated from intervals. That is why the sampling intervals should be selected not only to suit the embedded lithological information but also for the layers' dips in A-CMC.

In the simulation process, there are two different choices in the vertical (upward and downward) and horizontal (leftward and rightward) directions, which lead to four different propagation directions for the Markov Chain modelling. The choice of directions is important because of the directional artefacts appearing in the simulation process which can be compensated for by the simulation path just like the use of regular simulation path needed to improve the long-range continuity of patterns in multiple-point geostatistics (Hu and

Chugunova 2008). With the structural implications from seismic inversion (Figures 3.11 and 3.12), determination in the choice of simulation directions can be made. Other selections in the simulation path such as the forward and backward scheme or the alternate advancing (AA) path (Elfeki and Dekking 2005; Li 2007) have not been considered yet which could be a further study.

However, in reservoirs without complex structures, A-CMC can describe the facies architectures with the understanding of the subsurface geometries obtained from seismic interpretation or inversion. For reservoirs with different dip directions of the geological layers, the dominant one, or the one which has the most significance in reservoir characterization, has to be selected which shares similarities with the object model simulations (Daly and Caers 2010). This is also the rule when choosing the tolerance angle based on the layers' dips.

The confusion matrix and the *MCC* can be calculated only when the ground truth is known and they will help to quantify the simulation result numerically instead of visually. Here they have illustrated the improvements of A-CMC over CMC. However, when reservoir lithologies are simulated in the subsurface where the ground truth is not known, these tools are inapplicable. However, with an assistance of probability maps (Figure 3.14), the point-wise uncertainty can be assessed that will not be available in traditional interpretations in which geological layers are correlated by professional geologists.

Both horizontal and vertical matrices are needed in the simulation. The diagonal entries in the matrices are related to the lithologies' lateral extension and thickness, while the off-diagonal entries are proportional to the horizontal and vertical juxtapositional tendencies. If the plausible facies proportions, mean lengths, and juxtapositional tendencies, can be estimated, the Markov Chain model of spatial variability can be easily formulated (Carle 2000).

While it is straightforward to understand vertical stratal stacking patterns, it is not easy to obtain the facies extension in the lateral direction from borehole data, not only because of the typically sparse lateral spacing, but also because of unknown variations in depositional dips and strikes (Carle *et al*. 1998). Outcrop analog profiles and horizontal wells may mitigate this problem (Purkis *et al*. 2012). The principle of Walther's Law (Middleton 1973) has been implemented to estimate the facies frequency in the horizontal direction, with a diagonal ratio considering the subsurface structures revealed by the seismic inversion. The sensitivity of the dominant values has not been tested here which has been demonstrated by Elfeki and Dekking (2005).

3.7. Conclusions

In this chapter, a new approach is presented to incorporate different states to be used in the conditioning of Markov chain models, in order to account for the dipping effect of stratal layers in reservoirs. The new scheme (A-CMC) is an incremental improvement of CMC in which a tolerance angle is defined and can be extended to 3-D in which three chains are going to be coupled. Correlations of cored wells and interpretations on seismic inversion results will help the estimation of tolerance angles as well as the construction of horizontal matrices.

Simulation paths should be carefully chosen in order to make compensations for the directional artefacts. For reservoirs with very complex structures such as curvilinear channels, this method is not applicable because of the nullification of the small tolerance angle premise and dilemmas in the selection of simulation path and Markov Random Fields or multiple-point geostatistics may be the solution.

The minor role provided by seismic data has been shown and seismic inversions will be incorporated in order to predict the lithologies in Chapter 4 and Chapter 5.

References

Artun, E., Mohaghegh, S.D., Toro, J., Wilson, T., Sanchez, A. 2005. *Reservoir Characterization Using Intelligent Seismic Inversion*. SPE Eastern Regional Meeting.

Canny, J. 1986. A computational approach to edge detection. *IEEETrans. Pattern Analysis and Machine Intelligence*, **8**(6), 679-698.

Carle, S.F., Fogg, G.E. 1996. Transition probability-based indicator geostatistics. *Mathematical Geology*, **28**(4), 453-476.

Carle, S.F., Fogg, G.E. 1997. Modeling spatial variability with one and multidimensional continuous-lag Markov chains. *Mathematical Geology*, **29**(7), 891-918.

Carle, S.F., Labolle, E.M., Weissmann, G.S., Vanbrocklin, D., Fogg, G.E. 1998. Conditional simulation of hydrofacies architecture: a transition probability/Markov approach. *In*: Fraser GS, Davis JM (eds) *Hydrogeologic Models of Sedimentary Aquifers, Concepts in Hydrogeology and Environmental Geology*. SEPM, Special Publication, **1**, 147-170.

Carle, S.F. 2000. *Use of a transition probability/Markov approach to improve geostatistical simulation of facies architecture*. AAPG Hedberg Symposium, Applied Reservoir Characterization Using Geostatistics, The Woodlands, Texas.

Daly, C., Caers, J. 2010. Multi-point geostatistics – an introductory overview. *First Break*, **28**, 39-47.

Davis, J.C. 2002. *Statistics and data analysis in geology*, 3rd Edition. John Wiley and Sons, New York.

Elfeki, A., Dekking, M. 2001. A Markov chain model for subsurface characterization: theory and applications. *Mathematical Geology*, **33**(5), 569-589.

Elfeki, A., Dekking, M. 2005. Modelling subsurface heterogeneity by coupled Markov chains: Directional dependency, Walther's law and entropy. *Geotechnical and Geological Engineering*, **23**, 721-756.

Feng, R.H., Luthi, S.M., Gisolf, D. & Sharma, S. 2015. *Non-linear full-waveform inversion (FWI-res) of time-lapse seismic data on a higher-resolution geological and petrophysical model, Book Cliffs (Utah, USA)*. Presented at the 85th SEG Annual Meeting, New Orleans, 18-23 October.

Feng, R.H., Luthi, S.M., Gisolf, A., Siddharth, S. 2017. Obtaining a high-resolution geological and petrophysical model from the results of reservoir-orientated elastic wave-equation-based seismic inversion. *Petroleum Geoscience*. doi: 10.1144/petgeo2015-076.

Gisolf, A. & van den Berg, P.M. 2010a. *Target oriented non-linear inversion of seismic data*. P308, presented at the 72nd EAGE Annual Meeting, Barcelona, 14-17 June.

Gisolf, A. & van den Berg, P.M. 2010b. *Target-oriented non-linear inversion of time-lapse seismic data*. Presented at the 80th SEG Annual Meeting, Denver, 17-22 October.

Haldoren, H.H., Damsleth, E. 1990. Stochastic modelling. *Journal of Petroleum Technology*, **42**(4), 404-412.

Hu, L.Y., Chugunova, T. 2008. Multiple-point geostatistics for modeling subsurface heterogeneity: a comprehensive review. *Water Resources Research*, **44**, W11413, doi:10.1029/2008WR006993.

Jurman, G., Riccadonna, S., Furlanello, C. 2012. A Comparison of MCC and CEN Error Measures in Multi-Class Prediction. *PLoS ONE*, **7**(8). doi:10.1371/journal.pone.0041882

Kjønsberg, H., Kolbjørnsen, O. 2008. *Markov mesh simulations with data conditioning through indicator kriging*. Presented at the 8th International Geostatistics Congress, Santiago, Chile.

Kohavi, R., Provost, F. 1998. Glossary of terms, Special issue on applications of machine learning and the knowledge discovery process. *Machine Learning*, **30**. 271-274.

Li, W. 2007. Markov chain random fields for estimation of categorical variables. *Mathematical Geosciences*, **39**, 321-335.

Lin, C., Harbaugh, J.W. 1984. *Graphic display of two- and three-dimensional Markov computer models in geology*. Van Nostrand Reinhold, New York.

Matthews, B. 1975. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *BBA-Protein Struct M*. **405**, 442-451.

Middleton, G.V. 1973. Johannes Walther's Law of the correlation of facies. *Geological Society of America Bulletin*, **84**, 979-988.

O'Byrne, C.J., Flint, S. 1991. *Anatomy of a parasequence: Quantitative heterogeneity analysis for reservoir modeling, Book Cliffs, Utah*. Presented at the 74th AAPG Annual Meeting.

O'Byrne, C.J., Flint, S. 1993. High-resolution sequence stratigraphy of Cretaceous shallow marine sandstones, Book Cliffs outcrop, Utah, USA-application to reservoir modelling. *First Break*, **11**(10), 445-459.

Parks, K.P., Bentley, L.R., Crowe, A. 2000. Capturing geological realism in stochastic simulations of rock systems with markov statistics and simulated annealing. *Journal of sedimentary research*, **70**(4), 803-813.

Purkis, S.J., Vlaswinkel, B., Gracias, N. 2012. Vertical-to-lateral transitions among Cretaceous carbonate facies – a means to 3-D framework construction via Markov analysis. *Journal of Sedimentary Research*, **82**, 232-243.

Qi, X., Li, D., Phoon, K., Cao, Z., Tang, X. 2016. Simulation of geological uncertainty using coupled Markov chain. *Engineering Geology*, **207**, 129-140.

Rimstad, K., Avseth, P., Omre, H. 2012. Hierarchical Bayesian lithology/fluid prediction: A North Sea case study. *Geophysics*, **77**(2), B69-B85.

Sartore, L., Fabbri, P., Gaetan, C. 2016. spMC: an R-package for 3D lithological reconstructions based on spatial Markov chains. *Computers & Geosciences*, **94**, 40-47.

Stien, M., Kolbjørnsen, O. 2011. Facies modeling using a Markov mesh model specification. *Mathematical Geosciences*, **43**, 611-624.

Strebelle, S. 2002. Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology*, **34**(1), 1-21.

Tahmasebi, P., Hezarkhani, A., Sahimi, M. 2012. Multiple-point geostatistical modeling based on the cross-correlation functions. *Comput. Geosci.*, **16**, 1179-797.

Taylor, D.R., Lovell, R.W. 1991. Recognition of high-frequency sequences in the Kenilworth Member of the Blackhawk Formation, Book Cliffs, Utah. *AAPG Special Volumes*, **163**, 22-46

Tetyukhina, D., Luthi, S.M. & Gisolf, D. 2014. Acoustic non-linear full-waveform inversion on an outcrop-based detailed geological and petrophysical model (Book Cliffs, Utah). *American Association of Petroleum Geologist Bulletin*, **98**, 119-134.

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

Weissmann, G.S., Carle, S.F., Fogg, G.E. 1999. Three dimensional hydrofacies modeling based on soil surveys and transition probability geostatistics. *Water Resources Research*, **6**, 131-143.

Ziyan, U., Tuch, D., Westin, C. 2006. *Segmentation of thalamic nuclei from DTI using spectral clustering*. Presented at the 9th International Conference on Medical Image Computing and Computer-Assisted Intervention, Copenhagen.

Appendix A: Active Horizontal Conditioning with a Tolerance Angle

Similar as with the Coupled Markov Chain model with one conditioned future state in Elfeki and Dekking (2001), we perform the conditioning process in the horizontal chain first, then couple it with the vertical chain. The difference with before is that here there are more than one future states to be conditioned.

In the following, to keep things simple, we assume there are only two given future states (S_q and S_p) to be conditioned in the horizontal chain, and the states at the lattice ($Z_{i,j}$) need to be simulated in a two-dimensional domain of cells ($Nz \times Nx$). In total, N geological states have been given and are coded as numbers. The numerical expression is given by

$$\begin{aligned} &\Pr\left(Z_{i,j} = S_k \middle| Z_{i,j-1} = S_l, Z_{i-1,j} = S_m, Z_{i,Nx} = S_q, Z_{i',Nx} = S_p\right) \\ &= C' \Pr\left(Z_{i,j} = S_k \middle| Z_{i,j-1} = S_l, Z_{i,Nx} = S_q, Z_{i',Nx} = S_p\right) \cdot \Pr\left(Z_{i,j} = S_k \middle| Z_{i-1,j} = S_m\right) \end{aligned} \tag{3.4}$$

C' is the normalizing constant to exclude the transitions to different states in the two chains. $Z_{i,Nx}$ and $Z_{i',Nx}$ are in the well with the future states to be used for conditioning, if within the range defined by the tolerance angle. The subscripts of i and i' mean that the future states are at different heights in the well. Since the tolerance angle is very small, the two future states can be supposed at the same vertical height with the current cell even though they are not, and the same horizontal transition probability matrix (P^h) will be used.

We now define

$$\mathrm{A} = \Pr\left(Z_{i,j} = S_k \middle| Z_{i,j-1} = S_l, Z_{i,Nx} = S_q, Z_{i',Nx} = S_p\right) \tag{3.5}$$

and

$$\mathrm{B} = \Pr\left(Z_{i,j} = S_k \middle| Z_{i-1,j} = S_m\right) \tag{3.6}$$

After application of the joint probability theory, equation (3.5) can be rewritten as

$$\mathrm{A} = \frac{\Pr\left(Z_{i,j} = S_k, Z_{i,j-1} = S_l, Z_{i,Nx} = S_q, Z_{i',Nx} = S_p\right)}{\Pr\left(Z_{i,j-1} = S_l, Z_{i,Nx} = S_q, Z_{i',Nx} = S_p\right)} \tag{3.7}$$

The probabilities in the numerator and denominator of equation (3.7) can be further arranged in factors of conditional probabilities and marginal probabilities as:

$$\mathrm{A} = \frac{\Pr\left(Z_{i,Nx} = S_q \middle| Z_{i,j} = S_k, Z_{i,j-1} = S_l, Z_{i',Nx} = S_p\right) \cdot \Pr\left(Z_{i,j} = S_k, Z_{i,j-1} = S_l, Z_{i',Nx} = S_p\right)}{\Pr\left(Z_{i,Nx} = S_q \middle| Z_{i,j-1} = S_l, Z_{i',Nx} = S_p\right) \cdot \Pr\left(Z_{i,j-1} = S_l, Z_{i',Nx} = S_p\right)} \tag{3.8}$$

By applying the Markovian rule on the conditional probability and realizing that the known $Z_{i,Nx}$ and $Z_{i',Nx}$ are independent of each other, equation (3.8) can be simplified as:

$$\mathrm{A} = \frac{\Pr(Z_{i,Nx} = S_q | Z_{i,j} = S_k) \cdot \Pr(Z_{i,j} = S_k, Z_{i,j-1} = S_l, Z_{i',Nx} = S_p)}{\Pr\left(Z_{i,Nx} = S_q \middle| Z_{i,j-1} = S_l\right) \cdot \Pr\left(Z_{i,j-1} = S_l, Z_{i',Nx} = S_p\right)} \tag{3.9}$$

The theory of joint probability and Markovian rule can be applied repeatedly to obtain

$\mathrm{A} =$

$$\frac{\Pr(Z_{i,Nx} = S_q | Z_{i,j} = S_k) \cdot \Pr(Z_{i',Nx} = S_p | Z_{i,j} = S_k) \cdot \Pr(Z_{i,j} = S_k | Z_{i,j-1} = S_l) \cdot \Pr(Z_{i,j-1} = S_l)}{\Pr\left(Z_{i,Nx} = S_q \middle| Z_{i,j-1} = S_l\right) \cdot \Pr\left(Z_{i',Nx} = S_p | Z_{i,j-1} = S_l\right) \cdot \Pr(Z_{i,j-1} = S_l)}$$

$$= \frac{\Pr(Z_{i,Nx} = S_q | Z_{i,j} = S_k) \cdot \Pr(Z_{i',Nx} = S_p | Z_{i,j} = S_k) \cdot \Pr(Z_{i,j} = S_k | Z_{i,j-1} = S_l)}{\Pr\left(Z_{i,Nx} = S_q \middle| Z_{i,j-1} = S_l\right) \cdot \Pr\left(Z_{i',Nx} = S_p | Z_{i,j-1} = S_l\right)} \tag{3.10}$$

By substituting the transition probability matrix (Elfeki and Dekking 2001) for the conditional probabilities in equations (3.5) and (3.10), we get:

$$\mathrm{A} = \Pr\left(Z_{i,j} = S_k \middle| Z_{i,j-1} = S_l, Z_{i,Nx} = S_q, Z_{i',Nx} = S_p\right)$$

$$= \frac{p_{kq}^{h(Nx-j)} \cdot p_{kp}^{h'(Nx-j)} \cdot p_{lk}^{h}}{p_{lq}^{h(Nx-j+1)} \cdot p_{lp}^{h'(Nx-j+1)}} \tag{3.11}$$

in which $p_{kq}^{h(Nx-j)}$ and $p_{kp}^{h'(Nx-j)}$ are the $(Nx-j)$-step transition probabilities, $p_{lq}^{h(Nx-j+1)}$ and $p_{lp}^{h'(Nx-j+1)}$ are the $(Nx-j+1)$-step transition probabilities.

The probability matrices for p_{kq}^{h} , $p_{kp}^{h'}$ and p_{lq}^{h} , $p_{lp}^{h'}$ are approximately the same and equal the horizontal transition probability (P^h), because of the small tolerant angle (see Appendix B). This is also why the numbers of steps in p_{kq}^{h} , $p_{kp}^{h'}$ and p_{lq}^{h} , $p_{lp}^{h'}$ are the same.

Equation (3.6) can be formalized in the same way by using the vertical transition matrix (P^v)

$$\mathrm{B} = \Pr\left(Z_{i,j} = S_k \middle| Z_{i-1,j} = S_m\right) = p_{mk}^{v} \tag{3.12}$$

Combining Equation (3.11) and Equation (3.12), we find

$$\Pr\left(Z_{i,j} = S_k \middle| Z_{i,j-1} = S_l, Z_{i-1,j} = S_m, Z_{i,Nx} = S_q, Z_{i',Nx} = S_p\right)$$

$$= C' \frac{p_{kq}^{h(Nx-j)} \cdot p_{kp}^{h'(Nx-j)} \cdot p_{lk}^{h}}{p_{lq}^{h(Nx-j+1)} \cdot p_{lp}^{h'(Nx-j+1)}} \cdot p_{mk}^{v} \tag{3.13}$$

The normalizing constant C' can be computed in the same way as in Elfeki and Dekking (2001), and using the same horizontal matrix (P^h) and vertical matrix (P^v), finally we have

$$p_{lm,k|qp} = \Pr\left(Z_{i,j} = S_k \middle| Z_{i,j-1} = S_l, Z_{i-1,j} = S_m, Z_{i,Nx} = S_q, Z_{i',Nx} = S_p\right)$$

$$= \frac{p_{kq}^{h(Nx-j)} \cdot p_{kp}^{h(Nx-j)} \cdot p_{lk}^{h} \cdot p_{mk}^{v}}{\sum_{f=1}^{n} p_{fq}^{h(Nx-j)} \cdot p_{fp}^{h(Nx-j)} \cdot p_{lf}^{h} \cdot p_{mf}^{v}} \qquad k = 1, \cdots, N \tag{3.14}$$

Similarly, if there are more than two future states on the state space $\{S_1, S_2, \ldots, S_D\}$ to be conditioned, equation (3.14) will be

$$p_{lm,k|r} = \frac{[\prod_{r=1}^{d} p_{kr}^{h(Nx-j)}] \cdot p_{lk}^{h} \cdot p_{mk}^{v}}{\sum_{f=1}^{n} [\prod_{r=1}^{d} p_{fr}^{h(Nx-j)}] \cdot p_{lf}^{h} \cdot p_{mf}^{v}}$$

$$k = 1, \cdots, N; r = 1, \cdots, D; D \leq N \tag{3.15}$$

State space $\{S_1, S_2, \ldots, S_D\}$ is a subspace of state space $\{S_1, S_2, \ldots, S_N\}$.

Similar to the derivation of equation (3.15), in order to account for the dipping effect, the modification of equation (3.2) is going to be

$$p_{l,k|r} = \frac{[\prod_{r=1}^{d} p_{kr}^{h(Nx-j)}] \cdot p_{lk}^{h}}{\prod_{r=1}^{d} p_{lr}^{h(Nx-j+1)}}$$

$$k = 1, \cdots, N; r = 1, \cdots, D; D \leq N \tag{3.16}$$

Appendix B: Oblique Transition Probability Matrix

In equation (3.15), the probability matrices for p_{kq}^{h} and $p_{kp}^{h'}$ are assumed the same, as long as the condition of a small tolerance angle is satisfied. In the two examples of the Book Cliffs model (Chapter 2) the truth is already known and, therefore, the transition probability matrices at different oblique angles can be obtained easily and compared with the one in the horizontal direction. In real cases this is not possible, because the truth is not known.

Since there is only one future state to be used for conditioning at the very end of each row in the simulation process, the tolerance angle α and maximum oblique angle θ_o^{max} depend on the sampling intervals in the vertical and horizontal directions. The relationship between tolerance angle α, maximum oblique angle θ_o^{max} and sampling intervals (dz, dx) is the following:

$$\alpha \leq \theta_o^{max} = \tan^{-1}\left(\frac{dz}{dx}\right) \tag{3.17}$$

In the two Book Cliffs models, dz has been given different values (5 m and 0.4 m), while dx is the same (25 m). So the maximum oblique angle θ_o^{max} is 11° for the first example, and 0.9° for the second example, approximately.

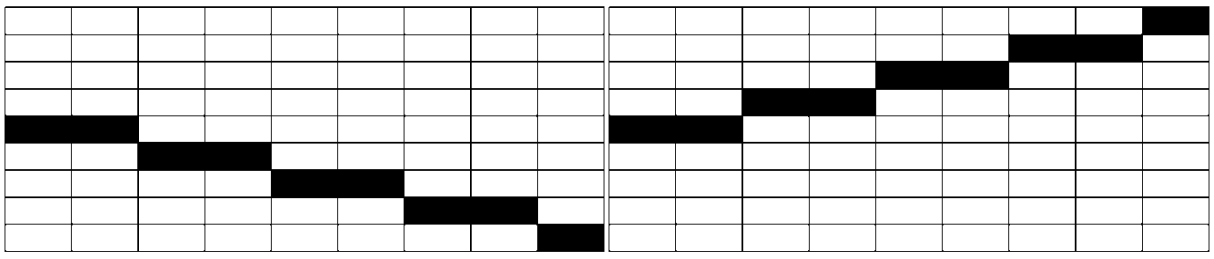
Equation (3.17) should be applied first, for the determination of tolerance angle, and at the same time the dip angle in stratal layers needs to be considered as well.

The Frobenius distance (Ziyan *et al.* 2006) is proposed as measure for the dissimilarity between the horizontal and oblique matrices:

$$d_F = \sqrt{tr((A-B)^2)} \tag{3.18}$$

where d_F is the Frobenius distance, tr is the trace function in linear algebra, and A and B are the two matrices.

There are two oblique directions, of which one is dipping downward (Figure 3.15a) and the other is upward (Figure 3.15b).



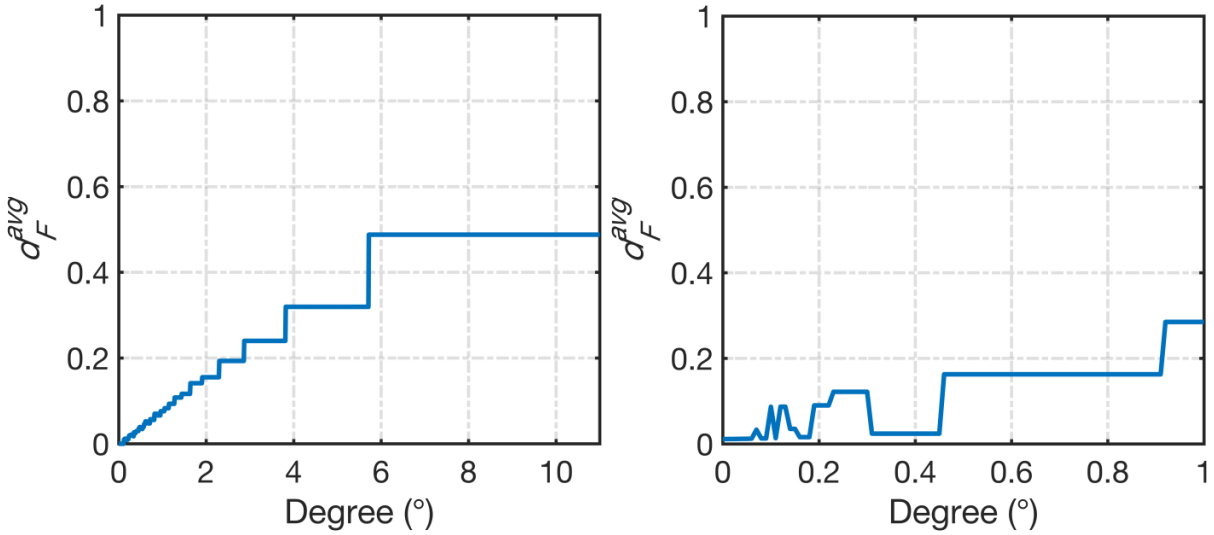
(a) (b)

Figure 3.15: Two oblique directions. (**a**) downward; (**b**) upward. Both are viewed from left to right.

