Important Notice

Integration of 3C-3D seismic data and well logs for rock property estimation

by

Todor I. Todorov

M.Sc. Thesis
University of Calgary
Department of Geology and Geophysics
July, 2000

Consortium for Research in
Elastic Wave Exploration Seismology

Thesis defense committee:

Supervisor, Robert.R. Stewart, Department of Geology and Geophysics
Larry R. Bentley, Department of Geology and Geophysics
Steven E. Franklin, Department of Geography

ABSTRACT

Well log measurements and 3C-3D seismic data are integrated for rock property estimation using three methodologies: inversion, geostatistics, and multi-attribute analysis. The 3C-3D seismic data set and well logs are from the Blackfoot field, Alberta.

Conventional model-based inversion is applied to the P-P data to estimate the acoustic impedance. 3-D converted-wave (P-S) inversion for shear velocity is developed that computes a P-S weighted-stack followed by conventional inversion algorithm. An approximation formula for the incident angle in the P-S case has been derived and tested successfully versus ray-tracing. Geostatistical methods of kriging, cokriging, and stochastic simulation are used for sand-shale distribution mapping and time-to-depth conversion. Linear multi-regression and neural networks are used to derive a relationship between porosity logs and a set of seismic attributes. Porosity, sand percentage and sand thickness are used to generate a hydrocarbon volume map.

ACKNOWLEDGEMENTS

My special thanks to Dr. Rob R. Stewart for his guidance. Thanks to the staff, students, and sponsors of the Consortium for Research in Elastic Wave Exploration Seismology (CREWES). Thanks to Dan Hampson and Brian Russell for their technical assistance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1
Introduction

1.1 Introduction

Mapping the physical properties of the subsurface is of key importance in the development of the hydrocarbon reservoirs. Those properties, such as P-wave velocity, S-wave velocity, density, porosity, permeability and so on, can be measured directly at the well locations using well logging tools or core samples. However, the geological model developed by 3-D interpolation of those measurements often cannot meet the need of the development team. The reasons may include sparseness of the wells, their location, or the complexities of the geology.

The 3-D seismic survey provides more complete coverage of the development area. However, the seismic data are band-limited and contaminated with noise and phase errors.

The description and application of the methods for integration of both sources of information is the objective of this thesis.

Post-stack seismic inversion methods (Russell, 1988) provide a picture of the acoustic impedance. Pre-stack (AVO) inversion techniques (Russell, 1988) attempt to determine the elastic properties of the subsurface. Both methods heavily rely on a theoretically derived relationship between the physical property and the seismic amplitude.

However, the influence of some properties, such as porosity and permeability, on the propagating elastic waves is complex and non-unique. Therefore great difficulty can arise in the attempting to develop a theoretical model. To overcome this problem, we may use statistical methods to derive relationships based on a particular data set. In the sixties, the mining geologists started to explore the spatial correlation between measured ore locations. An interpolation method, called kriging, based on the spatial correlation was developed. The method has been extended to incorporate a second variable and the method of cokriging was born. The cokriging method is able to integrate the sparse well

measurement and the dense seismic data to interpolate a desired rock property. Currently, the method is used widely for mapping of reservoir properties.

Regression analysis is used routinely to derive porosity from acoustic impedance inversion. In this thesis, I go beyond the concept of simple cross plotting. I use linear multi-regression to derive a relationship between a particular measured property at the well locations and some seismic attributes. Once found, this relationship is applied to the seismic volume and a cube of the desired rock property is generated. The new technology of artificial intelligence, or neural networks, is used to derive a non-linear relationship between the rock properties and the seismic attributes.

The invention of the three-component (3-C) seismic measurement put new challenges in front of the geophysical research. A 3-D converted-wave (P-S) model-based inversion algorithm is developed and tested. A model-based conversion of P-S data to P-P time is described.

1.2 Blackfoot 3C-3D data set

1.2.1 Geology of the Blackfoot field

The Blackfoot field is located about 15 km southeast of the town of Strathmore, Alberta, Canada, in Township 23, Range 23. A simplified chart of the stratigraphy of the study area is shown in Figure 1.1 (Miller, 1996). Within the study area, the Mannville Group unconformably overlies the Mississippian carbonates. The target rocks are incised, valley-fill sediments within the Glauconitic formation. Numerous Glauconitic incised valleys are presented in southern Alberta. The valleys are cut to varying depths through the underlying strata and thus their bases may be found directly over or within any one of the Ostracod, Sunburst, or Detrital formations. The Glauconitic member consists of very fine to medium grained quartz sandstone. The Ostracod beds underlying the Glauconitic are made up of brackish water shales, argillaceous, fossilliferous limestones and thin quartz sandstones and siltstones. The thin, low velocity Bantry Shale member underlies the Ostracod but is not laterally persistent. The Sunburst member contains ribbon and sheet sandstones made up of sub-litharenites and quartzarenites. The Detrital formation

has a highly heterogeneous lithology containing chert pebbles, lithic sandstone, siltstone and abundant shale.

In the Blackfoot area, the Glauconitic sandstone is encountered at a depth of approximately 1550 m and the valley-fill sediments vary from 0 to over 35 m in thickness. It is subdivided into three units corresponding to three phases of valley incision but they are not presented everywhere. The lower and upper members are made up of quartz sandstones with an average porosity of approximately 18%, while the middle member is tight lithic sandstone. The individual members range in thickness from 0 to 20 m. Hydrocarbon reservoirs are found in structural and stratigraphic traps where porous channel sands pinch out against non-reservoir regional strata or low-porosity channel sediments. The primary hydrocarbon at the field is oil, although some gas may be encountered.

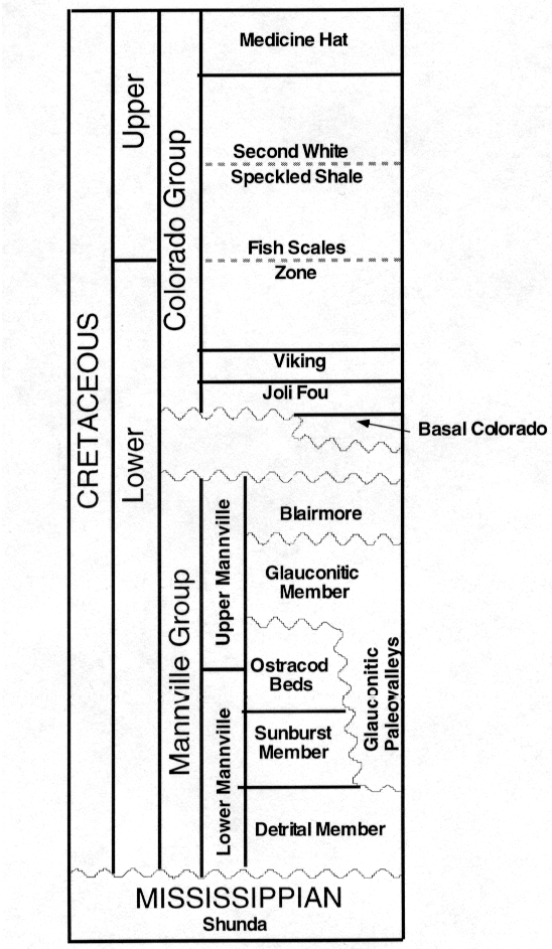

Figure 1.1: Stratigraphy chart of the Blackfoot area (from Miller, 1996).

1.2.2 Acquisition and processing of the Blackfoot 3C-3D data set

The design of the Blackfoot 3C-3D survey was in part based on the results obtained from the 3C-2D survey which was recorded in the area of the 3D survey in 1995 (Lawton et al., 1996). The acquisition geometry was established in order to maintain the number of source points to less than 1400, and record an active patch of up to 700 geophones (2100 channels). The acquisition parameters used in the final design were:

Source parameters:

Line orientation: north-south

Source interval: 60 m.

Source line interval: 210 m.

Number of source lines 24

Total number of source points: 1395

Receiver parameters:

Line orientation: east - west

Receiver interval: 60 m.

Receiver line interval: 255 m. (Glauconitic)

495 m. (Beaverhill Lake)

Number of receiver lines: 18

Total number of receivers: 903

An additional benefit of this geometry was that it provided smooth asymptotic fold for P-S data using the standard 30 x 30 m bin dimensions, with an average fold of 36 at the Glauconitic level.

The recorded 3C-3D seismic data were processed by the CREWES Project, University of Calgary (Lu and Margrave, 1998). The seismic processing software package "ProMAX" was used for data processing. The shot gathers were separated into vertical, north (H2), and east (H1) components. Deconvolution tests were performed including spiking deconvolution and surface consistent deconvolution with different operator lengths and pre-whitening parameters. Processing with or without refraction static correction and with or without spectral whitening were also tested. After evaluation, the chosen processing flow for the vertical component is as follows:

SEG-D FORMATTED DEMULTIPLEX INPUT
3D GEOMETRY ASSIGNMENT
TRACE EDITS
TRUE AMPLITUDE RECOVARY
SURACE CONSISTENT DECONVOLUTION
TIME VARIANT SPECTRAL WHITENINIG
ELEVATION AND REFRACTION STATIC CORRECTION
VELOCITY ANALYSIS
RESUDIAL SURFACE CONSISTENT STATICS
NORMAL MOVEOUT
FRONT END MUTING
CDP STACK
TIME VARIANT SPECTRAL WHITENING

TRACE EQUALIZATION
F-XY DECONVOLUTION
3D PHASE-SHIFT MIGRATION

Velocity analysis was performed using a grid of 600 m by 600 m. Phase-shift 3D migration with 100 % stacking velocities was applied.

The processing flow for the radial component is as follows:

SEG-D FORMATTED DEMULTIPLEX INPUT
COMPONENT SEPARATION
3D GEOMETRY ASSIGNMENT
TRACE EDITS
ASYMPTOTIC BINNING
SURFACE CONSISTENT DECONVOLUTION
TIME VARIANT SPECTRAL WHITENING
ELEVATION STATICS
APPLY FINAL REFRACTION AND RESIDUAL STATICS FROM P-P
CONSTRUCT INITIAL P-SV VELOCITY FROM FINAL P-P VELOCITY
VELOCITY ANALYSIS
RECEIVER RESIDUAL STATICS (HAND STATICS)
VELOCITY ANALYSIS
CONVENTIONAL RESIDUAL STATICS
VELOCITY ANALYSIS
NORMAL MOVEOUT
ACP TRIM STATICS
FRONT END MUTING
ACP STACK (DEPTH-VARIANT STACK AND DMO STACK)
TIME VARIANT SPECTRAL WHITENING
TRACE EQUALIZATION
F-XY DECONVOLUTION
3D PHASE-SHIFT MIGRATION

The conversion point binning was performed by the approximate binning method (Harrison, 1992), using an average Vp/Vs. The same Vp/Vs was used to construct the initial P-S stacking velocity. Velocity analysis was performed using a grid of 600 m by 300 m. After stacking, Vp/Vs for different time windows was extracted.

1.3 Hardware and software used

The work presented in this thesis was created on a *Sun Microsystems* network operated by the CREWES Project of the Department of Geology and Geophysics at the University of Calgary. The well log data were stored using *Geoview* from Hampson-Russell Software Ltd. The well logs were edited using *Matlab*. The P-P inversion was performed using *Strata* from Hampson-Russell Software Ltd. The P-S inversion algorithm was coded using the 'trace math' option in *Pro3D* software package from Hampson-Russell Software Ltd. The geostatistical analysis was performed using *Geostat* package and the well log prediction was done using *Emerge* from Hampson-Russell Software Ltd. The images in this thesis were screen captured using *XV*. *Microsoft Word* was used for word and image processing and thesis assembly.

Chapter 2
P-P and P-S inversion of the Blackfoot 3C-3D data set

2.1 Introduction

Seismic inversion can be described as 'a procedure for obtaining models which adequately describe a data set' (Treitel et al., 1988). In the case of seismic exploration, our recorded seismic traces show the effects of the rock physical properties on elastic waves propagating through the earth. The inversion process allows us to estimate these physical properties and so is of great interest to a geophysicist.

The inversion process often relies on forward modeling, which uses a mathematical relationship to generate the earth's response for a given set of model parameters. For example, we can generate a synthetic seismogram using the elastic wave equation and a model containing the wave velocity and density parameters. The inversion process can be seen as a 'reverse' of the forward modeling: for a given data set, find a model, which reproduces the observations.

We can denote the forward modeling process as a transformation: $s = F(x)$, where s is the model response, x is a vector containing model parameters, and F is the mathematical transformation which describes the physical process. Then, the process of inversion can be written as: $x' = F^{-1}(y)$, where x' is the set of estimated parameters derived from the data y and F^{-1} is the inverse transformation.

However, there are some difficulties, which make the inversion a challenging task. First, we have to describe adequately the physical phenomena by the mathematical transform F. Even, if we do so, F^{-1} may not exist. Often, there is more than one solution to $x' = F^{-1}(y)$, i.e. the solution is not unique. Another problem arises from the fact that the geophysical recordings are inevitably corrupted by noise. Unfortunately, noise can cause wide variations or instabilities in estimates of the model parameters and can destroy solution validity. Despite these difficulties, inversion has been successfully used to extract information from geophysical data (Lindseth, 1979; Cooke and Schneider, 1983; Oldenburg et al., 1983). Lines and Treitel (1984) give an excellent tutorial of the least-

squares inversion and its application. A comprehensive mathematical treatment of the problem can be found in Tarantola (1987).

The new technology of multi-component (3-C or 4-C) seismic measurement put new challenges in front of the researchers: developing algorithms and software for converted-wave (P-S) inversion. The method is of great interest since the changes in the P-wave reflectivity have only a partial dependence on the shear-wave velocity. By contrast, P-S reflectivity is more dependent on S-wave velocity. Stewart (1990) proposed a method for a joint P-P and P-S inversion based on the weighted-stacking technique. Vestrum and Stewart (1993) used synthetic data to show that the joint P-P and P-S inversion is effective in predicting the relative S-wave and P-wave velocities. Ferguson (1996) discusses the problem further and applies the method to the Blackfoot field data.

In the current chapter, a model-based inversion is performed to the Blackfoot P-P data using the software package 'STRATA' from Hampson-Russell Software Ltd. A 3-D model-based P-S inversion algorithm for estimating shear velocity is developed and applied to the P-S data from the Blackfoot area.

2.2 P-P inversion

2.2.1 Methods

Suppose that we have some initial guess, or estimate, of the model, characterized by the reflection coefficients, r^*_i, i = 1, ...,N. We could then calculate the model trace, M, using the convolutional model:

$$\mathbf{M_i} = \sum_{\mathbf{j=1}}^{\mathbf{N}} \mathbf{r^*_j w_{i-j+1}} \tag{2.1}$$

where r is the modeled earth reflectivity and w is the seismic wavelet.

This model trace would differ from the recorded trace, T, for two reasons. First the reflectivity r^* is different from the true value, r, and second, the recorded trace contains measurement noise. We may use least-squares optimization to find that value of r^* which makes the difference between T and M as small as possible. We may define the error trace as $E_i = T_i - M_i$, i = 1, ...,N.

Assume that the correct reflectivity is written as $r_i = r^*_i + \Delta r_i$, i = 1, ...,N. Then we wish to find the correction such that the squared error between the recorded trace and the final modeled trace is minimized:

$$J = \sum_{i=1}^{N} E^2{}_i = \sum_{i=1}^{N} \left(T_i - \sum_{j=1}^{N} \left(r^*{}_j + \Delta r_j \right) w(i-j+1) \right)^2 = \sum_{i=1}^{N} \left(E(i) - \sum \Delta r(j) w(i-j+1) \right)^2 \tag{2.2}$$

There is one set of reflection coefficients which minimizes the error J. A 'non-uniqueness' comes from the fact that there may be some other combinations of reflection coefficients which produce models almost as good as the one derived. One way to distinguish between a set of possible solutions is to use a constraint, which sets absolute boundaries on how far the final answer may deviate from the initial guess.

The discussed method is called 'constrained blocky inversion' in STRATA. Figure 2.1 describes the full flowchart of the inversion process in the program.

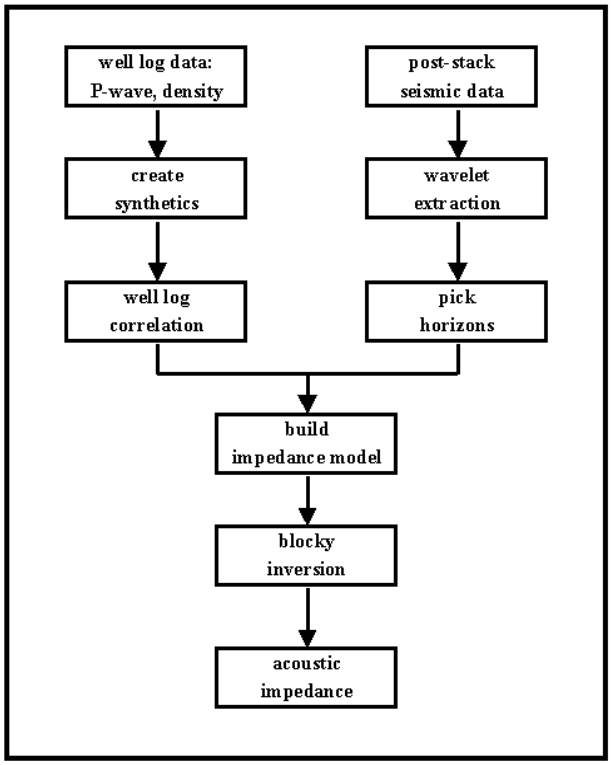

Figure 2.1: Post-stack inversion flowchart.

2.2.2 Inversion of the P-P Blackfoot seismic data

The inversion algorithms require information about the seismic wavelet to perform inversion. In the frequency domain, the wavelet extraction problem consists of two parts:

- determine the amplitude spectrum
- determine the phase spectrum

Determining the phase spectrum is the more difficult of the two parts and presents a major source of error in inversion. The extraction methods fall into three categories:

a) deterministic

   The wavelet is measured directly.

b) statistical

   The method estimates the wavelet from the seismic data alone. The method can not determine the phase spectrum reliably and must be supplied as a separate parameter. The amplitude spectrum is calculated as follows:

   - calculate the autocorrelation over a chosen window
   - calculate the amplitude spectrum of the autocorrelation
   - take the square root of the autocorrelation spectrum which approximates the amplitude spectrum of the wavelet
   - add the desired phase (zero, constant, minimum)
   - take the inverse FFT to produce the wavelet

c) using a well log

   The method combines the well log and seismic information. In theory, the method could provide exact phase information at the well locations. However, the method depends critically on the depth-to-time conversion and mis-ties degrade the result significantly.

The statistical wavelet extraction procedure was used to extract the wavelet with the following parameters (Figure 2.2):

- phase = 0
- start time = 800 ms
- end time = 1300 ms
- inline 70 to 120

- xline 110 to 140
- length 80 ms

Once the wavelet is extracted, well log correlation is performed. For each well, the process involves a synthetic trace generation and its comparison to the real data trace. Stretching and squeezing is applied to align the known events. Figure 2.3 shows the correlated 08-08 well.

The next step in the inversion process is to build an initial guess model by 3-D interpolation of the impedance logs (the multiplication of the sonic and density logs). Picked seismic horizons are used to introduce structural information in the interpolation. Figure 2.4 is a cross-section of the model so-built and Figure 2.5 is the average impedance from the model for the channel interval.

The ‘constrained blocky inversion' procedure is applied to the seismic data with the following parameters:

- maximum impedance change (deviation) from the model: 35%
- number of iterations: 10
- separate scaler for each trace

Figure 2.6 is a cross-section of the inversion result and Figure 2.7 is the average impedance of the channel interval. The known sand channel (oil wells 01-08, 08-08, 09-08) correlates with low-impedance values. The shale-plugged channel (dry well 12-16) is distinguished as high-impedance anomaly. The regional 09-17 is located in a relatively low-impedance area, which means that the P-P inversion result may be ambiguous in discriminating the sand channel from the regional geology.

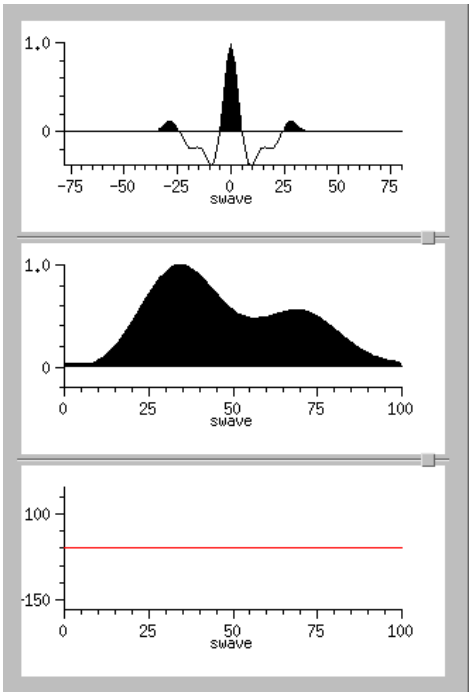

Figure 2.2: Extracted zero-phase wavelet.

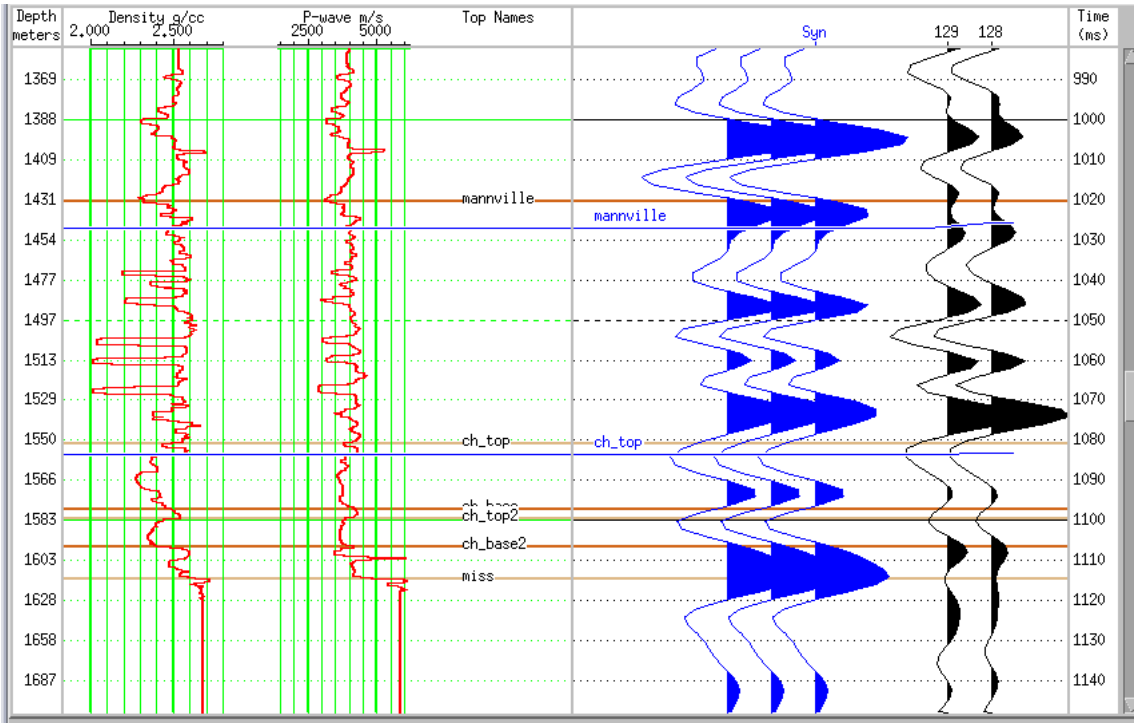

Figure 2.3: 08-08 well correlation.

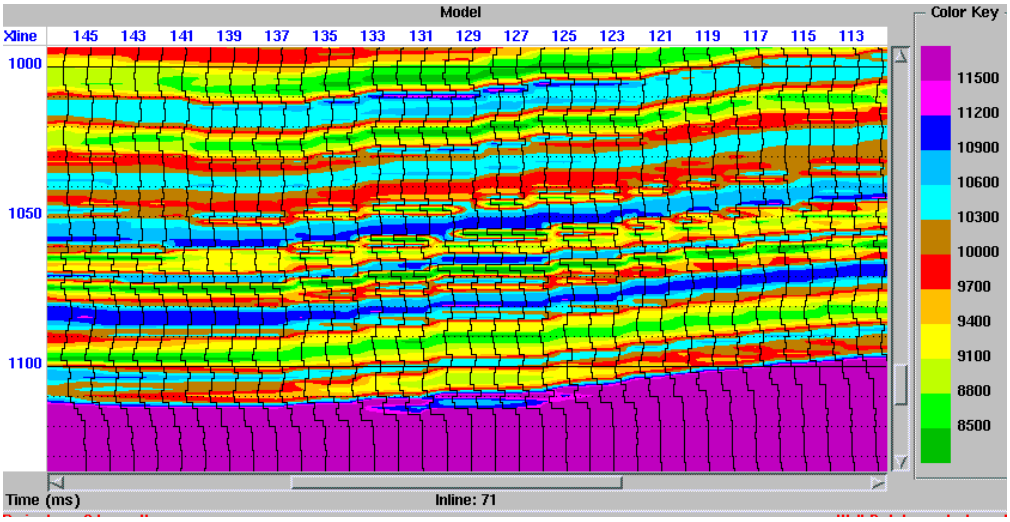

Figure 2.4: Cross-section of the initial impedance model derived by well log interpolation.

