CHARACTERIZATION AND MODELING OF STRUCTURED NOISE IN SEISMIC REFLECTION DATA

Master thesis in Earth and Space Physics and Engineering - Earth Physics and Exploration

by

Christian Lundmand Jensen

Submitted August 1st, 2013

Student number: s071278
Report serial number: IMM-MSc-2013-0082

Supervisors:
Professor Klaus Mosegaard and
Associate Professor Thomas Mejer Hansen
Department of Applied Mathematics and Computer Science
Technical University of Denmark

Technical University of Denmark
Department of Applied Mathematics and Computer Science
Building 324, DK-2800 Kongens Lyngby, Denmark
Reception: +45 45253370
reception@imm.dtu.dk
www.imm.dtu.dk IMM-MSc-2013-0082

Abstract

Obtaining reliable information about the noise in seismic reflection data can be difficult. Noise is often estimated from "expert knowledge" or other available information on which uncertainty is hard to quantify. I propose a method to characterize the noise directly from the data, without the need for external information, by using 4-D data. It is shown, that the covariance of the residual of two 3-D data sets (together constituting 4-D) with identical geological subsurface is a good representation of the noise in the summed data set. The noise is characterized in the form of a positive definite semi-variogram model, and hence can be used in any probabilistic inversion. The method is demonstrated on both synthetic and real data by solving the linear least-squares problem (LSQ, Tarantola [2005]) without change of parameters, effectively removing the modeled noise from the data. Tests on the synthetic data with realistic noise levels show almost perfect removal of additive noise. A real 4-D dataset from the Halfdan field in the North Sea with severe acquisition striping was also de-noised in this way. The data were mild- to moderately non-Gaussian, but the correlated noise was still convincingly attenuated. It is expected, that many commercial inversion algorithms use some form of linearized LSQ. Hence, the method developed here will easily integrate into existing software.

Resume (DK)

Indhentning af pålidelige oplysninger om støj i seismiske reflektionsdata kan være svært. Støj er ofte estimeret fra "ekspertviden" eller andre tilgængelige oplysninger hvis usikkerhed er svær at kvantificere. Jeg foreslår en metode til at karakterisere støj direkte fra data, uden behov for ekstern information, ved anvendelse af 4-D data. Det er vist, at kovariansen af residualet af to 3-D datasæt med identisk geologisk undergrund er en god repræsentation af støj i det summerede datasæt. Støjen er karakteriseret i form af en positiv definit semi-variogrammodel, og kan dermed anvendes i enhver probabilistisk inversion. Metoden er demonstreret på både syntetiske og ægte data ved at løse det lineære mindste kvadraters problem (LSQ, Tarantola [2005]) uden at ændre parametre mellem model- og datarum, hvilket i praksis fjerner den modelerede støj fra dataene. Test på syntetiske data med realistiske støjniveauer viser næsten perfekt fjernelse af additiv støj. Et ægte 4-D datasæt fra Halfdan feltet i Nordsøen blev også behandlet for støj ved denne metode. Dataene var mildt til moderat ikke-Gaussiske, hvilket vanskeliggører brugen af LSQ, men metoden fjernede alligevel meget af den korrolerede støj. Det forventes, at mange kommercielle inversionsalgoritmer bruger en form for lineariseret LSQ. Derfor vil metoden udvilket her nemt kunne integreres i eksisterende software.

Preface

This thesis was prepared at the department of Applied Mathematics and Computer Science at the Technical University of Denmark in fulfillment of the requirements for acquiring an M.Sc. in Geophysics and Space Technology.

The thesis consists of a methodology part, a theory part, a simulation part, and an application to a large, real data set. The methodology chapter presents an overview of the complete procedure, and in particular explains why the noise characterization is only possible with 4-D data. The basic theory upon which the method is developed is introduced more or less immediately before it is used, and is therefore scattered somewhat throughout the report. All simulation was done in MathWorks® MATLAB. An existing geostatistical software package, mGstat, along with MATLAB's build-in functions were used for generating models, simulating data and interpreting results. The inversion (the process of removing the noise) was written by myself. The full data set provided by Maersk Oil has not been certified for public release. However, the thesis includes only the statistical properties, results and conclusions on the data, and can thus be distributed freely. A set of learning objectives formulated in the initial phase of the project, can be found in the appendix.

Lyngby, August 1st, 2013

Acknowledgments

I would like to thank my contact at Maersk Oil and Gas, Adam Cherret, for providing the Halfdan data, and for acting as an external advisor without the obligation to do so. I would also like to thank my friends, family and girlfriend for distracting me while researching and writing this thesis. Lastly, I would like to thank my advisors Klaus Mosegaard and Thomas Mejer Hansen and their group of Ph.D. students for letting me take part in their weekly discussions and everyday life. And for making Italian espresso, exactly when I needed it.