The average distance d_F^{avg} will be the mean of the distances in the upward d_F^{up} and downward d_F^{down} directions:

$$d_F^{avg} = \frac{(d_F^{up} + d_F^{down})}{2} \tag{3.19}$$

Figure 3.16 is showing the distance results of the two Book Cliffs Models (Figure 3.16a for Model I; Figure 3.16b for Model II). Also, Figure 3.16a is showing the results when conditioning from right to left, while Figure 3.16b is showing the result when conditioning from left to right. These directions are consistent with the horizontal transition probability matrices in the Markov chain model.



(a) (b)

Figure 3.16: Distance results in the Book Cliffs models. (**a**) for Model I; (**b**) for Model II.

It can be seen that the distances or dissimilarities between the matrices are small, and even close to zero, when the predefined tolerance angles (0.8° and 0.2°) are used in Book Cliffs model I and II.

4

Reservoir Lithology Classification based on Seismic Inversion Results by Hidden Markov Models: Applying Prior Geological Information

Hidden Markov Models (HMMs) have been applied to predict reservoir lithologies using seismic inversion results as inputs. This approach takes into account the conditional probabilities between different lithologies, i.e. the vertical transitions in sedimentary sequences. These properties are used as prior geological information. In order to relate the seismic inversion results to the true well-log data, HMMs need to be trained based on the Expectation-Maximization theory. The application of the resulting model on a synthetic example from the Book Cliffs (Utah, USA) showed that most lithologies are classified correctly, even for some thin layers. A comparison with other point-wise methods in which data samples are treated independently from each other, such as k-Means and Fuzzy Logic classifiers, leads to the conclusion that the spatial correlation in HMMs allows better lithological predictions, because this prior information reflects the geological depositional processes. A real case study with data from the Vienna Basin (Austria) is performed, in which lithologies in a 3-D cube are obtained based on the seismic inversion results, via trained HMMs. While the vertical sequences are shown to be reasonably well predicted, the horizontal continuities are not. This indicates that a future research should focus on the lateral geological relationships.

4.1. Introduction

Lithology classification is one of the most important aspects in reservoir characterization, because it is the key to translation of rock-properties to relevant reservoir parameters. The distribution of reservoir lithologies, therefore, is needed to understand and predict a reservoir's production performance, through reservoir modeling and simulation. In order to separate lithologies into different classes, data obtained from cores and well-logs are usually used, which have a high vertical resolution. While these provide sufficient information on subsurface lithologies in a one-dimensional direction (depth), they provide little information about the lateral distribution of them. Additionally, in most cases, the density of boreholes is relatively sparse compared to the total reservoir volume. Seismic data, on the other hand, can provide two-dimensional or three-dimensional information over an area typically covering the extent of the target reservoir. It can, therefore, be used to obtain two-dimensional or three-dimensional models of the relevant reservoir properties. Here we present the results from an effort to extract reservoir lithologies, from the properties provided by full-elastic wave-equation based inversion of seismic data.

In general, full-waveform inversion, or FWI, is a challenging data-fitting procedure based on full-wavefield modelling to extract quantitative information from seismograms (Tarantola 1984; Virieux and Operto 2009). The information from different types of waves including refractions and diving waves, which used to be considered as noises in normal seismic data processing, are utilized in this optimization procedure. The goal is to obtain better seismic velocity models for seismic migration.

FWI became only feasible after a significant increase in computing power and longer offset seismic acquisition methods became available, because it is computationally intensive and sensitive to the structure of the intermediate/long wavelengths that could be obtained from long-offset data (Mora 1987). Pratt *et al*. (1996) used wide-angle seismic data to reconstruct a higher-resolution two-dimensional velocity model compared to pre-stack migration and traveltime tomography. Shipp and Singh (2002) used a two-dimensional elastic wave equation to generate all possible waves, including converted ones, in order to simulate a complex seismic wavefield in a marine environment. Plessix (2009) implemented a three-dimensional frequency-domain full-waveform inversion in which a multiscale approach with an iterative solver is adopted.

In contrast with the methods mentioned above, the scheme of wave-equation based inversion used here is based on the integral representation of the full-elastic wave equation. The orders of multiple scattering that are accounted for in the inversion are determined by the number of iterations. All internal transmission effects and internal multiple scattering/mode-conversion are considered, allowing recovery of broadband properties and providing a high resolution (Gisolf and van den Berg 2010a; 2010b). This makes the inversion results suitable for lithology determination. In the synthetic test presented here, PP and PS data are used as inputs together and the elastic parameters, such as compressibility, bulk density and shear compliance, are simultaneously inverted for. Since only PP data is available in the real case

study discussed later on, only compressibility and shear compliance are the resulting inversion outputs.

When there are several input properties, lithology determination is a multivariate problem. Hierarchical agglomerative clustering algorithms can be used to analyze the dataset, in which the procedure is iterated, based on previously established clusters, until a desired number of lithologies is obtained (Lindberg and Grana 2015). Typical examples of such algorithms include discriminant analysis, *k*-Means clustering, and others (Hastie *et al*. 2009). Intelligent systems such as artificial neural networks (ANN) and genetic algorithms have been proposed by Oldenziel *et al*. (2000) and Fang and Yang (2015). In order to take advantage of prior information, Bayes' theorem is used to calculate the posterior probabilities. A comparison between Bayes' classifier and discriminant analysis is given by Li (2006). Dubois *et al*. (2007) tested four different approaches including Bayes' classifier, Fuzzy Logic, *k*-nearest neighbor and ANN in the Panoma gas field, whereby the ANN outperformed the others. During the classification process, rock physical models resulting from previous experiences or lab experiments can be integrated into the Monte Carlo classification in a probabilistic manner to mitigate the posterior uncertainty (Grana *et al*. 2012). Seismic facies have also been classified by similar methods, in which three-dimensional models can be obtained based on seismic attributes and the lithologies from well logs at the borehole locations (Avseth *et al*. 2005).

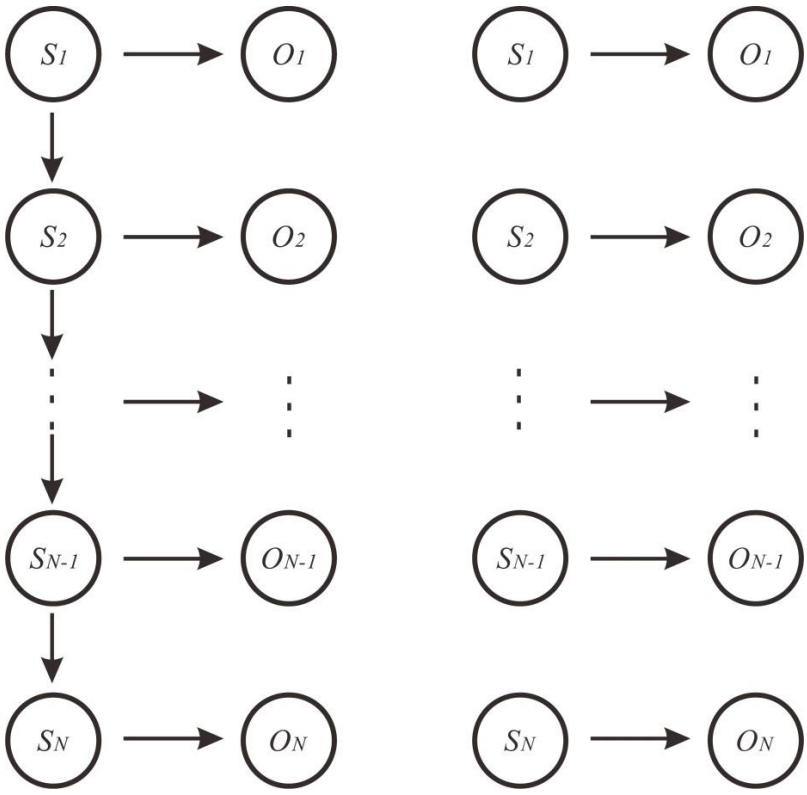
The methods mentioned above treat all data samples as independent from each other, i.e. without any typical vertical patterns as produced by sedimentary processes. Therefore, these can be described as point-wise approaches (Hammer *et al*. 2012) and some geologically, or petrophysically, unlikely transitions between lithologies can be obtained. In order to avoid this problem, a Markov prior model was applied in a Bayesian framework by Larsen *et al*. (2006). Hammer *et al*. (2012) adopted a more realistic Markov prior model to predict the lithologies and fluid saturations, than Kjønsberg *et al*. (2010), by using pre-stack seismic data without increasing the computational complexity.

In the present chapter, a new approach called Hidden Markov Models (HMMs) is introduced to lithologies classification (Rabiner 1989). Dymarski (2011) provided an overview of HMMs and their applications in speech recognition, computer science and other fields. Eidsvik *et al*. (2002; 2004) applied HMMs to translate well-log data into geological attributes. The Expectation-Maximization (EM) algorithm has been used by Lindberg and Grana (2015) to infer the parameters of HMMs and implement well-log inversion later. Here, instead of using the well-log data, the seismic inversion results are the inputs into the classification process, in which the prediction can be made without constraints by the well location. However, the parameters of HMMs need to be obtained first by training with well logs and seismic inversion results at the same location, in order to minimize the inversion error or noise.

In the following, the theory of HMMs is first outlined. A training procedure is demonstrated and performed to obtain suitable parameters. Lithology predictions with seismic inversion results as inputs are shown in the following, first on a synthetic case, then on a real data set. The final conclusions include discussions and future plans.

4.2. Methodology

The Hidden Markov Models (HMMs) is a sequence classifier with the purpose of allocating a state or lithology to every unit, or data sample, following a sequence rule, or a Markov process (Figure 4.1). The states, or lithologies, themselves are unobservable, or hidden, but indirect observations, or measurements, are available that are related to the states. For example, lithologies in the subsurface cannot be measured, but rock properties such as velocity, density etc., which depend on the lithologies, can be measured in wells. HMMs, can predict lithologies from observable measurements. Compared with other classification methods, the biggest advantage in HMMs is that the Markov prior knowledge is honored that excludes unlikely transitions such as a water-bearing sand on top of an oil-bearing sand (Figure 4.1).



(a) (b)

Figure 4.1: Comparison between HMMs and other point-wise methods. a) With vertical coupling. b) Without vertical connection. $S_1,\ S_2,\ \cdots,\ S_N$ are hidden variables while $O_1,\ O_2,\ \cdots, O_N$ are observations (modified from Lindberg and Grana 2015).

In HMMs, the states, or lithologies that we are trying to infer, follow a Markov process. It is a stochastic process that has a first-order Markov property: the future only depends on the present, not on the past. Following Elfeki and Dekking (2001), let $Z_1, Z_2, \cdots, Z_N$ be a sequence of random variables, which take values in the state space $S = \{S_1,\ S_2,\ \cdots,\ S_N\}$. The first-order Markov process or property then satisfies:

$$\Pr(Z_i = S_k | Z_{i-1} = S_l, Z_{i-2} = S_m, \cdots, Z_1 = S_r)$$

$$= \Pr(Z_i = S_k | Z_{i-1} = S_l) = p_{lk} \tag{4.1}$$

where p_{lk} is the conditional probability of S_k given S_l. In one dimension, the conditional probability can be described by a transition matrix which can be written in a square form:

$$P = \begin{bmatrix} p_{11} & \cdots & p_{1N} \\ \vdots & p_{ij} & \vdots \\ p_{N1} & \cdots & p_{NN} \end{bmatrix} \tag{4.2}$$

in which p_{ij} specifies the transition probability from state i to state j.

By taking a limit in the step of transitions (n) (Elfeki and Dekking 2001),

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j \tag{4.3}$$

a stationary probability of different states can be obtained: $\pi_j (j = 1,\ 2, \cdots,\ N)$, which is assumed to be the initial distributions of states and one of the inputs for HMMs. i disappears because the stationary probability does not depend on the starting states.

Given a state S_k at depth d, the probability density function of observation data at the same depth (O_d) is

$$\Pr(O_d | Z_d = S_k) = b_k(d) \tag{4.4}$$

Equation (4.4) reflects the dependency relationship in Figure 4.1, indicated by the horizontal arrows, and is called the likelihood function in general. The observations are continuous variables with Gaussian distributions. $B = \{b_k(d)\}$ denotes the emission probabilities of the observation given all the different states (Rabiner 1989).

Thus, a complete specification of HMMs parameters for a given model can be described by (Rabiner 1989; Lindberg and Grana 2015):

$$\boldsymbol{\lambda} = (P, B, \pi) \tag{4.5}$$

where P is the transition probability (equations (4.1) and (4.2)), B represents the probability distribution of observational data, while π is the stationary probability, which is the initial state distribution.

Three basic problems can be solved by HMMs:

1. Given the model $\boldsymbol{\lambda}$ and the observation data down to a certain depth $\mathbf{O} = O_1 O_2 \cdots O_d$, how to calculate $\Pr(\mathbf{O}|\boldsymbol{\lambda})$;

2. How to adjust the model parameters $\boldsymbol{\lambda} = (P, B, \pi)$ in order to maximize $\Pr(\mathbf{O}|\boldsymbol{\lambda})$;

3. Given the observation data $\mathbf{O} = O_1 O_2 \cdots O_d$ and the model $\boldsymbol{\lambda}$, how to choose a state sequence $\mathbf{Z} = Z_1 Z_2 \cdots Z_d$ that can explain the observations optimally.

Different methods can be invoked to solve these problems, and details can be found in Rabiner (1989).

In order to apply HMMs to classify lithologies by using seismic inversion results as inputs, first the model parameters $\boldsymbol{\lambda}$ need to be estimated from the log data at the well location. Then this model will be trained based on the inversion results, i.e. the parameters are optimized to best describe the observation data. This training step is crucial, because the inversion results will not always match the log readings and this procedure can take care of the deviations between them. It is similar to Problem 2, to which Problem 1 is related, because the probability of observation data given the model will be computed $\Pr(\mathbf{O}|\boldsymbol{\lambda})$. Then the trained model can be used to predict the lithologies at other locations, with inversion as the inputs, which is Problem 3. This is the step in which we are trying to uncover the hidden, or latent part of the model. There are no "correct" predictions since they are unknown. Different criteria could be introduced to solve this problem to obtain the most probable sequence of states (Rabiner 1989).

4.3. Synthetic Examples

The first application example is the synthetic Book Cliffs model created by Feng et al. (2017) (Chapter 2), in which more details have been added and more differentiation is put on the potential reservoir lithologies than in the original model. Figure 4.2 shows rock physical properties which can be calculated based on the well-logging data and lithologies at one single CMP (1130) that are used as inputs for the templates of lithologies and initial distributions of the likelihood functions in HMMs. Rock properties in terms of compressibility ($\kappa = 1/K$, with K being the bulk modulus) and shear compliance ($M = 1/\mu$, with μ being the shear modulus) are used for the lithology classification and will be obtained from seismic inversion.

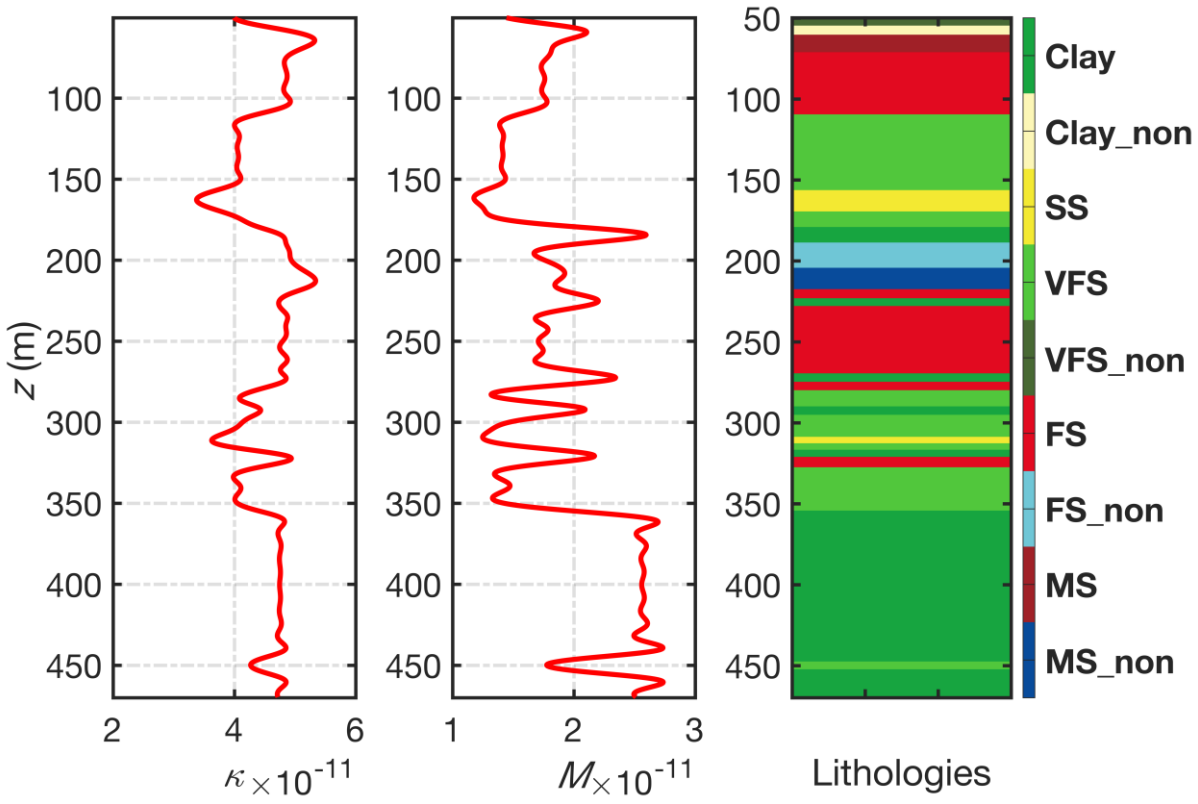


Figure 4.2: Rock physical properties and lithologies at one single CMP. MS−Medium-grained sandstones; FS−Fine-grained sandstones; VFS−Very fine-grained sandstones; SS−Siltstones, and each lithology has been divided into two parts−marine and non-marine (designated by non).

Inputs for HMMs (equation (4.5)) can be obtained based on the rock physical data and the known lithologies in the well. The transition matrix is derived by counting transitions between different lithologies and then normalizing by the row summations (equation (4.6)). Zero entries in the matrices have been set to very small values (0.0001) in order to avoid abrupt changes or absorbing lithologies (Lindberg and Grana 2015).

$$P = \begin{array}{c} \\ \text{MS\_non} \\ \text{MS} \\ \text{FS\_non} \\ \text{FS} \\ \text{VFS\_non} \\ \text{VFS} \\ \text{SS} \\ \text{Clay\_non} \\ \text{Clay} \end{array} \begin{array}{c} \begin{array}{ccccccccc} \text{MS\_non} & \text{MS} & \text{FS\_non} & \text{FS} & \text{VFS\_non} & \text{VFS} & \text{SS} & \text{Clay\_non} & \text{Clay} \end{array} \\ \begin{bmatrix} 0.6660 & 0.0001 & 0.0001 & 0.3333 & 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.0001 \\ 0.0001 & 0.6660 & 0.0001 & 0.3333 & 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.0001 \\ 0.3333 & 0.0001 & 0.6660 & 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.0001 \\ 0.0001 & 0.0001 & 0.0001 & 0.7216 & 0.0001 & 0.1667 & 0.0001 & 0.0001 & 0.1111 \\ 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.9992 & 0.0001 \\ 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.7077 & 0.0833 & 0.0001 & 0.2084 \\ 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.6660 & 0.3333 & 0.0001 & 0.0001 \\ 0.0001 & 0.9992 & 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.0001 & 0.0001 \\ 0.0001 & 0.0001 & 0.0370 & 0.1111 & 0.0001 & 0.0741 & 0.0001 & 0.0001 & 0.7773 \end{bmatrix} \end{array} \tag{4.6}$$

By applying equation (4.3), the initial distributions of different lithologies at this location can be calculated (equation (4.7))

$$\pi = \begin{bmatrix} \text{MS\_non} & 0.0423 \\ \text{MS} & 0.0009 \\ \text{FS\_non} & 0.0421 \\ \text{FS} & 0.2028 \\ \text{VFS\_non} & 0.0001 \\ \text{VFS} & 0.2962 \\ \text{SS} & 0.0371 \\ \text{Clay\_non} & 0.0002 \\ \text{Clay} & 0.3783 \end{bmatrix} \tag{4.7}$$

The last input is the emission probabilities where the assumption of bivariate Gaussian distributions is made for each lithology.

The parameter sets of HMMs can be trained, or updated, in order to account for the errors in the inversion results by an iterative Baum-Welch method (or equivalently the EM expectation-maximization method used by Rabiner (1989)). Here only the means and covariance matrices in the Gaussian likelihood functions are updated, because not all properties can be resolved fully by the seismic inversion (Figure 4.3). This is a partially supervised learning process. The transition matrices and initial distributions of different lithologies will be fixed during the training and classification process since they are not influenced by the inversion.

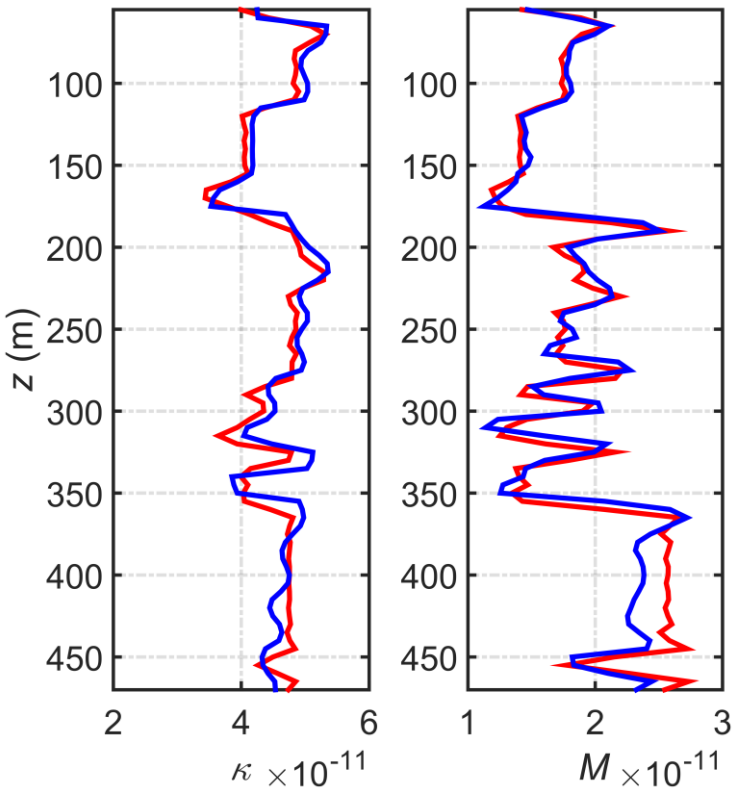
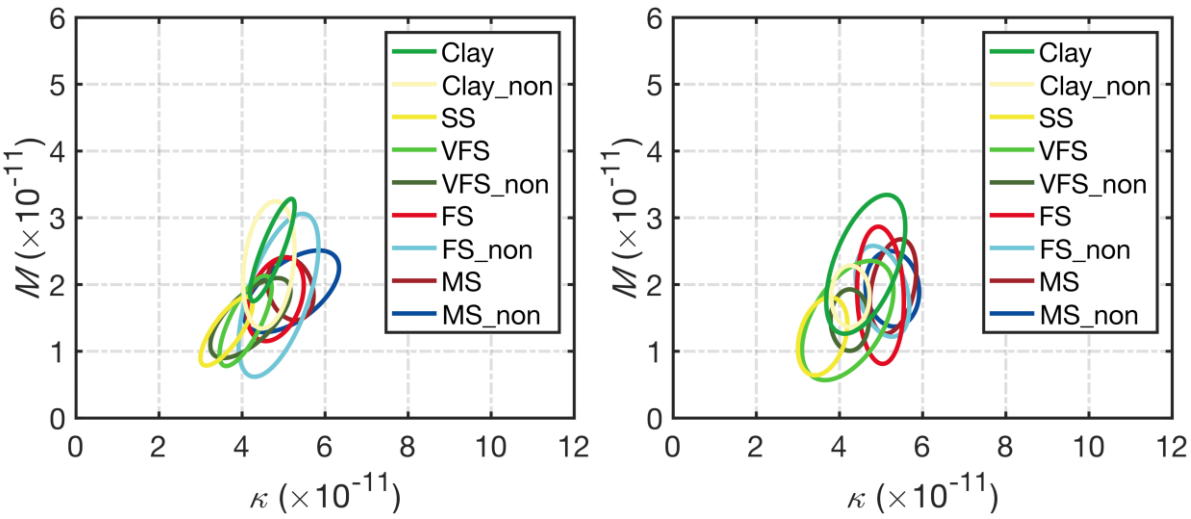


Figure 4.3: Truth (red curves) and inversion (blue curves). Rock properties have been resampled to the seismic grid size.