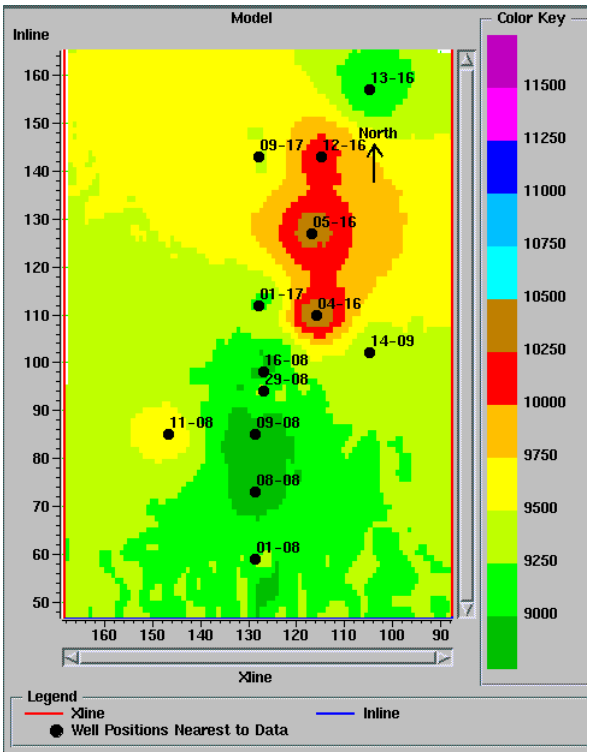

Figure 2.5: Average impedance for the channel interval from the initial model derived by well log interpolation.

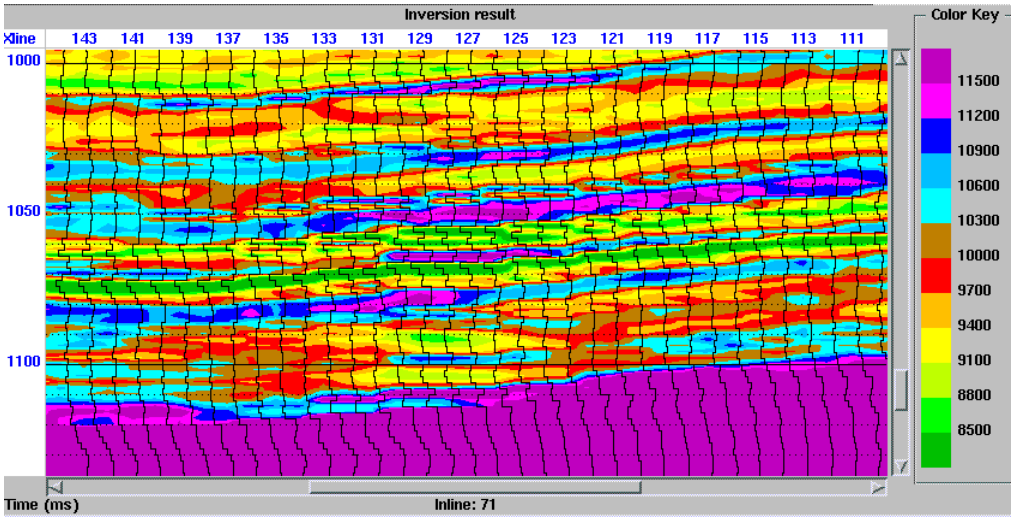

Figure 2.6: Cross-section of the inversion result. The sand channel is visible as low impedance anomaly at around 1100 ms.

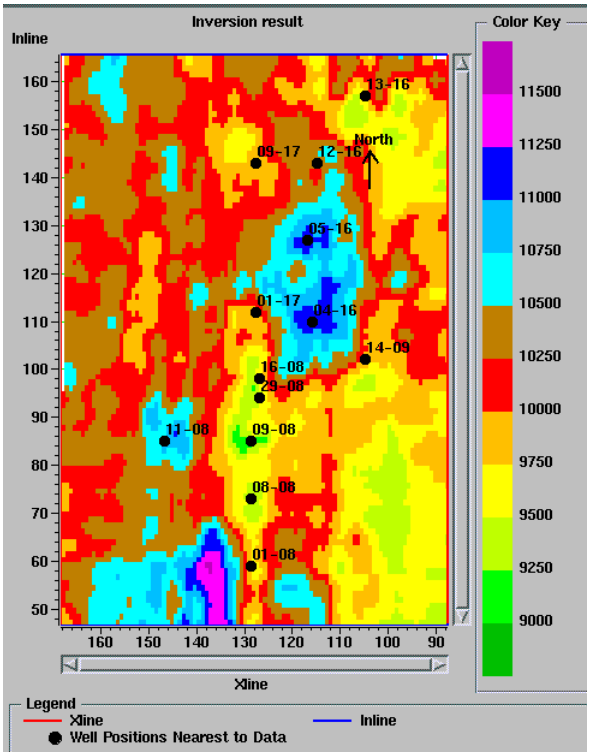

Figure 2.7: Inversion result, average impedance for the channel level. The oil wells are located in a low-impedance anomaly.

2.3 P-S inversion for shear velocity

2.3.1 Methods

2.3.1.1 P-S weighted stack

The Zoeppritz equations (Aki and Richards, 1980) allow us to derive the exact plane wave amplitude of a reflected, converted S-wave from an incident P-wave as a function of angle, but do not give us an intuitive understanding of how this amplitudes relate to the various physical parameters. Aki and Richards (1980) approximate the equation assuming small changes in elastic properties across an interface (Figure 2.8):

$$R^{PS} = c\frac{\Delta\rho}{\rho} + d\frac{\Delta\beta}{\beta} \tag{2.3}$$

where:

$$c = -\frac{\alpha\tan\varphi}{2\beta}\left(1 - \frac{2\beta^2}{\alpha^2}\sin^2\theta + \frac{2\beta}{\alpha}\cos\theta\cos\varphi\right) \tag{2.4}$$

$$d = \frac{\alpha\tan\varphi}{2\beta}\left(\frac{4\beta^2}{\alpha^2}\sin^2\theta - \frac{4\beta}{\alpha}\cos\theta\cos\varphi\right) \tag{2.5}$$

$\theta = (\theta_i + \theta_{i+1})/2,\ \varphi = (\varphi_i + \varphi_{i+1})/2$ - average P and S angles across the interface

α, β, ρ - average P-wave velocity, S-wave velocity, and density across the interface

$\Delta\beta/\beta, \Delta\rho/\rho$ - relative changes in S-wave velocity and density

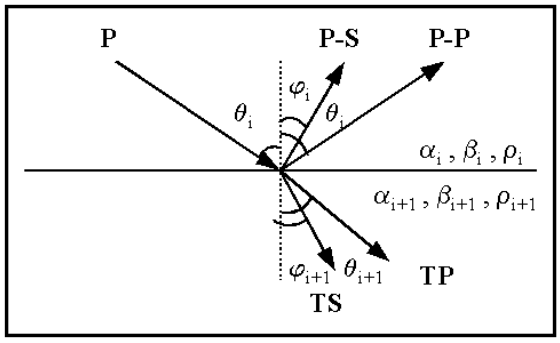

Figure 2.8: Incident P-wave partitioning at an interface. The reflected P-wave is denoted as P-P, the reflected S-wave as P-S, the transmitted P-wave as TP, and the transmitted S-wave as TS.

Equation (2.3) can be cast as a least-squares problem and solved for $\Delta\beta/\beta$ (Stewart, 1990). The sum of the squares of the error at a single interface is:

$$\varepsilon = \sum_{i=1}^{n} \left(R^{PS}{}_i - R_i \right)^2 \tag{2.6}$$

where R^{PS} is the recorded P-S reflectivity, R is the modeled P-S reflectivity and the summation is over the offsets in a seismic gather.

Equation 2.6 can be expanded:

$$\varepsilon = \sum_{i=1}^{n} \left(R^{PS}{}_i{}^2 - 2R^{PS} R_i + R_i{}^2 \right) =$$

$$\sum_{i=1}^{n} \left(R^{PS}{}_i{}^2 - 2R^{PS}{}_i \left(c_i \frac{\Delta\rho}{\rho} + d_i \frac{\Delta\beta}{\beta} \right) + \left(c_i \frac{\Delta\rho}{\rho} + d_i \frac{\Delta\beta}{\beta} \right)^2 \right) =$$

$$\sum_{i=1}^{n} \left(R^{PS}{}_i{}^2 - 2R^{PS}{}_i c_i \frac{\Delta\rho}{\rho} - 2R^{PS}{}_i d_i \frac{\Delta\beta}{\beta} + c_i{}^2 \left(\frac{\Delta\rho}{\rho} \right)^2 + 2c_i d_i \frac{\Delta\rho}{\rho} \frac{\Delta\beta}{\beta} + d_i{}^2 \left(\frac{\Delta\beta}{\beta} \right)^2 \right) \tag{2.7}$$

To find the value of $\Delta\beta/\beta$ that minimizes the error function ε, differentiate with respect to $\Delta\beta/\beta$:

$$\frac{\partial \varepsilon}{\partial \frac{\Delta\beta}{\beta}} = \sum_{i=1}^{n} \left(-2R^{PS}{}_i d_i + 2c_i d_i \frac{\Delta\rho}{\rho} + 2d_i{}^2 \frac{\Delta\beta}{\beta} \right) = 0 \tag{2.8}$$

Solving for $\Delta\beta/\beta$:

$$\frac{\Delta\beta}{\beta} = \frac{\sum_{i=1}^{n} R^{PS}{}_i d_i - \frac{\Delta\rho}{\rho} \sum_{i=1}^{n} c_i d_i}{\sum_{i=1}^{n} d_i{}^2} \tag{2.9}$$

P-wave, S-wave and density models in P-S time are required to obtain the $\Delta\beta/\beta$ weighted-stack.

2.3.1.2 Modeling

To create the P-S weighted-stack, we need a geological model containing P-wave velocity, S-wave velocity, and density. The model, in P-S time, can be built in the following way:

- at the well locations, compute the P-S pseudo-velocity logs, defined by:

$$\text{P - S pseudo - velocity log} = \mathbf{2}\left(\frac{V_P V_S}{V_P + V_S}\right) \tag{2.10}$$

  where V_P and V_S are the measured P-wave and S-wave velocity logs
- using the computed P-S pseudo-velocity log convert the V_P, V_S, and density logs into P-S time
- build 3-D V_P, V_S, and density volumes in P-S time by 3-D interpolation

2.3.1.3 Incident angle approximation

The $\Delta\beta/\beta$ weighted-stack calculation requires knowledge of the incident angle at any particular interface (the reflection and transmission angles can be found using Snell's law). The incident angle can be found using ray tracing, but in a complex model, as the one discussed above, the required time may be large and thus unattractive. The problem can be solved by deriving an approximation for the incident angle as a function of the offset (Todorov and Stewart, 1998).

First, we can look at the approximation of the incident angle for the P-P case (Figure 2.9).

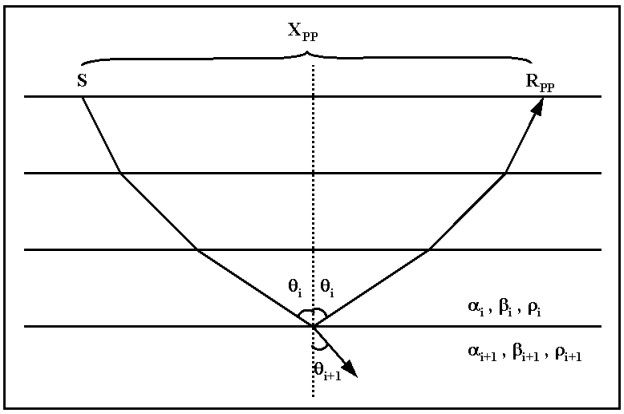

Figure 2.9: The raypath of a P-P wave in a horizontally layered medium.

The total two-way P-P travel time t_{PP} is:

$$t_{PP}^{\ 2} = t_{PP0}^{\ 2} + \frac{X_{PP}^{\ 2}}{\alpha_{RMS}^{\ 2}} \tag{2.10}$$

From Snell's law, $p = \sin\theta_i/\alpha_i$ (p-ray parameter) and since $p = dt/dx$ at any point, the incident angle to the interface i can be written as:

$$\sin\theta_i = \alpha_i \frac{dt}{dx} \tag{2.11}$$

Substitute the two-way travel time in equation (2.11) and solve:

$$\sin\theta_i = \alpha_i \frac{d}{dx}\left(\sqrt{t_{PP0}^{\ 2} + \frac{X_{PP}^{\ 2}}{\alpha_{RMS}^{\ 2}}}\right) = \frac{\alpha_i X_{PP}}{t_{PP}\alpha_{RMS}^{\ 2}} \tag{2.12}$$

Now let's look at the P-S case (Figure 2.10).

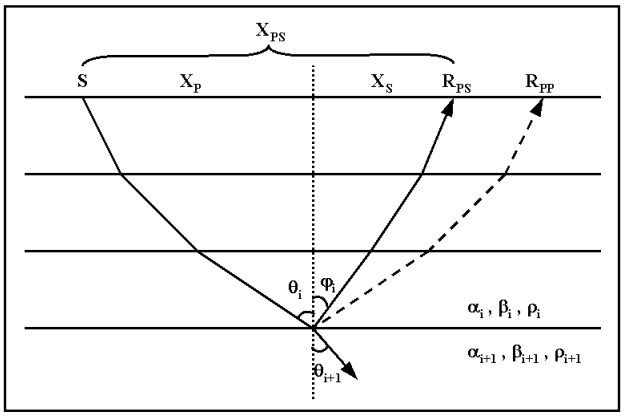

Figure 2.10: The raypath of a P-S wave.

From the P-P case, for the same incident angle:

$$\sin\theta_i = \frac{\alpha_i X_{PP}}{t_{PP}\alpha_{RMS}^{\,2}} = \frac{2X_P\alpha_i}{\alpha_{RMS}^{\,2}\sqrt{t_{PP0}^{\,2} + \frac{4X_P^{\,2}}{\alpha_{RMS}^{\,2}}}} \tag{2.13}$$

At zero offset:

$$t_{PS0} = t_P + t_S = t_P\left(1 + \frac{t_S}{t_P}\right) = \frac{t_{PP0}}{2}\left(1 + \frac{\overline{\alpha}}{\overline{\beta}}\right) = \frac{\overline{\alpha} + \overline{\beta}}{2\overline{\beta}}t_{PP0} \tag{2.14}$$

where: $\overline{\alpha}$, $\overline{\beta}$ - average P-wave and S-wave velocities.

Furthermore, from Tatham and McCormick (1991), we can convert the P-P offset, $X_{PP}=2X_P$, to P-S offset, $X_{PS}=X_P+X_S$:

$$X_P = gX_{PS}, \text{ where } g = \frac{1}{1 + \frac{\overline{\alpha}}{\overline{\beta}}\frac{\beta_{RMS}^{\,2}}{\alpha_{RMS}^{\,2}}} \tag{2.15}$$

And then, we write the approximation for the incident angle in the P-S case:

$$\sin\theta_i = \frac{2gX_{PS}\alpha_i}{\alpha_{RMS}^2\sqrt{\left(\frac{2\bar{\beta}}{\bar{\alpha}+\bar{\beta}}t_{PS0}\right)^2 + \frac{4g^2X_{PS}^2}{\alpha_{RMS}^2}}} \tag{2.16}$$

The angle goes into equation (2.9) to calculate the weights for the $\Delta\beta/\beta$ stack.

Java 2 was used to write a computer program to compare estimated incidence angles using ray tracing (Snell's law) and equation 2.16. Figure 2.11 shows the result using P-wave and S-wave velocities at the 08-08 location and P-S offset of 1230 m.

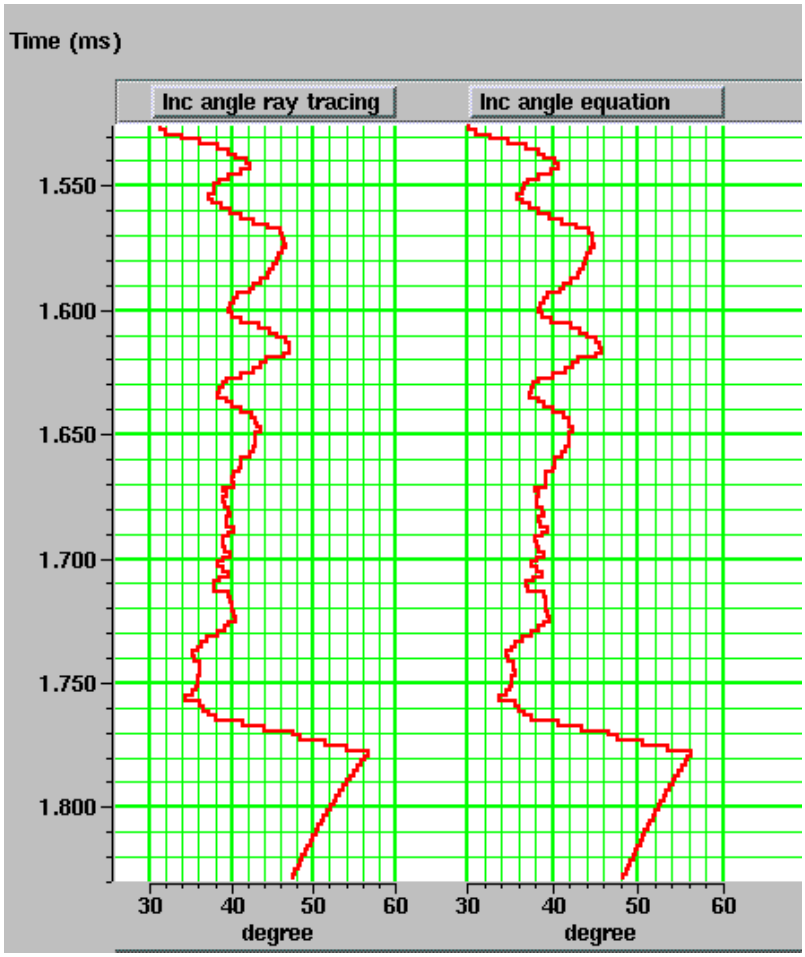

Figure 2.11: Incident angles computed using ray tracing and equation 2.16.

Now I will describe the P-S inversion flow. It begins with building the geological model in P-S time, containing P-wave, S-wave and density information for each seismic sample. Then using the model, we calculate the stacking weights for each NMO-corrected CCP gather and perform weighted stacking. The resulting $\Delta\beta/\beta$ volume can be inverted using any available P-P inversion routine to derive the shear velocity. Figure 2.12 shows the P-S inversion flowchart.

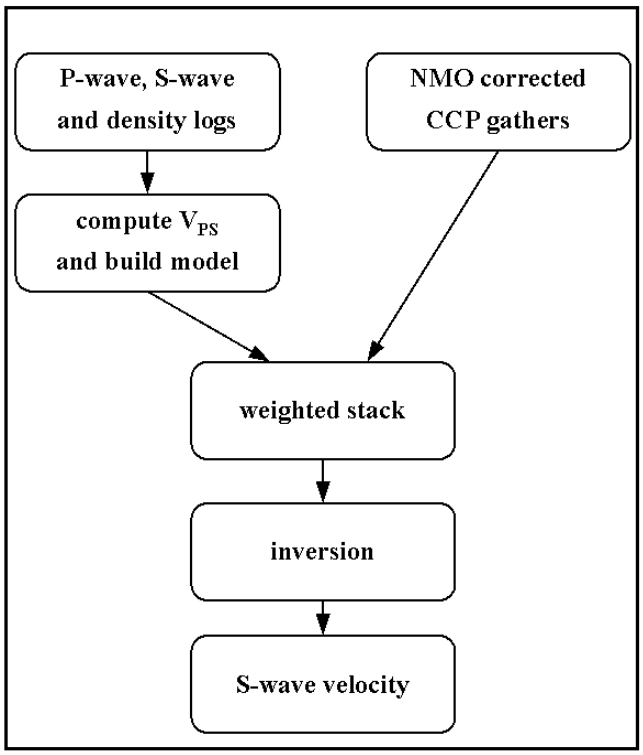

Figure 2.12: Converted-wave (P-S) inversion flowchart.

2.3.2 Inversion of the P-S Blackfoot data set

The S-wave velocity log has been measured at four well locations in the Blackfoot field: 08-08, 12-16, 09-17, and 04-16. Using equation (2.10) the P-S pseudo-velocity logs are computed for the later wells. Figure 2.13 shows the P-wave, S-wave, and computed P-S logs for 08-08 well. Note that the time on the right side is the P-S time derived from the P-S pseudo-velocity log. 3-D P-wave, S-wave, and density models are built in P-S time by 3-D interpolation (Figures 2.14 - 2.16). The computed P-S velocity logs are used for the depth-time mapping. Using equation (2.16) the incident angles for ten P-S offsets intervals, from 670 m to 1930 m, are computed. Figure 2.17 is a cross-section from the angle's volume for an offset of 1230 m. Using equations (2.4) and (2.5) the required c-weights and d-weights are computed. Figures 2.18 and 2.19 show cross-sections from the volumes for the offset of 1230 m. The computed weights are applied to the models and the seismic data, and the weighted-stack volume is computed. A 180-degree phase shift (polarity reversal) is applied to the weighted-stack before the inversion algorithm. Figure 2.20 is the derived weighted-stack from the NMO-corrected P-S CCP gathers.

Using the already discussed model-based inversion technique for zero-offset P-P inversion, the S-wave velocity volume is generated. The previously built S-wave model in P-S time is used as an initial guess model (Figure 2.15). Figure 2.21 is a cross-section from the inverted volume. Figure 2.22 is an S-wave data slice at the channel level.

As expected, the producing oil well, 08-08, is located in a high S-wave velocity anomaly. The regional well 09-17 is located in a low S-wave velocity anomaly, while shale-plugged 12-16 has mid-range S-wave velocity.

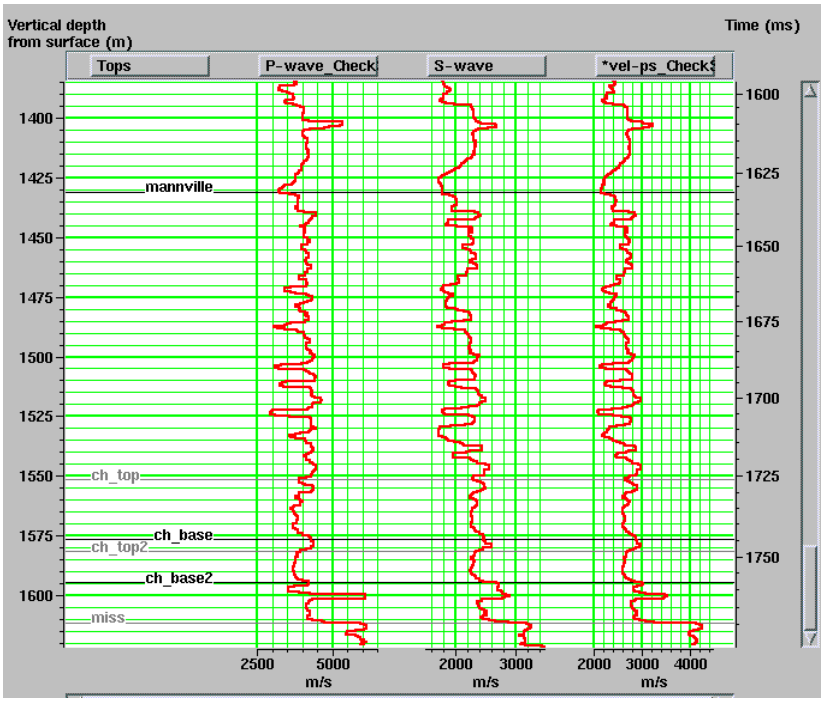

Figure 2.13: P-S pseudo-velocity log.