Contents

Chapter 1

Introduction

As the oil industry is exploring and exploiting natural resources in increasingly difficult environments, the need for very accurate subsurface parameter estimation arises accordingly. Higher accuracy can be achieved by minimizing noise in the data from which the parameters are derived. However, the degree to which noise can be minimized is directly linked to the quality of information available on the noise itself. Unfortunately, noise is an unavoidable part of measuring any real quantity. Despite this, noise is rarely investigated in detail. Therefore, it is the purpose of this thesis to develop a method which objectively estimates an accurate, detailed description of the statistical properties of the noise.

I begin by introducing the concept of noise and noise removal. The noise characterization and removal method will be explained in Chapter 2, followed by the theory of inverse problems in Chapter 3. As will be explained in the coming chapters, the variogram is central to the method. Hence, Chapter 4 is dedicated to variogram theory and simulation. In a synthetic case, noise will be estimated and removed in Chapter 5. The method will be repeated on real data in Chapter 7. The chapter in between is dedicated to investigating the statistical properties of the real data. The project is concluded with a discussion and conclusion in Chapter 8.

1.1 What is noise?

The concept of noise in data is to many a very intuitive way of characterizing the part of a signal or data collection which is not due to the property we would like to measure. E.g. wind noise on the cell phone. Or CMB radiation induced static on the radio. Noise is everywhere. However, most people are only concerned about what information (data) that lies behind the noise. The industries which have a commercial interest in the data strive to enhance the quality of these data by removing the noise, or rather, by *suppressing* or *reducing* whatever they think is not supposed to be there. Hence the term "noise reduction".

Noise can be reduced, even sometimes completely removed, if there exists, or one can estimate, a mathematical model of the noise. E.g. a probability density function (pdf) describing the statistical properties of the noise such as mean and variance. Unfortunately, obtaining a true description of the noise is complicated, since one usually does not know exactly what created that noise in the first place. Fortunately, the human brain is very good at picking out noise in an otherwise well behaving signal. In many cases, good (but not necessarily true) noise models can be estimated just by looking at the data. For instance, anyone listening to a radio station that is not perfectly tuned, will be able to tell that the constant "shh" sound in the background is noise, despite none of them knowing

the true origin of that noise. A mathematical description of this type of noise can then be estimated and the noise removed.

In a seismic survey, the picture is not so simple. Usually many types of noise have contaminated the signal, hence estimating the noise may be very difficult. Increasingly demanding environments (high temperature, high pressure, narrow reservoirs, etc.) call for detailed surveys and high confidence in the seismic processing. Increasing the accuracy of the noise model is a step in that direction.

1.2 Marine reflection seismics

Reflection seismology utilizes the fact, that different rock types in the Earth have different wave velocities. Hence, a signal (a pressure wave) propagating downwards will be reflected at different layer interfaces and received at the surface where it was originally emitted. Pressure, then, as a function of time, constitutes the raw seismic data. This is later heavily processed and converted into a time dependent reflection amplitude, which is how reflection seismics are usually presented. At sea, data are collected by dragging a tow line of hydrophones behind a vessel, while firing an air gun or other pressure wave inducing device at regular intervals. Figure 1.1 is a schematic drawing of the basic setup. A single pass above some feature of interest will yield a 2-D cross-section of the subsurface reminiscent of Figure 1.2.

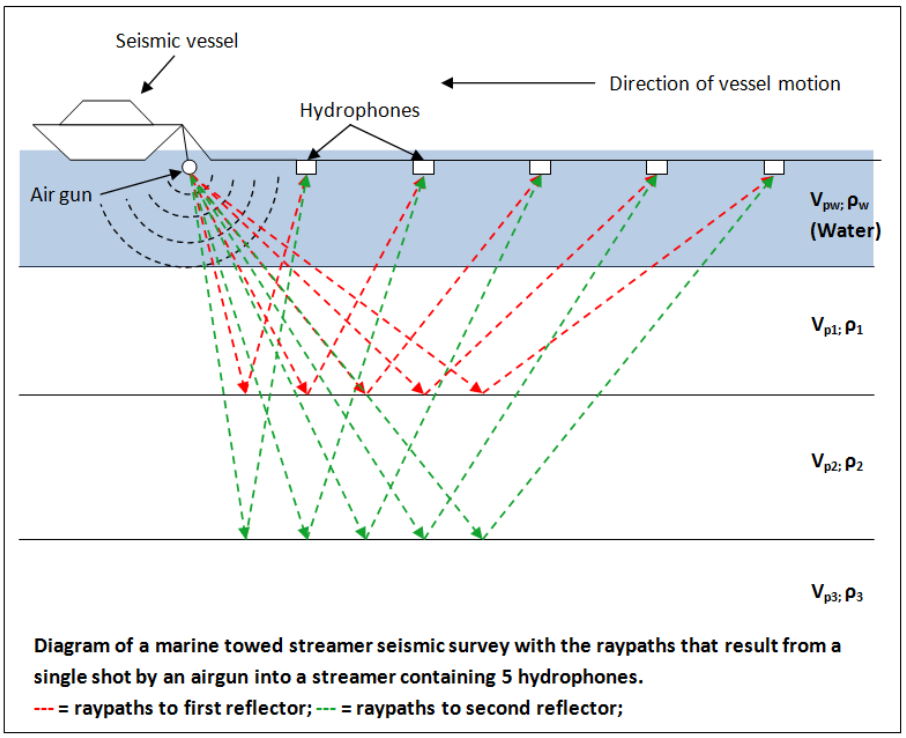


Figure 1.1: Diagram of a marine seismic survey.