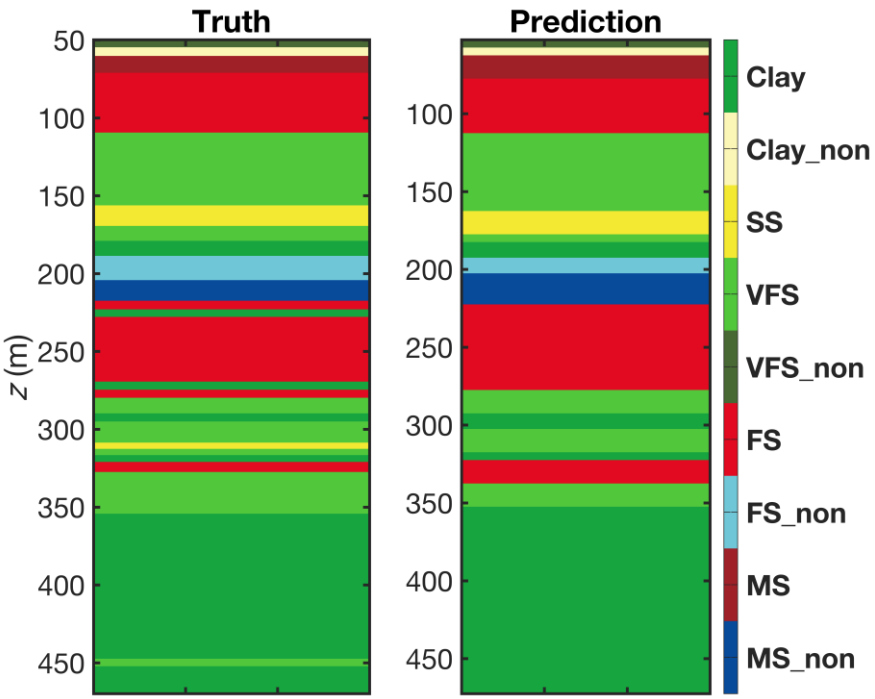
Figure 4.4 shows the 90% confidence regions of the Bivariate Gaussian likelihood model (κ, M) for the probability of each lithology, in which the left one is based on the well data and the right one has been updated according to the inversion results at the same location. It can be noticed that the shapes of the confidence regions have changed and some positive correlations between the variables become negative after training, such as FS_non which is reflected as a change in the shape direction, due to errors in κ and M of the inversion results (Figure 4.3).



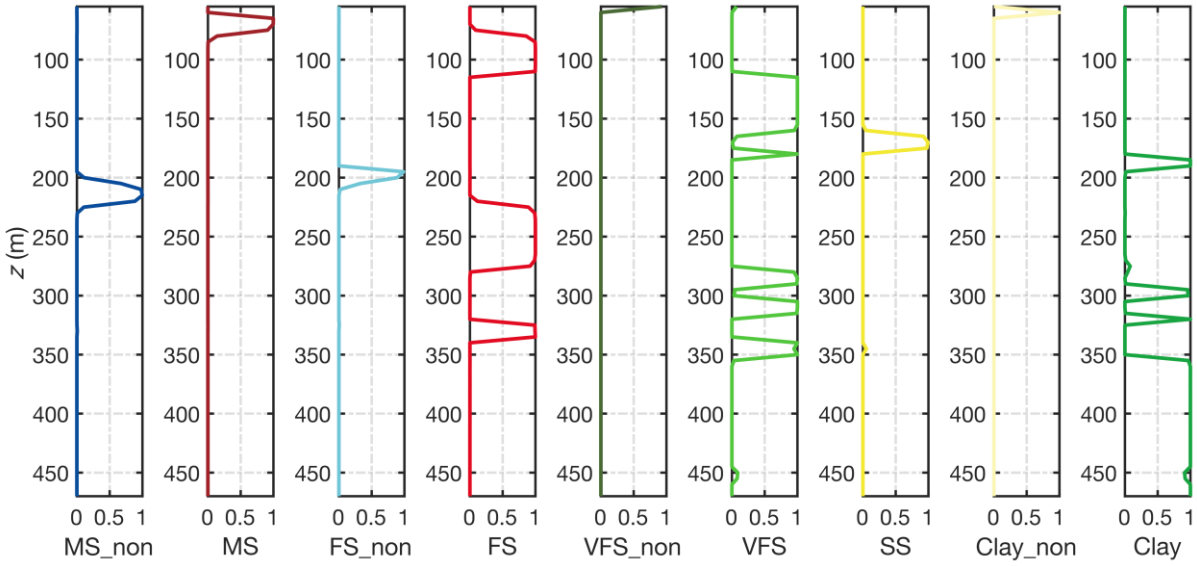
(a) (b)

Figure 4.4: 90% confidence regions of the Bivariate Gaussian likelihood model for the distribution of each lithology.

In order to test HMMs before application, the inversion results at the training location are used as input to predict the lithologies. The true and predicted results are shown in Figure 4.5, together with the Maximum *a posteriori* probabilities (MAP) for each state at depth.



(a)



(b)

Figure 4.5: a) True and predicted lithologies. b) Maximum *a posteriori* probabilities.

The most important observation is that the potential reservoir-quality lithologies (MS, FS whether marine or non-marine parts) are well predicted. Some thin layers such as Clay_non at the top are also predicted correctly. However, there are some thin layers missing. For example, Clay at depths 225m and 275m, and SS at 320m, which can be attributed to the difference between the true and inverted rock properties (Figure 4.3). The fact that VFS at depth 455m cannot be detected, is partly attributed to the lower transition probability from Clay to VFS as well as a small mismatch between truth and inversion results.

The trained means and covariances of the Gaussian likelihood model and initial distributions and transition matrices of different lithologies in HMMs can be used to make predictions at other locations where the inputs are limited to the inversion results only.

The truth and inversion properties at another CMP (1140) are shown in Figure 4.6. The predicted lithologies and MAP probabilities can be seen in Figure 4.7.

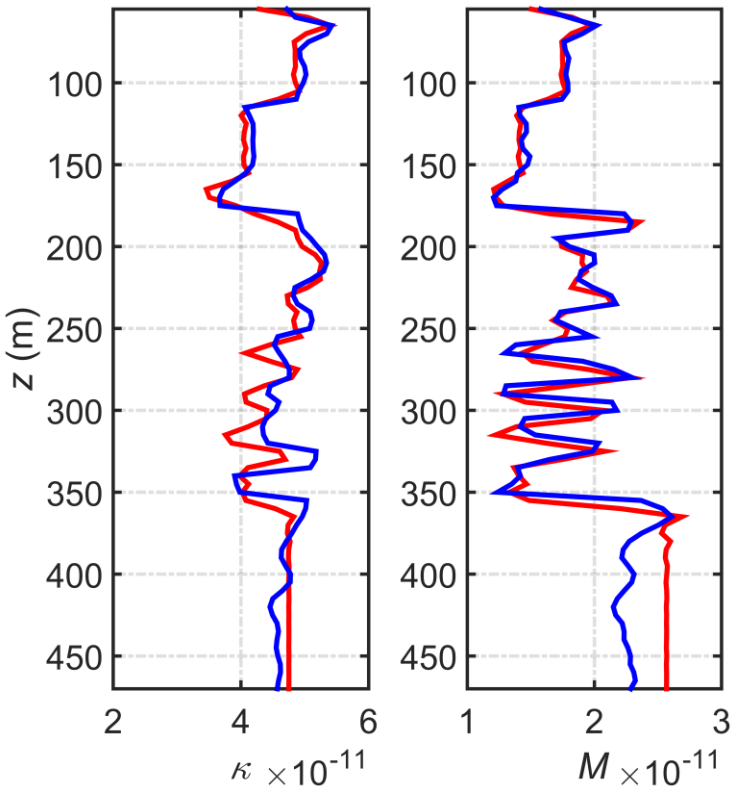
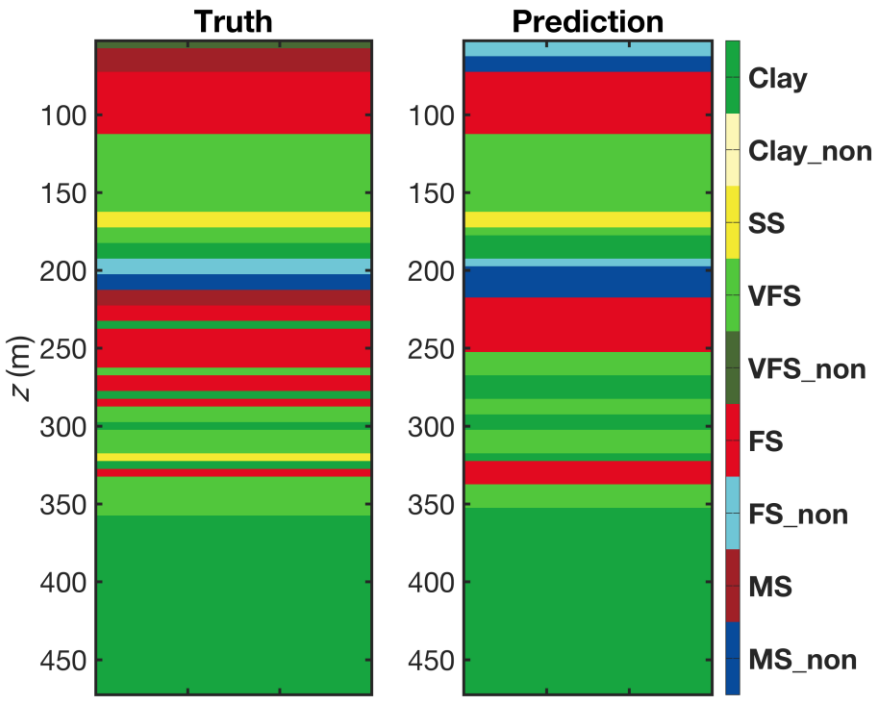
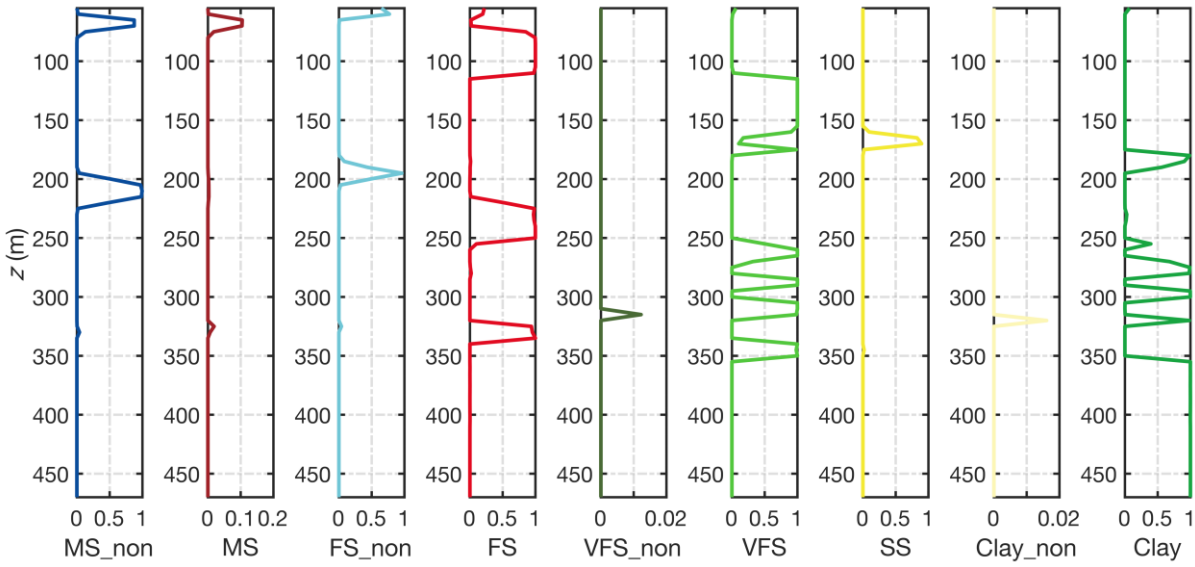


Figure 4.6: Truth (red curves) and inversion (blue curves) at another CMP (1140).



(a)



(b)

Figure 4.7: a) True and predicted lithologies. b) Maximum *a Posteriori* probabilities.

It can be see that the predicted results are quite good, because almost all lithologies have been predicted correctly, as well as the thin layers. There are some misclassified lithologies such as MS being predicted as MS_non at the top, which is due to the similarity of their properties.

In order to compare HMMs with other point-wise methods such as *k*-Means (Seber 1984) and Fuzzy Logic (Saggaf and Nebrija 2003; Feng *et al*. 2016), the results from these methods at another CMP (Figure 4.6) are shown in Figure 4.8.

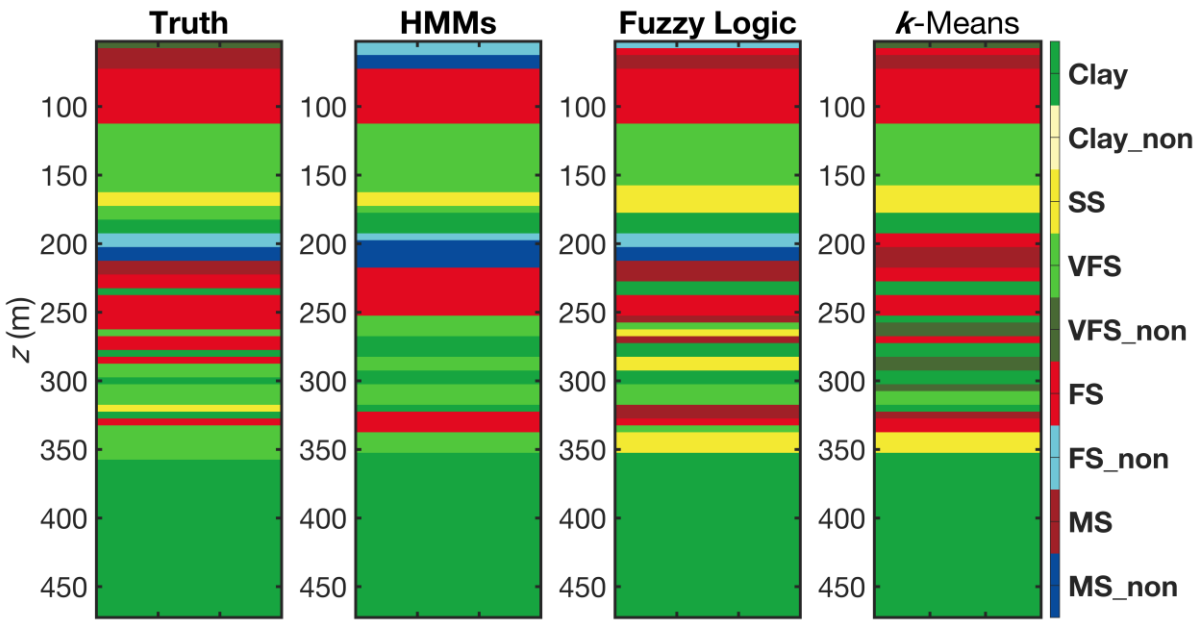


Figure 4.8: Comparison between HMMs, Fuzzy Logic and *k*-Means results. HMMs: *MCC* = 0.6989; Fuzzy Logic: *MCC* = 0.6887; k-Means: *MCC* = 0.6223. *MCC* is calculated according to equation (3.3).

Without taking into account of spatial coupling, the results of *k*-Means and Fuzzy Logic tend to be more random, especially in the middle part where they predict more thin layers than the HMMs. *MCC* is also showing that the result of HMMs is better than the other two.

4.4. Field Case Study

In this section, the methodology of HMMs is applied to a dataset from a clastic reservoir in the Vienna Basin which is an extensional basin between the Eastern Alps and the Western Carpathians (Strauss *et al*. 2006). Vintages of 3-D seismic surveys acquired in different years have been merged into a single dataset VBSM (Vienna Basin Super Merge). These pre-stack seismic data are used as input for the non-linear full-waveform inversion scheme (Gisolf *et al*. 2014). Several wells had been drilled in this area reaching the hydrocarbon zone, with log-suites including gamma ray, density, resistivity, neutron porosity, compressional and shear velocities etc. which are the basis of the standard lithological determinations made in petrophysical evaluation. However, this approach is limited to an application at the well locations. Therefore, the inverted rock properties from seismic data are proposed as inputs for the lithology discrimination.

The property values κ and M between the depths of 1700 m and 2200 m have been selected in a well together with three known lithologies (shale, shaly sandstone and sandstone) in the same interval (Figure 4.9).

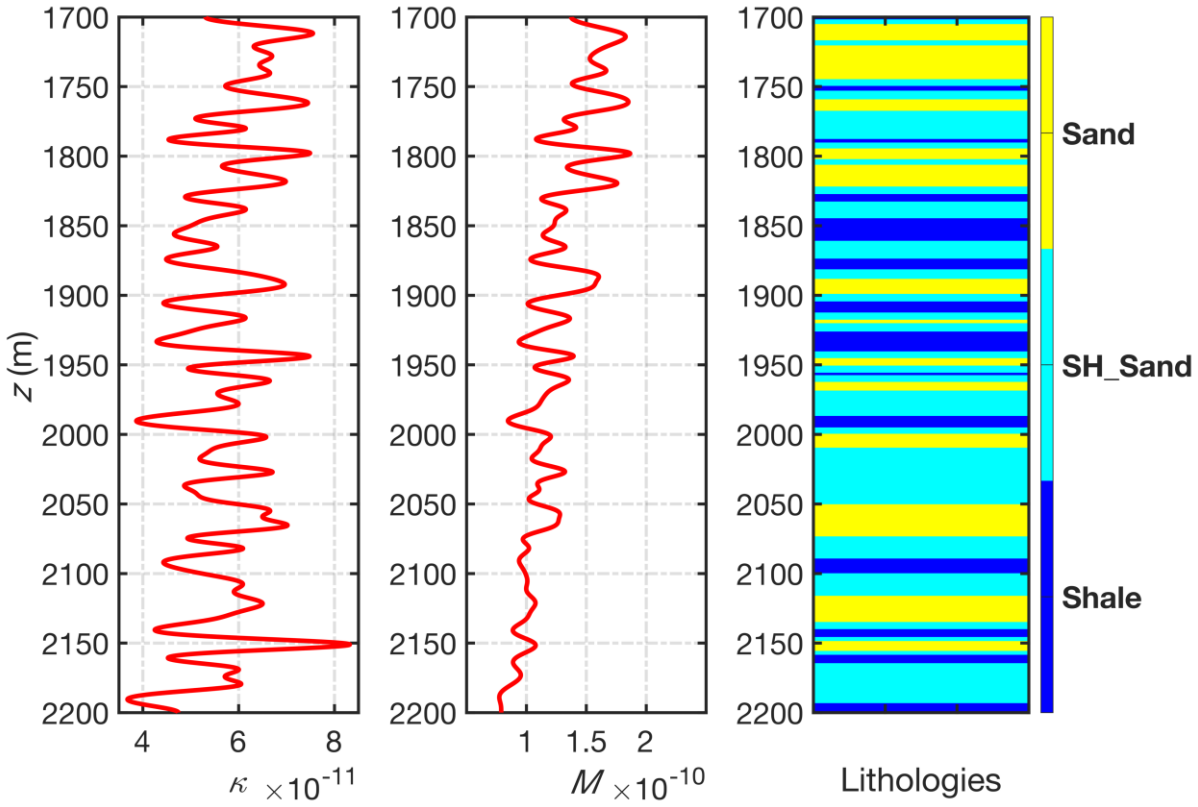


Figure 4.9: Well logs expressed as κ and M, together with predefined lithologies. All values have been projected onto the vertical direction.

With the property values and lithologies in this well, the input parameters of HMMs can be obtained (equations (4.8) and (4.9) as well as Figure 4.10).

$$P = \begin{array}{c} \\ \text{Shale} \\ \text{SH\_sand} \\ \text{Sand} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Shale} & \text{SH\_Sand} & \text{Sand} \end{array} \\ \begin{bmatrix} 0.9805 & 0.0194 & 0.0001 \\ 0.0081 & 0.9838 & 0.0081 \\ 0.0001 & 0.0123 & 0.9876 \end{bmatrix} \end{array} \tag{4.8}$$

$$\pi = \begin{array}{c} \\ \text{Shale} \\ \text{SH\_sand} \\ \text{Sand} \end{array} \begin{bmatrix} 0.2013 \\ 0.4819 \\ 0.3168 \end{bmatrix} \tag{4.9}$$

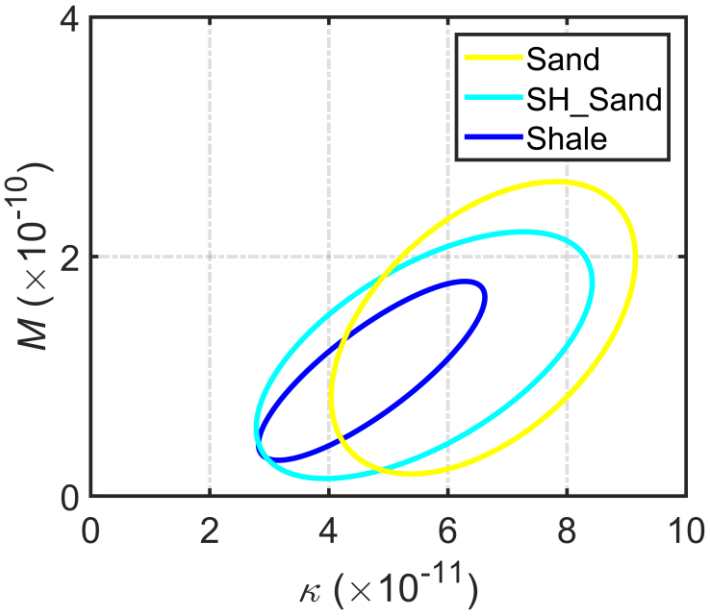


Figure 4.10: 90% confidence regions of the Bivariate Gaussian likelihood model of shale, shaly sandstone and sandstone in the well.

In equation (4.8), the zero transitional probabilities have been set to 0.0001, because all the transitions are theoretically possible. The confidence regions in Figure 4.10 indicate large overlaps between different lithologies, which means it will be difficult to distinguish the lithologies if only these properties are considered. In order to set a benchmark for the classification, property values from the well logs are used as inputs for HMMs, as well as for *k*-Means and Fuzzy Logic methods (Figure 4.11). In this case, all of them are supervised learning processes.

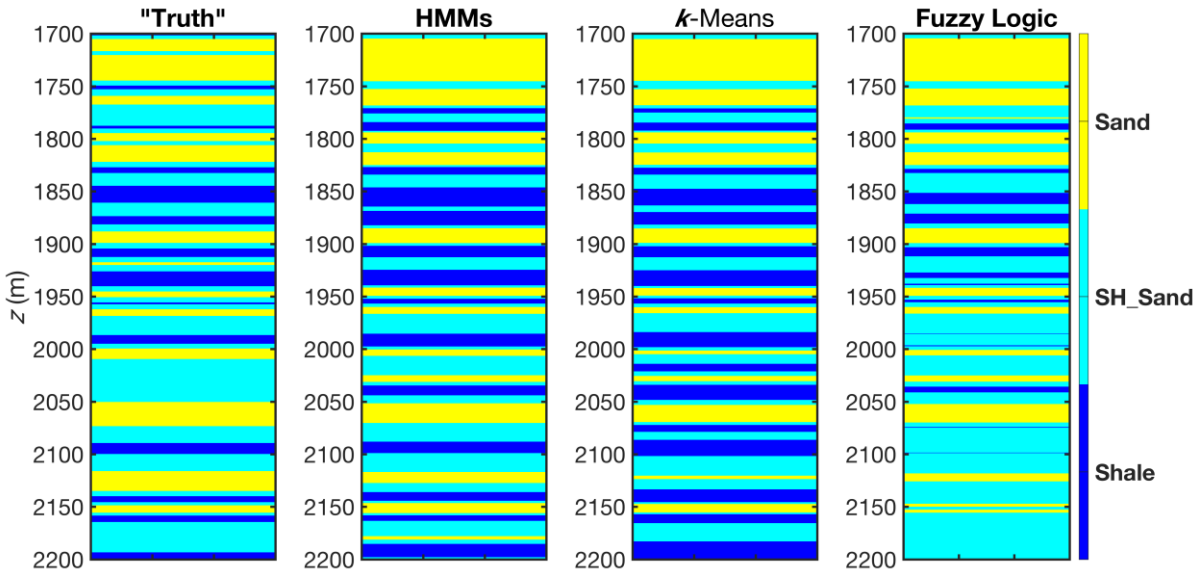


Figure 4.11: Predicted results by three different methods. HMMs: *MCC* = 0.5844; *k*-Means: *MCC* = 0.5202; Fuzzy Logic: *MCC* = 0.4911. *MCC* is calculated according to equation (3.3).

It can be seen that compared with the other two statistical methods, the result of HMMs is better (*MCC*), because almost all lithologies have been classified correctly, especially in the middle part. Without the geological prior information, which is the Markov sequence, the statistical or histogram-based approaches (*k*-Means and Fuzzy Logic) are unable to incorporate the internal transitional trends between different states, transitions that are the results of specific sedimentary processes.