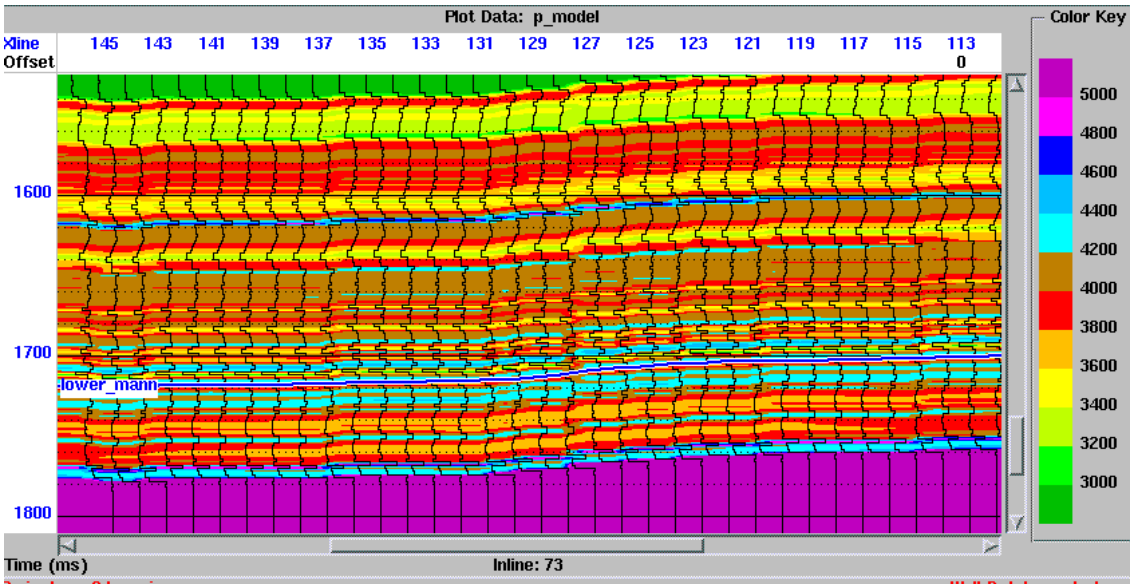

Figure 2.14: Cross-section of the P-wave velocity model in P-S time.

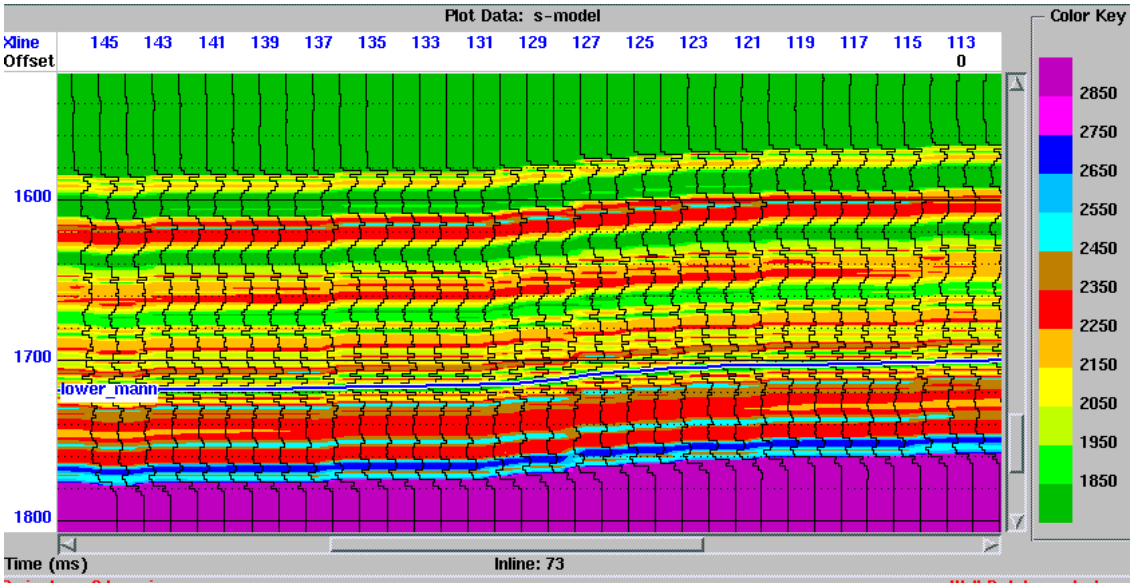

Figure 2.15: Cross-section of the S-wave velocity model in P-S time.

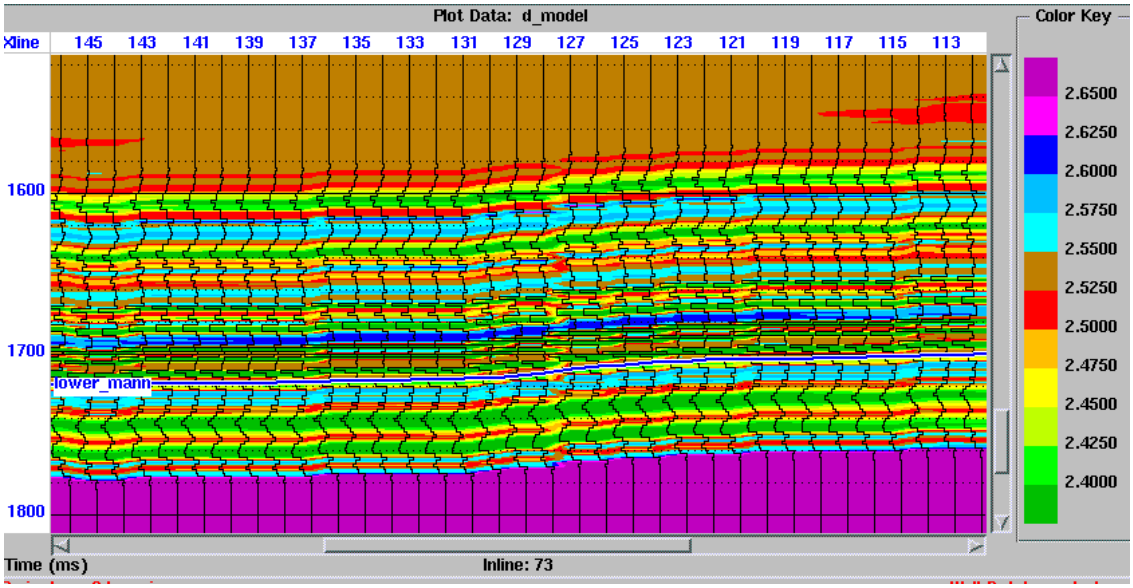

Figure 2.16: Cross-section of the density model in P-S time.

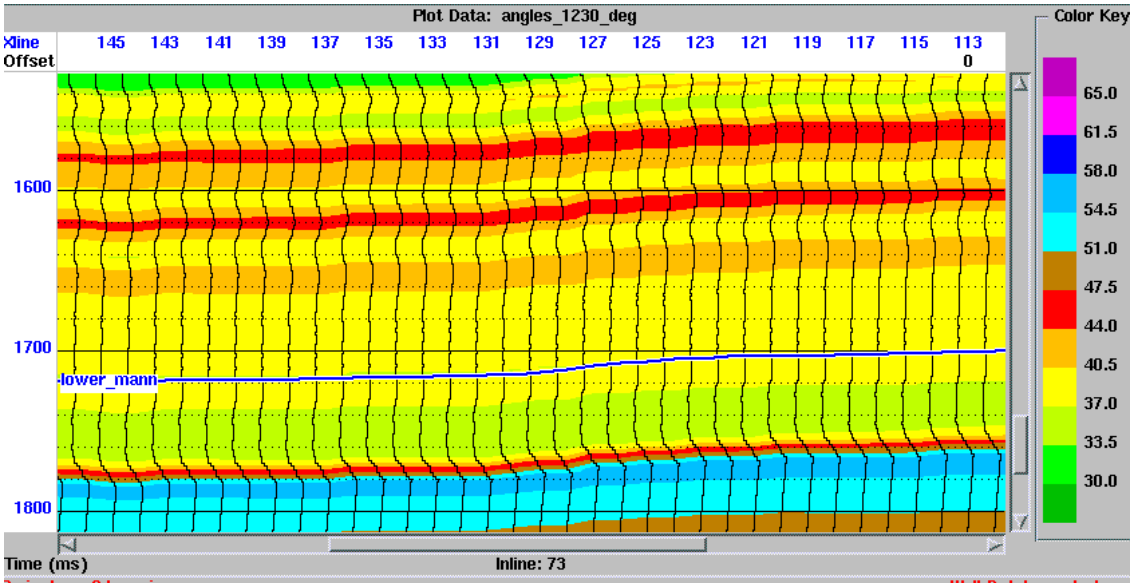

Figure 2.17: Cross-section of the computed incident angles for an offset of 1230 m.

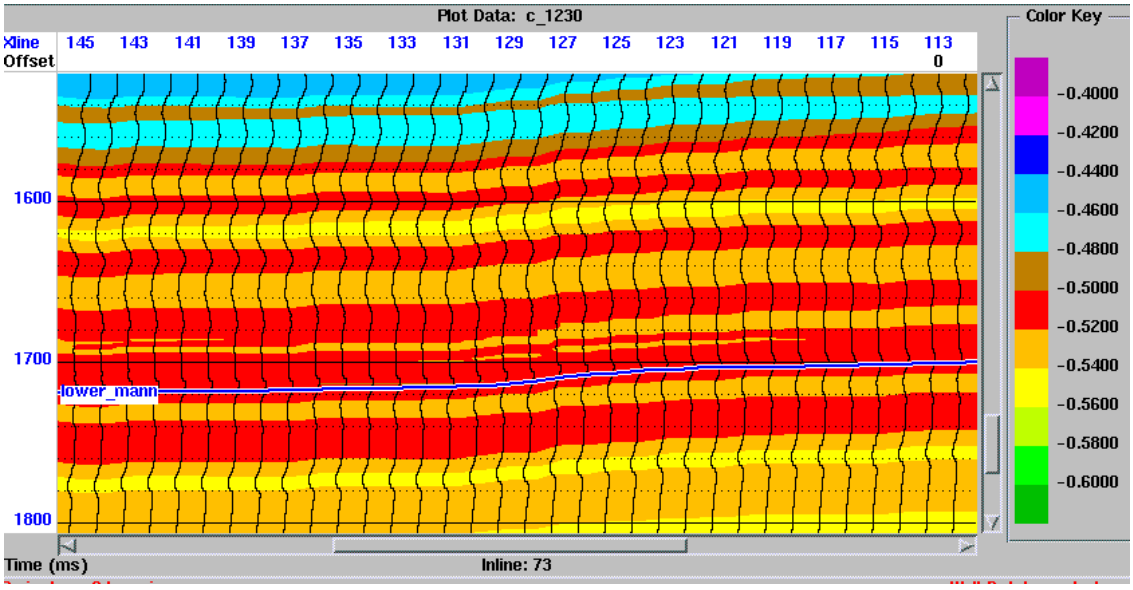

Figure 2.18: Cross-section of the computed c-weights for an offset of 1230 m.

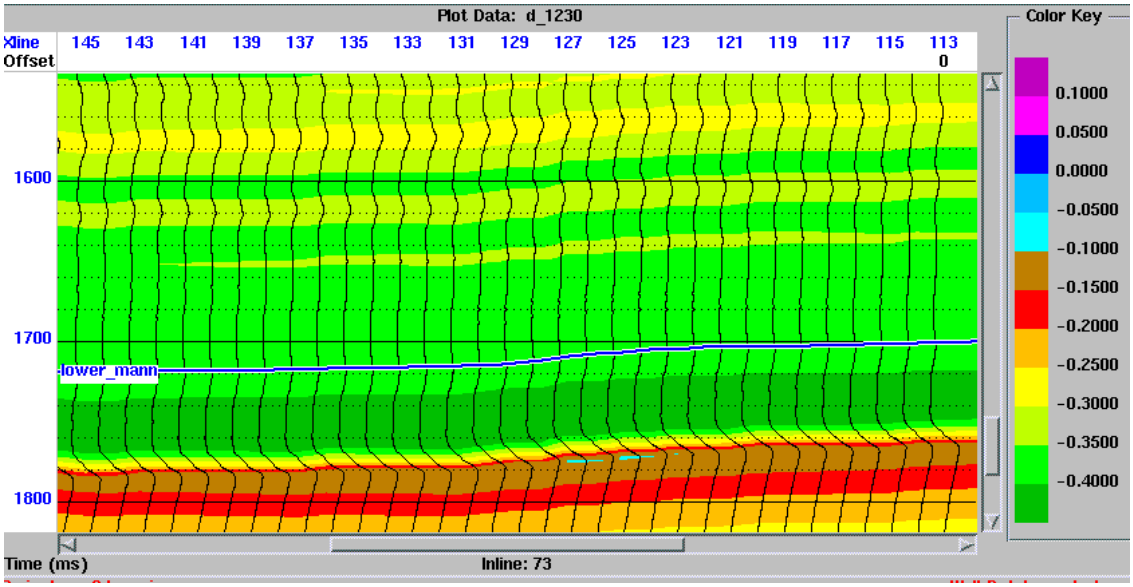

Figure 2.19: Cross-section of the computed d-weights for an offset of 1230.

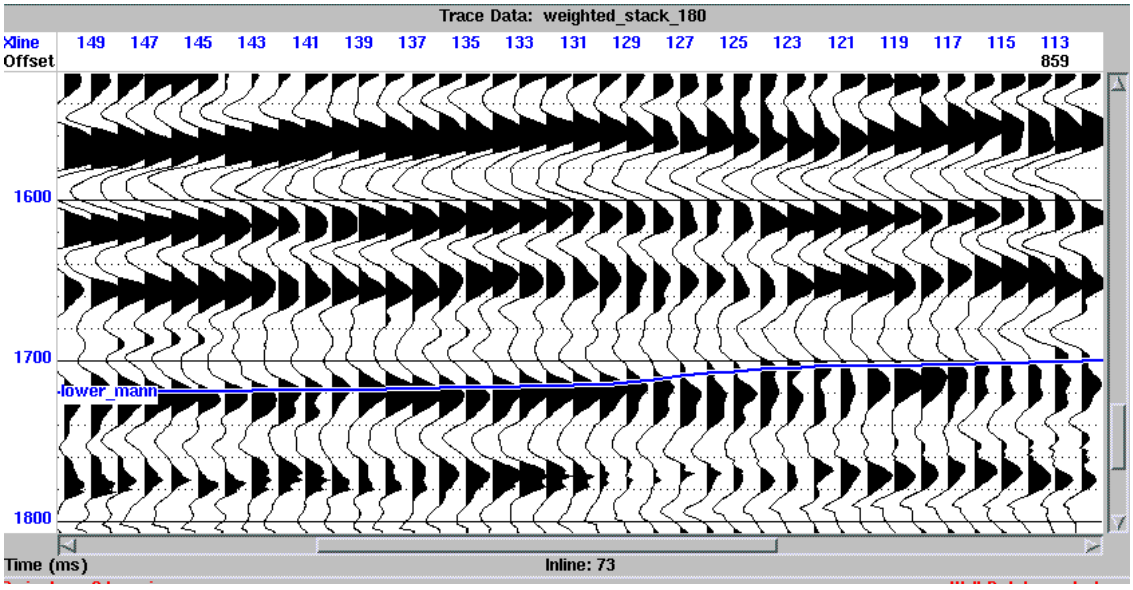

Figure 2.20: Cross-section of the weighted-stack volume computed from the NMO-corrected CCP P-S gathers.

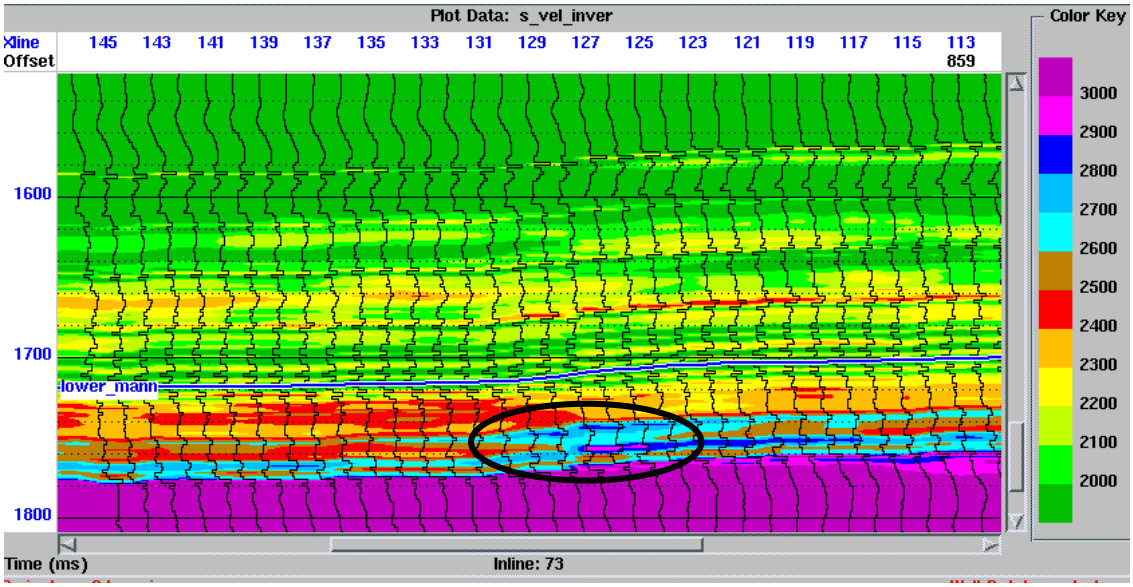

Figure 2.21: Cross-section of the shear velocity volume from P-S inversion. The ellipse shows the channel location.

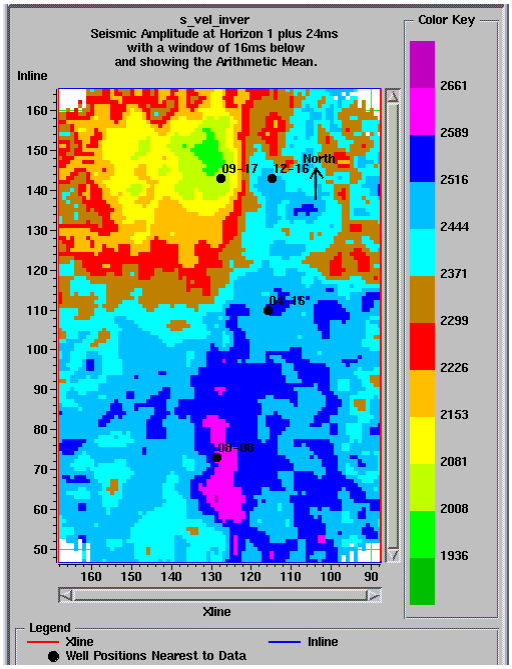

Figure 2.22: Shear velocity at the channel level.

2.4 Conclusions

Model-based P-P inversion has been performed. The producing oil wells in the sand-fill channel correlate with a low-impedance anomaly (~9500 g/cc*m/s), while the dry wells in the shale-plugged channel fall into high-impedance anomaly (~10500 g/cc*m/s). The inversion result can be used to discriminate the sand-fill from the shale-fill channel. However, the regional 09-17 is located in relatively low-impedance area which makes the differentiation of the sand-fill channel and the regional geology ambiguous.

3-D converted-wave (P-S) inversion algorithm has been developed and applied to the Blackfoot data set. An approximation formula for the incident angle in the P-S case has been derived and used to compute the P-S weighted-stack. The approximation has been tested versus ray tracing using Snell's law. The difference between the tow methods falls within 1.5 degree, i.e. the approximation is appropriate for practical use. A conventional (P-P) inversion algorithm has been applied to the weighted-stack and shear (S-wave) velocity volume has been derived. The producing oil wells are located in a relatively high shear velocity anomaly (~2650 m/s), while the dry and regional wells fall into lower shear velocity areas (~2000 - 2450 m/s).

Chapter 3:
Geostatistical integration of well logs and 3C-3D seismic data

3.1 Introduction

Predicting rock and reservoir properties is an essential task in today's hydrocarbon exploration and development. It is of crucial importance for estimating oil and gas reserves and planning future production operations.

Geophysical measurements in the hydrocarbon exploration and development can be divided into two major categories: measurements at the well locations (log curves and core samples) and seismic data. Geological and reservoir models based on sparse well data depend heavily on the well locations, since large areas are not covered by any points. They can lead to poor estimation, especially in geologically complex areas. Unlike well log data, seismic data can provide good coverage of the exploration area but have limitations associated with noise, band-limited frequency, phase error, and lower vertical resolution.

Conventional methods for estimating rock and reservoir properties from seismic data rely on empirical or regression formulas. Such approaches treat the data as spatially independent observations and ignore the existence of spatial patterns. Geostatistical methods have the important ability to integrate different types of information into a consistent subsurface model. They provide improved reservoir description, considering the spatial correlation of the geophysical data, and add the ability to assess the uncertainty in the estimation process.

Geostatistics was developed within the mineral industry to help with mineral reserve estimations. The South African mining engineer D. G. Krige used the probabilistic theory to improve the gold estimation. His ideas were developed further theoretically by the Georges Matherton, an engineer with Ecoles the Mines, who proposed a spatial interpolation method called kriging in honor of Dr. Krige (Matherton, 1963). The successful application in mining industry led to the introduction of the methods to hydrocarbon exploration. Doyen (1988) used the cokriging method to derive porosity. Hwang and McCorkindale (1994) used geostatistically derived average velocities for

depth conversion. Gorell (1995) applied the geostatistical techniques to improve the reservoir description. Numerous geostatistical case studies have been published since then (Hirsche et al., 1996; Tonn, 1998).

In this chapter, I review the basic geostatistical methods (kriging, cokriging, stochastic simulations). Three real data examples from the Blackfoot area are presented: sand/shale mapping, time-to-depth conversion, and thickness estimation.

3.2 Geostatistical methods

3.2.1 Basic statistical terms

A discrete random variable, Z, is a variable that can take a series of outcomes (realizations) z_i , i = 1, ..., N, with a given set of probability of occurence p_i ,i = 1, ..., N. The probability of occurence of N outcomes must satisfy the following conditions:

$p_i \leq 1$, for all i = 1, ...,N

$$\sum_{i=1}^{N} p_i = 1 \tag{3.1}$$

The expected value, E{Z}, or mean m, of a random variable Z is defined as the probability - weighted sum of all possible occurrences:

$$E\{Z\} = m = \sum_{i=1}^{N} p_i z_i \tag{3.2}$$

where E denotes the expectation operator.

If each outcome is equally possible, $p_i = 1/N$, the mean can be written as:

$$m = \frac{1}{N} \sum_{i=1}^{N} z_i \tag{3.3}$$

The variance, σ^2 ,of a random variable Z is defined as the expected squared deviation of Z about its mean:

$$Var\{Z\} = \sigma^2 = E\{(Z - m)^2\} = \sum_{i=1}^{N} p_i (z_i - m)^2 \tag{3.4}$$

If all the probabilities are equal, $p_i = 1/N$:

$$\mathrm{Var}\{Z\} = \frac{1}{N}\sum_{i=1}^{N}(z_i - m)^2 \tag{3.5}$$

The variance is a characteristic of a spread of a distribution around its mean.

The covariance, Cov{X, Y}, of two random variables, X and Y, is defined as:

$$\mathrm{Cov}\{X, Y\} = \sigma_{XY} = E\{(X - m_X)(Y - m_Y)\} = E\{XY\} - m_X m_Y \tag{3.6}$$

where m_X and m_Y are the mean values of the two random variables.

A covariance standardized to be dimensionless is called the correlation coefficient:

$$\rho_{XY} = \frac{\mathrm{Cov}\{X, Y\}}{\sqrt{\mathrm{Var}\{X\}\mathrm{Var}\{Y\}}} \tag{3.7}$$

The correlation coefficient is a measure of linear dependence between two random variables and is confined in the interval [-1, 1].

The variogram, γ_{XY} , is a measure of variability of two random variables, X and Y, and is defined as:

$$\gamma_{XY} = \frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2 \tag{3.8}$$

The greater the variogram value the less related the two variables are.

3.2.2 Variogram – a measure of spatial variability

If we view some measured data points on a map, we can often see that low values tend to be near other low values and high values tend to be near other high values, i.e. there is an existing spatial continuity in the data set. Two data points close to each other are more likely to have similar values than two points that are far apart. The dissimilarity between all the data pairs can be expressed by the variogram as a function of the distance h, often called offset or lag, between them:

$$\gamma^*(h) = \frac{1}{N(h)}\sum_{i=1}^{N(h)}[(Z\ (x_i) - Z\ (x_i + h)]^2 \tag{3.9}$$

where N(h) is the number of all pairs separated by distance h.

If two different properties have been measured, Z and Y, the measure of variability is the cross-variogram, defined by:

$$\gamma^*(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} [(Z\ (x_i) - Z\ (x_i + h)][Y_i(x_i) - Y\ (x_i + h)] \qquad (3.10)$$

The computation of the cross-variogram requires collocated measurements of the two properties.

In practice, we usually separate the data points into a number of offset bins and compute the variogram value in each bin. In this case, h defines the offset to the center of the bin.

The direct use of the variogram in a geostatistical technique can lead to some mathematical complications, for example singular matrices, multiple solutions, or negative mean-square errors (Isaaks and Strivastava, 1989). The variogram values are computed for all pairs of sample locations, but the geostatistical techniques require a variogram value between all the sampled locations and the locations we wish to estimate. Therefore, we require a variogram model which is a function of the offset h. The model cannot be an arbitrary function, it must obey certain rules (Isaaks and Strivastava, 1989). The model is defined by second-order stationarity (Wackernagel, 1995): the expected value is a constant over the area and the variogram depends on the length, and the orientation of the offset h, but not on its position x. Figure 3.1 is a typical example of a computed (experimental) variogram and a model (theoretical) variogram fit to it. The offset distance at which the model 'flattens' off to a 'plateau' is called range. This value defines the distance at which the difference between the wells becomes random, i.e. they are not spatially related. The value of the 'plateau' is called the sill. Note that a non-zero value, called nugget, can be added to the model at zero offset.