A 3-D seismic survey is shot much the same way as 2-D, but with numerous, closely space tow lines laid out in a grid to obtain high spatial resolution. After acquisition, the data must be processed in several ways to obtain what we refer to as seismic reflection data. Some common corrections are Normal Moveout correction (NMO), stacking by Common Midpoint gather (CMP), and low frequency noise filtering. Subsequently, the data can be

converted or inverted to yield more interesting parameters, such as time/depth conversion, acoustic impedance, porosity or oil saturation.

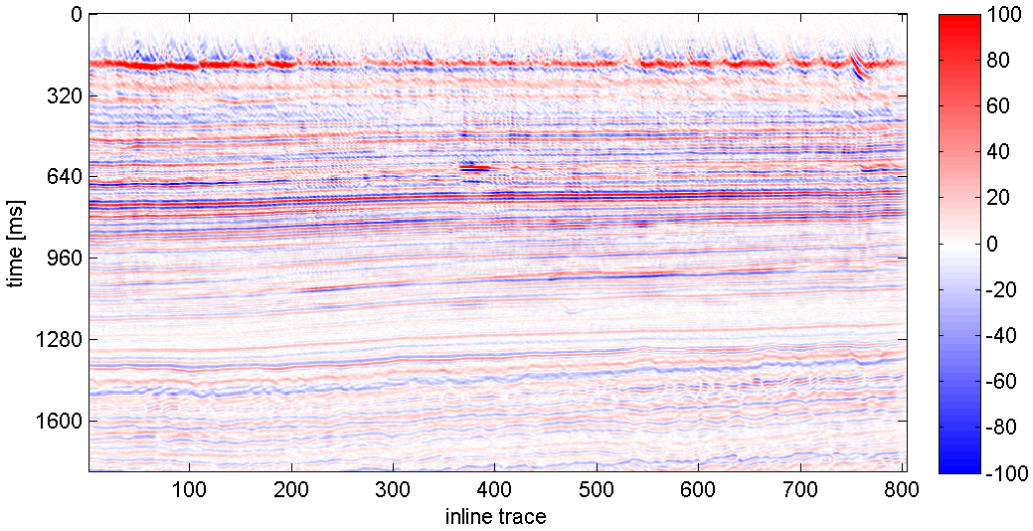


Figure 1.2: Seismic reflection amplitude versus time. 2-D cross-section of 3-D Halfdan data along random inline.

1.3 Noise in marine seismics

Noise can be divided into two groups: Coherent and random. In marine seismics, the most common source of noise is bad weather conditions OGP and IAGC [2011], manifesting itself as uncorrelated or random noise in the data. Coherent noise, such as acquisition noise, is generated by infrastructure, other vessels, and the vessel itself. The data you see in Figure 1.2 is affected by acquisition noise, especially in the upper layers. This is better illustrated in a time slice like Figure 1.3.
Understanding how these unwanted sources affect the seismic signal is key to eliminating them. Of course, the best way to ensure high quality data, is to use well calibrated, high precision equipment, "removing" noise before they enter the data. However, some types of noise cannot be attenuated in this way and must be mathematically removed post acquisition. Unfortunately, it is near impossible to have a complete understanding of the noise. Hence, noise processing tends to be based largely on intuition. The danger in this approach is that intuitive processing cannot be completely accounted for, leading to results which cannot be reproduced.

In reality, even though one seldom knows the exact type and magnitude of noise obscuring the data, trial and error, along with experience, is often enough to effectively remove what is believed to be noise. However, one can only wonder how much of the true signal is actually affected by the possibly inexact noise description? In particular, what is the impact on uncertainty?

The method developed here is an attempt to characterize noise and uncertainty as realistically and objectively as possible, by deriving information about the noise straight from the data. An objective noise estimate will increase trust in all results derived from the data. The method is developed specifically to remove 2-D acquisition noise, working on time slices (i.e. arrays of data as a function of x and y coordinates, for a constant value

of reflection travel time) like Figure 1.3, but will in principle work in any directions with any number of dimensions. The next chapter will explain the methodology.

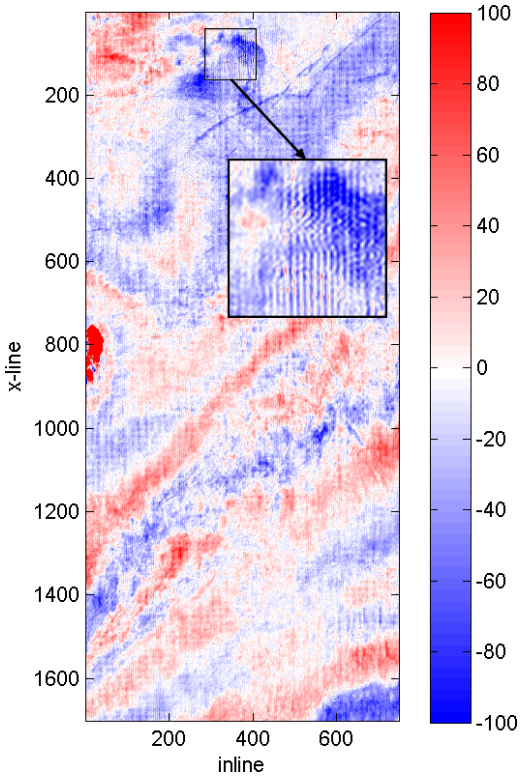


Figure 1.3: Time slice of Halfdan data (seismic reflection amplitude) in the shallow part of the overburden.