The ultimate goal here is to use seismic inversion results as inputs for the lithology classification. The inverted and true properties at same well location are shown in Figure 4.12 where it can be seen that inversion quality is not very good, probably attributable to some pre-processing problems of the pre-stack seismic data. Unlike in the synthetic case, in which PP and PS data are used for the inversion (Figures 4.3 and 4.6), the quality of the inverted M is worse than the one of κ which is due to the fact that only PP data are available here.

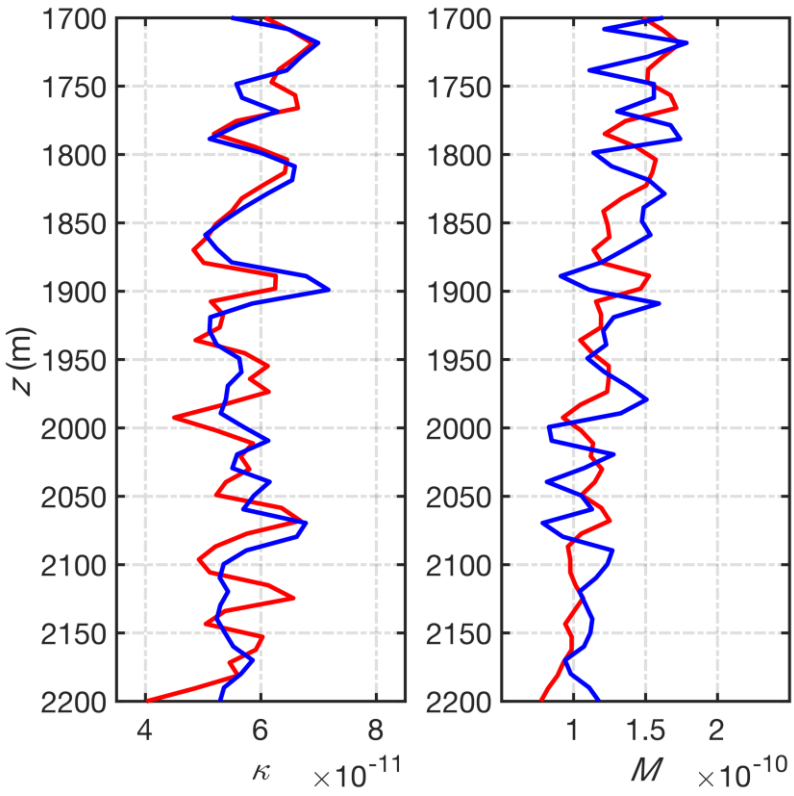


Figure 4.12: True (red curves) and inverted (blue curves) rock properties. Values in the well have been projected onto the vertical direction.

Only the means in the Gaussian likelihood functions will be updated in the training process, which is different from the process applied in the synthetic case, because the quality of the inversion here is much worse. After training, the 90% confidence regions for each lithology are displayed in Figure 4.13 and the predicted lithologies with the inversion as inputs are shown in Figure 4.14. Because the inversion does not match the truth very well, perfectly predicted lithologies cannot be expected, although the prediction is not bad either, especially in the upper part where the misfit is relatively small.

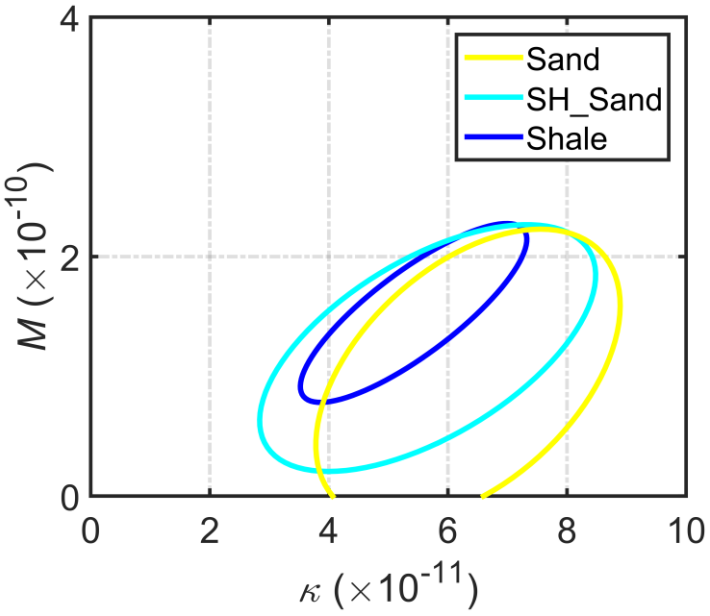


Figure 4.13: 90% confidence regions of the Bivariate Gaussian likelihood model based on the inversion.

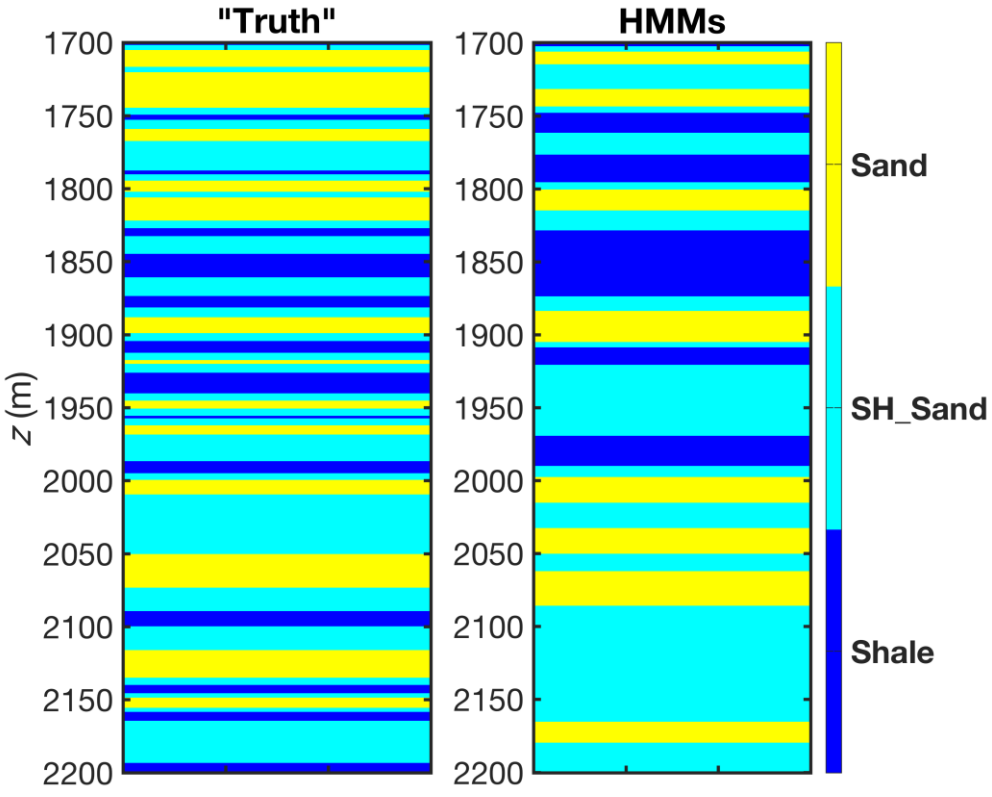


Figure 4.14: Predicted lithologies by HMMs based on the seismic inversion at the training well location.

In the next step, the inversion results at another well location are chosen as a blind test, in which the log properties have not been used for training beforehand (Figure 4.15).

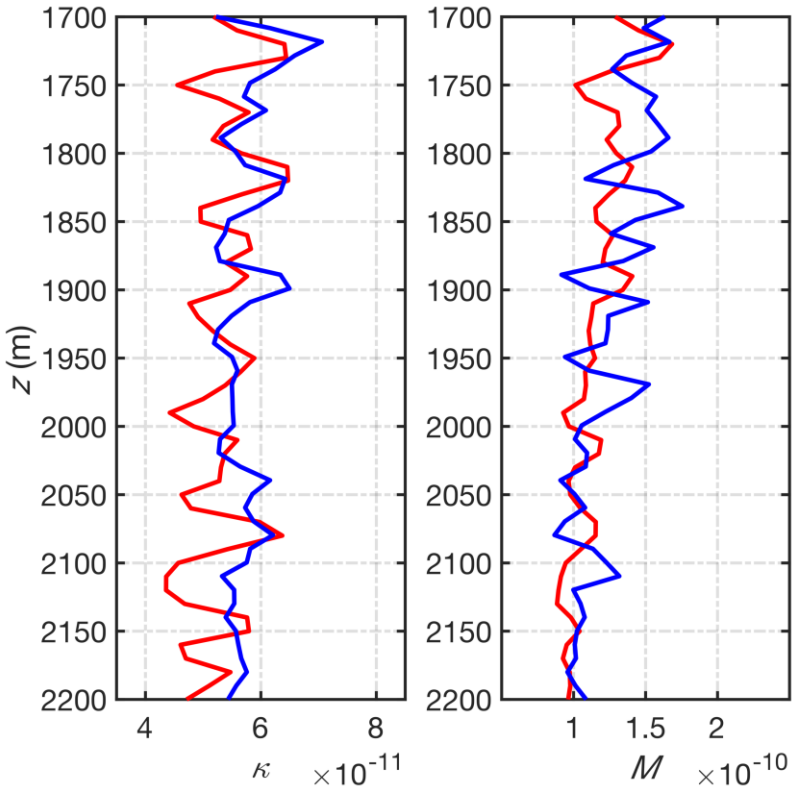


Figure 4.15: True (red curves) and inverted (blue curves) rock properties at a second well location used for a blind test.

As expected, the inversion results are worse because of the previously mentioned problems. The true lithologies and the predicted lithologies with the seismic inversion results as inputs, are shown in Figure 4.16. Some sandstone units have been successfully predicted, but many thin shale streaks have been incorrectly classified as shaly sandstone in the lower part (1900 m – 2200 m).

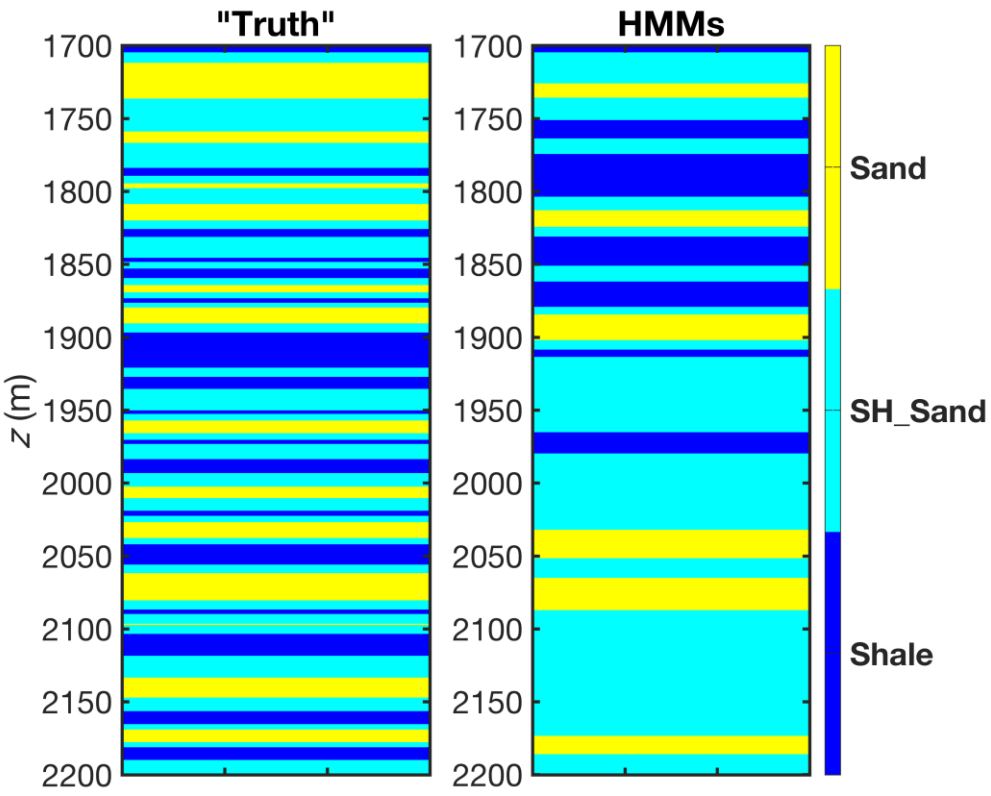
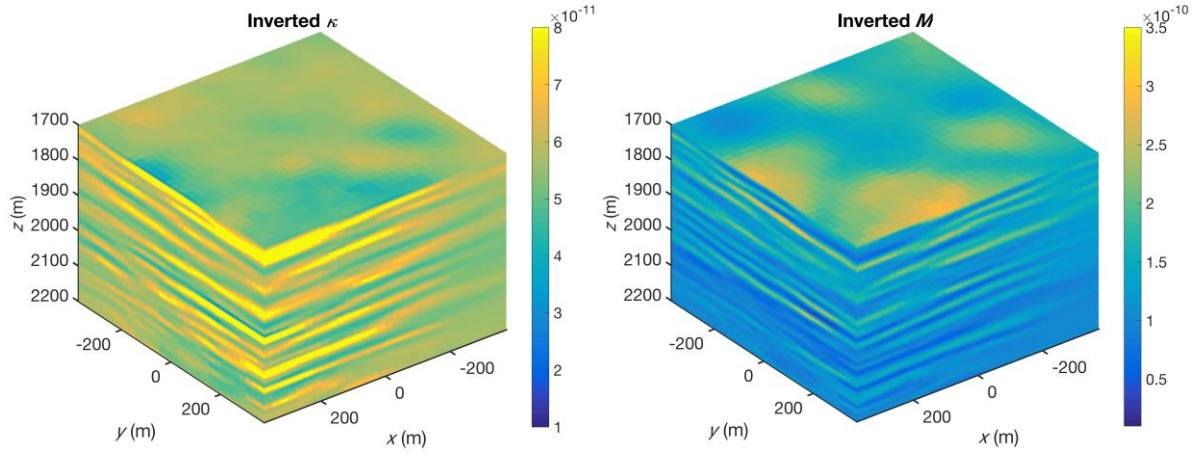


Figure 4.16: Predicted lithologies by HMMs based on the seismic inversion at a second well location in which the properties have not been used for training.

With the inversion results in a 3-D cube (Figure 4.17) as inputs for lithological classification and predefined parameters such as transitional matrices (equation (4.8)), initial state distributions (equation (4.9)) and means as well as covariance matrices (Figure 4.13) in the Gaussian likelihood model, the predicted lithologies are obtained as shown in Figure 4.18.



(a) (b)

Figure 4.17: Inversion results in a 3-D cube.

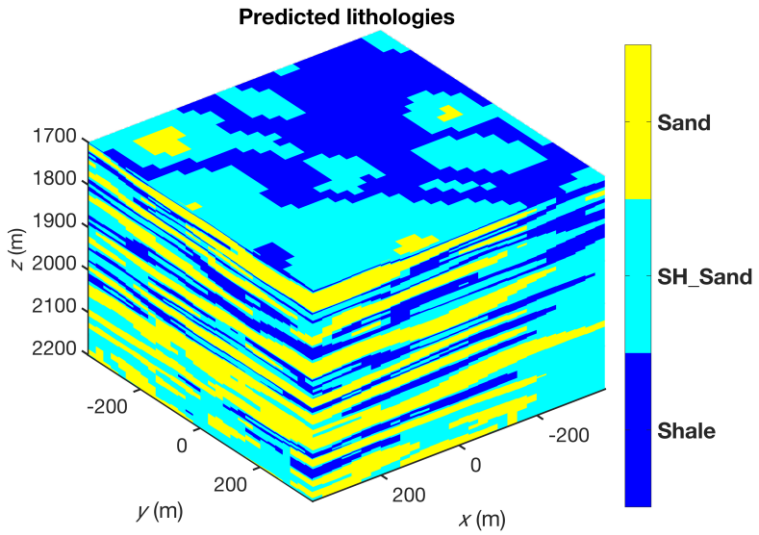
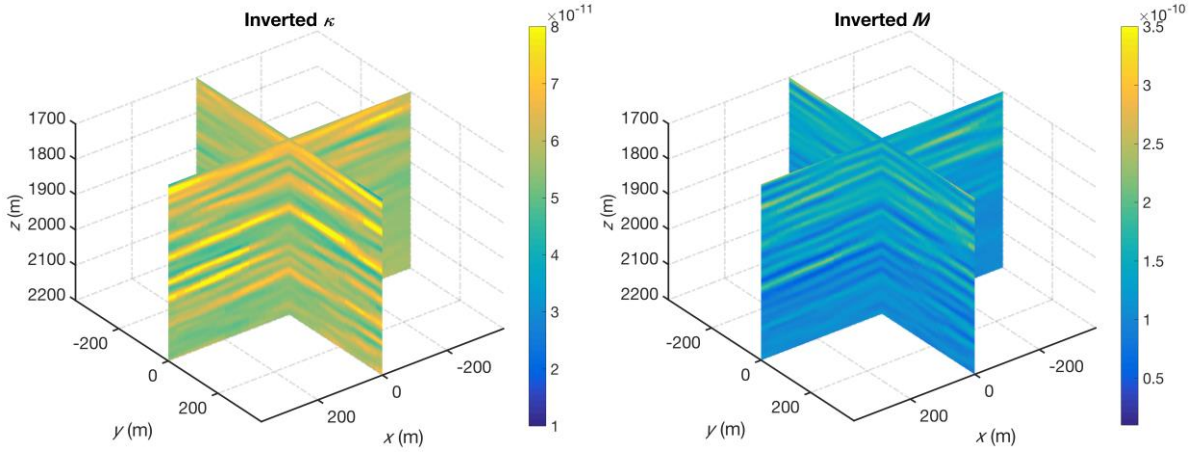


Figure 4.18: Predicted lithologies in a 3-D cube with seismic inversion as inputs shown in Figure 4.17.

The predicted layers are continuous in most parts of the cube. At the top, sand patches are obtained which are also indicated in the inversion results (Figure 4.17). In order to inspect the internal details, two slices along the inline and crossline directions are selected (Figures 4.19 and 4.20). The intersection of the slices is at the (first) training well location.



(a) (b)

Figure 4.19: Two slices of the inversion results along the inline and crossline directions. Their intersection is at the training well location.

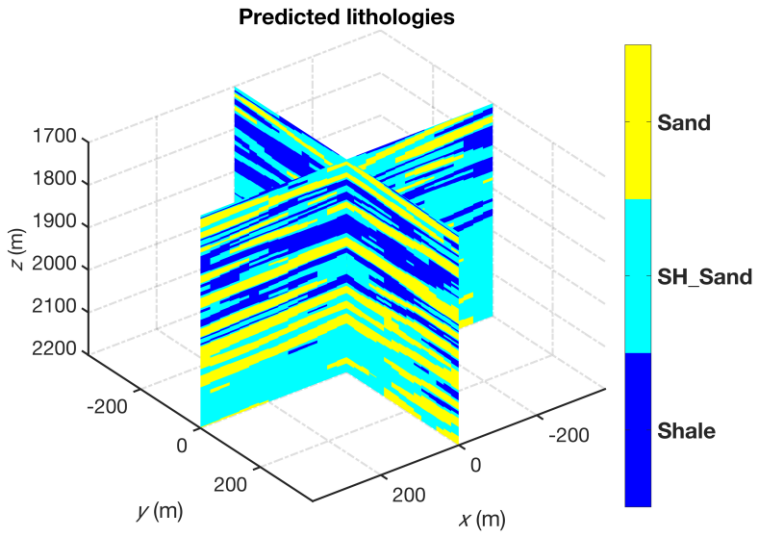


Figure 4.20: Two slices of the predicted lithologies along the inline and crossline directions. Their intersection is at the training well location.

Thus, with the help of the seismic inversion, the limitations of sparse well locations have been overcome and a cube of lithologies was produced by HMMs, albeit with significant uncertainties.

4.5. Discussion and Conclusion

In this chapter, geological prior information in the form of a transition probability matrix for the vertical coupling between different states, as a Markov sequence, has been taken into account in the determination of lithologies by HMMs. Compared with other point-wise classification methods, this model predicts lithologies based on input rock properties, either from wells, or from seismic inversion. The transition probabilities are representative for certain typical sedimentary sequences

In contrast to other methods in which the data from well logs are used, here the results from seismic data are used in order to produce 2-D or 3-D property cubes. In the approaches proposed by Ulvmoen and Omre (2010a; 2010b), only the amplitude variation versus offset (AVO) effect has been considered. In the present chapter and Chapter 2, the output from 1.5-D elastic wave-equation based inversion is used. This type of seismic inversion yields a high resolution properties, because also the internal multiple scattering energy, as well as multiple wave-mode conversions are taken into account.

Instead of inverting for velocities and densities, like other methods do, the scheme adopted here inverts the seismic data for compressibility (κ) and shear compliance (M), properties that are deemed to relate more directly to the rock types. For example, if the value of κ is large and M is small, the type of rock could be a sandstone, because it is more easily compressed due to its often higher porosity, while it is difficult to be sheared due to its strong rigidity. Conversely, if κ is small and M is large, it could be shale due to the lower porosity and weak rigidity. These parameters are therefore assumed to be reasonable reservoir

indicators, especially in time-lapse inversion because of the complementary property behavior (Feng *et al*. 2015).

However, these two properties have different inversion qualities, with κ often being better than M especially when only PP data are available, such as in the field case discussed here (Figures 4.12 and 4.15). Another aspect is the joint influence of the two parameters which can be clearly seen from the top slice of the 3-D cube in Figure 4.18, where some sand patches have been predicted even though the inverted M is large (Figure 4.17b). A likely explanation for this is that the inverted κ is also relatively large (Figure 4.17a). Another output of seismic inversions — bulk density could be introduced as an additional input for the classification since it would be related to lithology strongly, even though it often has a larger uncertainty.

In the synthetic and real examples presented here, two wells have been chosen, in which the first one is used for training, to obtain the means and/or covariances of the Gaussian likelihood model, while the initial distributions and transition matrices or covariances remain fixed. Therefore, this can be considered as a partially supervised learning process. HMMs with the predefined and trained parameters are then used to determine the lithologies at the same location as the well used for training. However, the input datasets for the classification are the inversion results instead of the well-log data. One advantage of using the inversion results from 3-D seismic data is that the result is a 3-D lithology cube (Figures 4.18 and 4.20), compared to having results only at the well locations. A disadvantage is the lower resolution of the seismic inversion compared to the well logs (Figures 4.12 and 4.15).

In the synthetic test, the inversion scheme used achieved a reliable prediction of the rock properties which made the classification of lithologies quite successful (Figure 4.5). However, it is still significant between the properties, especially for the marine and non-marine counterparts as demonstrated in Figure 4.4. These results (Figures 4.5 and 4.7) have less reliable predictions of pairs such as MS and MS_non.

For the real data case study, the seismic-to-well match is poor, either due to some problems in the pre-processing phase, or due to fault cut-out in the synthetic, or both. Therefore, the inversion result is not as good as hoped for, even though in the upper part of the well the match is quite reasonable (Figure 4.12). Specifically, quite a few sandstone layers have been predicted correctly even though there is a large overlap between the properties of the three lithologies (Figures 4.10 and 4.13).

In the cases where the partially trained HMMs are applied to a second well close to the training well location, it becomes a totally supervised learning procedure. The predictions are quite good in the synthetic Book Cliffs example, even though misclassified lithologies occur such as FS_non at the top (Figure 4.7), which is attributed to an overestimation of compressibility (Figure 4.6). In the real study, the prediction at a blind test location is not very good, possibly due to the reasons given above.

In order to compare the performance of HMMs with other point-wise methods such as *k*-Means and Fuzzy Logic inference, the classified results are displayed in Figures 4.8 and 4.11. The *k*-Means clustering approach only uses the input data at the same location and tries to cluster the data points into different groups by minimizing the mean distances and comparing the means with those of HMMs, which can be viewed as a partially supervised learning. In Fuzzy Logic, the membership functions are trained based on the first well and then applied to the second well. Thus it is a supervised learning process. Since HMMs takes lithological transition probabilities into account, it is seen to outperform the other two approaches. Unlike with the inversion as inputs in the Book Cliffs case (Figure 4.8), "true" well-log data of the first well are used for the comparison in order to set a benchmark in the real case example (Figure 4.11).

In all these methods, the number of lithologies to be determined was defined beforehand for the synthetic as well as for the real case study. In other cases, this information may be missing, which could be solved by a subtractive clustering method (Bagheripour and Asoodeh 2013) that can find an optimal number of clusters based on the similarity of the input datasets. However, an expert's intervention is necessary to avoid geologically meaningless clusters.

To conclude, prior information is used in HMMs, which provides a spatial coupling between different data points. Thus, the geological depositional process is implicitly honored in the classification process. The construction of Markov matrices could be based on a geological understanding of the regional settings which should provide likely transitions between different lithologies, or it can be based on a simple scan of the cored wells, if available.