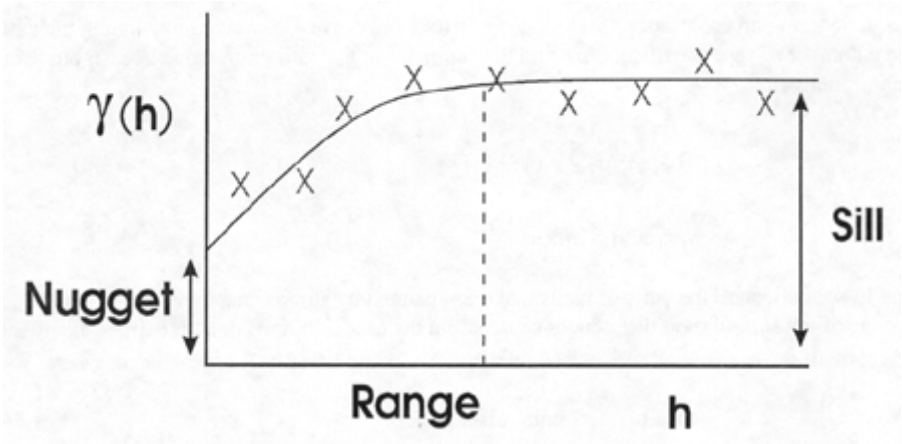

Figure 3.1: Experimental and theoretical variograms.

The most commonly used variogram models in practice are:

- Spherical, defined by:

$$\gamma(h) = C[1.5(h/a) - 0.5(h/a)^3], \quad \text{if } h \le a$$
$$= C, \quad \text{if } h > a \tag{3.11}$$

- Exponential, defined by:

$$\gamma(h) = C[1 - \exp(-h/a)] \tag{3.12}$$

- Gaussian, defined by:

$$\gamma(h) = C[1 - \exp(-h/a)^2] \tag{3.13}$$

- Power, defined by:

$$\gamma(h) = Ch^n \tag{3.14}$$

where a is the range, and C is the sill.

The covariance function can be computed from the variogram (Journel, 1989):

$$Cov(h) = \gamma(\infty) - \gamma(h) = Cov(0) - \gamma(h) \tag{3.15}$$

where $\gamma(\infty)$ is the sill value and C(0) is the zero-offset covariance.

3.2.3 Kriging

Kriging is a basic statistical interpolation technique. The goal is to estimate a particular property z_0^* at an unmeasured location by a linearly weighted sum of the known (measured) values. A linear estimator, written as an equation for N known values z_i , i=1,...,N, is:

$$z_0^* = \sum_{i=1}^{N} w_i z_i + w_0 \tag{3.16}$$

The estimation error at each location is the difference between the estimated value, z_0^*, and the true value, z_0. The mean of the error distribution is often referred to as the bias. A reasonable goal for any estimation method is to produce unbiased estimates, i.e the mean value of the estimation errors to be zero:

$$\frac{1}{K}\sum_{i=1}^{K}\left(z_{0i}^* - z_{0i}\right) = 0 \tag{3.17}$$

where K is the number of estimated points.

Unfortunately, we do not know the true values. The solution to this problem is to conceptualize the unknown values as the outcome of a random variable. For any point at which we attempt to estimate the unknown value, our model is a random function that consists of random variables, one for each of the sampled locations, Z_i, i = 1...N, and one for the point we are trying to estimate, Z_0^*:

$$Z_{\mathbf{0}}^{*} = \sum_{i=\mathbf{1}}^{N} w_i Z_i + w_{\mathbf{0}} \tag{3.18}$$

The estimation error is defined as the difference between the estimated Z^* and the random variable Z modeling the true value. The estimator is unbaised if the expected value of the estimation error is zero:

$$E\{Z_0^{*} - Z_0\} = 0 \tag{3.19}$$

Using that the expected value is a linear operator (Journel, 1989), we write:

$$E\{Z_0^{*} - Z_0\} = E\{Z_0^{*}\} - E\{Z_0\} = E\{\sum_{i=1}^{N} w_i Z_i + w_0\} - m = E\{\sum_{i=1}^{N} w_i Z_i\} + w_0 - m =$$
$$= \sum_{i=1}^{N} w_i E\{Z_i\} + w_0 - m = m\sum_{i=1}^{N} w_i + w_0 - m = m(\sum_{i=1}^{N} w_i - 1) + w_0 \tag{3.20}$$

Clearly, the unbaised estimator can be achieved only if:

$$\begin{cases} w_0 = 0 \\ \sum_{i=1}^{N} w_i = 1 \end{cases} \tag{3.21}$$

Thus, the Unbiased Linear Estimator is written as:

$$\begin{cases} Z^* = \sum_{i=1}^{N} w_i Z_i \\ \sum_{i=1}^{N} w_i = 1 \end{cases} \tag{3.22}$$

The kriging estimator is said to be 'best' because the kriging weights are determined by the minimization of the error variance:

$$\mathrm{Var}\{Z_0^* - Z_0\} \Rightarrow \min \tag{3.23}$$

The set of weights that minimize the error variance under the constraint that they sum to one satisfies the following N + 1 equations (Isaaks and Srivastava, 1989):

$$\begin{cases} \sum_{j=1}^{N} w_j \mathrm{Cov}\{Z_i, Z_j\} + \mu = \mathrm{Cov}\{Z_0, Z_i\} \qquad i = 1,\ldots,N \\ \sum_{i=1}^{N} w_i = 1 \end{cases} \tag{3.24}$$

where: μ - Lagrange parameter

Cov{ } - covariance function

The covariance values in the above equations are derived from the modeled variogram. The weights, which are solution of the above system (known as the ordinary kriging system) are used to compute the ordinary kriging estimate. Note that having the kriging weights, we can compute the actual minimum value of the error variance.

3.2.4 Cokriging

In oil exploration, we often have two independent measurements: well log curves or core samples and a seismic data set. At the well locations we can measure directly the physical property of the subsurface, for example the acoustic and shear velocities, the density, the neutron porosity, etc. Using the kriging method, we can interpolate and map those measurements between the wells. However, we would like to include the seismic data into the mapping process since they may provide very good spatial coverage of the exploration area.

The ordinary cokriging estimator Z_0^* is written as:

$$Z_0^* = \sum_{i=1}^{N} w_i Z_i + \sum_{j=1}^{M} v_j X_j \tag{3.25}$$

As in the kriging case, the estimator is unbiased if the expected value of the estimation error is zero, i.e.:

$$E\{Z_0^* - Z_0\} = E\{Z_0^*\} - E\{Z_0\} = E\left\{\sum_{i=1}^{N} w_i Z_i - \sum_{j=1}^{M} v_j X_j\right\} - m_Z = \\ = m_Z\left(\sum_{i=1}^{N} w_i - 1\right) - m_X \sum_{j=1}^{M} v_j \tag{3.26}$$

One way of guaranteeing unbiasedness is to ensure that the weights in the first term sum to 1 while those in the second term sum to 0 (Isaaks and Srivastava, 1989):

$$\begin{cases} \sum_{i=1}^{N} w_i = 1 \\ \sum_{j=1}^{M} v_j = 0 \end{cases} \tag{3.27}$$

It should be noted that other non-bias conditions are possible.

The development of the cokriging system is similar to the development of the ordinary kriging system. The ordinary cokriging system is derived by minimizing the estimation error variance (Isaaks and Srivastava, 1989):

$$\begin{cases} \sum_{i=1}^{N} w_i Cov\{Z_i, Z_j\} + \sum_{i=1}^{M} v_i Cov\{X_i, Z_j\} + \mu_1 = Cov\{Z_0, Z_j\}, & \text{for } j = 1,...,N \\ \sum_{i=1}^{N} w_i Cov\{Z_i, X_j\} + \sum_{i=1}^{M} v_i Cov\{X_i, X_j\} + \mu = Cov\{Z_0, X_j\}, & \text{for } j = 1,...,M \\ \sum_{i=1}^{N} w_i = 1 \\ \sum_{j=1}^{M} v_j = 0 \end{cases} \tag{3.28}$$

To derive the required covariance values, we need three variogram models: well-to-well, seismic-to-seismic, and well-to-seismic.

The derived weights are used in the cokriging estimation.

The kriging and cokriging algorithms assume that the used data are stationary. If a trend is present in the data (the mean value varies in the area) it violates the condition under which geostatistics operates best. A possible solution to this problem follows (GEOSTAT documentation, 1999):

- divide the dense data into two components, trend and residual:

Data(x, y)=Trend(x, y) + Residual(x, y)

- remove the trend using linear relationship between the dense and sparse data sets
- perform cokriging with the residual data (assumed to be stationary)
- restore the trend

Another possible way to integrate well and seismic data is called kriging with external drift (GEOSTAT documentation, 1999).

3.2.5 Stochastic simulation

Stochastic simulation is a process of generating alternative, equally probable realizations of a random variable. The realizations are also called stochastic images. Different stochastic simulation algorithms have been developed, but the most used one in the geophysical practice is called Sequential Gaussian simulation. The Sequential Gaussian simulation is constrained in such a way that all realizations honor the measured data values at their locations. The generated stochastic images are called conditional.

Conditional simulation is used to correct for the smoothing effect seen on the kriging maps. The kriging estimates are weighted moving averages of the original data and they have less spatial variability than the data. A smoothed map provided by the kriging is appropriate for showing global trends while conditionally simulated maps are more appropriate for studies that are sensitive to patterns of local variability (reservoir flow simulations for example). The process also provides a measure of uncertainty about the generated spatial distribution. By creating a large number of simulation maps, we try to reproduce the probability distribution at each point of the grid. From these probability distributions, we derive probabilities associated with ranges of the estimated parameter.

The main steps in the Sequential Gaussian simulation are:

- define a grid and insert the known data at there locations
- select a random point and estimate its value using kriging or cokriging
- the estimate consists of a value and a variance; assuming a Gaussian distribution around the estimated value, determine a new value for the point by a random number generator

- repeat the above steps for every point, treating the previously estimated ones as an exact ones, i.e. use them in the kriging (cokriging) estimation

3.3 Sand/shale distribution mapping

Gamma ray logs measure natural radioactivity in formation and because of that, they can be used for identifying different lithologies. Shale-free sandstones and carbonates have low concentration of radioactive material, and give a low gamma ray reading. In contrast, shales and clays show high gamma ray readings because of their high concentration of radioactive elements. Besides their use in identifying lithologies and correlating zones, gamma ray logs provide information for calculating the volume of shale in sandsones and carbonates. One such measure is the gamma ray index, defined by:

$$GR_{index} = \frac{GR_{log} - GR_{sand}}{GR_{shale} - GR_{sand}} \tag{3.29}$$

where:

GR_{log} – gamma ray reading

GR_{shale} – shale line (clean shale)

GR_{sand} – sand line (clean sand)

Using a sand line of 15 API and a shale line of 135 API the average gamma ray index is computed for the Channel – Mississippian interval (Table 3.1).

well	gamma ray index	well	gamma ray index
01-08	0.42	16-08	0.42
08-08	0.39	04-16	0.63
09-08	0.37	05-16	0.61
2/09-08	0.36	12-16	0.64
11-08	0.51	09-17	0.54

Table 3.1: Average gamma ray index computed for the Channel - Mississippian interval.

As expected, the producing oil wells (08-08, 09-08, etc.) have low gamma ray index values while the shale-plugged 12-16 has the highest value.

Through the analysis of isochron maps from multi-component seismic data, the Vp/Vs value can be computed (Garotta, 1987):

$$Vp/Vs = \mathbf{2}\frac{t_{PS}}{t_{PP}} - \mathbf{1} \tag{3.30}$$

where:

t_{PP} – P-P isochron

t_{PS} – P-S isochron

Figure 3.2 is the computed Vp/Vs ratio map for the Top Channel – Wabamun interval in the Blackfoot area. It is assumed that the carbonates below the Missippian have a constant effect on the map. The low Vp/Vs anomaly correlates with the producing oil wells. The cross plot of the gamma ray index versus derived Vp/Vs values shows high correlation of 0.94 (Figure 3.3). Figures 3.4 - 3.6 show the calculated experimental and fitted theoretical variograms. Table 3.2 contains the used parameters for the variogram modeling. The offset units are number of bins (1 bin = 30 m).

	well-to-well	seismic-to-seismic	well-to-seismic
type	spherical	spherical	spherical
range	66	93	73
sill	0.02	0.016	0.022
nugget	0	0	0

Table 3.2: Parameters used in the variogram modeling.

Using the kriging and cokriging algorithms sand/shale distribution maps are generated (Figures 3.7 and 3.9). In both maps, a low shale content anomaly is visible and it correlates well with the known sand channel, but the kriging result is considerably smoother. The absolute errors from cross-validation tests for both methods are shown on Figures 3.8 and 3.10. Although the cokriging method shows similar errors we may argue that the cokriging result has higher spatial resolution. Ten Gaussian simulations are performed and Figure 3.11 shows one of the results. Figure 3.12 is a map of the average sand/shale distribution values from the performed simulations. Figure 3.13 is a plot of the

probability to find a gamma ray index less then 0.35, i.e. the probability of finding reasonably clean sand. The result clearly identifies the sand channel.

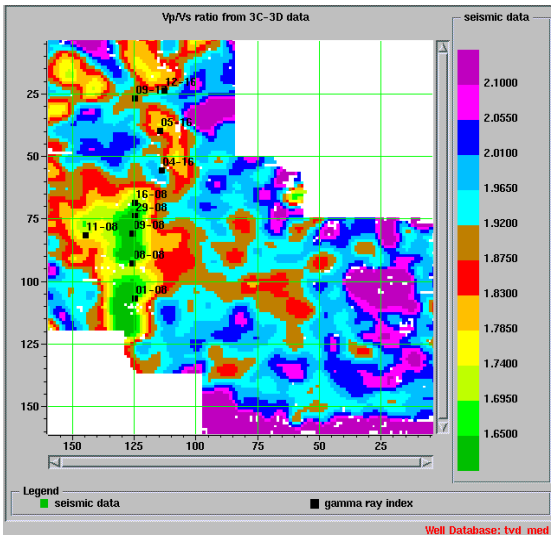

Figure 3.2: Input data: Vp/Vs map.

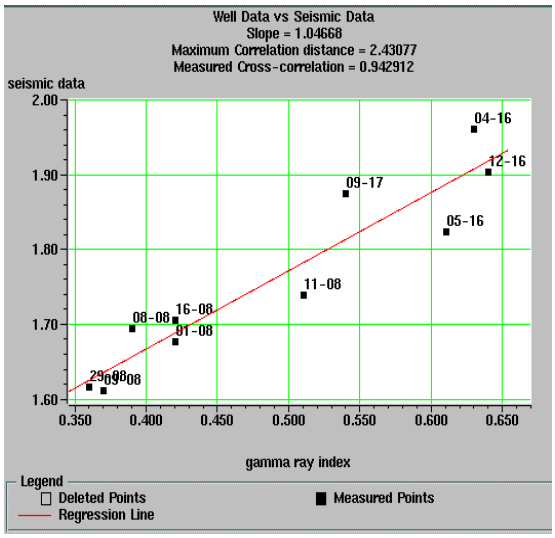

Figure 3.3: Cross plot, gamma ray index vs Vp/Vs.

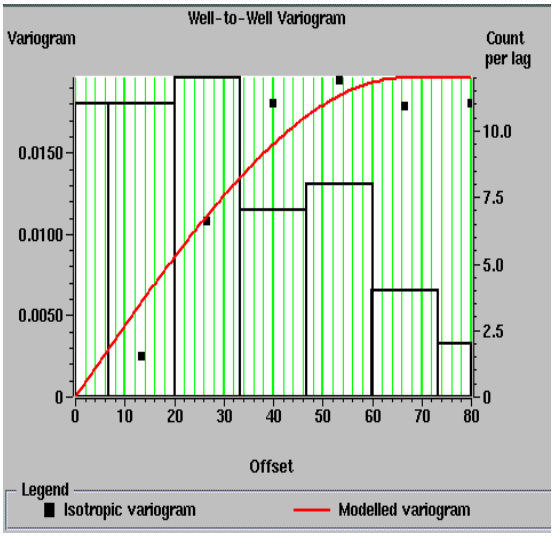

Figure 3.4: Well-to-well variogram.

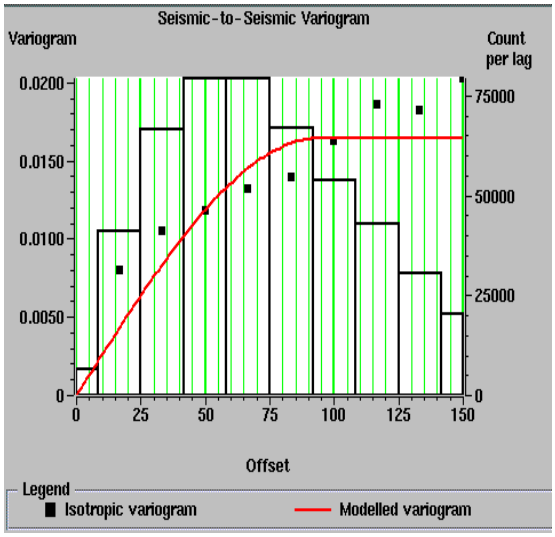

Figure 3.5: Seismic-to-seismic variogram.

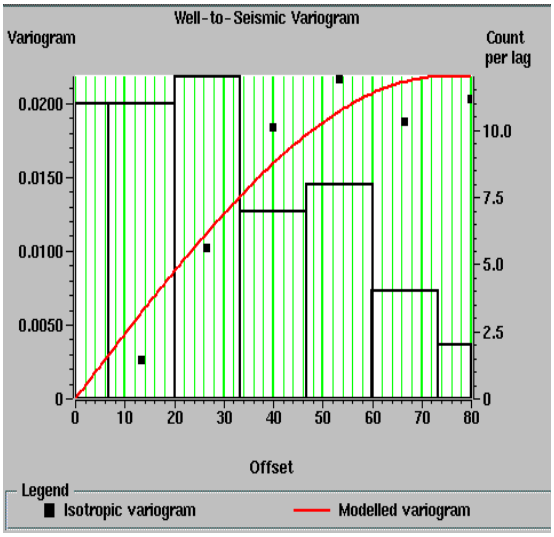

Figure 3.6: Well-to-seismic variogram.

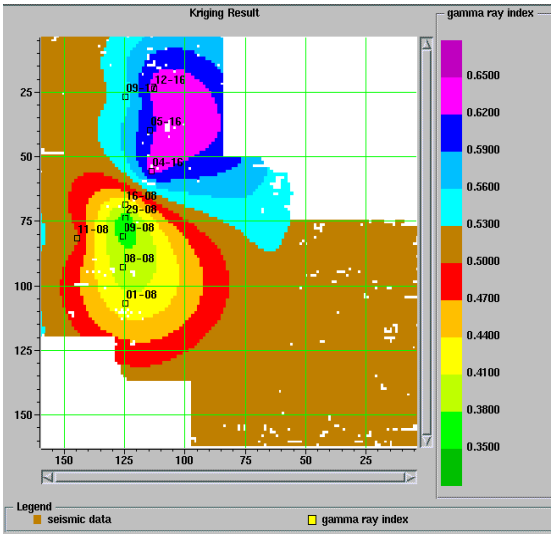

Figure 3.7: Gamma ray index, kriging result.

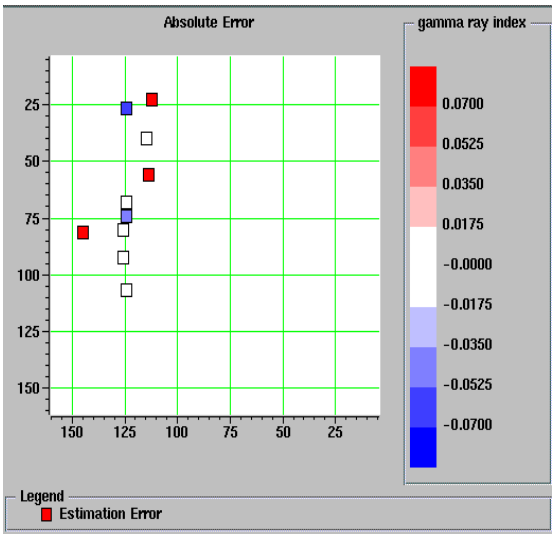

Figure 3.8: Cross-validation, absolute error.

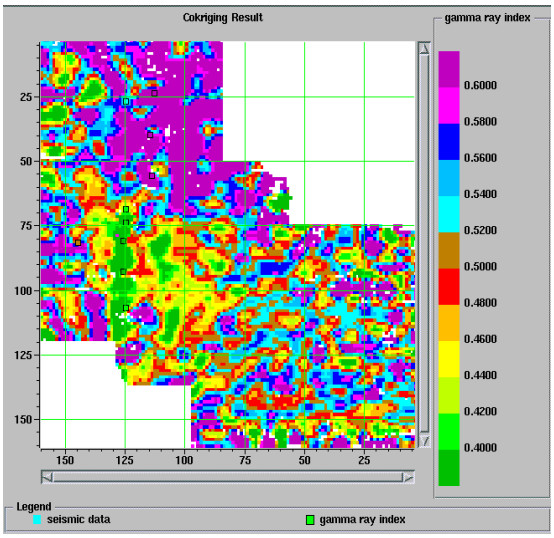

Figure 3.9: Gamma ray index, cokriging result.

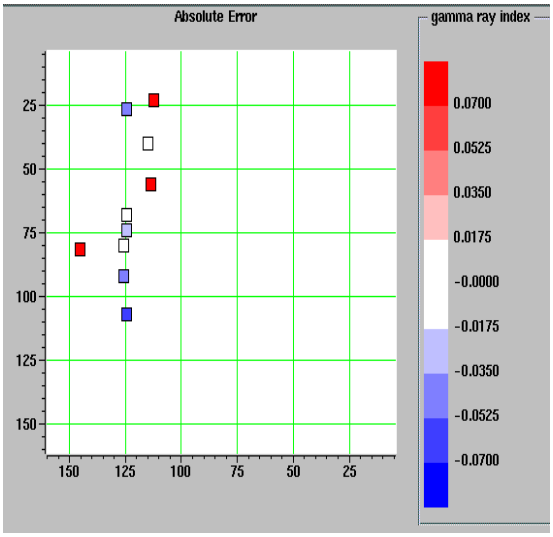

Figure 3.10: Cross-validation, absolute error.

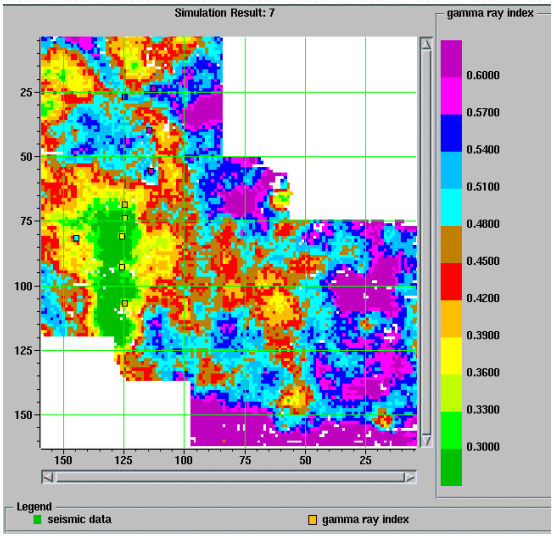

Figure 3.11: Gamma ray index, simulation result.

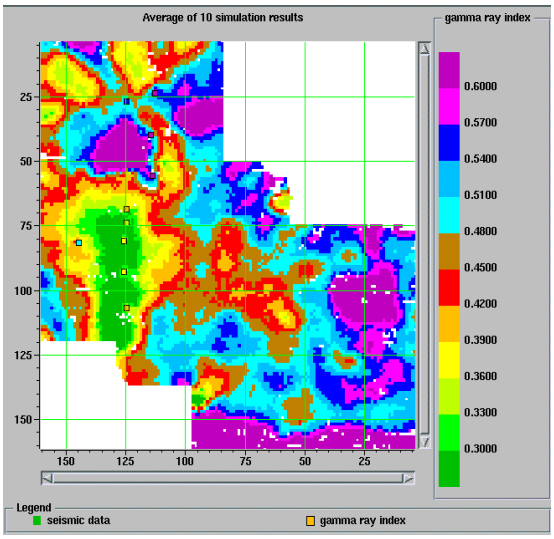

Figure 3.12: Gamma ray index, average of ten simulations.