Chapter 2

Methodology

In this chapter, I explain the method in small sections. It is not meant as a complete description of all aspects of the method. Only to serve as an overview. Each topic will be thoroughly investigated in due course. The complete procedure is illustrated in the flowchart figure 2.1. It has three main parts:

1) Characterizing and modeling the noise.
2) Modeling the (noise-free) geology.
3) Removing the noise.

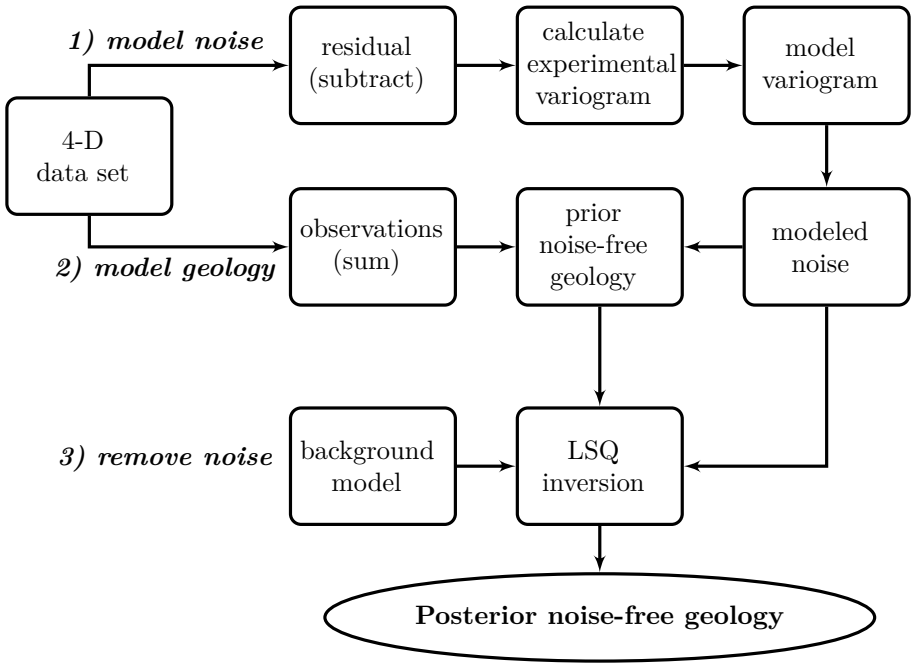


Figure 2.1: Flowchart of the general procedure. The main goal is to model the noise covariance from the residual of the data. Then remove the noise via LSQ inversion.

The focus of part 1, is to characterize and model the noise correctly in accordance with the observations, in order to obtain an objective probabilistic description of the noise.

Part 2 repeats part 1 but on the geology. Part 3 describes how the noise is removed and how to interpret the result. The chapter concludes with a brief description of the set of assumptions needed for this method to be valid.

Characterizing and modeling the noise correctly is central to the validity of any inversion results. The modeling procedure is therefore described in detail in Chapter 4. Modeling the geology (whatever signal that is not noise) is also required, but is not important for characterizing the noise and hence will not receive the same attention. The noise is removed by applying a special form of least-squares (LSQ) inversion where the forward model operator is the identity matrix. The inversion method is described in Chapter 3.

Some confusion, if not handled now, might arise in relation with the word "model": When one, in inverse problem theory, refers to a "model", it refers to the unknown set of parameters, in our case the noise-free data (geology). When mentioned in context with geostatistics, the "model" refers to a semi-variogram model or combination of semi-variogram models. The two should not be confused. I also use the verb *to model* as in fitting a semi-variogram model to data. The overlap is unfortunate but unavoidable when mixing inverse problem theory with geostatistics. When the two definitions occur together, I will attempt to clarify their meaning more literally.

2.1 The 4-D data

The method relies on two sets of 3-D data, acquired from the same area, but at different times, collectively constituting one 4-D data set. Each survey is assumed to include additive noise and hence can be separated into two parts: The noise-free data and the noise. The interesting part is of course the noise-free data, which depend only on the geological structure in the subsurface. I will refer to the noise-free data as *geology*. Whatever is not geology, is then considered and referred to as *noise*. Following this setup, the two surveys, referred to as *Base* and *Monitor*, can be described as

$$\text{Base: } s_1 = g + n_1 \qquad \text{Monitor: } s_2 = g + n_2. \tag{2.1}$$

Both surveys are assumed to include the same geology g. Hence, for this to hold, geology must be fairly undisturbed in between acquisitions; if production is presumed to cause structural changes to the formation in between the two surveys, results can become unreliable.