The inversion results from seismic data are used as input, instead of well-log data, which are locally limited and therefore always sparse. Using this approach, a 2-D section or a 3-D cube can be produced. However, in HMMs used here only the vertical coupling is taken into account. Horizontal correlations should be introduced in a next step in order to fully utilize the power of seismic data and make geologically more realistic 2-D or even 3-D volumes. A Markov random field could enforce these dependencies vertically and horizontally (Chapter 5).

References

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

Bagheripour, P. and Asoodeh, M. 2013. Fuzzy ruling between core porosity and petrophysical logs: Subtractive clustering vs. genetic algorithm – pattern search. *Journal of Applied Geophysics*, **99**: 35-41.

Dubois, M.K., Bohling, G.C., Chakrabarti, S. 2007. Comparison of four approaches to a rock facies classification problem. *Computers & Geosciences*, **33**(5), 599-617.

Dymarski, P. 2011. *Hidden Markov models. Theory and applications*. InTech.

Elfeki, A. and Dekking, M. 2001. A Markov chain model for subsurface characterization: theory and applications. *Mathematical Geology*, **33**(5), 569-589.

Eidsvik, J., Mukerji, T. and Switzer, P. 2002. Modeling lithofacies alternations from well logs using Hierarchical Markov Chains. *72nd SEG Meeting*, Salt Lake City, USA, Expanded Abstract.

Eidsvik, J., Mukerji, T. and Switzer, P. 2004. Estimation of geological attributes from a well log: An application of hidden Markov chains. *Mathematical Geology*, **36**(3), 379-397.

Fang, Z. and Yang, D. 2015. Inversion of reservoir porosity, saturation, and permeability based on a robust hybrid genetic algorithm. *Geophysics*, **80**(5), R265-R280.

Feng, R., Luthi, S.M. Gisolf, A. and Sharma, S. 2015. *Non-linear full-waveform inversion (FWI-res) of time-lapse seismic data on a higher-resolution geological and petrophysical model, Book Cliffs (Utah, USA).* 85th SEG Meeting, New Orleans, USA, Expanded Abstracts.

Feng, R. Luthi, S.M. Gisolf, A. and Sharma, S. 2016. *Lithology prediction from the results of full elastic wave-equation based inversion scheme*. 78th EAGE Conference and Exhibition, Vienna, Austria, Expanded Abstracts.

Feng, R., Luthi, S.M. Gisolf, A. and Sharma, S. 2017. Obtaining a high-resolution geological and petrophysical model from the results of reservoir-oriented elastic wave-equation based seismic inversion. *Petroleum Geoscience*, doi: 10.1144/petgeo2015-076.

Gisolf, A. and van den Berg P.M. 2010a. *Target oriented non-linear inversion of seismic data*. 72nd EAGE Meeting, Barcelona, Spain, Expanded Abstracts, P308.

Gisolf, A. and van den Berg P.M. 2010b. Target-oriented non-linear inversion of time-lapse seismic data. 80th SEG Meeting, Denver, USA, Expanded Abstracts.

Gisolf, A., Huis in't Veld R., Haffinger, P., Hanitzsch, C., Doulgeris, P. and Veeken, P.C.H. 2014. *Non-linear full wavefield inversion applied to carboniferous reservoirs in the Wingate Gas Field (SNS, Offshore UK)*. 76th EAGE Meeting, Amsterdam, The Netherlands, Expanded Abstracts, E106.

Grana, D., Pirrone, M. and Mukerji, T. 2012. Quantitative log interpretation and uncertainty propagation of petrophysical properties and facies classification from rock-physics modelling and formation evaluation analysis. *Geophysics*, **77**(3), WA45-WA63.

Hammer, H., Kolbjørnsen, O., Tjelmeland, H. and Buland, A. 2012. Lithology and fluid prediction from prestack seismic data using a Bayesian model with Markov process prior. *Geophysical Prospecting*, **60**: 500-515.

Hastie, T., Tibshirani, R. and Friedman, J. 2009. *The elements of statistical learning*, 2nd edition. Springer.

Kjønsberg, H., Hauge, R., Kolbjørnsen, O. and Buland, A. 2010. Bayesian Monte Carlo method for seismic predrill prospect assessment. *Geophysics*, **75**(2), O9-O19.

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

Li, Y. and Anderson-Sprecher, R. 2006. Facies identification from well logs: a comparison of discriminant analysis and naïve Bayes classifier. *Journal of Petroleum and Engineering*, **53**, 149-157.

Lindberg, D.V. and Grana, D. 2015. Petro-elastic log-facies classification using the expectation-maximization algorithm and hidden Markov models. *Mathematical Geosciences*, **47**, 719-752.

Mora, P.R. 1987. Nonlinear two-dimensional elastic inversion of multi-offset seismic data. *Geophysics*, **52**(9), 1211-1228.

Oldenziel, T., de Groot P. and Kvamme, L.B. 2000. Statfjord study demonstrates use of neural network to predict porosity and water saturation from time-lapse seismic. *First Break*, **18**(2), 65–69.

Plessix, R. 2009. Three-dimensional frequency-domain full-waveform inversion with an iterative solver. *Geophysics*, **74**(6), WCC149-WCC157.

Pratt, R.G., Song, Z.M., Williamson, P. and Warner, M. 1996. Two-dimensional velocity models from wide-angle seismic data by wavefield inversion. *Geophysical Journal International*, **124**, 323-340.

Rabiner, L. 1989. A tutorial on Hidden Markov Models and selected applications in speech recognition. *Proc. IEEE*, **77**(2), 257-286.

Saggaf, M.M. and Nebrija, E.L. 2003. A fuzzy logic approach for the estimation of facies from wire-line logs. *AAPG Bulletin*, **87**(7), 1223-1240.

Seber, G.A.F. 1984. *Multivariate Observations*. Hoboken, NJ: John Wiley & Sons, Inc..

Shipp, R.M. and Singh, S.C. 2002. Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data. *Geophysical Journal International*, **151**(2), 325-344.

Strauss, P., Harzhauser, M., Hinsch, R., Wagreich, M. 2006. Sequence stratigraphy in a classic pull-apart basin (Neogene, Vienna Basin). A 3D seismic based integrated approach. *Geologica Carpathica*, **57**(3), 185-197.

Tarantola, A. 1984. Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, **49**(8), 1259-1266.

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

Ulvmoen, M., Omre, H. and Buland A. 2010b. Improved resolution in Bayesian lithology/fluid inversion from prestack seismic data and well observations: Part 1 – Real case study. *Geophysics*, **75(2)**: B73-B82.

Virieux, J. and Operto, S. 2009. An overview of full-waveform inversion in exploration geophysics. *Geophysics*, **74**(6): WCC1-WCC26.

5

Determination of Reservoir Lithology by a 2-D Hidden Markov Random Field model

In this study, geological prior information is incorporated in the classification of reservoir lithologies after the adoption of Markov Random Field (MRF). The prediction of hidden lithologies is based on measured observations such as seismic inversion results, which are associated with the latent categorical variables, based on the distribution of Gaussian assumptions. Compared with other statistical methods such as the Gaussian Mixture Model (GMM) or k-Means, which do not take spatial relationships into account, the Hidden Markov Random Field (HMRF) approach can connect the same, or similar lithologies horizontally while ensure a geologically reasonable vertical ordering. It is, therefore, able to exclude randomly appearing lithologies caused by errors in the inversion. The prior information consist of a Gibbs distribution function and transition probability matrices. The Gibbs distribution connects the same or similar lithologies which does not need a geological definition from the outside. The transition matrices provide preferential transitions between different lithologies and an estimation of them implicitly depends on the depositional environments and juxtaposition rules between different lithologies. Analogue cross-sections from the subsurface or outcrop studies can contribute to the construction of these matrices by a simple counting procedure.

5.1. Introduction

The classification of lithologies is an essential step in reservoir characterization and in the building of a static reservoir model. The definition of lithologies, and the number and types of lithologies can be provided by geologists. Preliminary analysis of well-log data will identify various lithologies and most of the time the number of lithologies will be kept constant afterwards. Other sources of information, such as seismic data, can provide a larger areal coverage, thus overcoming the limitations provided by sparse well locations.

Inference of lithologies from seismic data is a challenging task and actually an ill-posed inverse problem, because a variety of different lithological characteristics may result in identical or similar seismic responses (Larsen *et al.* 2006). The Bayesian concept is usually applied to mitigate this problem as applied by Mukerji *et al.* (2001) and Houck (2002) who identified lithology/fluid (LF) classes based on amplitude-versus-offset (AVO) analysis. Buland and Omre (2003) developed a linearized AVO inversion approach under the Bayesian framework. Subsequently, Buland *et al.* (2008) proposed a fast Bayesian inversion approach for 3-D lithology and fluid prediction from prestack data.

However, the approach mentioned above is point- or location-based, which means that the spatial coupling between data points is not considered. In order to address this problem, prior information can be included in which a Markov chain or a Markov Random Field is applied. Eidsvik *et al.* (2004) translated well-log data to geological attributes by hidden Markov chains. Larsen *et al.* (2006) incorporated a stationary Markov-chain prior model to simulating vertical continuity of lithology/fluid classes. Ulvmoen and Hammer (2010) compared two algorithms — approximated and exact likelihood models for the inversion of lithologies and fluids - in which the Markov *a priori* knowledge is incorporated in a Bayesian setting. Ulvmoen and Omre (2010) and Ulvmoen *et al.* (2010) adopted a profile Markov Random Field to model the *a priori* information of the lithology/fluid classes in order to improve the resolution in Bayesian lithology/fluid inversion from prestack seismic data. Hammer *et al.* (2012) inverted a vertical profile of rock properties based on seismic amplitude data in which a Markov process prior is included to guarantee that vertical dependencies are honored. Other reservoir parameters such as porosity and saturation could also be inferred from seismic data combined with well observations as has been done by Bosch *et al.* (2009).

In this chapter, as in the previous one, instead of deriving the lithologies from prestack or stacked seismic data, the efforts are geared towards the usage of inversion results in reservoir description. The wave-equation based inversion scheme can provide high resolution results because the intrinsic relationship between rock properties and the seismic data has been fully exploited (Gisolf *et al.* 2014). This feature makes the approach suitable as a potential input for the reservoir characterization process.

The 3-D distribution of lithologies in the subsurface is not directly observable, with only limited information provided by wells. However, indirect observations in the form of measurements are available that contain information on them. Hidden Markov Models (HMMs) are trying to uncover these latent states under this concept, but only in 1-D, i.e.

vertical direction (Eidsvik *et al*. 2004; Lindberg and Grana 2015). Here we present a 2-D method, in which the horizontal prior information is also incorporated through the introduction of Markov Random Field.

The rock properties obtained by the inversion of seismic data are assumed to be distributed according to multivariate Gaussian functions (Avseth *et al*. 2005). Thus a Gaussian Mixture Model (GMM) is used to describe the conditional probabilities of the inverted properties from seismic data, given the different lithologies.

In this chapter, first a short introduction of the Markov Random Field is given, then the theory of the GMM-based Hidden Markov Random Field (GMM-HMRF) is described. Finally, some synthetic examples and a field case study from Vienna Basin will be shown together with discussions and some conclusions.

5.2. Markov Random Field

First introduced by Ising (Ising 1925), a Markov Random Field (MRF) is an undirected graphical model and can be described by a group of random variables that possess a Markov property. This Markov property can be defined by a joint probability distribution, which is determined by a local conditional distribution. Figure 5.1 illustrates this concept in which the white node is independent of all other black nodes given the red nodes.

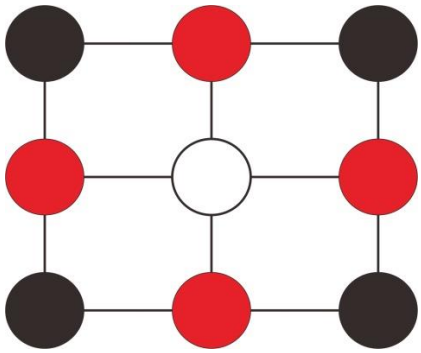


Figure 5.1: Schematic view of the dependency between nodes.

The following equation describes the conditional distribution of Z_n:

$$\Pr(Z_n|Z_m, m \neq n) = \Pr(Z_n|Z_m, m \in \varepsilon) \tag{5.1}$$

where Z_n can take a value in the set of categorical variables such as lithology, which is associated with the node n; ε represents the local neighborhood set of nodes that share an edge with node n in the graph.

However, it is not easy to construct the joint distribution of a MRF based on the local conditional distribution of $\Pr(Z_n|Z_m, m \in \varepsilon)$. The Hammersley-Clifford theorem builds the equality between the joint distribution of any MRF and a Gibbs distribution; the joint distribution of a MRF can be defined by a clique potential (see below) (Besag 1974; Winkler 1995). A Gibbs distribution is taking the form:

$$\Pr(\mathbf{Z}) = \frac{1}{X} e^{-U(\mathbf{Z})} \tag{5.2}$$

in which

$$U(\mathbf{Z}) = \sum_{c \in C} V_c(\mathbf{Z}) \tag{5.3}$$

where $\Pr(\mathbf{Z})$ is the probability distribution of random variables $\mathbf{Z}$, $U(\mathbf{Z})$ is the energy function, X is the partition function, c is a clique which is a subset of nodes satisfying the demand that every node is linked to every other one, and C is the set of c; $V_c(\mathbf{Z})$ can be referred as clique potential functions (Wang *et al*. 2016).

5.3. Gaussian Mixture Model based Hidden Markov Random Field (GMM-HMRF)

Similar to HMMs (Rabiner 1989; Eidsvik *et al*. 2004), Hidden Markov Random Field (HMRF) are also trying to uncover the categorical variables that are hidden to the observers (Figure 5.2). The difference with HMMs is that the theory of the MRF is applied, which has no limitation in 1-D (depth; see Chapter 4). That is why it is more suitable for quantifying reservoir properties in 2-D or even 3-D.

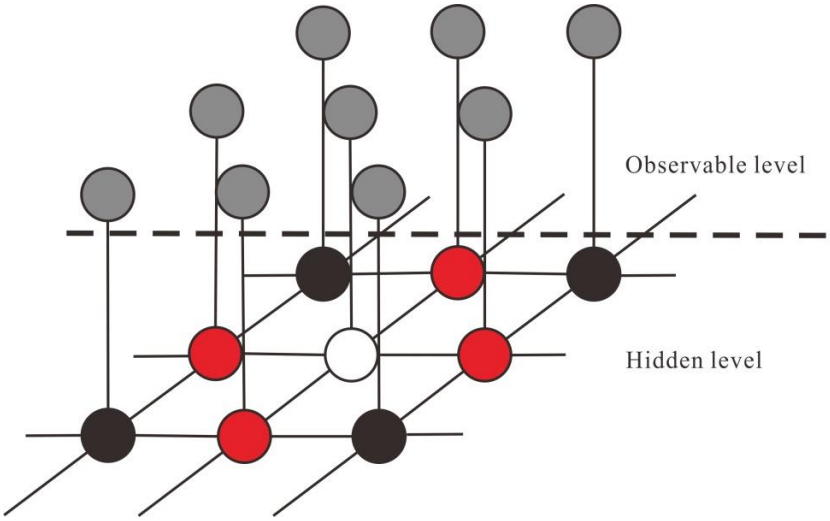


Figure 5.2: Hidden Markov Random Field with observable and hidden levels.

In HMRF, the hidden level is associated with categorical variables in physical space, while in the observable level the data can be obtained in statistical space. Hence, a HMRF model is defined as a stochastic process derived by a MRF whose label configuration cannot be observed directly (Wang *et al*. 2016). However, a MRF could generate measurable datasets that are assumed to honor certain probability distribution functions known as the emission probability functions (Rabiner 1989; Wang *et al*. 2016).

In HMRF, according to the Maximum a Posterior (MAP) criterion, the purpose is to seek the states $\hat{\mathbf{Z}}$ that satisfy (Zhang *et al*. 2001; Wang 2012):

$$\hat{\mathbf{Z}} = \underset{\mathbf{Z}}{\operatorname{argmax}}\{\Pr(\mathbf{Y}|\mathbf{Z},\boldsymbol{\theta})\Pr(\mathbf{Z})\} \tag{5.4}$$

where $\Pr(\mathbf{Z})$ is the prior probability, which is a Gibbs distribution in equation (5.2); $\Pr(\mathbf{Y}|\mathbf{Z},\boldsymbol{\theta})$ is the joint likelihood probability of the observation $\mathbf{Y}$.

A typical characteristic of $\Pr(\mathbf{Y}|\mathbf{Z},\boldsymbol{\theta})$ is the conditional independence (Zhang *et al*. 2001; Wang 2012):

$$\Pr(\mathbf{Y}|\mathbf{Z},\boldsymbol{\theta}) = \prod_i \Pr(Y_i|Z_i,\theta_{Z_i}) \tag{5.5}$$

$\Pr(Y_i|Z_i,\theta_{Z_i})$ is the emission distribution of the observation Y_i, with parameters θ_{Z_i}. Different probability functions can be applied to describe it, but to keep analytical tractability, a Gaussian assumption is made (Lindberg and Grana 2015).

Given the observation data $\mathbf{Y}$, for a certain state Z_i, which takes a value in the state space $\mathbf{S} = \{S_1,\ S_2,\ \cdots,\ S_N\}$, the Gaussian distribution has the following form with the means μ_j and the covariance matrices σ_j:

$$\Pr(Y_i|Z_i = S_j,\theta_j) = f(Y_i;\mu_j,\sigma_j) \tag{5.6}$$

where $\theta_j = (\mu_j,\sigma_j)$ which is the specified θ_{Z_i} when $Z_i = S_j$ and

$$f(Y_i;\mu_j,\sigma_j) = \frac{1}{\sigma_j\sqrt{2\pi}}exp(-\frac{(Y_i-\mu_j)^2}{2\sigma_j^2}) \tag{5.7}$$

In equation (5.7), the intensity distribution of each state, or lithology to be classified, is a Gaussian distribution with the parameter sets $\theta_j = (\mu_j,\sigma_j)$. However, sometimes it is insufficient to describe the complexity in the distribution of the observation data, especially for multimodal distributions. Thus a GMM is more powerful than a single Gaussian function to model the complexity and can be described in the following with the parameter sets θ_j in which there are k components:

$$\theta_j = \{(\mu_{j,1},\sigma_{j,1},\omega_{j,1}),\cdots,(\mu_{j,k},\sigma_{j,k},\omega_{j,k})\} \tag{5.8}$$

where $\omega_{j,k}$ is a mixture weight of the k^{th} component given a specific state S_j.

Accordingly, equation (5.7) will have a weighted probability form:

$$f(Y_i;\theta_j) = \sum_{n=1}^{k} \omega_{j,k}\, f(Y_i;\mu_{j,k},\sigma_{j,k}) \tag{5.9}$$

Without the spatial correlation of Pr($\mathbf{Z}$), equation (5.4) will become a degenerated case of HMRF, in which the GMM is defined and can specified fully by the histogram of the data (Zhang *et al*. 2001). After the incorporation of Pr($\mathbf{Z}$) as a prior, the classification problem is then approached statistically as well as spatially.

However, the prior Pr($\mathbf{Z}$) only considers the spatial correlation of the neighbors and tries to make the same (constant image) or similar (continuous image) prediction with the contextual constraints (equation (5.4)), and it does not need a specification of geological knowledge which can be considered as an internal prior. Thus, this is not sufficient because some unrealistic classifications could happen, such as a water sand on top of a gas or oil sand in a given reservoir. This could be due to measurement errors or misleading neighbors (Lindberg and Grana 2015). Therefore, a MRF matrix as an external prior is proposed in which another constraint on lithological transitions will be introduced.

Unlike a traditional Markov chain matrix, which is obtained by the procedure of counting the transitions and normalizing in a vertical direction (depth) (Elfeki and Dekking 2001), this new profile Markov matrix is going to be used, in which the usual counting and normalizing in the vertical-upward direction will be kept and at the same time the left and right neighbors in the lateral direction of the future state will be considered (Ulvmoen and Omre 2010; Ulvmoen *et al*. 2010). A detailed description of this prior matrix $P_{(:,:)}$ for lithologies is provided in the Appendices.

Thus, equation (5.4) has to be formatted in order to take the matrix $P_{(:,:)}$ into account:

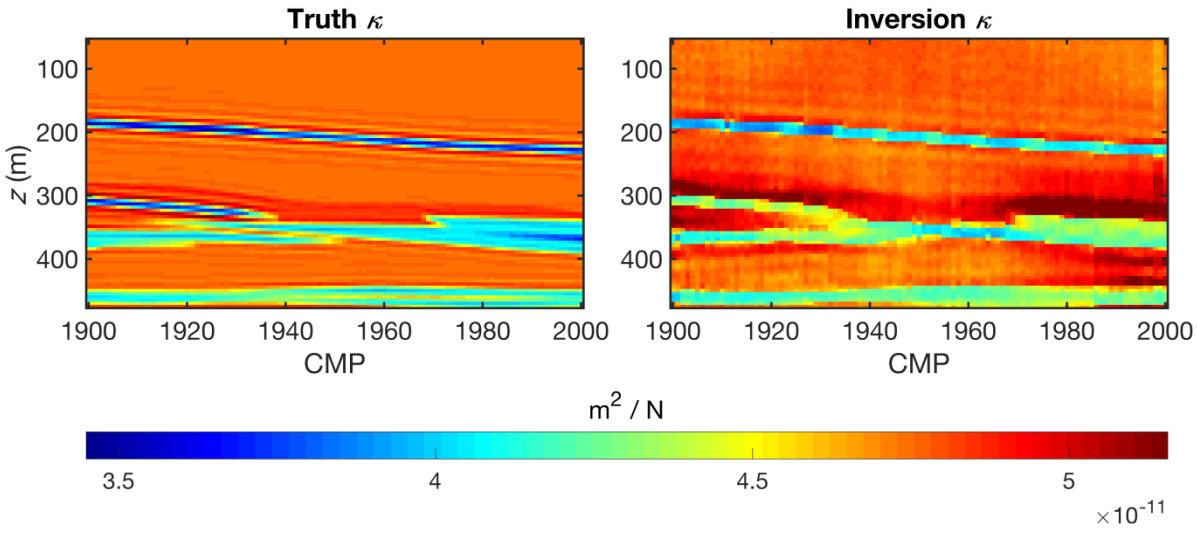
$$\hat{\mathbf{Z}} = \underset{\mathbf{Z}}{\operatorname{argmax}}\{\Pr(\mathbf{Y}|\mathbf{Z},\boldsymbol{\theta})\Pr(\mathbf{Z})P_{(:,:)}\} \tag{5.10}$$

In order to find an estimated $\hat{\mathbf{Z}}$, equation (5.10) is invoked, in which both the states and the parameter sets in GMM are unknown, as described above. Furthermore, they are connected with each other. Different techniques have been introduced to solve this problem, in which the Expectation-Maximization (EM) method is the one most widely used (Zhang *et al*. 2001). The strategy in the EM approach is as follows: Given the current estimated $\boldsymbol{\theta}$, predict the hidden variables $\hat{\mathbf{Z}}$; then $\boldsymbol{\theta}$ can be updated by maximizing the expectation of the complete-data likelihood function $\mathrm{E}[\Pr(\mathbf{Z},\mathbf{Y}|\boldsymbol{\theta})P_{(:,:)}]$ (Zhang *et al*. 2001; Wang *et al*. 2016). This process will be iterated until certain conditions are met. For the mathematical details, the readers are referred to Zhang *et al*. (2001).

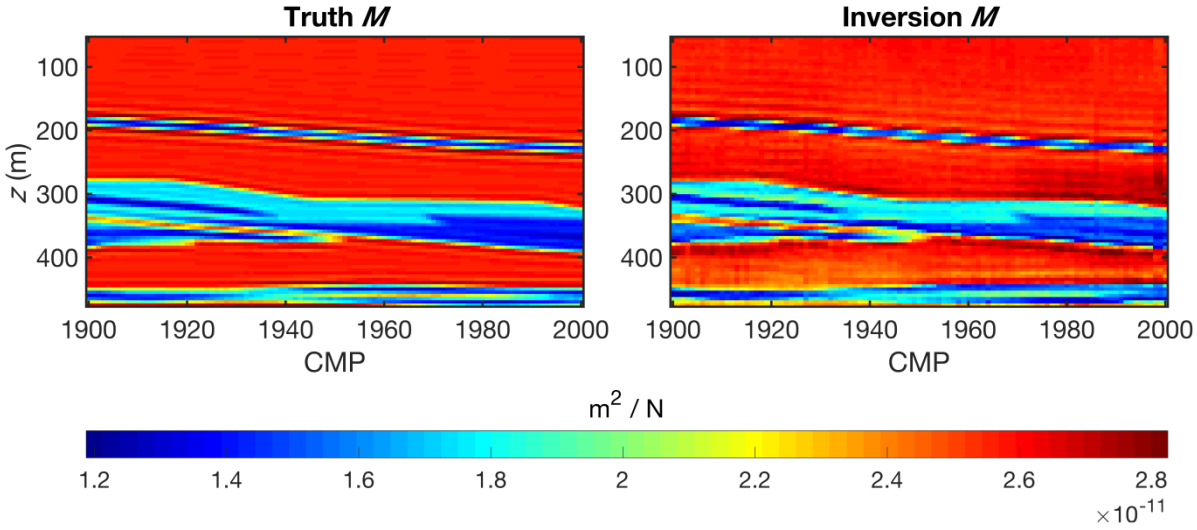
5.4. Book Cliffs Example

The first example for applying this approach is the synthetic Book Cliffs model created by Feng *et al*. (2017) (Chapter 2) in which more details have been added and more

differentiation is put on the potential reservoir lithologies than in the original. As a test, only a subset of the whole 2-D section has been selected, and Figure 5.3 shows the true and inverted properties in terms of κ and M ($\kappa = 1/K$, with K being the bulk modulus; $M = 1/\mu$, with μ being the shear modulus).