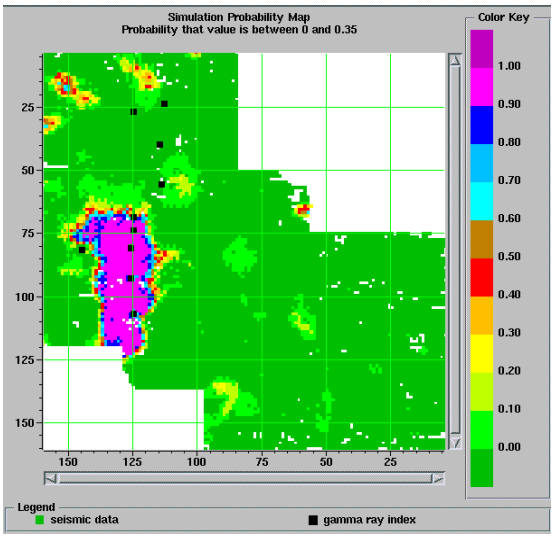

Figure 3.13: Probability of finding clean sand.

3.4 Time-to-depth conversion

Seismic data are recorded in time. To obtain a depth image of the earth subsurface, a correct velocity model is needed. It can be developed by interpolation of measured sonic logs, calculated from velocity analysis, or combining both sources. Well information is

often sparse and interpolation does not reflect possible velocity variations in the regions between the wells. Developing a model from velocity analysis could be problematic due to the large number of variables that influence the velocity and structural complexity may add more difficulties. So, except for some simple cases, it is very difficult to derive the correct velocity solution and thus the subsurface depth image. The errors may result in mispositioning of future wells or miscalculation of potential reserves.

A possible solution to the problem is to integrate the known depths to a particular geological top (at the well locations), and the measured two-way traveltime to the corresponding horizon.

Figure 3.15 is the P-P two-way traveltime of the Mannville event. Since the two-way traveltime is represented from the seismic processing datum at 1000 m., all true vertical depths to the Mannville top in the wells were adjusted to the datum (Figure 3.14):

$$h_a = h_w + h_c \tag{3.31}$$

where:

h_a - adjusted to the seismic processing datum depth

h_w - true vertical depth (tvd) to the Mannville

h_c - correction taking into account the difference between the seismic datum and the earth surface

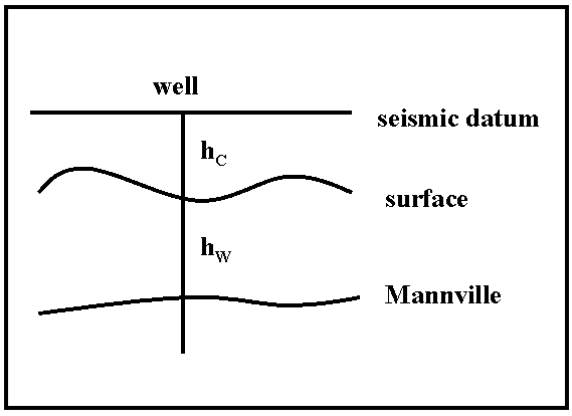

Figure 3.14: Depth correction diagram, showing the relative position of the Mannville top, earth surface, and seismic datum.

Table 3.3 shows the wells, used in the geostatistical analysis (column one), the measured true vertical depth to the Mannville top at the well locations (column two), the calculated correction (column three), and the corresponding adjusted depth (column four).

well	tvd depth	correction	adjusted depth
01-08	1424	83	1507
08-08	1431	82	1513
09-08	1433	82	1515
2/09-08	1432	75	1507
11-08	1437	82	1519
16-08	1433	75	1508
14-09	1433	73	1506
04-16	1433	75	1509
05-16	1444	61	1508
12-16	1445	62	1507
13-16	1467	31	1498
01-17	1435	75	1510
09-17	1450	62	1512
14-03	1465	25	1490

Table 3.3: Calculation of the adjusted Mannville depth.

Figure 3.16 is a cross-plot of the Mannville P-P two-way traveltime versus the adjusted depth. The measure cross-correlation coefficient is 0.96. Since the seismic data contain a trend of decreasing values from east to west the trend is calculated (Figure 3.17) and removed from the data. Figures 3.18 - 3.20 show the calculated experimental and fitted theoretical variograms. Table 3.4 contains the parameters used for the variogram modeling.

	well-to-well	seismic-to-seismic	well-to-seismic
type	spherical	spherical	spherical
range	22	90	28
sill	15	17	9
nugget	0	0	0

Table 3.4: Parameters used in the variogram modeling.

Using the cokriging method, a Mannville depth structure map is generated (Figure 3.21). The estimated depth decreases from west to east and has a value of 1505 – 1510 meters in the productive area. A cross-validation test is performed and Figure 3.22 shows the calculated absolute error. Since a relatively small absolute error is achieved, the generated Mannville depth map can be considered a reliable result. The use of cokriging method shows good performance and does not require a velocity model. Ten Gaussian simulations are performed. Figure 3.23 shows one of the generated maps and Figure 3.24 is the average of the performed simulations.

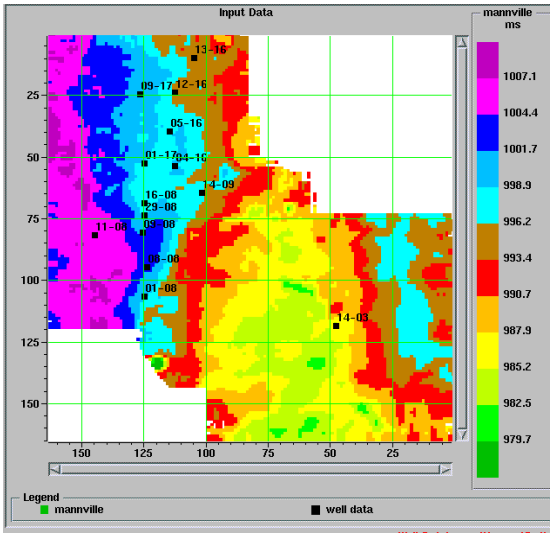

Figure 3.15: Input data, Mannville two-way traveltime.

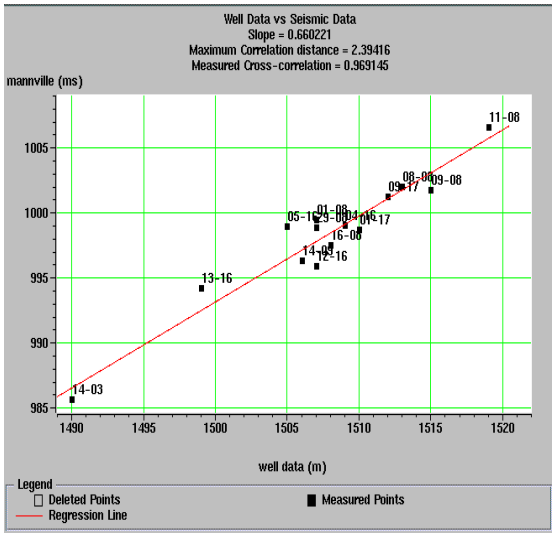

Figure 3.16: Cross plot, tvd depth vs two-way traveltime.

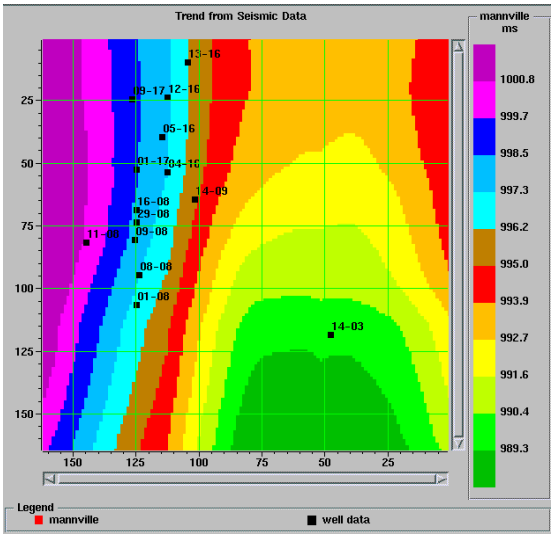

Figure 3.17: Computed trend.

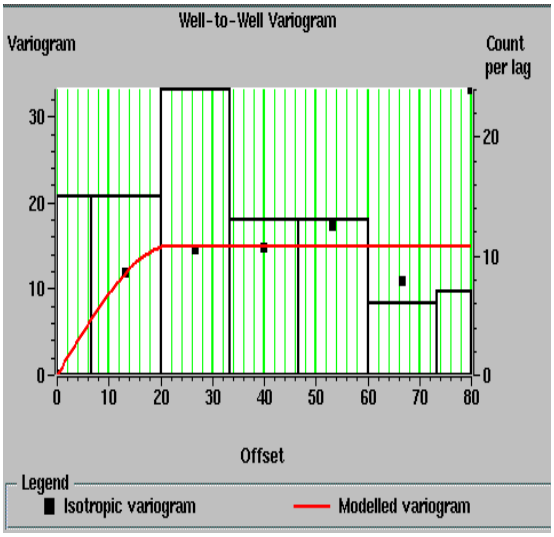

Figure 3.18: Well-to-well variogram.

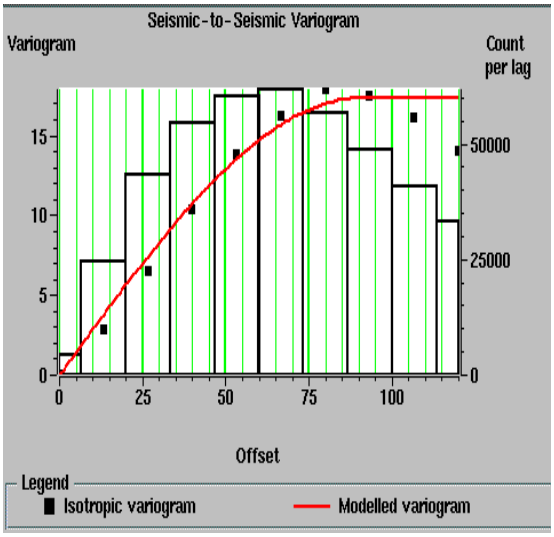

Figure 3.19: Seismic-to-seismic variogram.

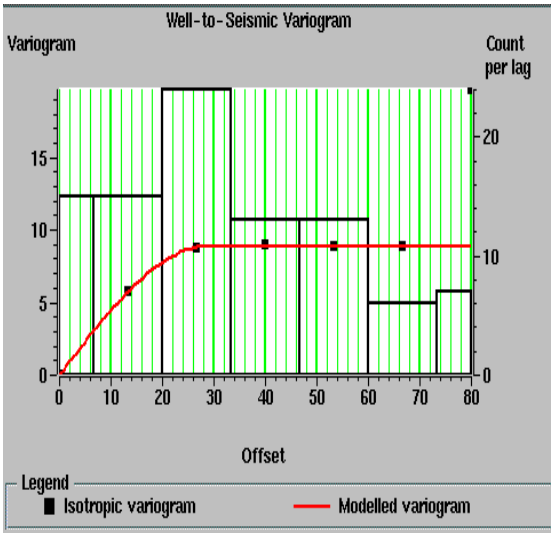

Figure 3.20: Well-to-seismic variogram.

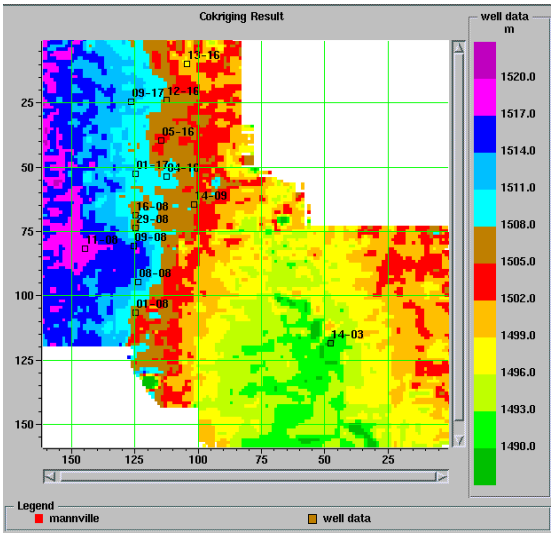

Figure 3.21: Mannville depth structure, cokriging result.

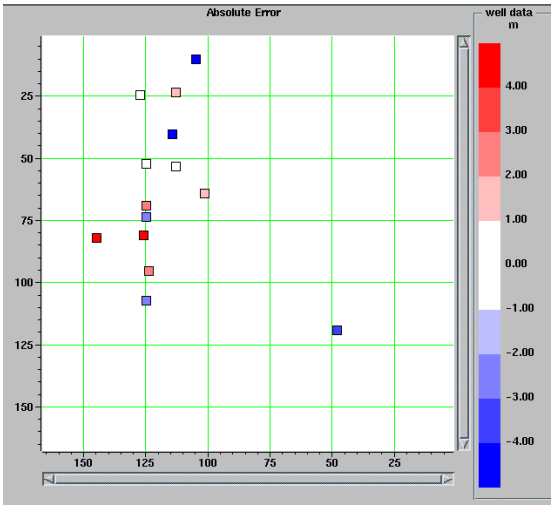

Figure 3.22: Cross-validation, absolute error.

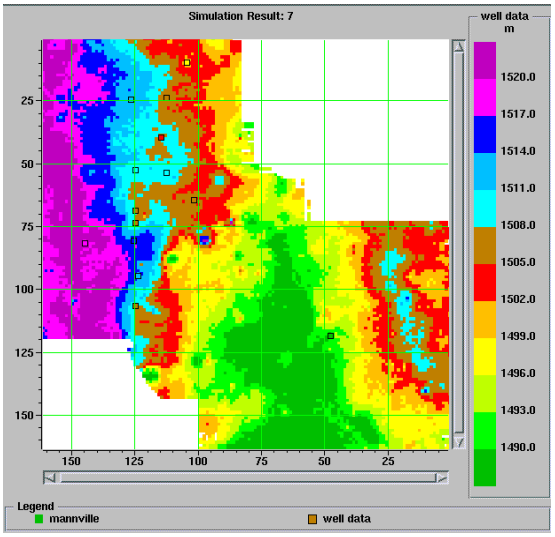

Figure 3.23: Mannville depth structure, simulation result.

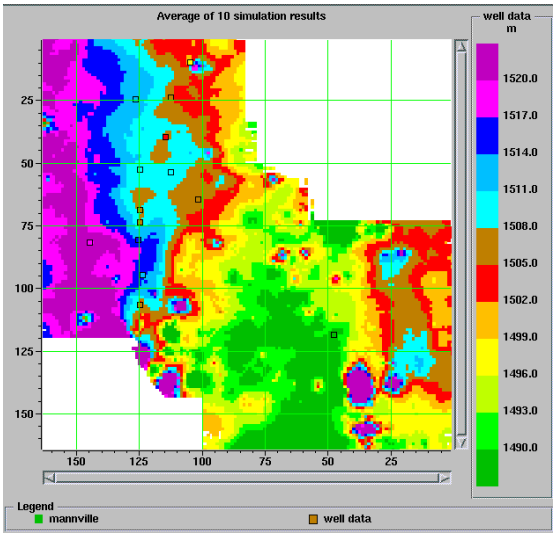

Figure 3.24: Mannville depth structure, average of ten simulations.

3.4 Thickness estimation

The problem of thickness estimation is similar to that of time-to-depth conversion. The common approach requires the average velocity over the studied interval. The problems with obtaining a correct velocity model were already discussed. Another possible solution

for obtaining isopach maps is to integrate in geostatistical fashion a seismic isochron map derived from the 3-D survey and the corresponding thickness estimated from well data. Figure 3.25 is a P-P isochron map from Mississippian to Mannville events. The thickness of the Mannville – Mississippian interval is calculated from well log data at various well locations (Table 3.5).

well	thickness (m)	well	thickness (m)
01-08	176.0	13-16	167.5
08-08	180.5	05-16	194.0
2/09-08	194.0	01-17	175.0
11-08	177.0	09-17	187.0
14-09	166.5	14-03	180
12-16	165.5		

Table 3.5: Thickness of the Mannville-Mississippisn interval.

Figure 3.26 is a cross-plot of the isochron values versus calculated thickness at the well locations. The cross-correlation is 0.945.

Due to difficulties to fit a model to the calculated well-to-well variogram points, only the seismic-to-seismic variogram is generated (Figure 3.27). Cokriging is performed assuming linear relationship between the well and seismic data, i.e. well-to-well and well-to-seismic variograms are computed from the seismic-to-seismic variograms based on that linear relationship. Figure 3.28 is the generated cokriging isopach map and figure 3.29 is the absolute error from the cross- validation test. Ten simulations are performed and Figure 3.30 is the average result.

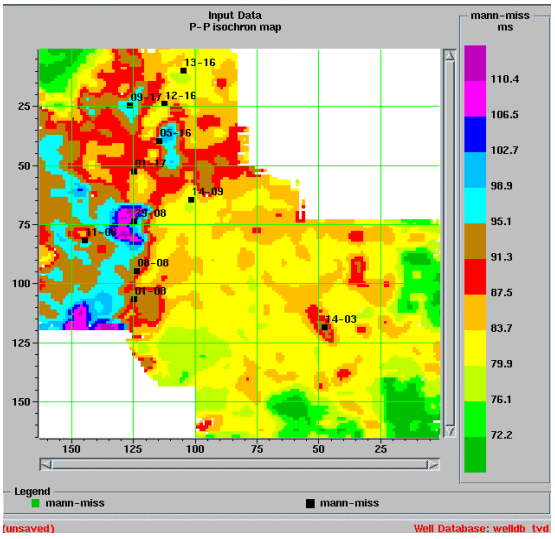

Figure 3.25: Input data, Mannville-Mississippian isochron map.

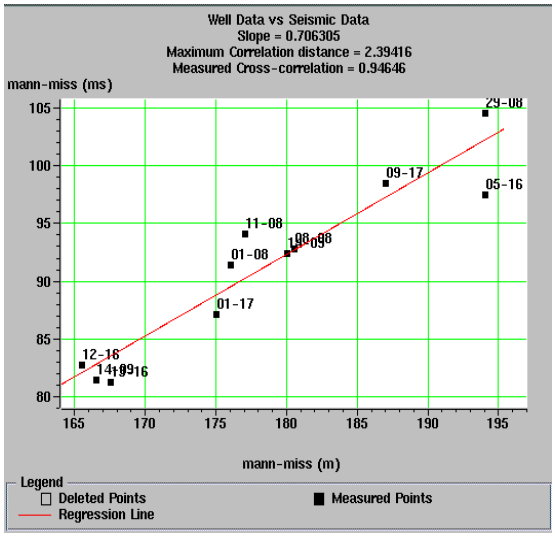

Figure 3.26: Cross-plot, isochron values vs thickness.

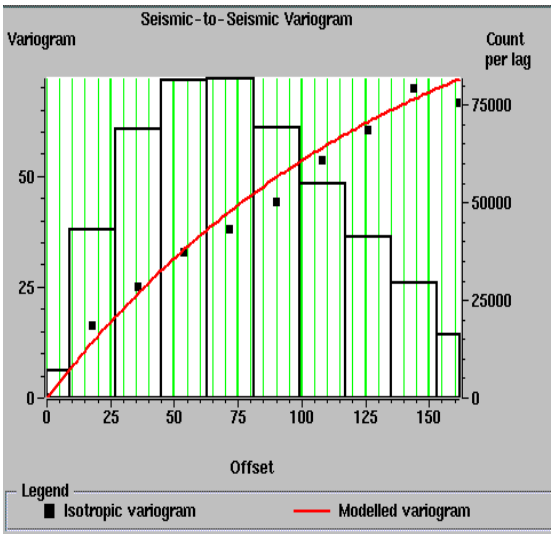

Figure 3.27: Seismic-to-seismic variogram.

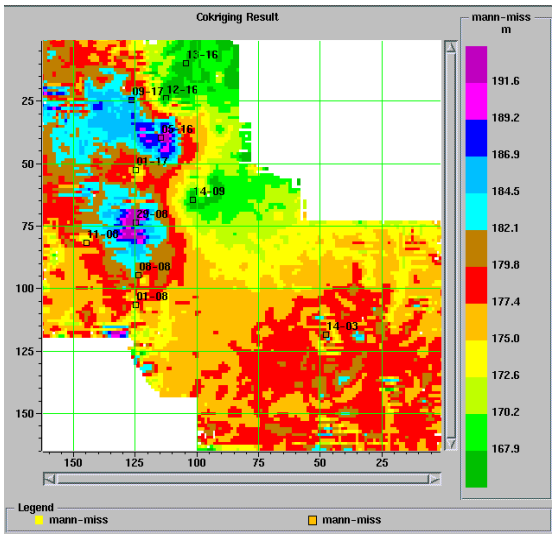

Figure 3.28: Mannville-Mississippian isopach map, cokriging result.

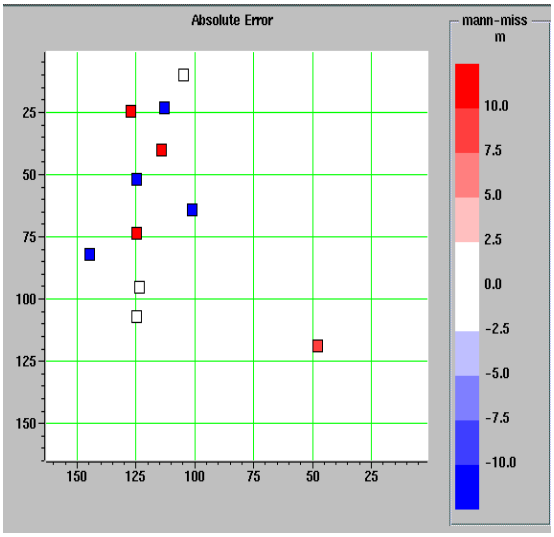

Figure 3.29: Cross-validation, absolute error.

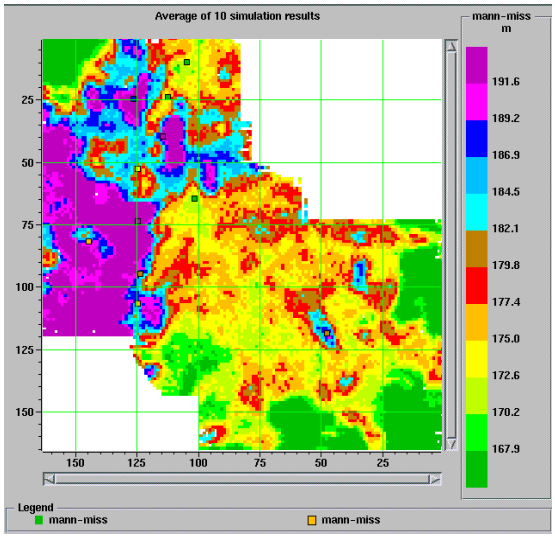

Figure 3.30: Mannville-Mississippian isopach map, average of ten simulations.

3.5 Conclusions

A brief review of the geostatistical methods has been given. Three real data examples from the Blackfoot area are shown. Vp/Vs map, derived from 3C-3D seismic measurements has been integrated with sand/shale values computed from gamma ray logs to generate a sand/shale distribution map. The low shale content anomaly on the cokriging map correlates with the producing oil wells. New prospect areas can be identified on the map. A picked seismic horizon, Mannville, has been converted to depth by cokriging with the corresponding geological depth. The time interval (isochron map) between two seismic horizons, Mannville and Mississippian, has been converted to depth interval (isopach map). The performed cross-validation tests show relatively low absolute errors, i.e. the geostatistical integration provides reliable pictures of the subsurface.

Chapter 4:
Well log estimates using elastic seismic attributes

4.1 Introduction

The derivation of reservoir and rock properties from seismic data is a major task in exploration geophysics. The complex response of the earth to the wave propagation makes this task difficult and challenging. Post-stack seismic inversion attempts to deduce the subsurface impedance. AVO inversion is used to derive elastic properties. However, those methods rely on theoretical relationships usually based upon some assumptions. A different approach is to use statistical, rather than deterministic method to derive the subsurface properties. Those methods are also known as data-driven methodologies (Schultz et al., 1994).

3-D seismic data are usually presented in a form of stacked volume. Taner et al. (1979) introduced the complex seismic trace and instantaneous attributes. Since then, seismic attribute technology has seen great interest and new seismic attributes are introduced routinely. Chen and Sidney (1997) give a good summary of the existing seismic attributes. However, the interpretation of the seismic attribute volumes is ambiguous and it is difficult to derive quantitative information from them.

In this chapter, I present a method for deducing rock properties based on the integration of seismic attributes and measured well log curves (or core samples). The basic idea is to find a relationship between the measured rock properties and some seismic attributes at the well location. Once derived it can be applied to the seismic volume and a predicted log property volume is generated. The derived relationship can be linear (using linear multi-regression analysis) or non-linear (using neural networks). The reliability of the derived relationship is determined by cross-validation tests. The theoretical foundations of the methods are discussed and two real data examples are presented involving prediction of impedance and porosity logs. In the second example, attributes derived from converted (P-S) data are included in the prediction alongside the conventional P-P attributes. To do so, a model-based conversion of P-S data to P-P time is developed.