Throughout this report it is assumed that both geology and noise can be described by a stochastic, stationary Gaussian process. Hence, g, the true, noise-free geology, and the noise n_1 and n_2 are all realizations of their individual stochastic processes (see full set of assumptions in section 2.6). This assumption is necessary when we isolate and describe the noise from its covariance in the next section.

2.2 Residual and Observations

To filter the noise correctly, the noise must be removed from the *summed* signal $s_1 + s_2$ and characterized by the *differenced* signal $s_1 - s_2$. Why this is will become clear in a moment. I will refer to the summed signal as the *Observations* and the differenced signal as the *Residual*. Then we have

$$\text{Observations: } d_{obs} = s_1 + s_2 = 2g + n_1 + n_2 \qquad \text{Residual: } d_{res} = s_1 - s_2 = n_1 - n_2. \tag{2.2}$$

Note that the Observations and the Residual distributions are also Gaussian if g, n_1 and n_2 are Gaussian. The Residual then only includes (Gaussian) noise, as the geology cancels out. Specifically, it includes noise in the form $n_1 - n_2$. The Observations on the other hand, includes noise (also Gaussian) in the form $n_1 + n_2$. Thus, Observations and Residual contain different realizations of noise. However, if N_1 and N_2 are independent, stochastic variables, with dimension k, mean μ_1 and μ_2, and variance Σ_1 and Σ_2 of a Gaussian distribution:

$$N_1 \sim \mathcal{N}_k(\mu_1, \Sigma_1) \qquad \text{and} \qquad N_2 \sim \mathcal{N}_k(\mu_2, \Sigma_2). \tag{2.3}$$

Then it holds, that

$$N_1 \pm N_2 \sim \mathcal{N}_k(\mu_1 \pm \mu_2, \Sigma_1 + \Sigma_2). \tag{2.4}$$

It follows (see Appendix A), that

$$Cov(N_1 + N_2) = Cov(N_1 - N_2). \tag{2.5}$$

Thus, assuming n_1 and n_2 are realizations of N_1 and N_2, the covariance of the Residual $(n_1 - n_2)$, is identical to the covariance of the noise in the Observations $(n_1 + n_2)$:

$$Cov(n_1 + n_2) = Cov(n_1 - n_2), \tag{2.6}$$

assuming a reasonable number of data points in n_1 and n_2. This means that the covariance describing the noise in $s_1 + s_2$ equals the covariance of $s_1 - s_2$. Note, this is only possible if all stochastic processes are Gaussian, and if the geology cancels out in the Residual. The geology will only cancel, if the subsurface properties are unchanged in between surveys.

2.3 Modeling the noise

We will choose to remove the noise on d_{obs} via linear LSQ inversion. This type of inversion requires information on the covariance on the *model* and the *data*. (More on LSQ inversion in the next chapter). In our case this refers to the noise-free geology (model) and the noise (data). With the noise characterization method described in the former section, an experimental data covariance matrix can now be calculated from $n_1 - n_2$ describing the noise in d_{obs}. (We forget about the geology covariance for now). However, the LSQ inversion technique requires that the inverse of the covariance matrix exists. This is guaranteed if the covariance matrix is positive definite. Hence, we must somehow make the noise covariance matrix positive definite. This is achieved by fitting a known positive definite covariance model, or rather a positive definite semi-variogram model[a], to the experimental covariance. Thus ensuring an invertible matrix in the LSQ inversion.

2.4 Modeling the geology

When the noise covariance has been estimated, one should assess the covariance of the geology. The geology covariance should reflect the prior knowledge of the noise-free geology. As with the noise covariance, the geology covariance should also be positive definite and must be modeled using a similar procedure. Using previous assumptions, and taking the covariance on the Observations, we get

$$Cov(s_1 + s_2) = Cov(2g) + Cov(n_1 + n_2) = 2^2 Cov(g) + Cov(n_1) + Cov(n_2). \tag{2.7}$$

[a]The relationship between the covariance and the semi-variogram is stated in Chapter 4

Knowing $Cov(s_1 + s_2)$ and having just modeled $Cov(n_1 + n_2)$ (the covariance of the noise)[b] we simply find $Cov(2g)$ by rearranging Equation 2.7. Notice, that $Cov(2g) \neq Cov(g) + Cov(g)$. g and g are obviously not independent, and the law for adding variances of uncorrelated stochastic variables does not apply. Rather, $Cov(2g) = 2^2 Cov(g)$[c]. Thus, when modeling the geology on the Observations we would expect about four times the variance of g. This is not important for the modeling or the inversion. Only when the inversion result is compared to the original geology g or any other single survey from the data, i.e. s_1 or s_2.