(a)



(b)

Figure 5.3: True and inverted properties of a selected part from the Book Cliffs model. (a) κ; (b) M.

The quality of inversion results is quite good when compared with the truth, since most geometries have been recovered correctly, as well as the properties, which is due to the fact that the nonlinear relationship between rock properties and seismic data has been fully exploited (Gisolf *et al*. 2014).

Observations from wells are needed for inferring the prior Markov models (Ulvmoen *et al*. 2010) as well as for building the lithological templates. As a starting point of the

classification process, two pseudo-wells have been "drilled" at the left- and right-most locations of the selected section (CMP = 1900 and 2000) (Figure 5.4).

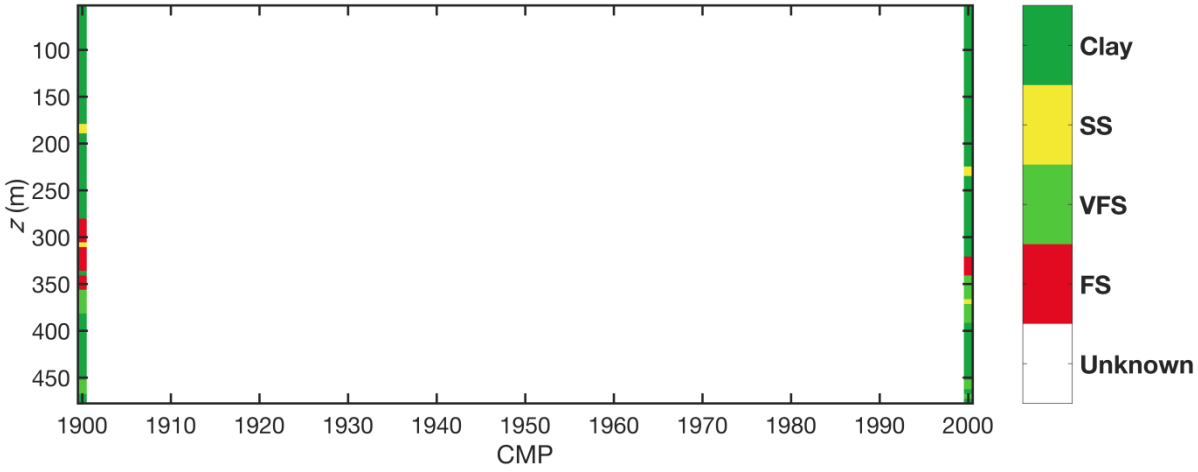


Figure 5.4: Two pseudo-wells at CMP = 1900 and 2000. Lithologies are known at the well locations with FS (fine-grained sandstone), VFS (very fine-grained sandstone), SS (siltstone) and Clay.

The true and inverted properties at the well locations can be seen in Figure 5.5. Figure 5.6 shows 90% confidence regions of the Bivariate Gaussian likelihood model in terms of κ and M.

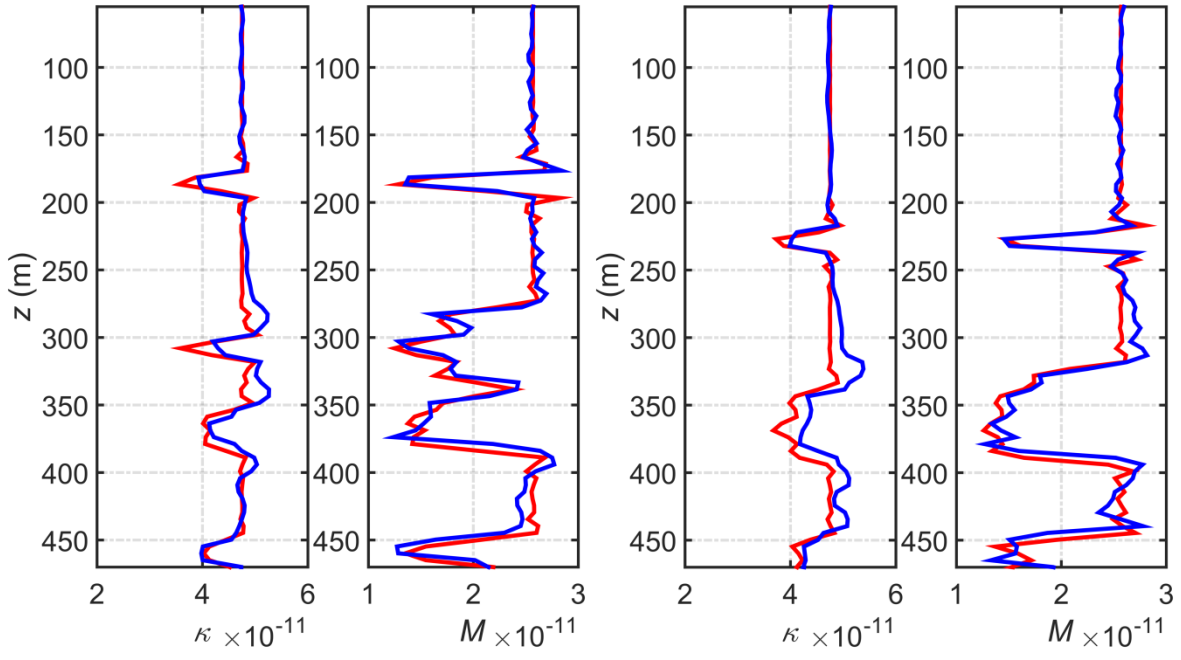


Figure 5.5: True (red curves) and inverted (blue curves) properties at CMP = 1900 (a) and 2000 (b). Note that the values in the wells have been up-scaled to the seismic grid interval.

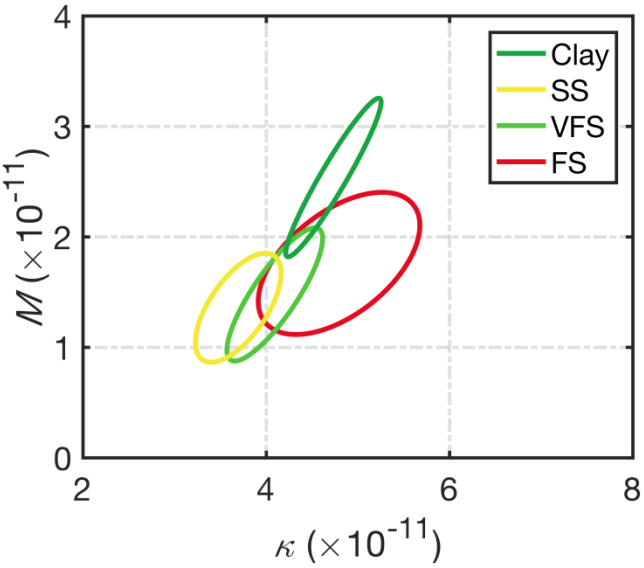


Figure 5.6: 90% confidence regions of the Bivariate Gaussian likelihood model for the distribution of each lithology.

From Figure 5.6, it can be seen that there are some overlapping areas between different lithologies, especially for SS and VFS, which makes the differentiation difficult when only the property values are considered. This Gaussian likelihood model (Figure 5.6) is adopted which is the emission probability of properties given a specific lithology (Rabiner 1989) or the termed likelihood function generally (Ulvmoen and Omre 2010). However, in contrast with other statistical methods such as GMM or *k*-Means (Seber 1984), two additional parameter sets are introduced in GMM-HMRF: The Gibbs prior, and the profile Markov matrix. These ensure that the geological information is implicitly incorporated during the classification process. Figure 5.7 displays the cross-section "truth" in the subsurface. The starting model of the classification is shown in Figure 5.8, which is derived from a non-iterative histogram-based statistical approach with the two “drilled” wells as lithological templates and inversion results as inputs (Figures 5.4, 5.5 and 5.6).

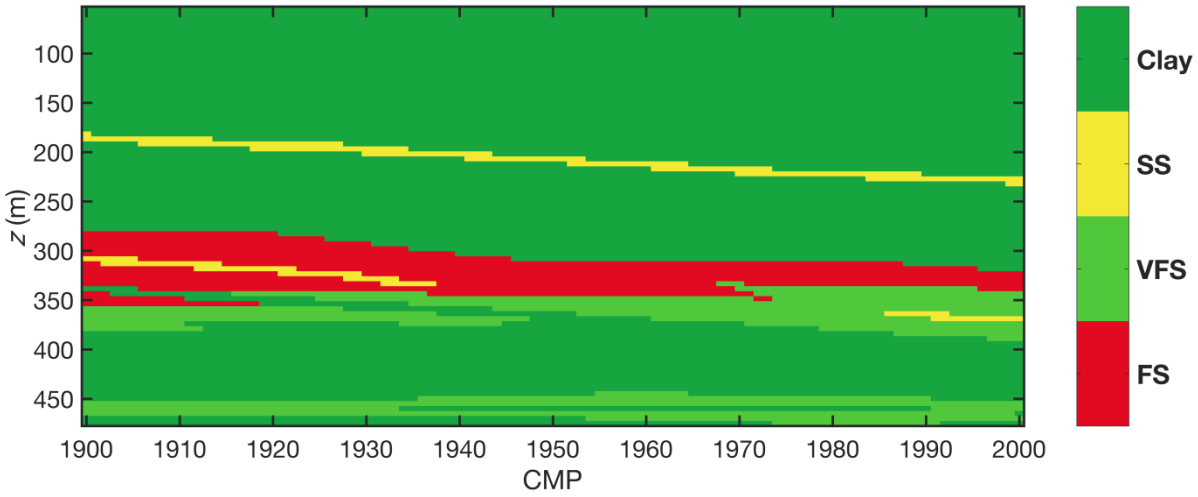


Figure 5.7: Subsurface cross section in terms of lithologies.

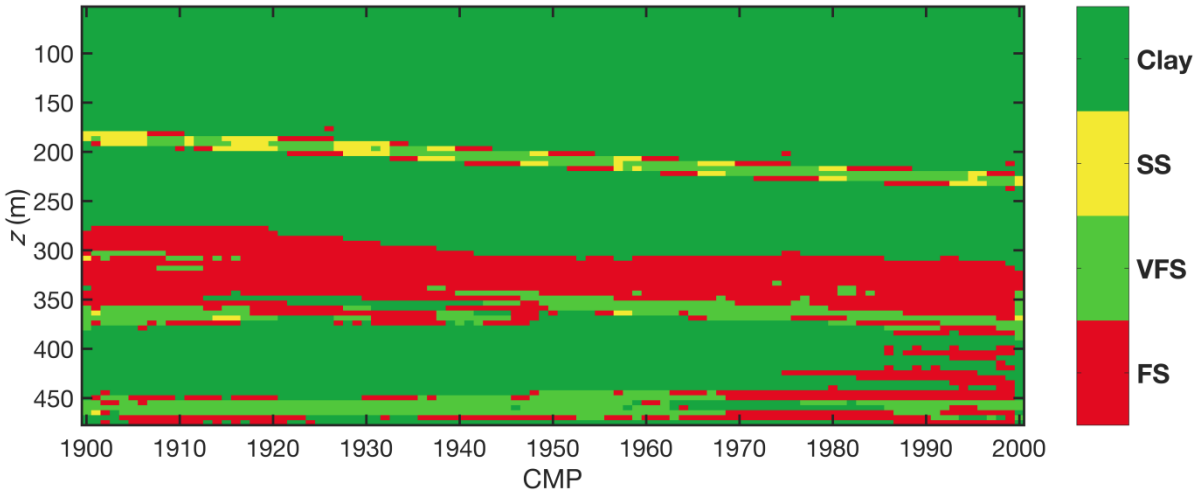


Figure 5.8: Starting model in terms of lithologies.

Since only the inverted property values are used, there is no spatial correlation between the sample points, which makes the prediction unreliable in the form of randomly appearing lithologies, particularly at the layer boundaries. Subsequently the GMM methodology is applied with the lithological distributions in Figure 5.8 as inputs. By contrast with the simple histogram-based and non-iterative approach used in Figure 5.8, the GMM clusters sample points into different groups (lithologies) by applying an iterative procedure called expectation-maximization (EM), as discussed above.

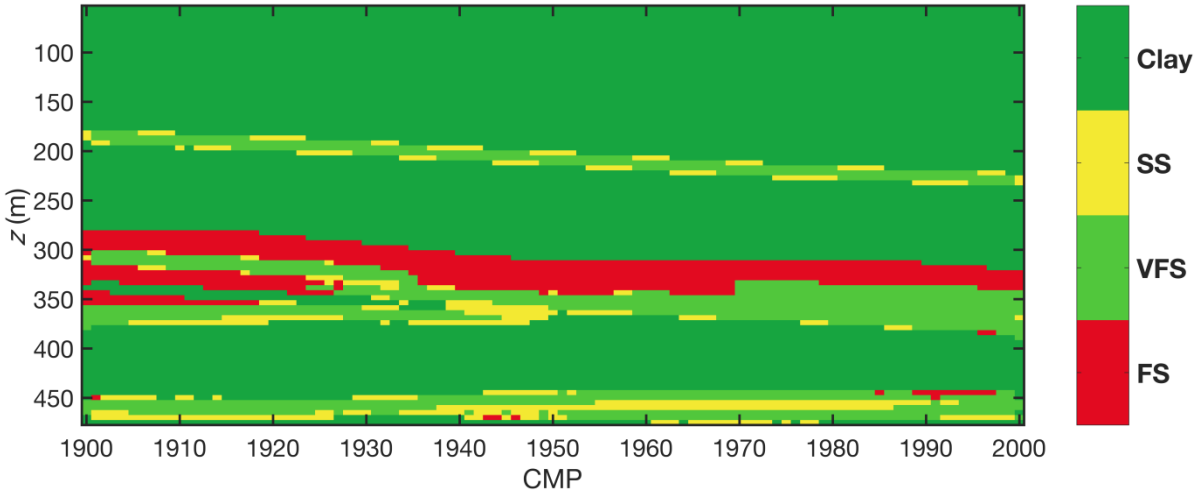


Figure 5.9: Result of GMM with the starting model shown in Figure 5.8.

The EM algorithm finds the maximum likelihood estimates of parameters in probabilistic models in the presence of missing data, which in this context means that the lithologies are unknown (Lindberg and Grana 2015). An iterative scheme is performed in which the expectation step calculates the probability of every sample point belonging to each lithology, while in the maximization step, it maximizes the means and covariance matrices according to the probabilities computed in the expectation step. As in Grana and Della Rossa (2010), the spatial correlation is ignored in the estimation of the parameters here in which some random

states or lithologies still occur, although the result (Figure 5.9) is improved compared with the initial one in Figure 5.8.

In order to test the different roles of the prior information, i.e. the Gibbs function (equation (5.2)) and the Markov matrix (Appendix A), results incorporated with the prior information separately are shown (Figures 5.10 and 5.11). With only the Gibbs prior used (Figure 5.10), some lithologies are better connected than with the GMM (Figure 5.9). However, some random lithologies still exist such as the SS in the VFS layer in the upper part and the FS in the lower-middle area has been separated because of locally connected VFS at the depths between 300 and 350m and CMPs between 1925 and 1940.

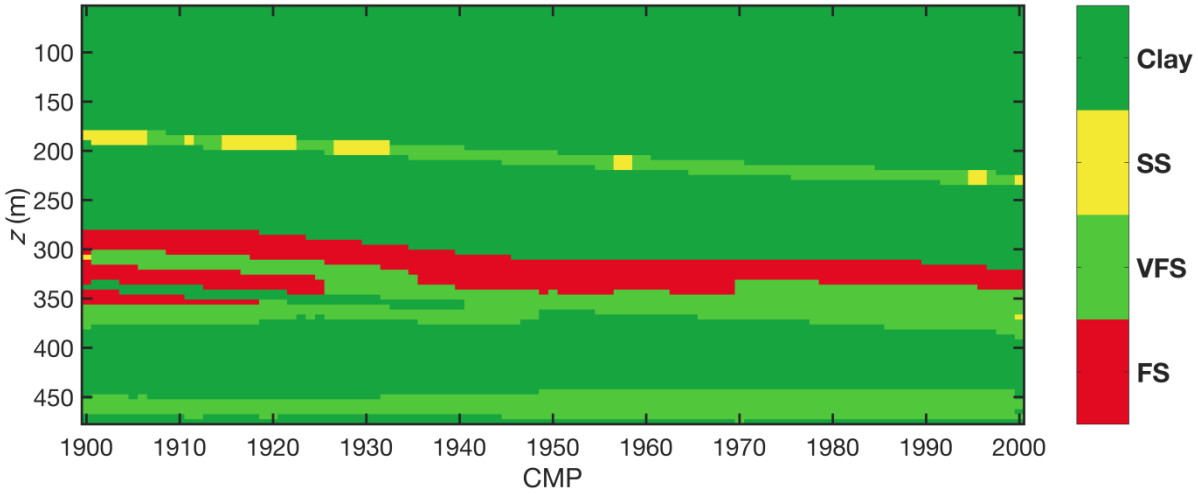


Figure 5.10: Result of GMM-HMRF incorporated with the Gibbs prior only.

With only the Markov matrix used, the distribution of lithologies is preferential in the horizontal direction (Figure 5.11) since the transition has been governed by the left and right neighbors, demonstrated as larger values, that provide information on the lateral continuity or is under the consideration of layered formations (Appendix A). However, these matrices have to be modified in order to simulate transitions in a given reservoir which will be discussed in the real case study below.

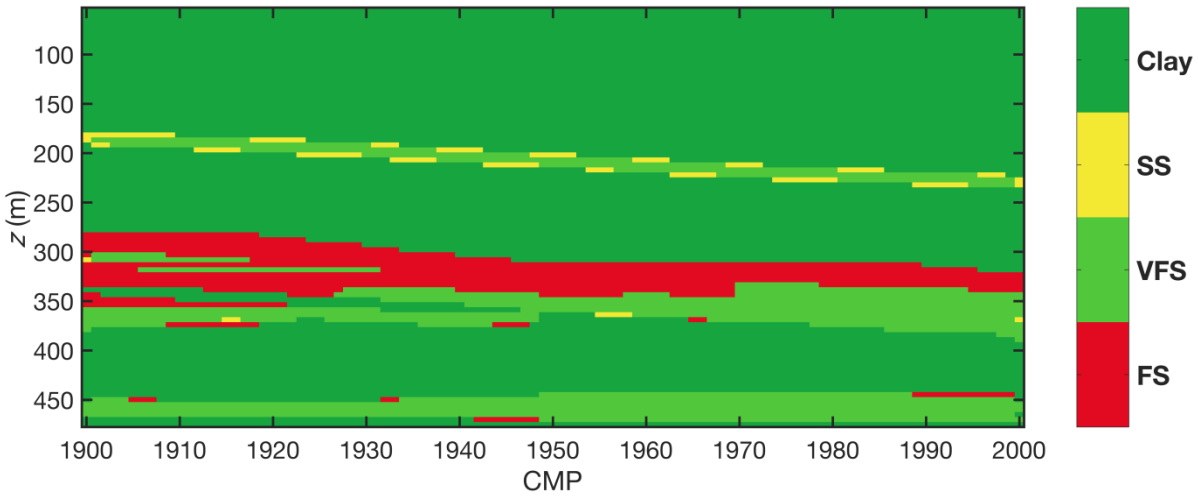


Figure 5.11: Result of GMM-HMRF incorporated with the Markov matrix only.

By applying both prior information, the spatial correlation is considered in the vertical and horizontal directions. Compared with the result by the GMM (Figure 5.9) and others with either one of the prior information incorporated (Figures 5.10 and 5.11), the classified lithologies (Figure 5.12) are distributed more orderly and closer to the truth (Figure 5.7).

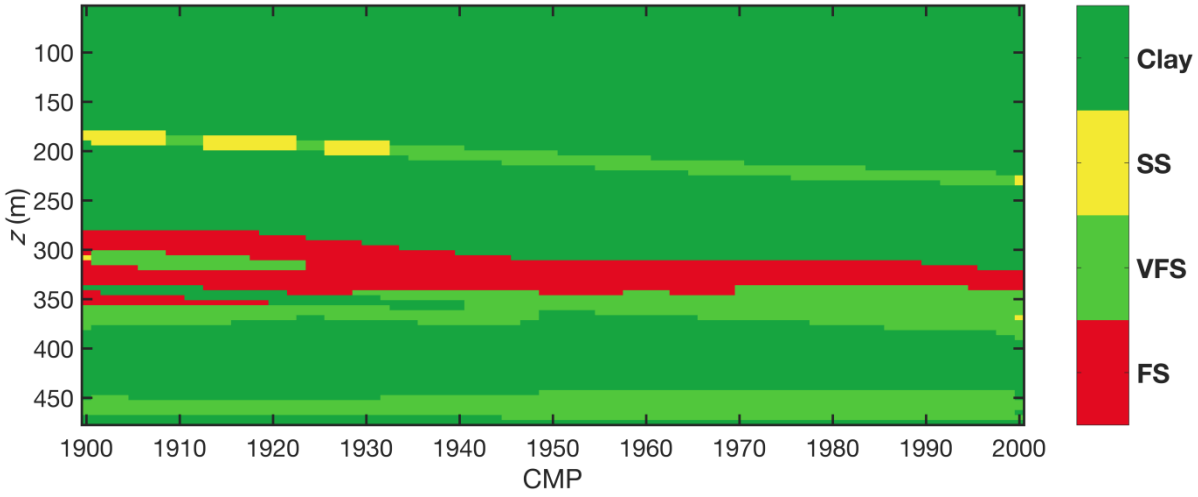
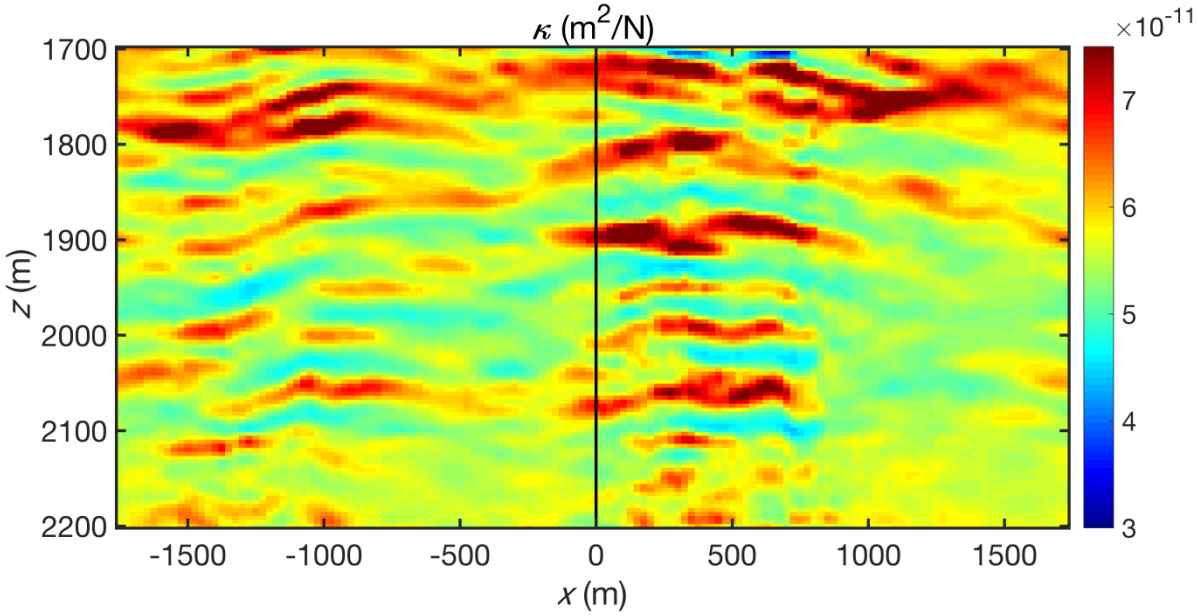


Figure 5.12: Result of GMM-HMRF incorporated with both prior information.