4.2 Methods

In general, the relationship between the log property and the seismic attributes (in time) can be written in the following form:

$$P(x, y, t) = F[A_1(x, y, t), A_2(x, y, t), \ldots, A_M(x, y, t)] \quad (4.1)$$

where:

P(x, y, t) - log property as a function of space and time coordinates x, y, t

F[...] - functional relationship

A_i - seismic attributes from i = 1, ..., M at coordinates (x, y, t)

4.2.1 Multi-regression analysis

The functional relationship can be found using linear multi-regression analysis. For N measured log property values at a particular location at different times, we have:

$$\begin{aligned} P_1 &= W_1A_{11} + W_2A_{21} + \ldots + W_MA_{M1} + C \\ P_2 &= W_1A_{12} + W_2A_{22} + \ldots + W_MA_{M2} + C \\ &\ldots \\ P_N &= W_1A_{1N} + W_2A_{2N} + \ldots + W_MA_{MN} + C \end{aligned} \quad (4.2)$$

where:

P_j - well log values as function of time increment, j = 1, ..., N

W_i - unknown weights, i = 1, ..., M

A_{ij} - attribute samples, i = 1, ..., M, number of attributes, j = 1, ..., N, number of time samples

C - constant

A more advanced approach is to use time convolution operators instead of single weight in the regression analysis:

$$P = \mathbf{W_1} * \mathbf{A_1} + \mathbf{W_2} * \mathbf{A_2} + \ldots + \mathbf{W_M} * \mathbf{A_M} + C \quad (4.3)$$

where:

$\mathbf{W_i}$ - convolution operators (vectors), i = 1, ..., M

If we have 3-point convolutional operator and two attributes (Figure 4.1), the j^{th} sample is computed by the following equation (i=1,2):

$$P_j = W_{1,-1}A_{1,j-1} + W_{1.0}A_{1,j} + W_{1,1}A_{1,j+1} + W_{2,-1}A_{2,j-1} + W_{2,0}A_{2,j} + W_{2,1}A_{2,j+1} + C \quad (4.4)$$

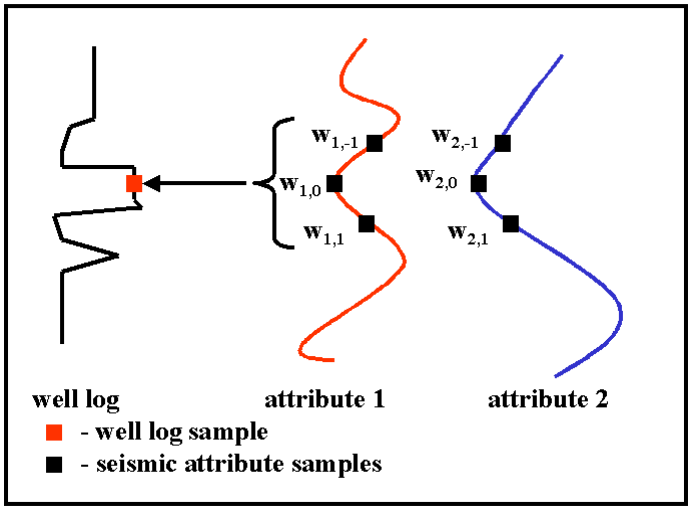

Figure 4.1: Using 3-point convolutional operator.

In the case of L-point convolution operator, there are L•M unknown weights to be determined by minimizing the mean squared prediction error:

$$e^2 = \frac{1}{N}\sum_{j=1}^{N}\left(P_j - \mathbf{W_1} * \mathbf{A_1} - \mathbf{W_2} * \mathbf{A_2} - \ldots - \mathbf{W_M} * \mathbf{A_M} - C\right)^2 \quad (4.5)$$

A number of attributes have been used recently. So the question, "How many and which attributes to use?" is a major one. One may want to find the best attributes, K of them, out of the total M. One possible procedure is the exhaustive search, which means to try all possible K combinations from the M attributes. Then the combination with the lowest prediction error is selected. The problem with the method is computational time, which becomes unacceptably long for practical work.

A much faster procedure is step-wise regression (Draper and Smith, 1981). The procedure can be described as follows:

- from all M attributes find the one with the smallest prediction error, i.e. the best attribute {A1}
- Find the best pair of attributes, assuming that the first attribute is A1, i.e. {A1, A2}

- find the best triplet of attributes assuming that the first two are A1, A2, i.e. {A1, A2, A3}
- repeat until the best K attributes are found, i.e. {A1, A2, A3, ..., AK}

From a theoretical point of view, the combination of K+1 attributes will have smaller prediction error than the one of K attributes. The step-wise regression can tell us which are the best K attributes, but can not answer the question "How many attributes to use?", i.e. the quantity of K.

The process that may be used to determine the value of K is called cross-validation. In cross-validation analysis, we exclude a well from the regression analysis, compute the weights, and predict the values of the excluded well. Since we know the actual measured value of the well, we may compute the validation error between the real log and the predicted one:

$$e_V = \frac{1}{N}\sum_{j=1}^{N}\left(m_j - p_j\right)^2 \tag{4.6}$$

where:

m - measured log samples

p - predicted log samples

N - number of samples

We can repeat the process for all the wells used in the analysis and compute the average validation error E_V:

$$E_V = \frac{1}{L}\sum_{k=1}^{L} e_{Vk} \tag{4.7}$$

where L is the number of the wells.

Now the procedure for determining the meaningful attributes for a particular length of the convolutional operator can be described:

- find the best single attribute and compute the validation error, VE(1)
- find the best pair of attributes and compute the validation error, VE(2)
- if VE(2) < VE(1), find the best tripled of attributes and compute the validation error, (VE3)
- if VE(3) < VE(2), find the best four attributes and compute the validation error, VE4

- do so until VE(K+1) > VE(K)

If VE(K+1) > VE(K), the K+1 attribute introduces noise in the prediction process and the combination of the first K attributes should be used in the prediction process.

4.2.2 Neural networks

The discussed method can perform well if the functional relationship between the predicted log property and the seismic attribute is linear. In the case of a non-linear relationship, we may apply a non-linear transform prior to the least-squares optimization or we may use artificial neural networks (Haykin, 1994) as a prediction tool.

In its most general form, an artificial neural network is a set of electronic components or computer program that is designed to model the way in which the brain is thought to perform. The brain has been described as a highly complex, nonlinear, and parallel information-processing system. The structural constituents of the brain are nerve cells called neurons, which are linked by a large number of connections called synapses. This complex system has the great ability to build up its own rules and store information through what we usually refer to as 'experience'.

The neural network resembles the brain in two respects:

- knowledge is acquired by the network through a learning process;
- inter-neuron connection strengths known as synaptic weights are used to store the knowledge.

The procedure used to perform the learning process is called a learning algorithm. Its function is to modify the synaptic weights of the network in an orderly fashion to attain a desired design objective.

Although neural networks are relatively new to the petroleum industry, their origins can be traced back to the 1940's, when psychologists began developing models of human learning. With the advent of the computer, researchers began to program neural network models to simulate the complex behavior of the brain. However, in 1969, Marvin Minsky proved that one-layer perceptrons, a simple neural network being studied at that time, are incapable of solving many simple problems. Optimism soared again in 1986 when Rumelhart and McClelland published a two-volume book 'Parallel Distributed

Processing'. The book presented the back-propagation algorithm, which has become one of the most popular learning algorithms for the multi-layer feedforward neural network. Since then, the effort to develop and implement different architectures and learning algorithms has been enormous. In 1990, Donald Specht published the idea of the probabilistic neural network, which has its roots in the probability theory. The trend was picked up by research geophysicists and a number of successful applications were reported in the geophysical literature (Huang et al., 1996; Todorov et al., 1998).

4.2.2.1 Multilayered feedforward neural network (MLFN)

Figure 4.2 shows schematically the basic neural network architecture of multilayered feedforward neural network. It consists of a set of neurons that are arranged into two or more layers. There is an input layer and an output layer, each containing at least one neuron. Between them there are one or more 'hidden' layers. The neurons are connected in the following fashion: inputs to neurons in each layer come from outputs of previous layer, and outputs from these neurons are passed to neurons in the next layer. Each connection represents a weight. In the example shown on Figure 4.2, we have four inputs (four seismic attribute samples: A1, A2, A3, A4), one 'hidden layer' containing three neurons and an output neuron (the measured log sample). The number of connections is 15, i.e. we have 15 weights.

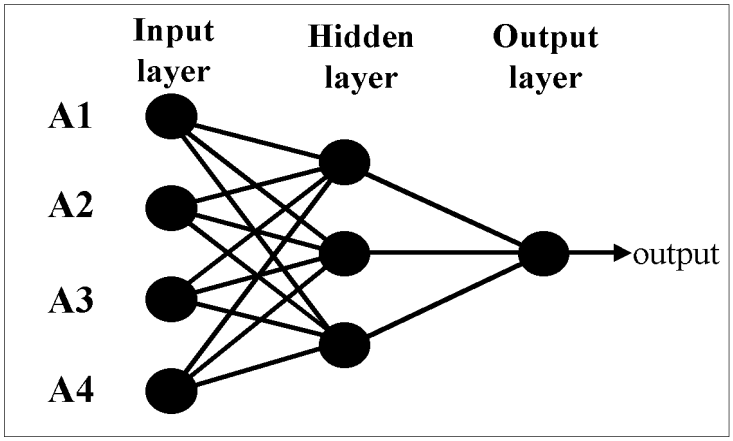

Figure 4.2: Basic neural network architecture.

A neuron is an information-processing unit that is fundamental to the operation of the neural network. Figure 4.3 shows the model of a neuron. We may identify three basic processes of the neuron model:

- each of the input signals x_j is multiplied by the corresponding synaptic weight w_j
- the weighted input signals are summed
- a nonlinear function, called the activation function, is applied to the sum

Mathematically, the process is written as:

$$\text{neuron's output} = y = f(\sum_{i=0}^{n-1} x_i w_i + w_n) \tag{4.8}$$

where:

y - neuron output

w_i - connection synaptic weights, i = 1, ..., n-1

w_n - constant called bias

x_i - neuron inputs, i = 1, ..., n-1

f - activation function

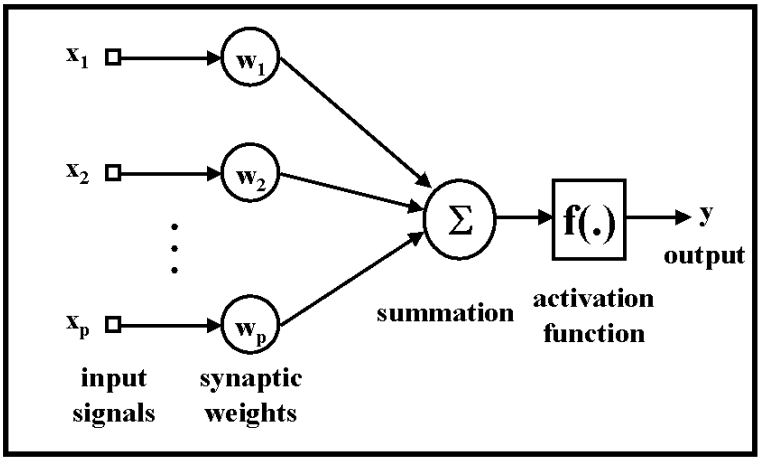

Figure 4.3: Model of a neuron.

The activation function defines the output of a neuron in terms of the activity level associated with its input. The sigmoid function is by far the most common form of

activation function used in the construction of artificial neural networks. It is defined as a strictly increasing function that exhibits smoothness and asymptotic properties. An example of the sigmoid function is the logistic function, defined by:

$$f(x) = \frac{1}{1+e^{-x}} \tag{4.9}$$

The logistic function assumes a continuous range of values from 0 to 1. It is sometimes desirable to have the activation function range from -1 to 1, in which case the activation function assumes anti-symmetric form with respect to the origin. An example is the hyperbolic tangent function, defined by:

$$f(x) = \tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \tag{4.10}$$

A neural network is completely defined by the number of layers, neurons in each layer, and the connection weights. The process of weight estimation is called training.

During the process of training, the neural network builds a model by presenting examples. Each example consists of an input-output pair: an input signal and the corresponding desired response for the neural network. Thus, a set of examples represents the knowledge. For each example, we compare the outputs obtained by the network with the desired outputs. If $y = [y_1, y_2, ..., y_p]$ is a vector containing the outputs (note that p is the number of neurons in the output layer), and $d = [d_1, d_2, ..., d_p]$ is a vector containing the desired response, we can compute the error for the example k:

$$e_k = \frac{1}{p}\sum_{j=1}^{p}(y_j - d_j)^2 \tag{4.11}$$

If we have n examples, the total error is:

$$e = \frac{1}{n}\sum_{k=1}^{n} e_k \tag{4.12}$$

Obviously, our goal is to reduce the error. It can be done by updating the weights to minimize the error. Thus, in its basic form a neural network training algorithm is an optimization algorithm which minimizes the error with respect to the network weights. Masters (1995) combines conjugate-gradient algorithm with simulated annealing for search of the global minimum of the error function.

As training carries on, the error based on the training data set gets smaller. Theoretically, given enough neurons and iterations, the error based on the training set will approach zero. However, this is undesirable since the neural net will be fitting random noise and some irrelevant details of the individual cases. This pitfall is called 'overfitting' or 'overtraining'. The problem of 'overfitting' versus 'generalization' is similar to the one of fitting a function to known points and later use the function for prediction. If we use a high-enough-degree polynomial, we may fit the points exactly (Figure 4.4, the solid line). However, if we use a 'smoother' function, the prediction of the unknown points is better (Figure 4.4, the dashed line).

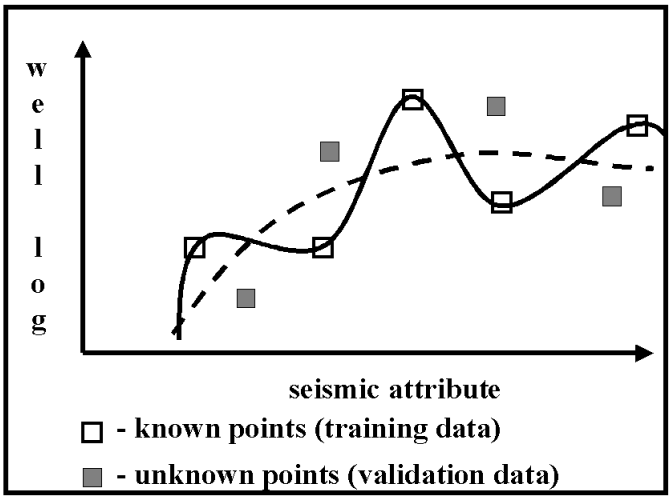

Figure 4.4: Overfitting versus generalization. The solid line fits the known points exactly, but the 'smoother' dashed line predicts the unknown points better.

To overcome the problem, we can divide the data into two sets: training and testing. The first one is used to train the neural network and the second one to evaluate its performance. The training is performed in the following fashion:

- hidden neurons are added one at a time
- training is performed and tested
- construction is stopped when the correlation on the test data shows no further improvement

During training the network builds a nonlinear mathematical model which later is applied to the seismic attributes to generate a predicted well log property cube.

4.2.2.2 Probabilistic neural network (PNN)

The basic idea behind the general regression probabilistic neural network (Specht, 1991; Masters, 1995) is to use a set of one or more measured values, called independent variables, to predict the value of a single dependent variable. The independent variable can be represented by a vector $\mathbf{x} = [x_1, x_2, ..., x_p]$, where p is the number of independent variables. The dependent variable, y, is a scalar. The inputs to the neural network are the independent variables, $x_1, x_2, ..., x_p$, and the output is the dependent variable, y. The goal is to estimate the unknown dependent variable, y', at a location where the independent variables are known. This estimation is based on the fundamental equation of the general regression probabilistic neural network:

$$y'(x) = \frac{\sum_{i=1}^{n} y_i \exp(-D(x, x_i))}{\sum_{i=1}^{n} \exp(-D(x, x_i))} \tag{4.13}$$

where n is the number of examples and $D(x, x_i)$ is defined by:

$$D(x, x_i) = \sum_{j=1}^{p} \left(\frac{x_j - x_{ij}}{\sigma_j} \right) \tag{4.14}$$

$D(x, x_i)$ is actually the scaled 'distance' between the point we are trying to estimate, x, and the training points, x_i. The 'distance' is scaled by the quantity σ_j, called the smoothing parameter, which may be different for each independent variable.

The actual training of the neural network consists of determining the optimal set of smoothing parameters, σ_j. The criterion for optimization is minimization of the validation error. For the m^{th} example, the prediction is:

$$y'_m(x_m) = \frac{\sum_{i \neq m}^{n} y_i \exp(-D(x_m, x_i))}{\sum_{i \neq m}^{n} \exp(-D(x_m, x_i))} \tag{4.15}$$

So the predicted value of the m^{th} sample is y'_m. Since we know the actual value, y_m, we can calculate the validation error:

$$e_m = (y_m - y'_m)^2 \tag{4.16}$$

The total validation error for the n examples is:

$$e = \sum_{i=1}^{n} (y_i - y'_i)^2 \tag{4.17}$$

The validation error than is minimized with respect to the smoothing parameters using conjugate-gradient algorithm.

4.3 Model-based conversion of P-S data to P-P time

To simultaneously use the seismic attributes extracted from P-P and P-S data, we have to convert the P-S data to P-P time. The two-way, zero-offset P-S time t_{PS} to a particular depth z is:

$$t_{PS} = \frac{z}{\overline{V}_P} + \frac{z}{\overline{V}_S} = z\left(\frac{\overline{V}_P + \overline{V}_S}{\overline{V}_P \overline{V}_S}\right) \tag{4.18}$$

where $\overline{V}_P$ and $\overline{V}_S$ are the P-wave and S-wave average velocities.

The depth z can be written as a function of the two-way, zero-offset P-P time t_{PP}:

$$z = \frac{t_{PP}}{2\overline{V}_P} \tag{4.19}$$

Using equations (4.18) and (4.19), we can write the two-way, zero-offset P-S time:

$$t_{PS} = \frac{t_{PP}}{2\overline{V}_P}\left(\frac{\overline{V}_P + \overline{V}_S}{\overline{V}_P \overline{V}_S}\right) \tag{4.20}$$

Solving for t_{PP}:

$$t_{PP} = 2\left(\frac{\overline{V}_S}{\overline{V}_P + \overline{V}_S}\right) t_{PS} \tag{4.21}$$

The model-based conversion scheme is done in the following way:

- at the well locations compute the P-S pseudo-velocity logs (equation 2.10)

- using the computed P-S pseudo-velocity logs, convert the V_P and V_S logs into P-S time
- build a 3-D V_P and V_S model in P-S time by 3-D interpolation
- compute the P-P time for each P-S sample using equation (4.21)

Note that the sampling rate in the resulting seismic trace is not a constant due to varying V_{PS}/V_P ratio (Figure 4.5).

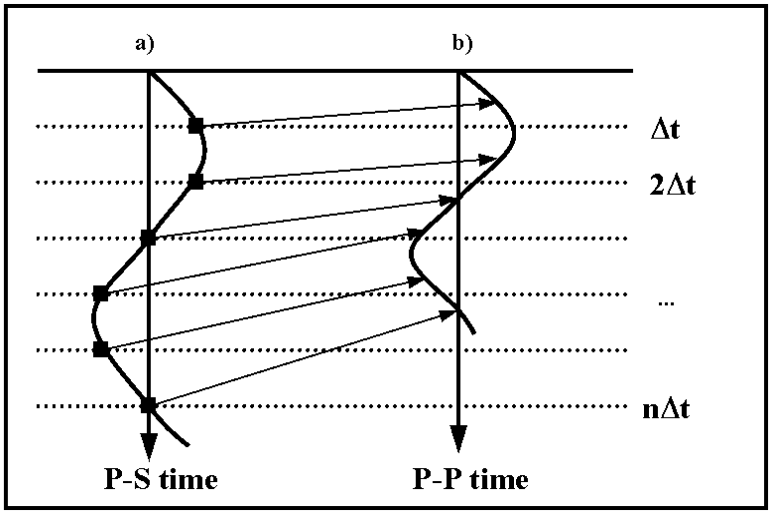

Figure 4.5: Conversion of P-S data to P-P time. Note that the regularly sampled P-S seismic trace (a) (Δt – sampling rate) becomes an irregularly sampled trace in P-P time.

The next step is to resample the converted-wave trace (now in P-P time) in a regularly sampled sequence with a constant sampling rate Δt. Let's compute the amplitude at the location nΔt (Figure 4.6). We find the actual amplitude before and after this time, A_i and A_{i+1}, and then we compute the amplitude at time location nΔt by linear interpolation:

$$A_{n\Delta t} = A_i + \frac{(A_{i+1} - A_i)(n\Delta t - PP_i)}{PP_{i+1} - PP_i} \qquad (4.20)$$

where PP_i is the computed P-P time for the i-th P-S sample.

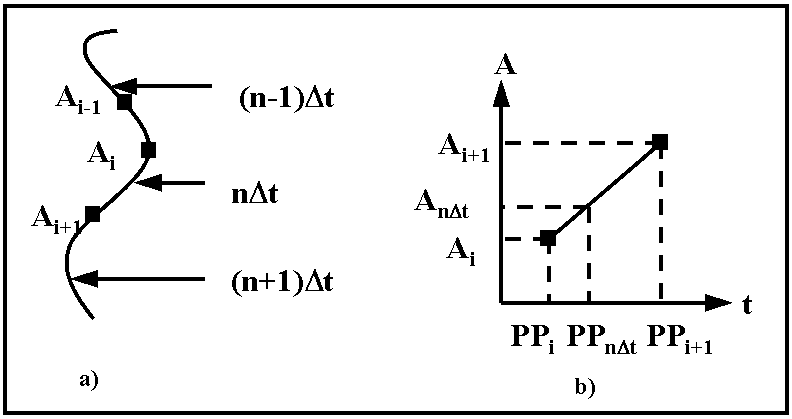

Figure 4.6: Resampling the converted P-S data to regularly sampled P-P time.

Using the previously described procedure, the Blackfoot converted-wave (P-S) data volume (Figure 4.7) is converted to P-P time (Figure 4.8).

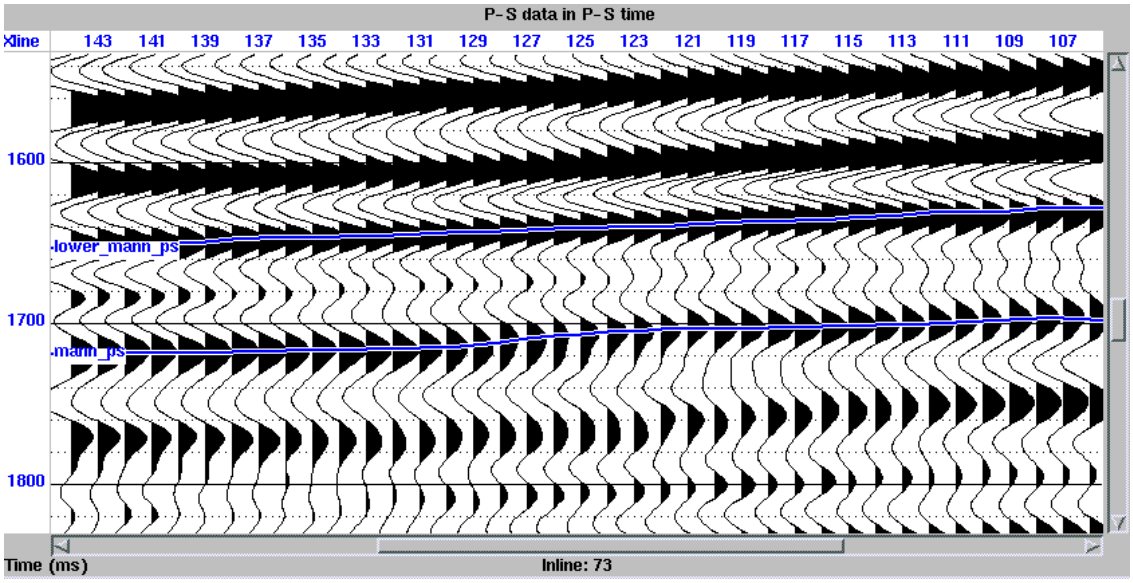

Figure 4.7: Traces from the P-S data volume in P-S time.