2.5 Inversion

As an appetizer to the coming theory chapter, I present the noise removal method of choice; Linear Least-Squares (LSQ) inversion. The mean posterior solution $\tilde{\mathbf{m}}$ and covariance $\tilde{\mathbf{C}}_m$ are stated here:

$$\boxed{\begin{aligned} \tilde{\mathbf{m}} &= \mathbf{m}_{prior} + \mathbf{C}_m \mathbf{G}^t (\mathbf{G}\mathbf{C}_m\mathbf{G}^t + \mathbf{C}_d)^{-1}(\mathbf{d}_{obs} - \mathbf{G}\mathbf{m}_{prior}) \\ \tilde{\mathbf{C}}_m &= \mathbf{C}_m - \mathbf{C}_m \mathbf{G}^t (\mathbf{G}\mathbf{C}_m\mathbf{G}^t + \mathbf{C}_d)^{-1}\mathbf{G}\mathbf{C}_m. \end{aligned}} \tag{2.8}$$

Here, $\mathbf{C}_d$ is the covariance matrix of the noise we have just modeled on the Residual. $\mathbf{C}_m$ is the covariance matrix for the prior geology. And $\mathbf{d}_{obs}$ is the vector of Observations defined in Equation 2.2. Note, that the noise must then be removed from the *summed* signal, the Observations, only. Not a single survey. Removing noise from either s_1 or s_2 independently with noise based on the combined signal $s_1 - s_2$ will overestimate the noise and result in unnecessary high posterior uncertainty. However, one can get around this by assuming the stochastic noise process is the same in both surveys[d]. The operator $\mathbf{G}$ is, in this case, just the identity matrix. I will elaborate on this in the next chapter. $\mathbf{m}_{prior}$ is the background model vector and is throughout this thesis assumed to be zero, unless otherwise stated.

It should be noted that, as the noise is removed from the *sum* of two surveys $(s_1 + s_2)$ with identical geology, a realization of the posterior distribution will have about twice the mean and four times the variance than either s_1 or s_2. Hence, when comparing a posterior realization to either the Base or Monitor survey, one should remember to divide the posterior mean by 2 and the posterior covariance by 4 to obtain comparable amplitudes. The 4 is due to the rule of adding completely correlated variables (or rather, the variance of a constant times a stochastic variable), as mentioned in section 2.4. The 2 is due to the fact, that the noise is removed from the summed signal d_{obs} with double the magnitude and double the mean of the original s_1 or s_2 signal.

2.6 Assumptions

A number of assumptions have been made along the way:

- **Stationary, additive geology and noise:** The method relies on stationary, additive noise and geology. Otherwise the problem will become non-linear. It is likely to encounter non-stationary, non-additive noise. However, it is often possible to find an interval in which the problem becomes linear or only mildly non-linear.

[b] We actually modeled $Cov(n_1 - n_2)$, but I proved earlier, that the two are equivalent.
[c] If X is a stochastic variable and a is a constant, then $Cov(aX) = a^2 Cov(X)$.
[d] Divide the mean and variance of the estimated noise process by 2, and let $\mathbf{d}_{obs} = s_1$ or s_2.

- **Geological stability between surveys:** Base and Monitor surveys must contain the same subsurface formation for the geology to cancel out. If geological activity is present in the area of interest, due to production or natural causes, the method developed here, might not be applicable. If changes in geology does occur between surveys, steps should be taken to mitigate the effect.
- **Pixel to pixel correspondence in data sets:** The two data sets must be properly aligned to ensure the geology cancels out. Hence, care must be taken, that each position in each data cube corresponds to the same depth and horizontal position in all surveys.
- **Data must be normally distributed:** The linear least squares inverse problem implicitly assumes normally distributed data and models. Hence, geology and noise must be Gaussian. The data should be tested to certify that they comply reasonably well with this assumption.

Each item will be discussed later in context with the data to which they apply.

Chapter 3

The inverse problem and noise filtering technique

In this chapter I will describe the inverse problem and present the general solution as well as the linearized one. For more detailed account, please refer to Tarantola [2005].

Inverse problem theory is a very important part of science and mathematics. The category holds problems in which the parameters of interest can only be measured indirectly. One might measure the gravity field of the Earth from a satellite and ask: What does the acquired data tell me about the mass density distribution of the interior of the Earth? It is not possible to drill to the center of the Earth and take samples, but given gravity data and the relation between gravity and density with distance, we can calculate an estimate of the density profile. This is called an inverse problem. In geophysics, this type of problem is encountered everyday. However, the problem with inverse problems, is that due to our ignorance about the complete problem (exact temperature, pressure, uncertainty in gravity measurements, etc.), it is usually not possible to find a unique solution. One may calculate an optimal solution, e.g. the solution with minimum variance, but without any prior knowledge about the subsurface, like a bore hole, many different earth models may explain the same data equally well.

Since there exists a host solutions (often infinite) to the inverse problem, it is desired to introduce some knowledge about the problem that will constrain the solution. This type of information is called *a priori* information. The combination of a priori information with observations was formulated by Thomas Bayes in 1761 in terms of probability. His contribution to probability theory was vital and influences the work of statisticians among many others to this day. Tarantola and Valette [1982] formulated a more general approach to solve the inverse problem, from which Bayes' theorem can be derived as a special case (see Appendix B). All probability theory and combination of states of information in this project is based on their work. The a priori information in particular will be discussed extensively, as building an unbiased, scientific prior probability distribution is essential to the validity of the solution of the inverse problem.