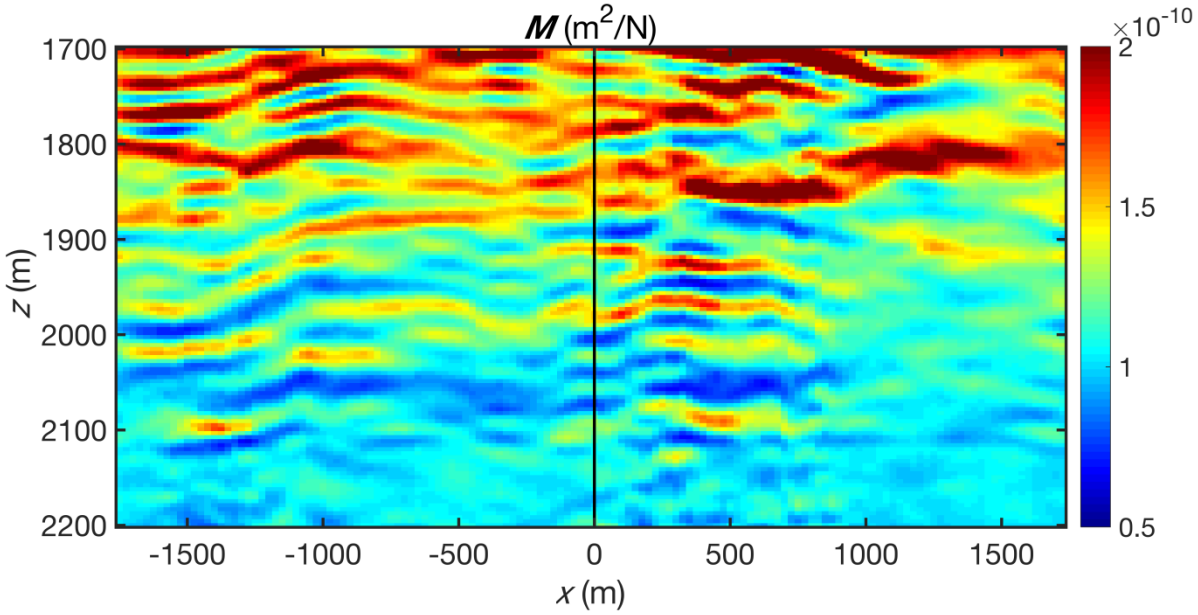
5.5. Real Case Study in the Vienna Basin

In order to further test the ability of the proposed GMM-HMRF, a real field data set from the Vienna Basin is used. Vintages of 3-D seismic surveys acquired in different years have been merged into a single dataset VBSM (Vienna Basin Super Merge) which are used as inputs for the non-linear full-waveform inversion scheme proposed by Gisolf *et al*. (2014). The inverted rock properties from the seismic data are then used as inputs for the lithology prediction as illustrated above.

A single cross-section of inverted rock properties (κ and M) has been selected from the available data set, which is traversed by a logged well in the middle (Figure 5.13).



(a)



(b)

Figure 5.13: Inverted rock properties (a) κ; (b) M. The black line represents the location of the logged well.

Due to problems in the pre-processing phase, or fault cut-outs, or both, the inversion result is of suboptimal quality, even though in the upper part of the well the match is quite reasonable (Figure 5.14).

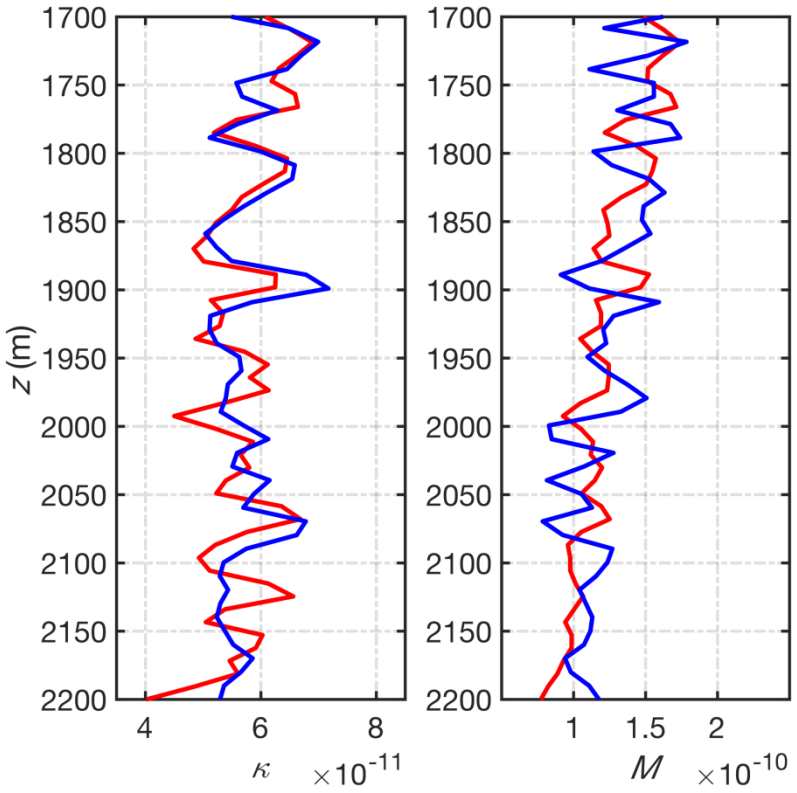


Figure 5.14: True (red curves) and inverted (blue curves) rock properties at the logged location (Figure 5.13).

In order to perform the new classification methodology, the starting model in the form of lithologies (Figure 5.15) is obtained by applying the *k*-Means method (Seber 1984) and a comparison of Euclidean distances (Elena and Michel 2009) between the cluster centroid locations and the known lithologies properties in the well. This is different from the synthetic example above since there is only one well here and the inversion quality is lower.

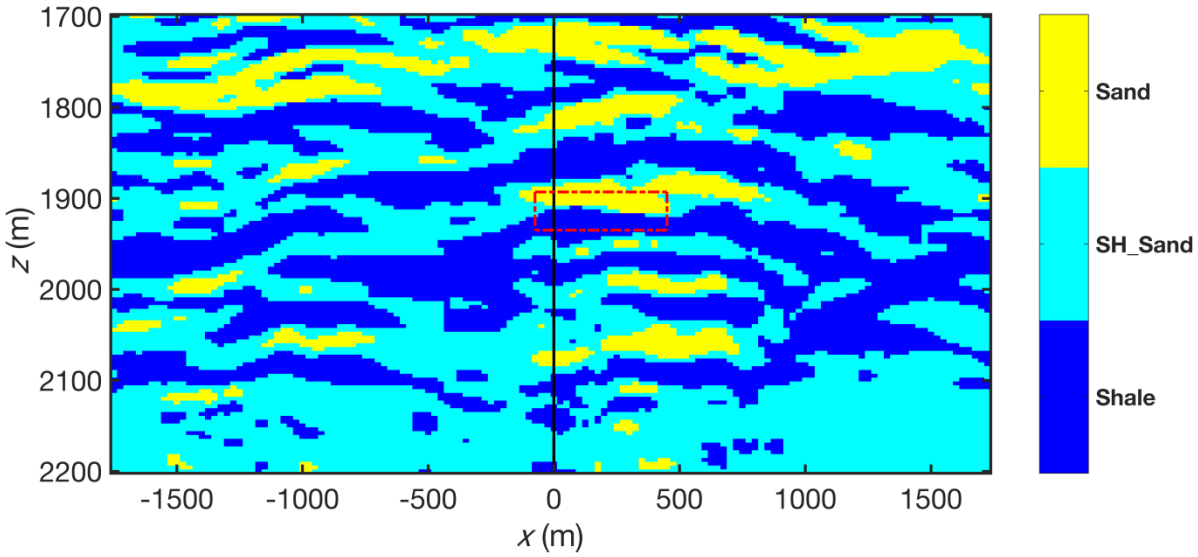


Figure 5.15: Starting model in terms of lithologies by *k*-Means (Sand: Sandstone; SH_Sand: Shaly Sandstone).

The classified and "true" lithologies at the well location is shown in Figure 5.16 in which it can be seen that almost all of the sandstone units have been predicted well even though the thin shale streaks have been missed due to the low seismic resolution and inversion quality.

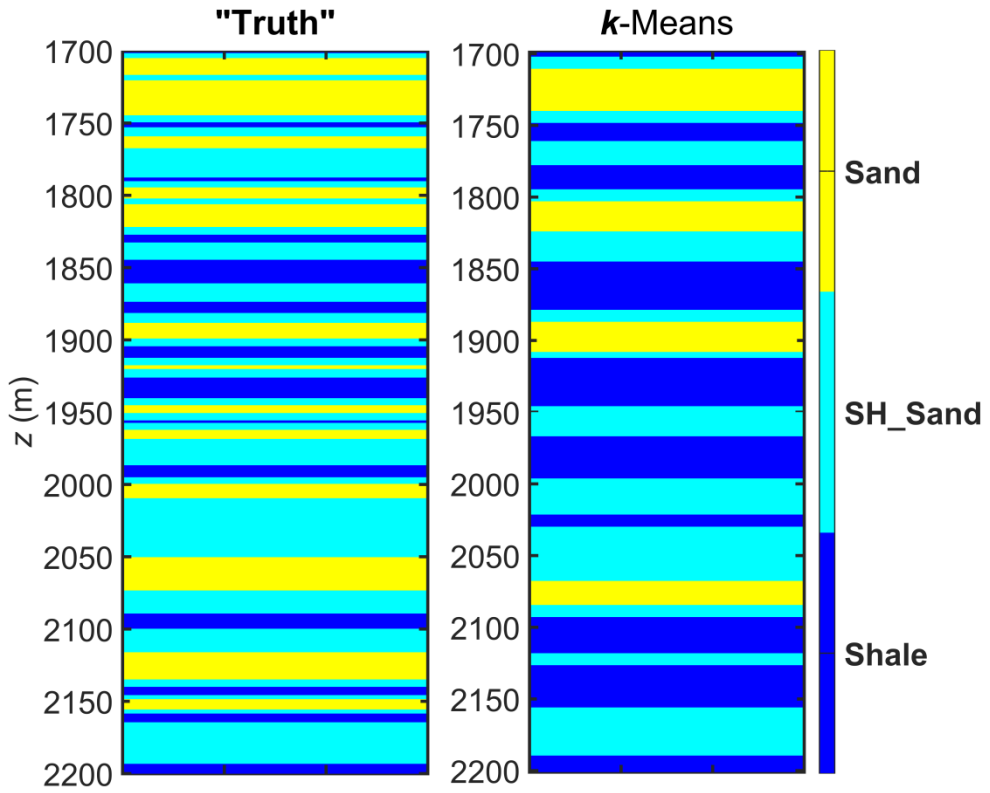


Figure 5.16: "Truth" and Prediction of *k*-Means at the well location (Black line in Figure 5.15).

After applying the proposed GMM-HMRF in equation (5.10) and the profile Markov matrices in Appendix B, the result is shown in Figure 5.17.

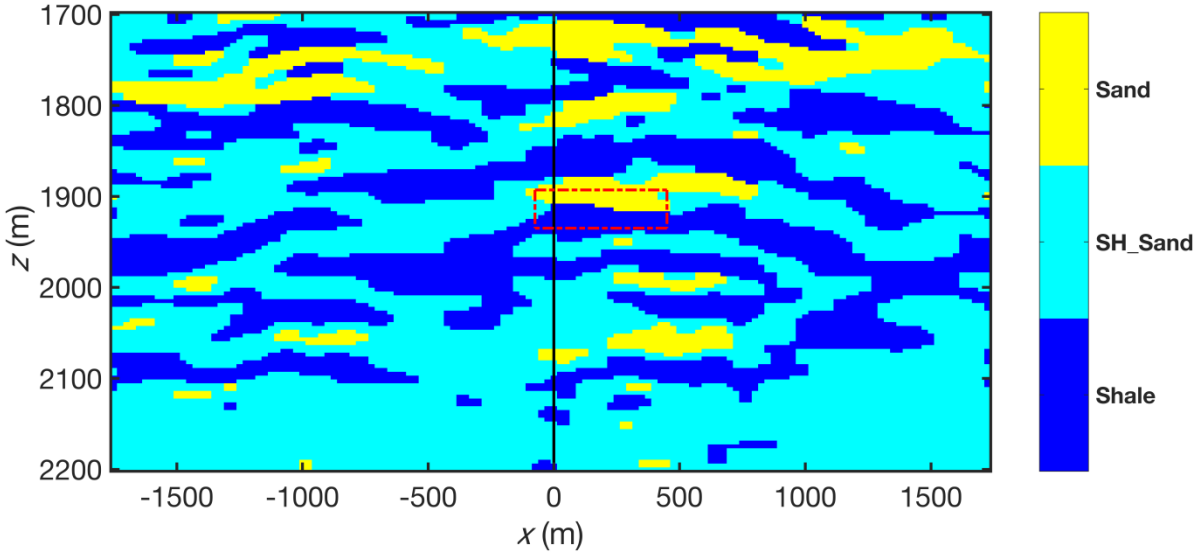


Figure 5.17: Classified result of GMM-HMRF with the Gibbs prior and Markov matrices in Appendix B.

Compared with the result in Figure 5.15, the distributions of lithologies in Figure 5.17 are more compact because of the incorporated priors which try to connect and simplify the lithology transitions, therefore reducing the “noise” in the lithological model (Figure 5.15). However, there are some “unlikely” transitions such as Sand on top of Shale (Figure 5.18) (this could happen in many geological settings, but here it is assessed as “unlikely” since it does not occur in the cored "truth" well, as seen in Figure 5.16). The reason for this is that small (but not zero) transition probabilities (0.0001) have been assigned in the Markov matrices (Appendix B).

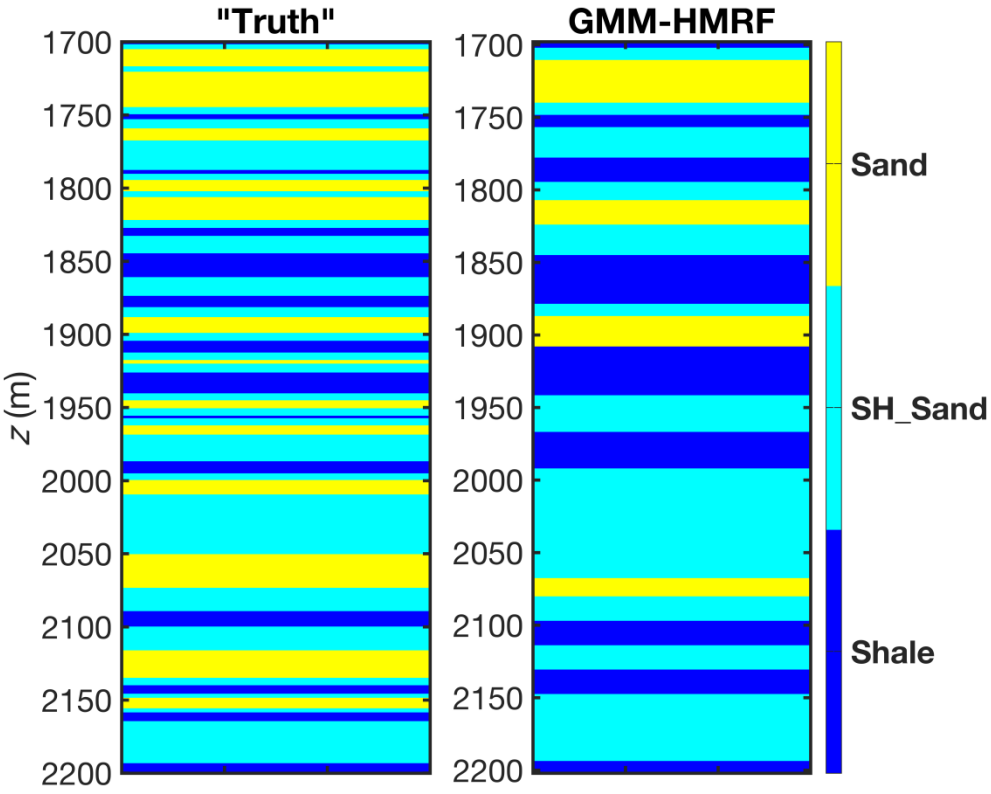


Figure 5.18: “Truth” and Prediction of GMM-HMRF at the well location (Figure 5.17).

In order to exclude this transition and simulate a typical transition in separated reservoirs such as a water sand not overlying an oil or gas sand because of gravity segregation, Markov matrices are modified to the ones as shown in Appendix C. After applying these new matrices, the transitions between Sand and Shale have been removed (Figures 5.19 and 5.20).

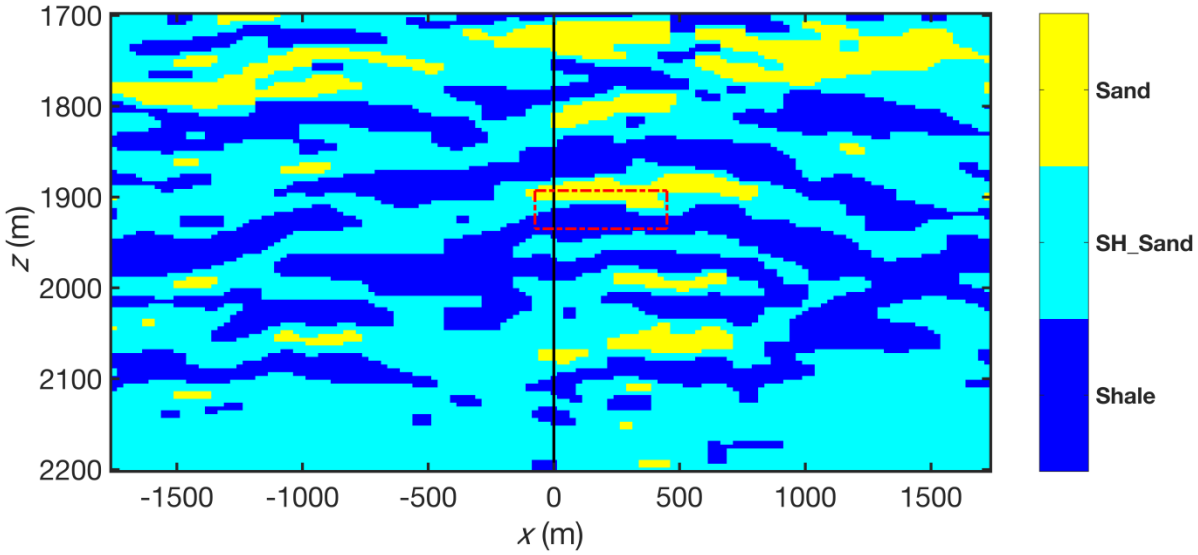


Figure 5.19: Classified result of GMM-HMRF with the Gibbs prior and modified Markov matrices in Appendix C.

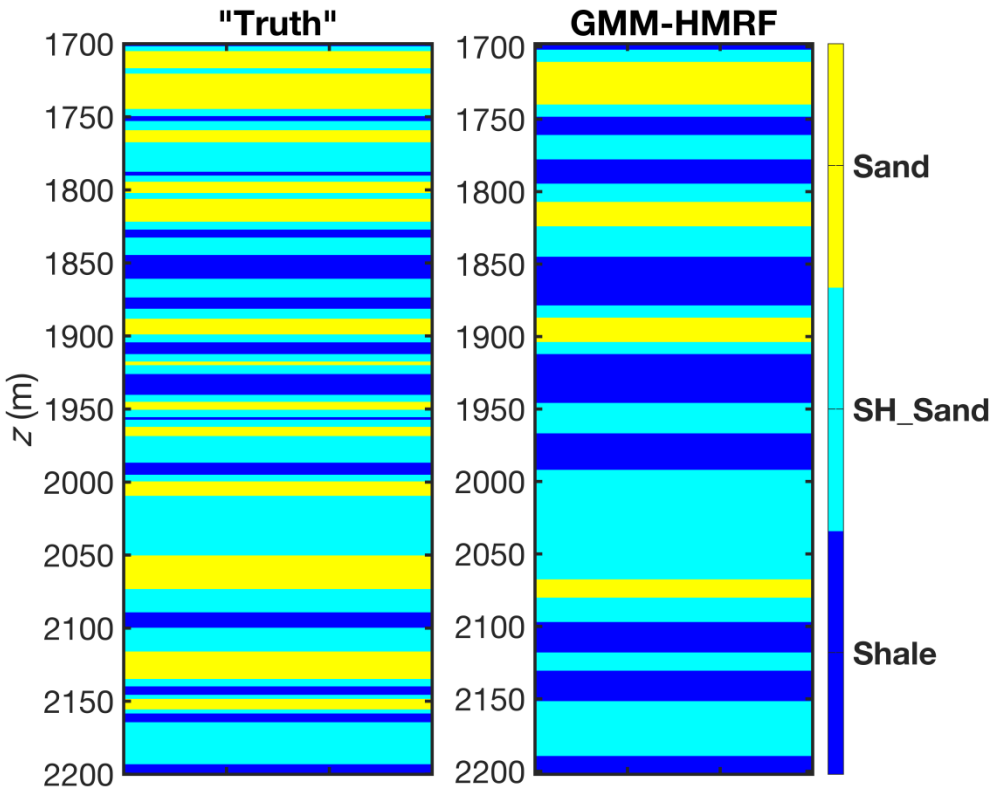


Figure 5.20: "Truth" and Prediction of GMM-HMRF at the well location (Figure 5.19).

5.6. Discussion

In this study, the spatial correlation during the lithological classification process is taken into account through the concept of MRF, in which the Gibbs prior and the profile Markov matrix are incorporated. In contrast to GMM or *k*-Means, which do not use the geological spatial prior knowledge, the proposed GMM-HMRF is able to produce better images of the

categorized variables, and each lithology tends to connect with the same or similar lithology horizontally and vertically based on preferential transitions.

The input data for the classification are the FWI seismic inversion results. Compared with other inversion methods, FWI can provide high resolution since the nonlinear relationship between the rock properties and seismic data has been exploited by utilizing wave-mode conversions and multiple scattering. In contrast to rock properties such as bulk density and acoustic velocity, the compressibility (κ) and shear compliance (M) are used here because they appear naturally in the elastic wave-equations and are more closely related to rock types.

However, the classified result shown in Figure 5.12 is not perfect compared with the truth in Figure 5.7 especially for SS in the upper part, which has been clustered as VFS. From CMPs 1900 and 2000, the lithology of SS attempts to be continuous in the beginning. However, when moving away from the controlled information of the "drilled" wells, the wrong prediction of VFS emerges and observations from the wells stand out, which is due to errors in the inversion results (Figure 5.3) as well as highly overlaps in the properties of lithologies (Figure 5.6). The same problem happens for SS which has been classified as VFS at the depth 300 m on the left part, even though the starting lithology in the well (CMP = 1900) is correct. Thus, the classification method highly depends on the quality of the input dataset and property difference between lithologies, which is also the case for every other classification method. If the inputs cannot provide a good description of the subsurface in terms of rock properties and structures, or if the properties of the various lithologies are highly overlapping, perfect prediction of lithologies cannot be expected, even with good geological prior information.

Since the EM algorithm converges locally, the initial model including the starting lithological section and the parameter sets such as means and covariance matrices in the GMM are important. Compared with the real case, the inversion quality is better and more "wells" are available in the synthetic example, thus a simple histogram-based method based on the inverted and known properties at the well locations is used in order to provide an initial model (Figure 5.8).

The histogram analysis helps to estimate the means and covariance matrices in the absence of prior information. According to the criteria of classification, the states should be separated widely from each other in terms of their properties, and at the same time, the intra-state variances need to be as low as possible (Figure 5.6). Other information about these parameters should be brought in, such as a regional or empirical model between the rock properties and lithologies.

In the real case study of the Vienna Basin, the inversion quality is low because of problems during the processing phase. Thus, a different approach is adopted in the form of the *k*-Means to provide the initial model (Figure 5.15). The histogram-based method used in the synthetic example has also been tested but the final result is relatively poor since only one well is available and low-quality means and covariance matrices are provided by the poor inversion at the well location.

The number of lithologies to be classified is determined from the cored wells in the cross section and has been maintained constant during the classification process in order to facilitate the prediction problem. However, some lithologies in the middle part of the section may not occur in the wells because of pinch-outs in the layers or 3-D complexities. In the synthetic case, this problem has been avoided since the well locations can be selected such that all lithologies occur in the two "wells". In real cases this is not possible and additional geological knowledge, for example from other fields or well outside the line of section, should be used to address this problem.

In the 2-D profile Markov transition matrix, the number of matrices is related to the number of states or lithologies in the system. If there are N lithologies, there will be $N(N+1)/2$ matrices which will not be easy to construct since training images are needed that may not always be available. In the synthetic and real cases presented here, in order to make the matrix construction feasible, values in the transitions are more strongly controlled by the neighbors (Appendices) and some unlikely transitions can be set to zero as simplified by the transitions between Sand and Shale in the real case study (red rectangle in Figures 5.15, 5.17 and 5.19).

The geological prior information that can be incorporated can be divided into two groups: One is the Gibbs distribution, which can be obtained from the energy function (equations (5.2) and (5.3)), and the other one is the transition matrix (Appendices). Different roles played by these two types of prior information have been shown in the synthetic example, whereby the first one tries to connect lithologies horizontally and vertically and the second one gives preferential transitions between different lithologies, while excluding unlikely transitions. The transition matrix can be derived using a general understanding of the depositional environment and the juxtapositions between different lithologies. Application of Walther's Law (Middleton 1973) can help in making the construction of these transition matrices easier. As a further research, different scenarios in terms of the Markov matrices could be designed in order to gain more confidence.