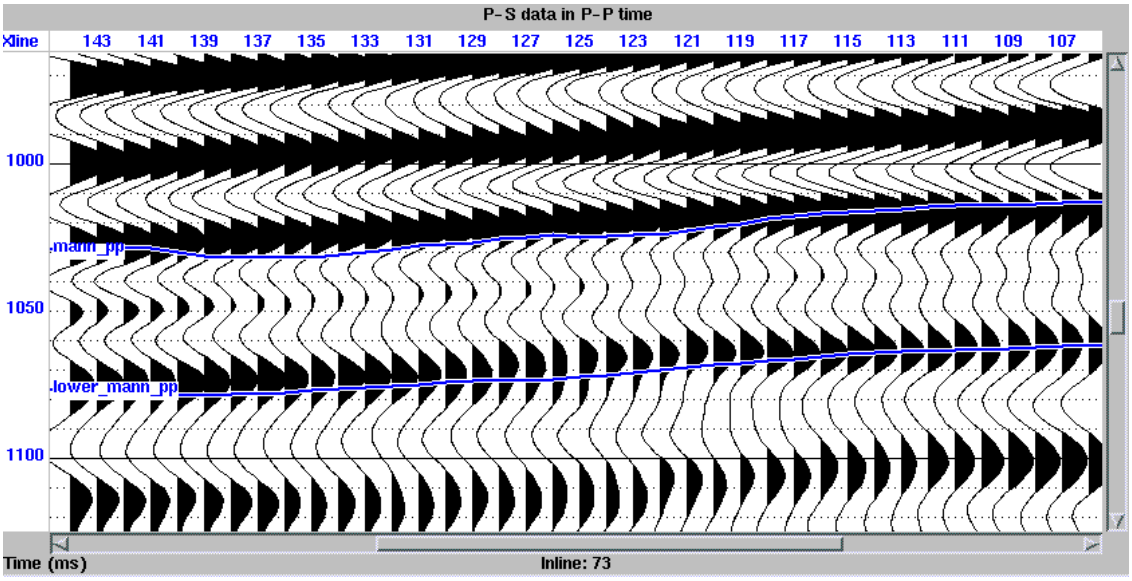

Figure 4.8: Traces from the P-S data volume converted to P-P time. The shown horizons are picked on the P-P volume.

4.4 Prediction of impedance logs

In the following section, I present an example of predicting acoustic impedance logs in the Blackfoot area. The logs are computed by multiplying the measured sonic and density logs. The first step is to convert the logs in depth to seismic time and resample them to the seismic sampling rate (2 ms). A number of seismic attributes are extracted from the seismic trace. However, because of the bandlimited nature of the seismic signal we need additional information for the low frequencies. The impedance model, built to perform the model-based inversion in chapter 2, was filtered with high-cut frequency of 20 Hz and used as an additional attribute. Figure 4.9 shows the input data for the 08-08 well: the target acoustic impedance log, the seismic trace at the well location (from which the attributes are extracted), and the low-frequency model used as an attribute. The horizontal red lines show the chosen time window for the analysis, the Mannville to Mississippian levels.

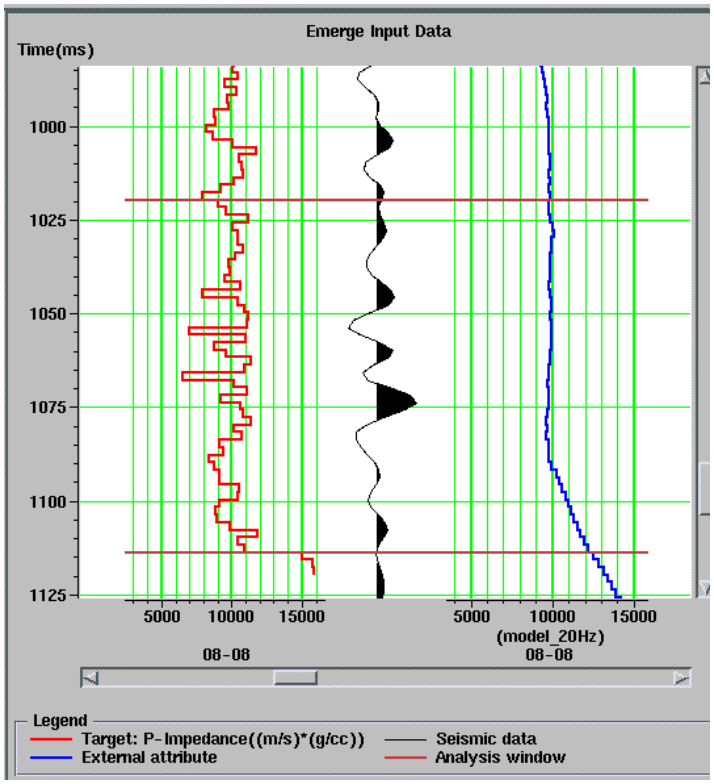

Figure 4.9: Input data for well 08-08.

Altogether 13 wells and the corresponding seismic traces (extracted at the well locations) are used in the analysis.

Table 4.1 shows the results of the performed step-wise regression using 9-point convolution operator. Note that the shown RMS error corresponds to a combination of the attribute with the ones above it. The 'Validation' column represents the cross-validation error. From a theoretical point of view, the error in the 'RMS error' column decrease as we add new attributes, but we see that by adding the fifth attribute, Average frequency, the validation error increases. So we choose to use the first four attributes in the prediction process. Figure 4.10 is graphical representation of the table 4.1. The lower black line is the error using all wells in the calculation and the upper red line is the validation error.

Attribute	RMS error (m/s).(g/cc)	Validation (m/s).(g/cc)
Integrated trace	1077	1095
Low frequency model 20 Hz	944	975
Amplitude weighted phase	916	965
Instantaneous phase	902	959
Average frequency	889	961

Table 4.1: Results from the step-wise regression.

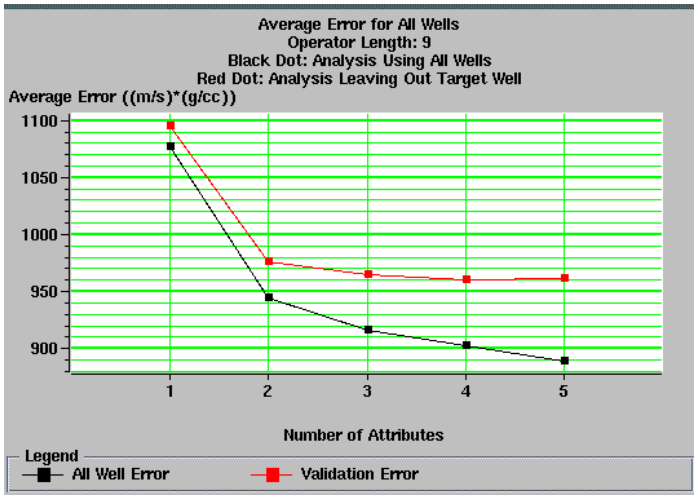

Figure 4.10: Average error as a function of the number of seismic attributes.

The two types of neural networks are trained using the same four attributes with a 9-point convolutional operator. MLFN prediction error is 846 m/s * g/cc and the validation error is 1273 m/s * g/cc. The PNN shows superior results, i.e. lower prediction error of 607 m/s * g/cc and lower validation error of 934 m/s * g/cc. Figures 4.11 and 4.12 show the measured (in black) and the predicted (in red) impedance logs at 16-08 and 29-08 locations. We see that the probabilistic neural network predicts the logs with higher accuracy. The multi-regression analysis predicted the logs with correlation 0.65 while the neural network predicted them with correlation 0.87. Figures 4.13 and 4.14 show the results from the validation analysis.

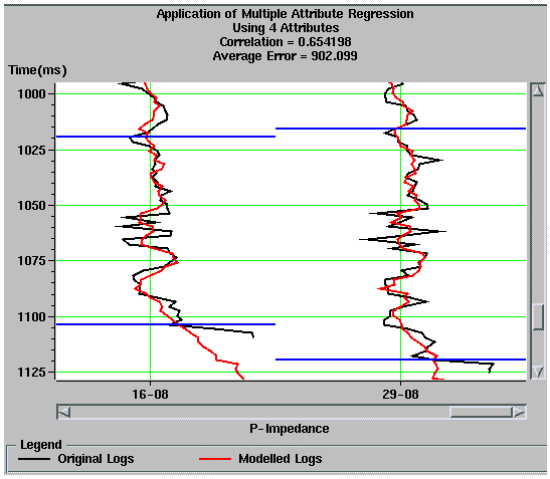

Figure 4.11: Measured impedance logs (in black) and the predicted ones (in red) using multi-regression. The correlation is 0.65.

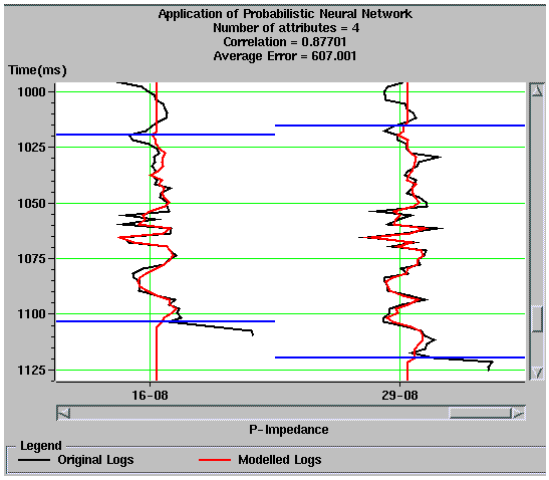

Figure 4.12: Measured impedance logs (in black) and the predicted ones (in red) using neural network. The correlation is 0.87.

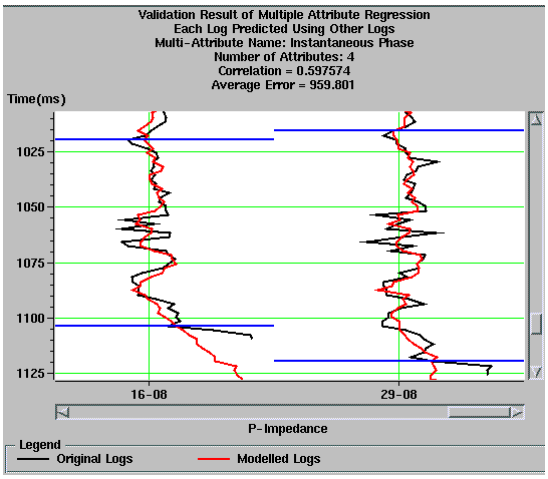

Figure 4.13: Validation result using multi-regression. The correlation is 0.59.

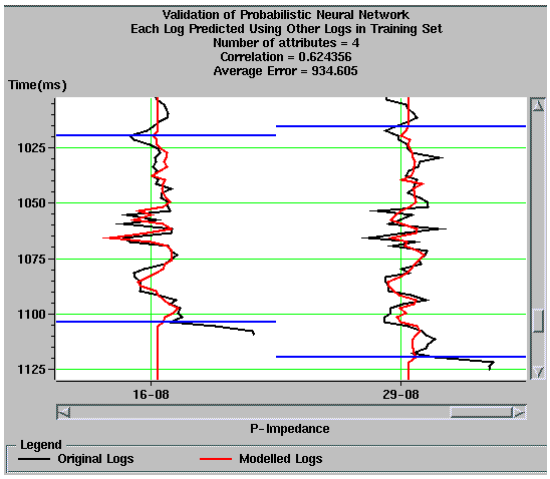

Figure 4.14: Validation result using neural network. The correlation is 0.63.

Once the relationship between the seismic attributes and the impedance logs has been determined, it is applied to the data volumes. Figures 4.15 and 4.16 show a cross-line from the predicted impedance cube. The measured impedance log at the 08-08 location is inserted. The sand channel is visible as a low impedance anomaly. Note the higher resolution achieved using the neural network.

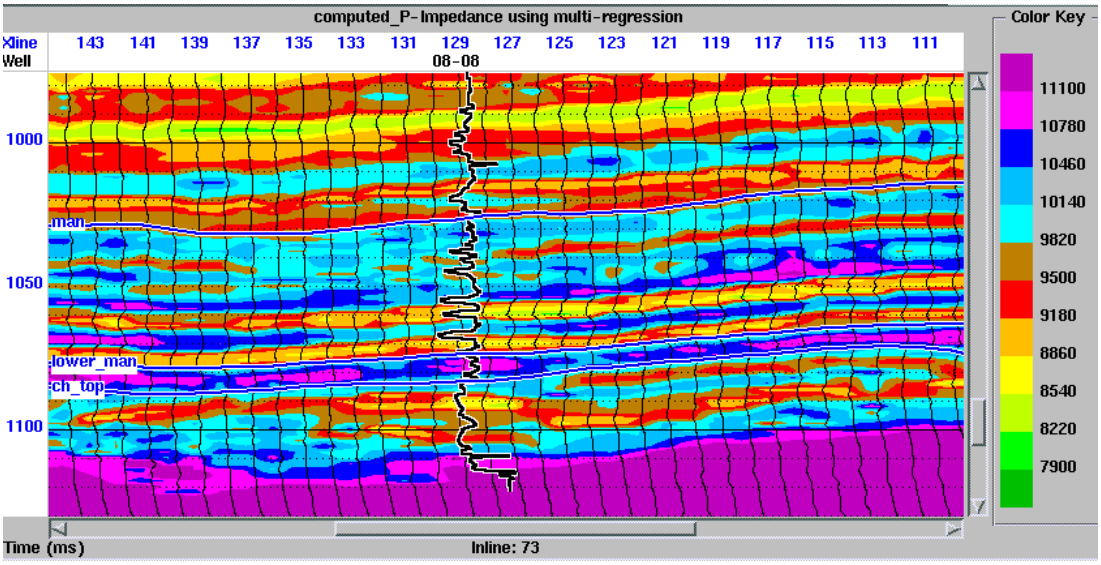

Figure 4.15: Cross-line from the predicted impedance volume using multi-regression analysis.

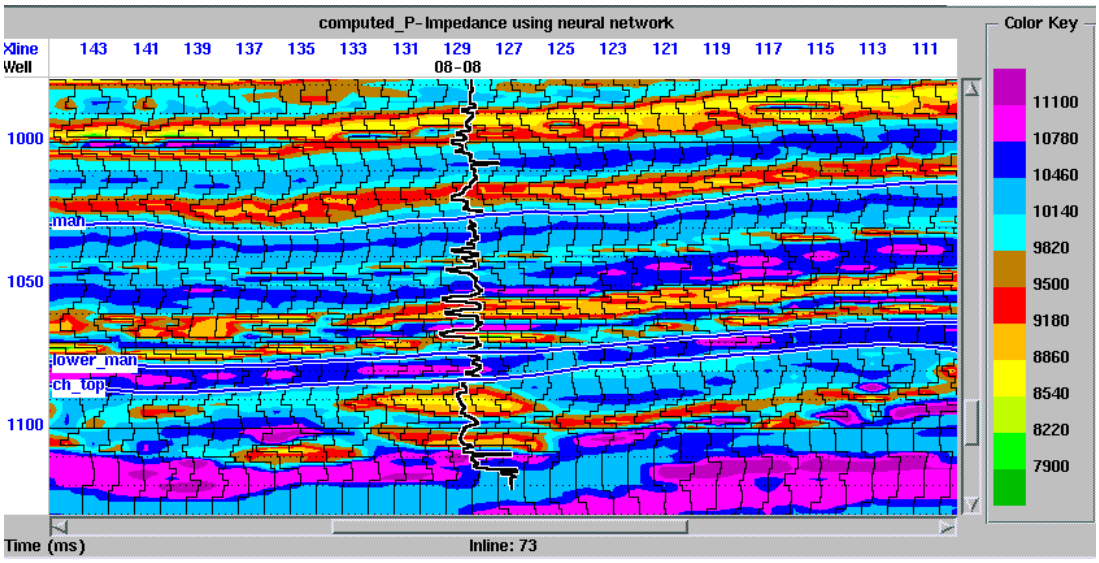

Figure 4.16: Cross-line from the predicted impedance volume using neural network.

Figures 4.17 and 4.18 are impedance data slices at the sand channel level. As in the conventional inversion, performed in chapter 2, the oil wells coincide with the low-impedance anomaly. The regional well 09-17 is similarly located in a low-impedance anomaly while the shale-fill channel has high impedance.

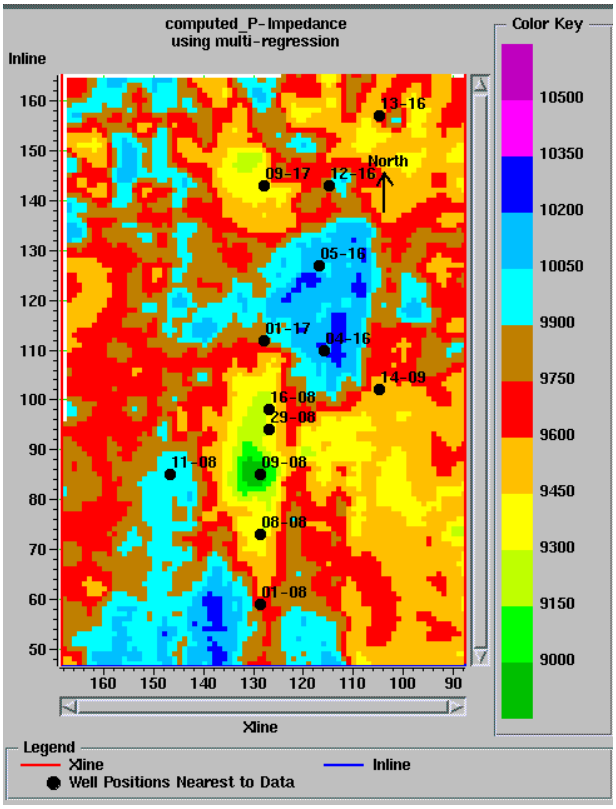

Figure 4.17: Data slice at the channel level using multi-regression analysis.

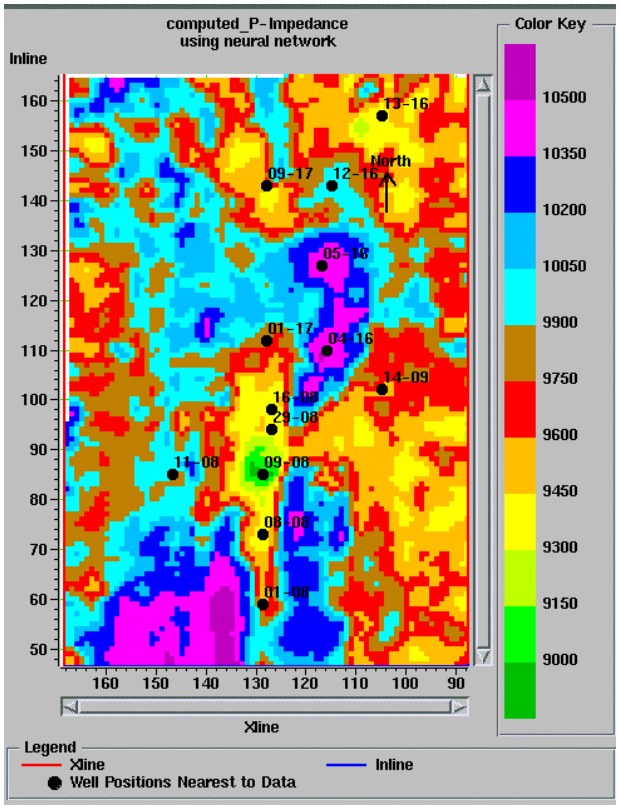

Figure 4.18: Data slice at the channel level using neural network.

4.5 Prediction of porosity logs

Porosity mapping is a major task in the exploration and development work. The second example in this chapter involves prediction of porosity logs simultaneously using attributes from P-P and P-S data. The calculation of the attributes for both data sets is the same; however, we have to convert the P-S data to P-P time. Figure 4.19 shows some of the input data: the measured porosity log, the P-S seismic trace, used to extract the P-S attributes, and some of the P-P attributes. Since the porosity is very often correlated with the impedance, the results from the model-based P-P and P-S inversions are used as additional attributes.

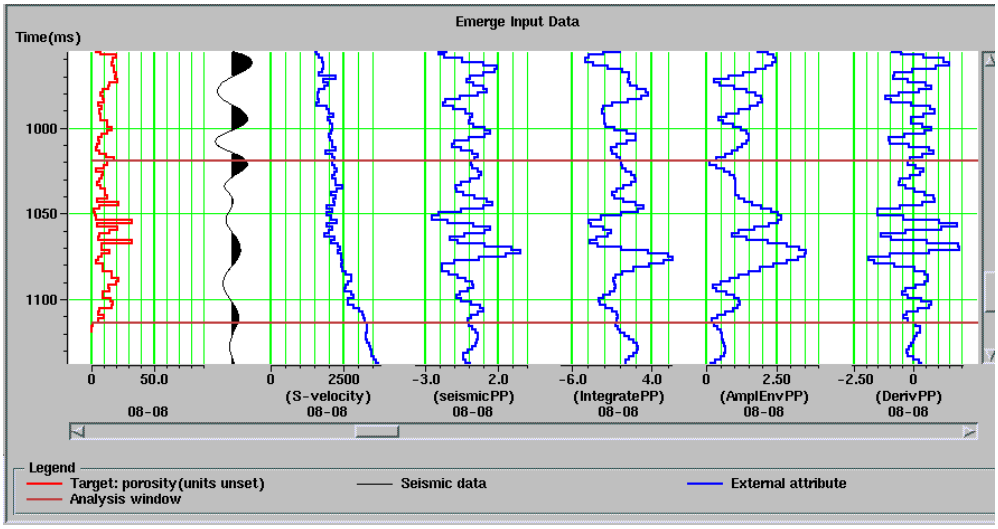

Figure 4.19: Input data.

Table 4.2 shows the results of the step-wise regression performed using a 3-point convolution operator (similar to table 4.1). In the current example, we see that by adding the seventh attribute, Seismic amplitude of the P-P trace, the validation error increases. So we choose to use the first six attributes in the prediction process. Figure 4.20 is graphical representation of the table 4.2. The lower black line is the error using all wells in the calculation and the upper red line is the validation error.

Attribute	**RMS error %**	**Validation %**
Impedance from P-P inversion	4.377	4.440
S-velocity from P-S inversion	4.315	4.401
Integrated trace (P-P)	4.256	4.396
Amplitude envelope (P-S)	4.228	4.395
Integrated trace (P-S)	4.192	4.389
Cosine instantaneous phase (P-S)	4.123	4.347
Seismic amplitude (P-P)	4.086	4.363

Table 4.2: Results from the step-wise regression.

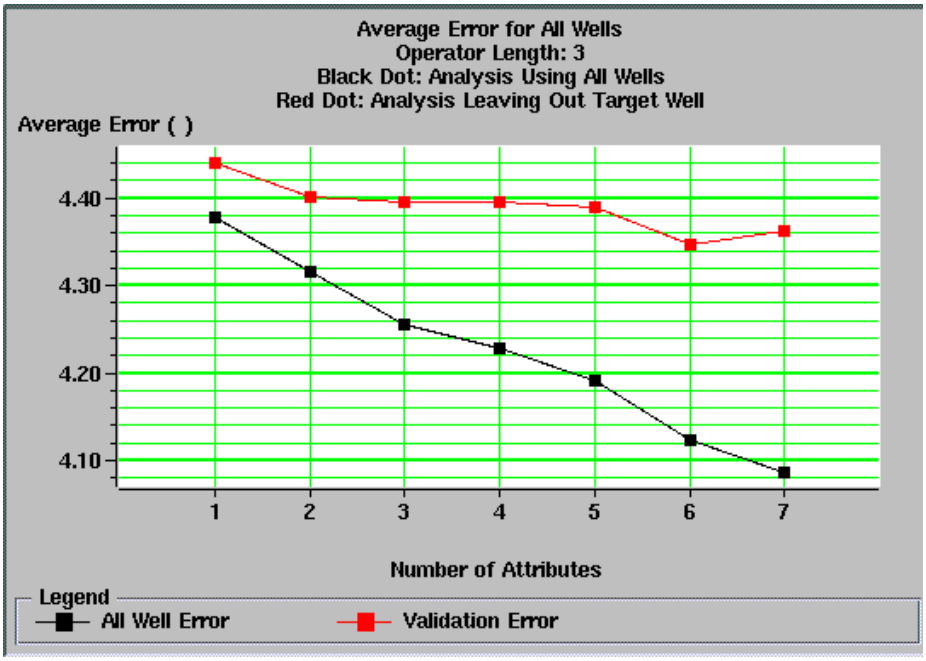

Figure 4.20: Average error as a function of the number of attributes.

As in the previous example, the two types of neural networks are trained using the same six attributes with 3-point convolutional operator. MLFN prediction error is 3.79 % and the validation error is 4.65 %. The PNN shows superior results, i.e. lower prediction error of 2.15 % and lower validation error of 3.98 %. Figures 4.21 and 4.22 show the measured (in black) and the predicted (in red) porosity logs at 08-08 and 09-08 locations. Again, we can see that the neural network predicts the logs with higher accuracy. The multi-regression predicted the logs with correlation 0.77 while the neural network predicted them with correlation 0.95. Figures 4.23 and 4.24 show the results from the validation analysis.