I present the essential results of Tarantola and Valette's theory below and follow their application to least-squares inversion. An interesting, though rather technical presentation and discussion of the probabilistic approach to the inverse problem, can be found in Mosegaard and Tarantola [2002].

3.1 The general inverse problem

The general inverse problem relates the observed data, $\mathbf{d}$, to the model parameters $\mathbf{m}$, by the possibly non-linear relation g:

$$\mathbf{d} = g(\mathbf{m}). \tag{3.1}$$

The model $\mathbf{m}$ is the unknown, and $\mathbf{d}$ are the observed data. In our case, $\mathbf{m}$ is the noise-free geology, and $\mathbf{d}$ the noise. If one accepts the mathematically friendly, but for some intuitively difficult statement, that the solution can be represented as a probability density distribution, and not a single unique set of parameters, the solution, including uncertainty, can be obtained in the form of an *a posteriori* probability distribution:

$$\sigma_m(\mathbf{m}) = k\rho_m(\mathbf{m})L(\mathbf{m}) \tag{3.2}$$

$L(\mathbf{m})$ is the *likelihood function* which gives a measure of how well the model $\mathbf{m}$ explains the data, and $\rho_m(\mathbf{m})$ is the *a priori* pdf describing the data-independent prior knowledge of the model parameters.

Please take a moment to dwell on the simplicity of Equation (3.2). This is *the* solution to the inverse problem. From the perspective of probability densities, the solution is unique, and from it, all manner of results can be generated, including mean, median, maximum likelihood and uncertainty.

3.2 Linear least-squares inversion

Many problems are naturally linear, or approximately linear over a certain interval. In that case, the model $\mathbf{m}$ becomes independent of the forward operator g, and the problem can be expressed as

$$\mathbf{d} = \mathbf{G}\mathbf{m}. \tag{3.3}$$

The linear inverse problem can be solved by minimizing the sum of squared differences between the data values and their corresponding modeled values, also known as *least-squares inversion*, but only if all "input" probabilities in Equation 3.2 are Gaussian:

If an observation, $\mathbf{d}_{obs}$, is made, and one assumes negligible modelization errors (as is done in many applications), and if the observational errors are assumed Gaussian, the likelihood function is Gaussian and can be written

$$L(\mathbf{m}) = const.\ exp\left\{-\frac{1}{2}(\mathbf{d}-\mathbf{d}_{obs})^t\mathbf{C}_d^{-1}(\mathbf{d}-\mathbf{d}_{obs})\right\}. \tag{3.4}$$

The matrix $\mathbf{C}_d$ is the data covariance matrix. Or in our case, the *noise covariance matrix*. It describes the errors associated with each observation. A completely uncorrelated and noise-free dataset will yield a covariance matrix with ones in the diagonal and zeros everywhere else. Most of the time this however is not the case, and some consideration should go into constructing this matrix. It is the purpose of this project to model the noise on the data, so as to make the noise covariance matrix as objective as possible.

The Gaussian a priori model pdf can be written in a similar manner:

$$\rho_m(\mathbf{m}) = const.\ exp\left\{-\frac{1}{2}(\mathbf{m}-\mathbf{m}_{prior})^t\mathbf{C}_m^{-1}(\mathbf{m}-\mathbf{m}_{prior})\right\}. \tag{3.5}$$

In our case, this will represent the geology pdf with mean $\mathbf{m}_{prior}$ and covariance $\mathbf{C}_m$. $\mathbf{C}_m$ is the model covariance matrix representing the covariance of the noise-free data; the

geology. In Chapter 5 I will inspect how various geology and noise covariance models affect the inversion result.

For a linear problem, in which the likelihood and the a priori pdf can be described by a Gaussian distribution, the a posteriori distribution is also Gaussian and is completely described by its mean and covariance. Hence, the posterior probability density can be written as:

$$\sigma_m(\mathbf{m}) = \text{const. } \exp(-\frac{1}{2}(\mathbf{m}-\tilde{\mathbf{m}})^t\tilde{\mathbf{C}}_m^{-1}(\mathbf{m}-\tilde{\mathbf{m}}). \tag{3.6}$$

If $\tilde{\mathbf{m}}$ is defined as the center of the Gaussian posterior probability density with posterior covariance $\tilde{\mathbf{C}}_m$, one can solve for each of them. In the end, we obtain (also shown in Equation (2.8)).

$$\begin{aligned}\tilde{\mathbf{m}} &= \mathbf{m}_{prior} + \mathbf{C}_m\mathbf{G}^t(\mathbf{G}\mathbf{C}_m\mathbf{G}^t + \mathbf{C}_d)^{-1}(\mathbf{d}_{obs} - \mathbf{G}\mathbf{m}_{prior}) \\ \tilde{\mathbf{C}}_m &= \mathbf{C}_m - \mathbf{C}_m\mathbf{G}^t(\mathbf{G}\mathbf{C}_m\mathbf{G}^t + \mathbf{C}_d)^{-1}\mathbf{G}\mathbf{C}_m.\end{aligned} \tag{3.7}$$