5.7. Conclusions

To conclude, in this chapter, geological prior knowledge has been introduced for the classification of reservoir lithologies by an application of the MRF which could guarantee the vertical and horizontal couplings. Compared with statistical or histogram-based methods, GMM-HMRF can help make the prediction more geologically reasonable, since lithologies are connected with each other both in the lateral and vertical directions using known preferential probabilities. In this way, a more realistic reservoir architecture will be obtained.

References:

Avseth, P., Mukerji, T., Mavko, G. 2005. *Quantitative seismic interpretation: Applying rock physics tools to reduce interpretation risk*. Cambridge University Press.

Besag, J. 1974. Spatial interaction and the statistical analysis of lattice systems. *J R Stat Soc Ser B (Methodol)*, **36**, 192–236.

Bosch, M., Carvajal, C., Rodrigues, J., Torres, A., Aldana, M., Sierra, J. 2009. Petrophysical seismic inversion conditioned to well-log data: Methods and application to a gas reservoir. *Geophysics*, **74**(2), O01–O15.

Buland, A., Kolbjørnsen, O., Hauge, R., Skjæveland, Ø., Duffaut, K. 2008. Bayesian lithology and fluid prediction from seismic prestack data. *Geophysics*, **73**(3), 13-21.

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

Eidsvik, J., Mukerji, T., Switzer, P. 2004. Estimation of geological attributes from a well log: An application of hidden Markov chains. *Mathematical Geology*, **36**(3), 379-397.

Elena, D. and Michel, M. D. 2009. Encyclopedia of Distance, Springer.

Elfeki, A. and Dekking, M. 2001. Markov chain model for subsurface characterization: theory and applications. *Mathematical Geology*, **33**(5), 569-589.

Feng R., Luthi S.M., Gisolf A. and Sharma, S. 2017. Obtaining a high-resolution geological and petrophysical model from the results of reservoir-oriented elastic wave-equation based seismic inversion. *Petroleum Geoscience*, doi: 10.1144/petgeo2015-076.

Gisolf, A., Huis in't Veld, R., Haffinger, P., Hanitzsch, C., Doulgeris, P. & Veeken, P.C.H. 2014. *Non-linear full wavefield inversion applied to carboniferous reservoirs in the Wingate Gas Field (SNS, Offshore UK)*. Presented at the 76th EAGE Annual Meeting, Amsterdam, 16-19 June.

Grana, D. and Della Rossa, E. 2010. Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion. *Geophysics*, **75**(3), 21-37.

Hammer, H., Kolbjørnsen, O., Tjelmeland, H., Buland, A. 2012. Lithology and fluid prediction from prestack seismic data using a Bayesian model with Markov process prior. *Geophysical Prospecting*, **60**, 500-515.

Houck, R. T. 2002. Quantifying the uncertainty in an AVO interpretation. *Geophysics*, **67**, 117–125.

Ising, E. 1925. Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik A Hadrons Nuclei*, **31**, 253–258.

Larsen, A.L., Ulvmoen, M., Omre, H., Buland, A. 2006. Bayesian lithology/fluid prediction and simulation on the basis of a Markov-chain prior model. *Geophysics*, **71**(5), 69-78.

Lindberg, D.V. and Grana, D. 2015. Petro-elastic log-facies classification using the expectation-maximization algorithm and hidden Markov models. *Mathematical Geosciences*, **47**, 719-752.

Middleton, G.V. 1973. Johannes Walther's law of the correlation of facies. *GSA Bulletin*, **84**(3), 979-988.

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

Rabiner, L. 1989. A tutorial on Hidden Markov Models and selected applications in speech recognition. *Proc. IEEE*, **77**(2), 257-286.

Seber G.A.F. 1984. Multivariate Observations. Hoboken, NJ: John Wiley & Sons, Inc..

Ulvmoen, M., Hammer, H. 2010. Bayesian lithology/fluid inversion — comparison of two algorithms. *Computational Geosciences*, **14**, 357-367.

Ulvmoen, M., Omre, H. 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., Omre, H., Buland, A. 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.

Wang, H., Wellmann, J.F., Li, Z., Wang, X., Liang, R.Y. 2016. A segmentation approach for stochastic geological modelling using hidden Markov random field. *Mathematical Geosciences*, doi:10.1007/s11004-016-9663-9.

Wang, Q. 2012. GMM-based hidden Markov random field for color image and 3D volume segmentation. arXiv:1212.4527 [cs.CV].

Winkler, G. 1995. Image analysis, random field and Markov chain Monte Carlo methods. Springer.

Zhang, Y., Brady, M., Smith, S. 2001. Segmentation of brain MR images through a hidden Markov random filed model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, **20**(1), 45-57.

Appendix A: Profile Markov transition matrix in synthetic example

Adopting the idea from Ulvmoen and Omre (2010) and Ulvmoen *et al*. (2010), the transition matrices of the 2-D Markov Random Field in the synthetic study are as follows:

$$P_{FS,FS} = \begin{array}{c} \\ \text{FS} \\ \text{VFS} \\ \text{SS} \\ \text{Clay} \end{array} \begin{array}{c} \begin{array}{cccc} \text{FS} & \text{VFS} & \text{SS} & \text{Clay} \end{array} \\ \begin{bmatrix} 0.9997 & 0.0001 & 0.0001 & 0.0001 \\ 0.9997 & 0.0001 & 0.0001 & 0.0001 \\ 0.9997 & 0.0001 & 0.0001 & 0.0001 \\ 0.9997 & 0.0001 & 0.0001 & 0.0001 \end{bmatrix} \end{array} \tag{5.11}$$

$$P_{FS,VFS} = \begin{array}{c} \\ \text{FS} \\ \text{VFS} \\ \text{SS} \\ \text{Clay} \end{array} \begin{array}{c} \begin{array}{cccc} \text{FS} & \text{VFS} & \text{SS} & \text{Clay} \end{array} \\ \begin{bmatrix} 0.4998 & 0.5000 & 0.0001 & 0.0001 \\ 0.4998 & 0.5000 & 0.0001 & 0.0001 \\ 0.4998 & 0.5000 & 0.0001 & 0.0001 \\ 0.4998 & 0.5000 & 0.0001 & 0.0001 \end{bmatrix} \end{array} \tag{5.12}$$

$$P_{FS,SS} = \begin{array}{c} \\ \text{FS} \\ \text{VFS} \\ \text{SS} \\ \text{Clay} \end{array} \begin{array}{c} \begin{array}{cccc} \text{FS} & \text{VFS} & \text{SS} & \text{Clay} \end{array} \\ \begin{bmatrix} 0.4998 & 0.0001 & 0.5000 & 0.0001 \\ 0.4998 & 0.0001 & 0.5000 & 0.0001 \\ 0.4998 & 0.0001 & 0.5000 & 0.0001 \\ 0.4998 & 0.0001 & 0.5000 & 0.0001 \end{bmatrix} \end{array} \tag{5.13}$$

$$P_{FS,Clay} = \begin{array}{c} \\ \text{FS} \\ \text{VFS} \\ \text{SS} \\ \text{Clay} \end{array} \begin{array}{c} \begin{array}{cccc} \text{FS} & \text{VFS} & \text{SS} & \text{Clay} \end{array} \\ \begin{bmatrix} 0.4998 & 0.0001 & 0.0001 & 0.5000 \\ 0.4998 & 0.0001 & 0.0001 & 0.5000 \\ 0.4998 & 0.0001 & 0.0001 & 0.5000 \\ 0.4998 & 0.0001 & 0.0001 & 0.5000 \end{bmatrix} \end{array} \tag{5.14}$$

$$P_{VFS,VFS} = \begin{array}{c} \\ FS \\ VFS \\ SS \\ Clay \end{array} \begin{array}{c} \begin{array}{cccc} FS & VFS & SS & Clay \end{array} \\ \begin{bmatrix} 0.0001 & 0.9997 & 0.0001 & 0.0001 \\ 0.0001 & 0.9997 & 0.0001 & 0.0001 \\ 0.0001 & 0.9997 & 0.0001 & 0.0001 \\ 0.0001 & 0.9997 & 0.0001 & 0.0001 \end{bmatrix} \end{array} \quad (5.15)$$

$$P_{VFS,SS} = \begin{array}{c} \\ FS \\ VFS \\ SS \\ Clay \end{array} \begin{array}{c} \begin{array}{cccc} FS & VFS & SS & Clay \end{array} \\ \begin{bmatrix} 0.0001 & 0.4998 & 0.5000 & 0.0001 \\ 0.0001 & 0.4998 & 0.5000 & 0.0001 \\ 0.0001 & 0.4998 & 0.5000 & 0.0001 \\ 0.0001 & 0.4998 & 0.5000 & 0.0001 \end{bmatrix} \end{array} \quad (5.16)$$

$$P_{VFS,Clay} = \begin{array}{c} \\ FS \\ VFS \\ SS \\ Clay \end{array} \begin{array}{c} \begin{array}{cccc} FS & VFS & SS & Clay \end{array} \\ \begin{bmatrix} 0.0001 & 0.4998 & 0.0001 & 0.5000 \\ 0.0001 & 0.4998 & 0.0001 & 0.5000 \\ 0.0001 & 0.4998 & 0.0001 & 0.5000 \\ 0.0001 & 0.4998 & 0.0001 & 0.5000 \end{bmatrix} \end{array} \quad (5.17)$$

$$P_{SS,SS} = \begin{array}{c} \\ FS \\ VFS \\ SS \\ Clay \end{array} \begin{array}{c} \begin{array}{cccc} FS & VFS & SS & Clay \end{array} \\ \begin{bmatrix} 0.0001 & 0.0001 & 0.9997 & 0.0001 \\ 0.0001 & 0.0001 & 0.9997 & 0.0001 \\ 0.0001 & 0.0001 & 0.9997 & 0.0001 \\ 0.0001 & 0.0001 & 0.9997 & 0.0001 \end{bmatrix} \end{array} \quad (5.18)$$

$$P_{SS,Clay} = \begin{array}{c} \\ FS \\ VFS \\ SS \\ Clay \end{array} \begin{array}{c} \begin{array}{cccc} FS & VFS & SS & Clay \end{array} \\ \begin{bmatrix} 0.0001 & 0.0001 & 0.4998 & 0.5000 \\ 0.0001 & 0.0001 & 0.4998 & 0.5000 \\ 0.0001 & 0.0001 & 0.4998 & 0.5000 \\ 0.0001 & 0.0001 & 0.4998 & 0.5000 \end{bmatrix} \end{array} \quad (5.19)$$

$$P_{\text{Clay,Clay}} = \begin{array}{c} \\ \text{FS} \\ \text{VFS} \\ \text{SS} \\ \text{Clay} \end{array} \begin{array}{c} \begin{array}{cccc} \text{FS} & \text{VFS} & \text{SS} & \text{Clay} \end{array} \\ \begin{bmatrix} 0.0001 & 0.0001 & 0.0001 & 0.9997 \\ 0.0001 & 0.0001 & 0.0001 & 0.9997 \\ 0.0001 & 0.0001 & 0.0001 & 0.9997 \\ 0.0001 & 0.0001 & 0.0001 & 0.9997 \end{bmatrix} \end{array} \tag{5.20}$$

$P_{(:,:)}$ is the transition matrix with horizontal neighbors of different lithologies (FS, VFS, SS, Clay). The vertical transitions between lithologies are controlled by the neighbors which is attributed to the lateral extension of layers.

Appendix B: Profile Markov transition matrix in the real case study

The 2-D Markov matrices in the real case study are as follows:

$$P_{\text{Shale,Shale}} = \begin{array}{c} \\ \text{Shale} \\ \text{SH\_Sand} \\ \text{Sand} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Shale} & \text{SH\_Sand} & \text{Sand} \end{array} \\ \begin{bmatrix} 0.9998 & 0.0001 & 0.0001 \\ 0.9998 & 0.0001 & 0.0001 \\ 0.9998 & 0.0001 & 0.0001 \end{bmatrix} \end{array} \tag{5.21}$$

$$P_{\text{Shale,SH\_Sand}} = \begin{array}{c} \\ \text{Shale} \\ \text{SH\_Sand} \\ \text{Sand} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Shale} & \text{SH\_Sand} & \text{Sand} \end{array} \\ \begin{bmatrix} 0.4999 & 0.5000 & 0.0001 \\ 0.4999 & 0.5000 & 0.0001 \\ 0.4999 & 0.5000 & 0.0001 \end{bmatrix} \end{array} \tag{5.22}$$

$$P_{\text{Shale,Sand}} = \begin{array}{c} \\ \text{Shale} \\ \text{SH\_Sand} \\ \text{Sand} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Shale} & \text{SH\_Sand} & \text{Sand} \end{array} \\ \begin{bmatrix} 0.4999 & 0.0001 & 0.5000 \\ 0.4999 & 0.0001 & 0.5000 \\ 0.4999 & 0.0001 & 0.5000 \end{bmatrix} \end{array} \tag{5.23}$$

$$P_{\text{SH\_Sand,SH\_Sand}} = \begin{array}{c} \\ \text{Shale} \\ \text{SH\_Sand} \\ \text{Sand} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Shale} & \text{SH\_Sand} & \text{Sand} \end{array} \\ \begin{bmatrix} 0.0001 & 0.9998 & 0.0001 \\ 0.0001 & 0.9998 & 0.0001 \\ 0.0001 & 0.9998 & 0.0001 \end{bmatrix} \end{array} \tag{5.24}$$

$$P_{\text{SH\_Sand,Sand}} = \begin{array}{c} \\ \text{Shale} \\ \text{SH\_Sand} \\ \text{Sand} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Shale} & \text{SH\_Sand} & \text{Sand} \end{array} \\ \begin{bmatrix} 0.0001 & 0.4999 & 0.5000 \\ 0.0001 & 0.4999 & 0.5000 \\ 0.0001 & 0.4999 & 0.5000 \end{bmatrix} \end{array} \tag{5.25}$$

$$P_{\text{Sand,Sand}} = \begin{array}{c} \\ \text{Shale} \\ \text{SH\_Sand} \\ \text{Sand} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Shale} & \text{SH\_Sand} & \text{Sand} \end{array} \\ \begin{bmatrix} 0.0001 & 0.0001 & 0.9998 \\ 0.0001 & 0.0001 & 0.9998 \\ 0.0001 & 0.0001 & 0.9998 \end{bmatrix} \end{array} \tag{5.26}$$

Appendix C: Modified Profile Markov transition matrix in the real case study

The 2-D modified Markov matrices of the real case study are as follows:

$$P_{Shale,Shale} = \begin{array}{c} \\ Shale \\ SH\_Sand \\ Sand \end{array} \begin{array}{c} \begin{array}{ccc} Shale & SH\_Sand & Sand \end{array} \\ \begin{bmatrix} 0.9999 & 0.0001 & 0 \\ 0.9998 & 0.0001 & 0.0001 \\ 0 & 0.5000 & 0.5000 \end{bmatrix} \end{array} \quad (5.27)$$

$$P_{Shale,SH\_Sand} = \begin{array}{c} \\ Shale \\ SH\_Sand \\ Sand \end{array} \begin{array}{c} \begin{array}{ccc} Shale & SH\_Sand & Sand \end{array} \\ \begin{bmatrix} 0.5000 & 0.5000 & 0 \\ 0.4999 & 0.5000 & 0.0001 \\ 0 & 0.9999 & 0.0001 \end{bmatrix} \end{array} \quad (5.28)$$

$$P_{Shale,Sand} = \begin{array}{c} \\ Shale \\ SH\_Sand \\ Sand \end{array} \begin{array}{c} \begin{array}{ccc} Shale & SH\_Sand & Sand \end{array} \\ \begin{bmatrix} 0.9999 & 0.0001 & 0 \\ 0.4999 & 0.0001 & 0.5000 \\ 0 & 0.0001 & 0.9999 \end{bmatrix} \end{array} \quad (5.29)$$

$$P_{SH\_Sand,SH\_Sand} = \begin{array}{c} \\ Shale \\ SH\_Sand \\ Sand \end{array} \begin{array}{c} \begin{array}{ccc} Shale & SH\_Sand & Sand \end{array} \\ \begin{bmatrix} 0.0001 & 0.9999 & 0 \\ 0.0001 & 0.9998 & 0.0001 \\ 0 & 0.9999 & 0.0001 \end{bmatrix} \end{array} \quad (5.30)$$

$$P_{SH\_Sand,Sand} = \begin{array}{c} \\ Shale \\ SH\_Sand \\ Sand \end{array} \begin{array}{c} \begin{array}{ccc} Shale & SH\_Sand & Sand \end{array} \\ \begin{bmatrix} 0.0001 & 0.9999 & 0 \\ 0.0001 & 0.4999 & 0.5000 \\ 0 & 0.5000 & 0.5000 \end{bmatrix} \end{array} \quad (5.31)$$

$$P_{\text{Sand,Sand}} = \begin{array}{c} \\ \text{Shale} \\ \text{SH\_Sand} \\ \text{Sand} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Shale} & \text{SH\_Sand} & \text{Sand} \end{array} \\ \begin{bmatrix} 0.5000 & 0.5000 & 0 \\ 0.0001 & 0.0001 & 0.9998 \\ 0 & 0.0001 & 0.9999 \end{bmatrix} \end{array} \tag{5.32}$$

6

General Conclusions

In this thesis, several approaches are investigated to determine one of the key reservoir parameters — lithology based on the inputs of seismic inversion results. Compared with well-logging or core data, which have limitations due to their sparse sampling relative to the reservoir volume, seismic surveys provide a more complete coverage over the target reservoir. From seismic inversion the rock properties, either as compressional and shear velocities or as compressibility and shear compliance, can be obtained. These properties are implicitly related to different rock types. Therefore, in this study, the inversion results are proposed as inputs for the determination of reservoir lithologies, thus circumventing the spatial limitations of well data.

However, instead of using the seismic data or inversion results directly, in Chapter 3 a new Markov Chain model (A-CMC) is proposed to simulate the reservoir lithologies based on data obtained from wells. In this method, as an extension of the Coupled Markov Chain model proposed by Elfeki and Dekking (2005), a search within a predefined angle is performed in order to account for dips in the stratal lithologies. The angle and dipping direction can be derived from the interpreted seismic data or using a boundary detection technique (Canny 1986). This method has only been applied to the synthetic and improved Book Cliffs model (Chapter 2) and the results show an improved performance compared to the Coupled Markov Chain. The method has yet to be applied to actual field data.

Subsequently, in an effort to classify reservoir lithologies based on the seismic inversion results and to avoid the spatial limitations of well data, two different classification methods are proposed in which the main difference is whether the mutual dependencies between lithologies are considered or not (Chapter 4). As a consequence of sedimentary depositional processes, reservoir lithologies in the subsurface are often characterized by typical sequences, and this relationship can be taken into account by Hidden Markov Models (HMMs). Compared with other point-wise methods such as *k*-Means and Fuzzy Logic, the prediction of lithologies by HMMs is better and spurious lithologies are often eliminated (see e.g. Figure 4.8). In HMMs, only the vertical coupling is considered, but lithologies are also related horizontally according to typical sedimentary processes.

To address this problem, a 2-D classification method is designed in which the concept of Markov Random Field is adopted (Chapter 5). Two types of prior information are incorporated in which the first one is the Gibbs function (Besag 1974; Winkler 1995) and the second one is the Markov random matrix (Ulvmoen and Omre 2010; Ulvmoen *et al*. 2010). Together they are abbreviated as GMM-HMRF here. The Gibbs function can be considered as an internal prior since it tries to connect the same or similar lithologies spatially and it is only based on the available data. The Markov random matrix, on the other hand, can introduce geological *a priori* information obtained from training images or analogue outcrop sections. Both of these obviously, have to be relevant to the reservoir being analyzed.

The results predicted by HMMs and GMM-HMRF at the well location (Figure 4.14 and Figure 5.20) are quite different, especially at the lower part. This is caused by the different parameters in the likelihood function that are obtained at the well location (1-D) and from the

whole cross section (2-D) respectively. These differences are then also evident across the whole profile.

To conclude, different methods are used in this thesis to determine the reservoir lithologies in the subsurface wherein seismic and well data play different roles (Table 6.1). Geological prior information can be incorporated in all the three methods and can be derived from local geological knowledge or from training images. Thus, the geological processes are implicitly introduced into the estimation process. The output lithologies are thought to provide constraints for other reservoir parameters such as porosity and permeability (Mahgoub *et al*. 2017). Additionally, the interpretation of trapping mechanisms might be facilitated by the results (Jennette *et al*. 2003).

Methods	Dimension	Seismic Data	Well Data
A-CMC	2-D	Weak/Mild	Strong
HMMs	1-D	Strong	Strong
GMM-HMRF	2-D	Strong	Strong

Table 6.1: Comparison between A-CMC, HMMs and GMM-HMRF and their dependence on seismic and well data.

References:

Besag, J. 1974. Spatial interaction and the statistical analysis of lattice systems. J R Stat Soc Ser B (Methodol), **36**, 192–236

Canny, J. 1986. A computational approach to edge detection. *IEEETrans. Pattern Analysis and Machine Intelligence*, **8**(6), 679-698.

Elfeki, A., Dekking, M. 2005. Modelling subsurface heterogeneity by coupled Markov chains: Directional dependency, Walther's law and entropy. *Geotechnical and Geological Engineering*, **23**, 721-756.

Jennette, D. *et al*. 2003. Traps and turbidite reservoir characteristics from a complex and evolving tectonic setting, Veracruz Basin, southeastern Mexico. *AAPG Bulletin*, **87**(10), 1599-1622.

Mahgoub, M.I., Padmanabhan, E., Abdullatif, O.M. 2017. Seismic inversion as a predictive tool for porosity and facies delineation in Paleocene fluvial/lacustrine reservoirs, Melut Basin, Sudan. *Marine and Petroleum Geology*, **86**, 213-227.

Ulvmoen, M., Omre, H. 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., Omre, H., Buland, A. 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.

Winkler, G. 1995. Image analysis, random field and Markov chain Monte Carlo methods: Springer.

Curriculum Vitae

Runhai Feng was born in 1984 in Jiangsu of the P.R. of China. He received a bachelor degree from the China University of Geosciences (Wuhan) in Prospecting Engineering. He then worked as an assistant engineer in the Jiangsu Geological Bureau for one year. His main responsibilities there included exploration, mapping and reconnaissance. He then went to Nanjing University for a master degree in Geophysics. His thesis subject was on the inference of the crustal density structure and the Moho topography in Northeastern China. In 2013 he joined the Delft University of Technology in order to pursue a PhD degree. His research is sponsored by the Delphi consortium and focuses on reservoir characterization from seismic data in which geological prior information is utilized to mitigate the lithological uncertainties. After receiving his PhD, Runhai will pursue his studies as a Postdoc for one year in Delft.

Acknowledgements

During the past four years, I have felt great pleasures in Delft and would like to express gratitude to my promotor Stefan and co-promotor Dries for their invaluable guidance and continuous support. In the first year, we had lots of discussion about the project and how to build the detailed Book Cliffs model. I got lots of help from Stefan about the building of geological and petrophysical models. Every time I had new ideas, Stefan was very encouraging and helped me to evaluate these ideas to see whether they can be implemented or not.

I also express my thanks to Dries who helped a lot in geophysics. Every Friday morning, he came to the office to have the meeting. Even though Dries is already retired, he still has lots of passions in the research. He tutored me on how to use the FWI codes which was a very enjoyable process to learn. For the new ideas or concepts of the geological priors proposed, he was very supportive and tried to help me to incorporate them into the inversion.

It was a great pleasure to talk with Eric, even though it only happened in one week before the Delphi Meeting in Houston and Denhagg. But still I am very grateful to your suggestions for the presentations as well as the comments for the Delphi reports.

A special thanks goes to Guy who is my co-promotor in the last two years. We didn't have lots of discussions on this project, but still thanks for your reviewing and suggestions on the thesis and supports during the go-no-go meeting.

As a geologist, I enjoyed a lot to be a part of the Applied Geology section. Thanks Giovanni for the helps and advices. I also thank my colleagues: Kevin, Quinten, Rémi, Thais, Navid, Santosh, Liang, Rahul, Koen, Helena, Youwei, Cees. And very special thanks goes to Siddharth who is the research partner for the past four years. It was very enjoyable to work with you.

I also would like to thank the friends from the Geophysics: Gil, Boris, Joeri, Lisanne, Iris, Max, Jianhuan, Pawan, Christian, Karlien, Ranjani, Myrna, Bingkun, Lele, Rueben, Yohei.

As a member of Delphi Consortium, I also enjoyed the meetings with the Delphi friends: Mikhail, Apostolos, Abdulrahman, Bouchaib, Hussain, Aparajita, Shogo, Jan-Willem, Özkan, Shan, Aayush, Prabu, Ewoud, Sixue, Matteo, Nick, Alok, Tomohide, Gabriel.

Special thanks go to Jan-Willem who helped me with the translation of the summary into Dutch.

This work is sponsored by the DELPHI Consortium under the project of Characterization and Monitoring of Reservoirs (C&M). The sponsors of this project are also appreciated.