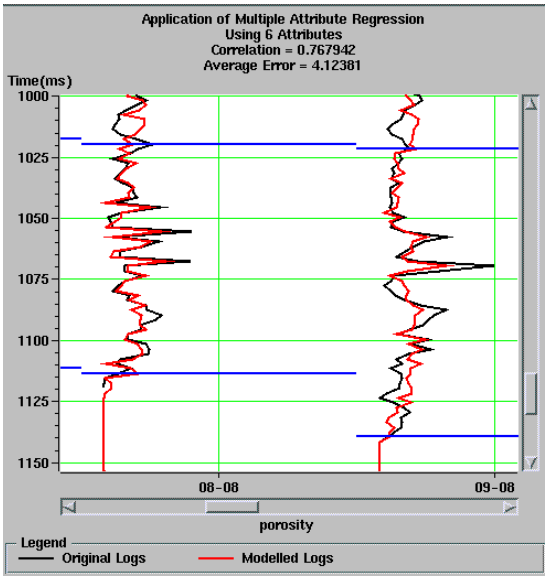

Figure 4.21: Measured porosity logs (in black) and the predicted ones (in red) using multi-regression. The correlation is 0.77.

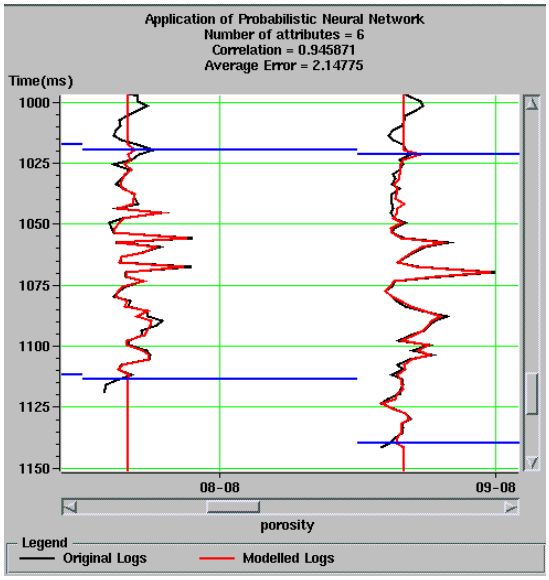

Figure 4.22: Measured porosity logs (in black) and the predicted ones (in red) using neural network. The correlation is 0.95.

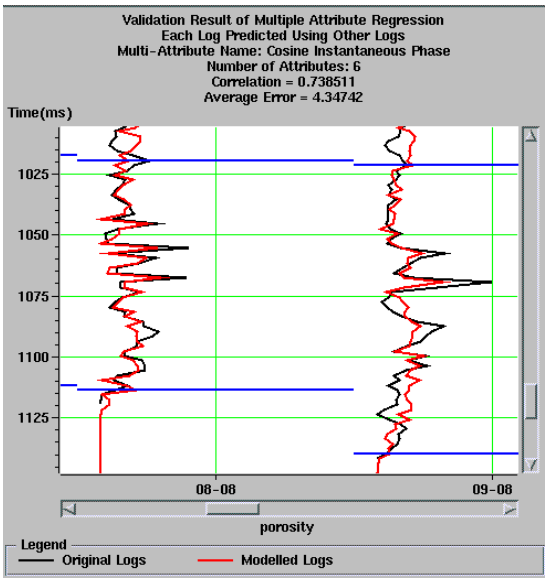

Figure 4.23: Validation result using multi-regression. The correlation is 0.74.

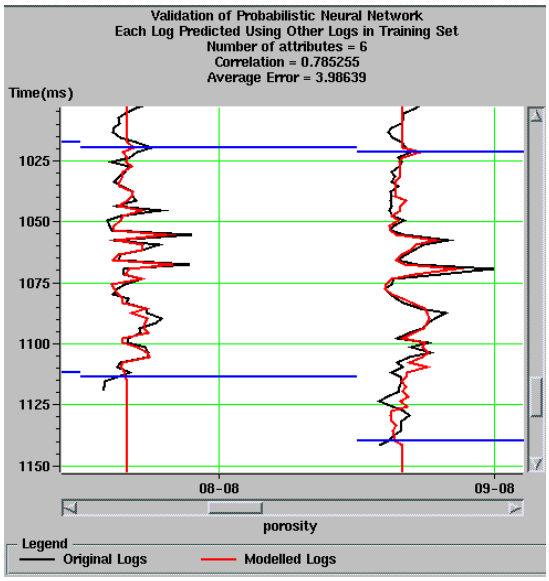

Figure 4.24: Validation result using neural network. The correlation is 0.79.

Once the relationship between the seismic attributes and the porosity logs has been determined it is applied to the data volumes. Figures 4.25 and 4.26 show a cross-line

from the predicted porosity cube. The sand channel can be distinguished very well as a high porosity anomaly. Again, note the higher resolution achieved using the neural network.

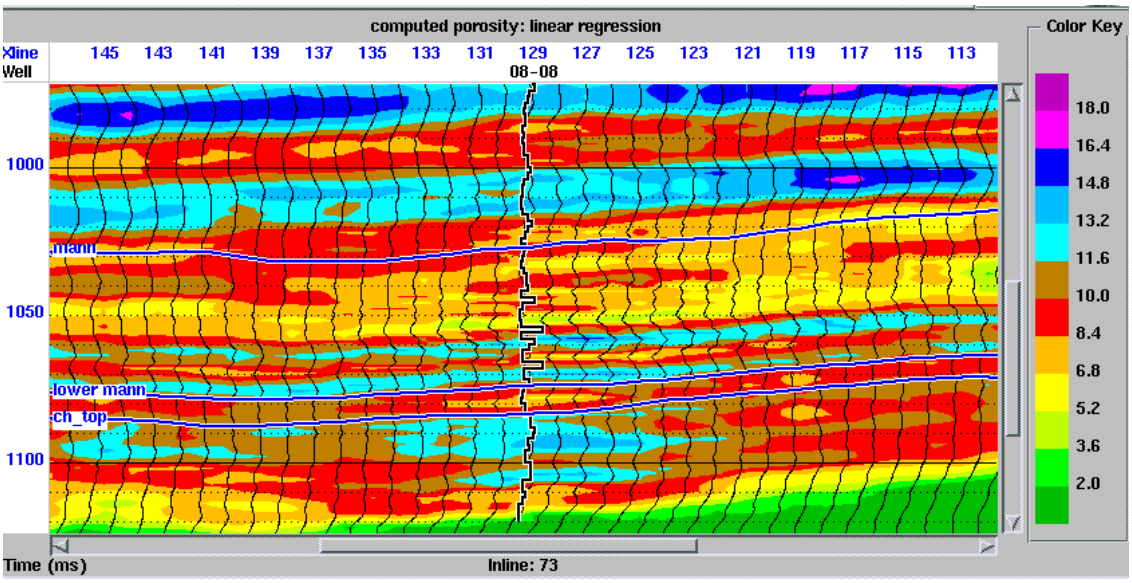

Figure 4.25: Cross-line from the predicted porosity cube using multi-regression analysis.

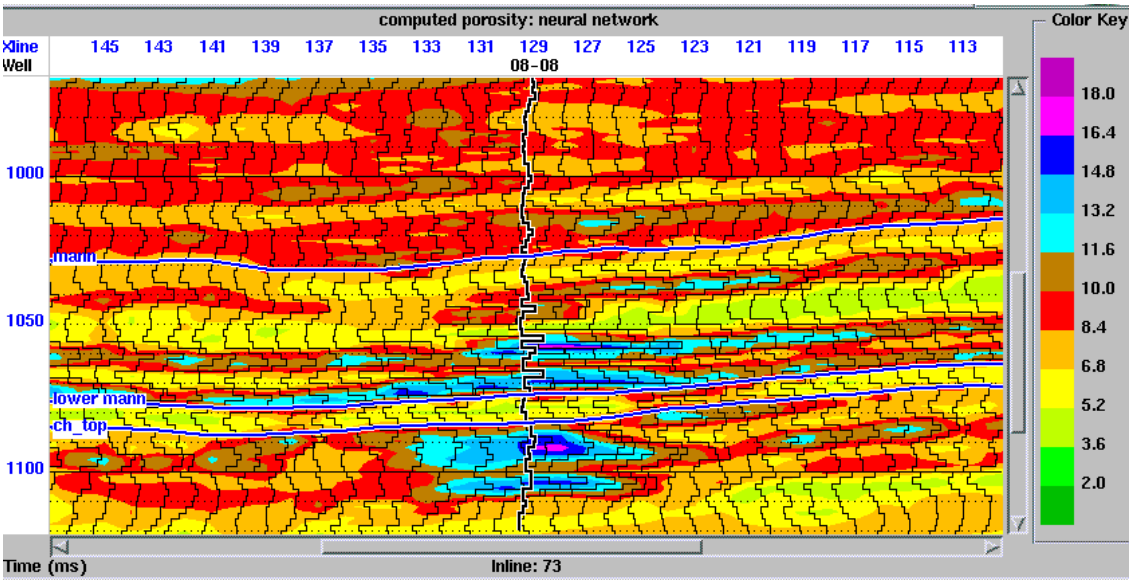

Figure 4.26: Cross-line from the predicted porosity cube using neural network.

Figures 4.27 and 4.28 are porosity data slices at the sand channel level. The oil wells coincide with the high porosity anomaly. The result from the neural network prediction identifies the channel better.

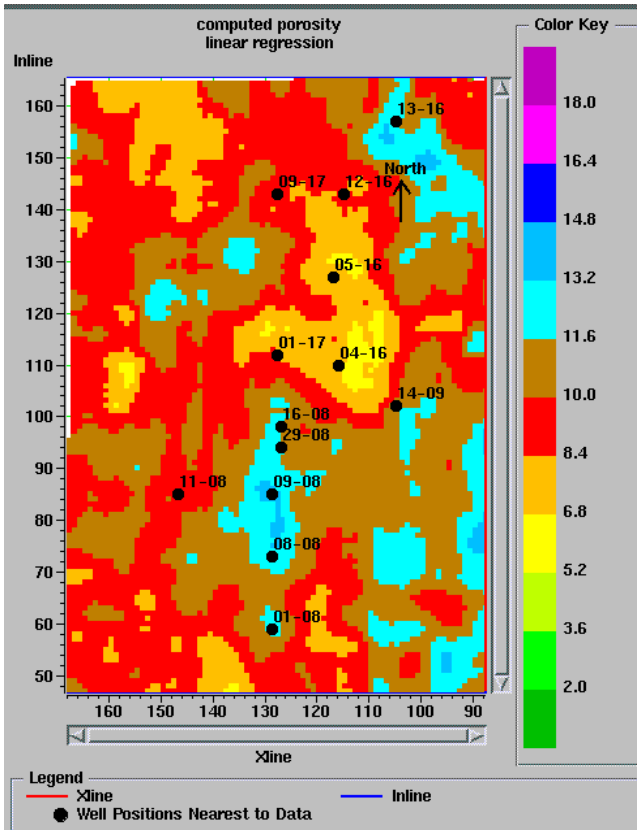

Figure 4.27: Porosity data slice at the channel level using multi-regression.

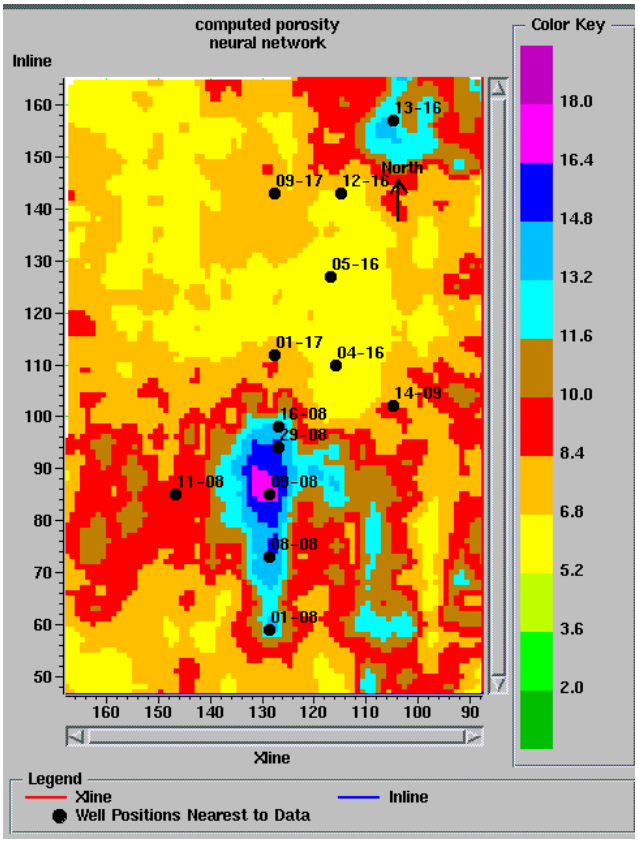

Figure 4.28: Porosity data slice at the channel level using neural networks.

4.6 Hydrocarbon reserves estimation

By multiplying isopach values, sand percentage, and porosity we can estimate the sand pore volume within the channel interval. The isopach map for the Mannville - Mississippian interval has been generated in chapter three. Since the Mannville - Mississipian interval is relatively constant, subtracting a constant value of 120 m from the isopach map can give the Channel top - Mississippian isopach map. By multiplying the sand porosity column map by the oil saturation, we can generate an oil column map. Figure 4.29 is the estimated oil column map using oil saturation of 75%. The reservoir area is then multiplied with the oil column map to estimate the oil reserves. Assuming 550 000 m^2 reservoir area with 3.3 m average oil thickness, the hydrocarbon reserves has been estimated at 11 340 000 barrels of oil.

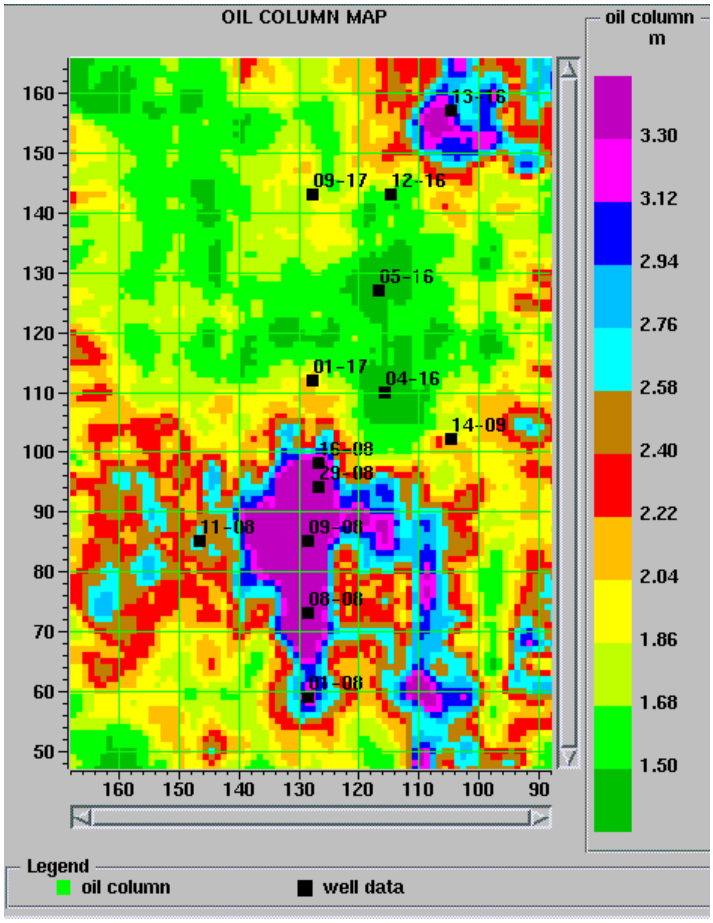

Figure 4.29: Oil column map.

4.7 Conclusions

Statistical methods have been applied successfully to estimate measured log properties from seismic attributes. Step-wise multi-regression analysis and cross-validation tests are used to determine the best attributes. Multi-regression analysis is used to find a linear relationship between the seismic attributes and the measured rock property at the well locations. A non-linear relationship is also derived using neural networks. The relationships are applied and a predicted rock property volume is generated. Cross-validation tests show the various levels of confidence in the prediction process.

Two types of neural networks, feed-forward and probabilistic, have been tested. The following conclusions can be made:

- the probabilistic general regression neural network showed lowest validation error, i.e. gives better results

- due to the random number generator, used in the simulated annealing in the feed-forward neural network, training with identical parameters may produce different results
- the training process in PNN is reproducible
- the application of the probabilistic neural network to a large data set is slow

Two real data examples have been presented. The prediction of impedance logs can be seen as a statistical algorithm for deriving the acoustic impedance of the subsurface. The derived result is similar to the traditional model-based inversion from Chapter 2. However, no theoretical assumptions have been made (like the convolutional model for example). No wavelet estimation is required, which is a major problem in traditional inversion algorithms.

Converted-wave (P-S) seismic attributes have been used simultaneously with P-P data to predict porosity. Although the best attribute to predict porosity is the acoustic impedance, four out of the best six attributes used are from P-S data.

Using the derived impedance result, the discrimination of the sand channel from the regional stratigraphy is ambiguous. However, the predicted porosity volume, integrating P-P and P-S data clearly discriminate the sand channel from the shale-plugged channel, and the regional stratigraphy.

By multiplying isopach values, sand percentage, porosity, and oil saturation, an oil column map has been generated and used to estimate the oil reserves in the field: 300 000 barrels of oil.

Chapter 5:
Conclusions

5.1 Conclusions

Integrating well log measurements and 3C-3D seismic data for improved description of the subsurface has been described. Three different approaches have been discussed: inversion, geostatistics, and multi-attribute analysis.

The goal of the post-stack, acoustic (P-P) inversion is to derive the acoustic impedance of the subsurface. The implementation of multi-component measurements leads to estimation of elastic properties. A flow for the inversion of 3-D converted-wave (P-S) data for shear velocity has been developed. It involves the computation of P-S weighted-stack from NMO-corrected CCP gathers followed by conventional inversion algorithm. An approximation formula for the incident angle in the P-S case has been derived and used in the P-S weighted-stack. The P-P and P-S inversion techniques have been applied to the Blackfot 3C-3D data set. The producing oil sand correlates with low-impedance anomalies (~9500 g/cc*m/s) from the P-P inversion while the shale-fill channel has higher impedance values (~10500 g/cc*m/s). However, some relatively low-impedance areas fall in regional geology, which may lead to an ambiguous interpretation. The ambiguity may be resolved by using the result from the P-S inversion since the regional geology has lower shear velocity than the reservoir sands.

3C-3D seismic measurements can be used to derive Vp/Vs values, which are a lithological indicator. High-correlation (0.94) between the sand/shale and Vp/Vs is found in the Blackfoot area. Geostatistics has been used to integrate the sparse sand/shale well measurements and the dense seismic data. As a result, a cokriging map of the sand/shale distribution has been generated with a relatively low absolute error. Geostatistical techniques have been used to perform depth-time conversion for the Mannville interval (~1490 - 1520 m) and thickness estimation for the Mannville-Mississippian interval (~160 - 190 m).

Statistical methods and artificial neural networks provide powerful tools for rock property estimation from seismic attributes. The theory and some practical consideration

have been discussed. Examples of impedance prediction from P-P attributes and porosity prediction from both, P-P and P-S attributes, have been shown. The predicted impedance confirms the conclusions derived from the conventional inversion. The sand-channel is identified as high-porosity anomaly (~18%). The ability of neural networks to find a non-linear relationship leads to lower prediction and validation errors compared with the linear-regression.

Although the three methods have similar objectives, they differ in a very basic level. The conventional inversion methods are based on existing physical models, i.e. the forward problem has been solved (often with assumptions and approximations). The geostatistical and multi-attribute analysis are based on statistical relationships derived from the existing data sets. They do not require an a priori physical model. Geostatistical methods explore the spatial correlation of the data, i.e. they require a variogram function. In the multi-attribute analysis we use regression or neural network to determine the relationship, which is not spatially dependent (although X, Y coordinates may be used as attributes). Note that the statistical methods depend heavily on the data quality and its representation of the physical phenomena.

By integrating some of the results from the three methods (isopach values, sand percentage, and porosity), an oil column map has been generated and used to estimate the oil reserves in the field: 11 340 000 barrels of oil.

5.2 Future work

Some future directions:

- P-S inversion

Post-stack processing of the P-S weighted-stack may improve the inversion result.

- geostatistics

The presented cokriging method can incorporate only one seismic attribute. Regression analysis can be used to predict the desired rock property using more then one attribute. Then the predicted property is used as a second variable in the cokriging method.

- neural networks

Although the back-propagation neural network is probably the most used type of neural network, in my work it has been outperformed by the probabilistic general-regression network. Further research in the area of 'overfitting versus generalization' should be done. Possible solution can be found in developing a back-propagation neural network with regularization operator.

References:

Aki, K., and Richards, P., 1980, Quantitative seismology: theory and methods: W. H. Freeman and Co.

Chen, Q., and Sidney, S., 1997, Seismic attribute technology for reservoir forecasting and monitoring: The Leading Edge, 16, No. 5.

Cooke, D., and Schneider, W., 1983, Generalized linear inversion of reflection seismic data: Geophysics, 48, No. 10.

Doyen, P., 1988, Porosity from seismic data: a geostatistical approach: Geophysics, 53, No.10.

Draper, N., and Smith, H., 1981, Applied regression analysis: John Wiley & Sons Inc.

Ferguson, R., 1996, P-S seismic inversion: Modeling, Processing, and Field Examples: M.Sc. thesis, University of Calgary.

Garrota, R., 1987, Two-component acquisition as a routine procedure, in Danbom, S., and Domenico, S., Eds., Shear-wave exploration: Geophysical development series, 1.

GEOSTAT Documentation, 1999: Hampson-Russell Software Ltd.

Gorell, S., 1995, Using geostatistics to aid in reservoir characterization: The Leading Edge, 14, No. 9.

Haykin, S., 1994, Neural Networks: Prentice Hall.

Hirsche, K., Mewhort, L., Porter-Hirsche, J., and Davis, R., 1996, Geostatistical reservoir characterization of a Canadian reef or the use and abuse of geostatistics: CSEG Recorder 21, No. 12.

Huang, Z., Shimeld, J., Williamson, M., and Katsube, J., 1996, Permeability prediction with artificial neural network modeling in the Venture gas field, offshore eastern Canada: Geophysics, 61, No. 2.

Hwang, L., and McCorkindale, D., 1994, Troll field depth conversion using geostatistically derived average velocities: The Leading Edge, 13, No. 4.

Isaaks, E., and Srivastava, R., 1989, Applied geostatistics: Oxford University Press.

Journel, A., 1989, Fundamentals of Geostatistics in five lessons: American Geophysical Union.

Lawton, D., Stewart, R., Cordsen, A., and Hrycak, S., 1996, Design review of the Blackfoot 3C-3D seismic program: The CREWES Project Research Report, 8.

Lindseth, R., 1979, Synthetic sonic logs - a process for stratigraphic interpretation: Geophysics, 44, No. 1.

Lines, L., and Treitel, S., 1984, A review of least-squares inversion and its application to geophysical problems: Geophysical Prospecting, 32, No. 2.

Lu, H., and Margrave, G., 1988, Reprocessing the Blackfoot 3C-3D seismic data: The CREWES Project Research Report, 10.

Masters, T., 1995, Advanced algorithms for neural networks: John Wiley & Sons, Inc.

Matherton, G., 1963, Principles of geostatistics, Economic Geology, 58.

Miller, S., 1996, Multicomponent seismic data interpretation: M.Sc. thesis, University of Calgary.

Oldenburg, D., Scheuer, T., and Levy, S., 1983, Recovery of the acoustic impedance from reflection seismograms: Geophysics, 48, No. 10.

Russell, B., 1988, Introduction to seismic inversion methods: The SEG course notes series, 2.

Schultz, P., Shuki, R., Hattori, M., and Corbett, C., 1994, Seismic-guided estimation of log properties: The Leading Edge, 13, No. 5.

Specht, D., 1991, A general regression neural network: IEEE Transactions on Neural Networks, 2.

Specht, D., 1990, Probabilistic neural networks: Neural Networks, 3.

Stewart, R., 1990, Joint P and P-SV inversion: The CREWES Project Research Report, 2.

Taner, M., Koehler, F., Sheriff, R., 1979, Complex seismic trace analysis: Geophysics, 44, No. 6.

Tarantola, A., 1987, Inverse problem theory: Elsevier .

Tatham, R., and McCormick, M., 1991, Multicomponent seismology in petroleum exploration: Society of Exploration Geophysicist.

Todorov, T., Stewart, R., Hampson, D., and Russell, B., 1988, Well log prediction using attributes from 3C-3D seismic data: 68th Annual SEG Meeting, Expended Abstracts.

Todorov, T., and Stewart, R., 1998, 3-D converted-wave inversion for shear velocity: The CREWES Project Research Report, 10.

Tonn, R., 1998, Seismic reservoir characterization of Montney sand in the Peace River Arch area, Canada: The Leading Edge, 17, No. 5.

Treitel, S., Lines, L., and Ruckgaber, G., 1988, Geophysical inversion and its applications: Course notes.

Vestrum, R., and Stewart, R., 1993, Joint P and P-SV inversion: application and testing: The CREWES Project Research Report, 5.

Wackernagel, H., 1995, Multivariate geostatistics: Springer.