(3.7) is the solution to the linear inverse problem. It provides a complete probabilistic description of the noise-attenuated signal, including uncertainty. The analytical form makes it fast and easy to compute. However, the covariance matrices $\mathbf{C}_m$ and $\mathbf{C}_d$ are required to have an inverse. The method also becomes painstakingly time and memory consuming for large $\mathbf{C}_m$ and $\mathbf{C}_d$ since matrix inversion complexity goes up as $\mathcal{O}(\mathrm{N}^3)$ for an N x N matrix. How I will ensure the inverses to these matrices exist, will be discussed in the next chapter. $\mathbf{m}_{pri}$ denotes the background model which is assumed zero, unless otherwise stated. $\tilde{\mathbf{m}}$ is the mean of the Gaussian posterior, also called the *maximum a posteriori* solution (MAP), and is often given as *the* solution. However, it is just the mean of a Gaussian distribution, and often much smoother than any random realization of the posterior. One can even construct a movie of realizations, in which each frame will capture the essential random fluctuations of the posterior. Hence, a series of Gaussian realizations shall accompany all solutions to the inverse problem in this report. Also, one should be confident that the data are Gaussian and the problem is linear. For problems where these criterions are not satisfied, it is discouraged to make use of this technique. However, the advantages of the approach are diverse and manifold, and are hence used extensively in one form or another, also in mildly non-Gaussian cases.

In the general nonlinear inverse problem, $g(\cdot)$ is called the forward operator, or the mapping operator. It expresses our mathematical model of the physical system we are studying and is used to predict the outcome of some observation. When the problem can be considered linear, the forward operator becomes linear and can be represented by a matrix, $\mathbf{G}$. Normally, $\mathbf{G}$ will be mapping the model parameters from the *model space* into the *data space*. However, in our case, we do not want to map between spaces. Only to reduce noise. I will not go into detail about the definition of model space and data space (see Tarantola [2005]). We only need to understand, that to filter noise using this particular method, $\mathbf{G}$ needs to map into *itself*. This is easily done by letting $\mathbf{G}$ equal the identity matrix.

Hence, assuming $\mathbf{G} = \mathbf{I}$ and $\mathbf{m}_{pri} = 0$, Equation (3.7) reduces to a compact noise filtering technique, with a Gaussian posterior distribution describing the LSQ estimate of the noise-free geology:

$$\boxed{\begin{aligned}\tilde{\mathbf{m}} &= \mathbf{C}_m(\mathbf{C}_m + \mathbf{C}_d)^{-1}\mathbf{d}_{obs} \\ \tilde{\mathbf{C}}_m &= \mathbf{C}_m - \mathbf{C}_m(\mathbf{C}_m + \mathbf{C}_d)^{-1}\mathbf{C}_m.\end{aligned}} \tag{3.8}$$

Chapter 4

Variogram investigation and simulation

In Chapter 3 I explained the linear inverse problem and the need for Gaussian pdfs and invertible covariance matrices. In this chapter I will investigate the semi-variogram models and the modeling procedure used to obtain these covariance matrices. I begin the variogram investigation with a short theory section.

4.1 Theory

Say for each position $\mathbf{r}$ in the domain $\mathcal{D}$, which is typically a subset of $\mathcal{R}^2$ or $\mathcal{R}^3$, we have a set of observed values $z(\mathbf{r})$, which all are realizations of the stochastic variable $Z(\mathbf{r})$. Assuming first and second order stationarity (i.e. the mean value and the variance are independent of $\mathbf{r}$), $Z(\mathbf{r})$ has the expectation value

$$E\{Z(\mathbf{r})\} = \mu(\mathbf{r}) = \mu, \tag{4.1}$$

and auto-covariance

$$Cov\{Z(\mathbf{r}), Z(\mathbf{r}+\mathbf{h})\} = C(\mathbf{r}, \mathbf{h}) = C(\mathbf{h}), \tag{4.2}$$

where $\mathbf{h}$ is termed the displacement vector. It is customary in the geophysical community to analyze and describe the statistical properties of data with a semi-variogram γ, which is related to the (stationary) auto-covariance in the following way:

$$\gamma(\mathbf{h}) = C(\mathbf{0}) - C(\mathbf{h}). \tag{4.3}$$

Note, that $C(\mathbf{0}) = \sigma^2$, the variance of the stochastic variable. The semi-variogram can be understood as a measure for the expected difference squared between stochastic variables as a function of the displacement vector. When given a set of measurements the semi-variogram may be calculated by means of the following estimator, which calculates (half) the mean value of the squared differences between all pairs of measurements $z(\mathbf{r}_k)$ and $z(\mathbf{r}_k + \mathbf{h})$ separated by the displacement vector $\mathbf{h}$:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{k=1}^{N(h)} [z(\mathbf{r}_k) - z(\mathbf{r}_k - \mathbf{h})]^2. \tag{4.4}$$

$N(\mathbf{h})$ is the number of point pairs separated by $\mathbf{h}$. $\hat{\gamma}$ is termed the *experimental semi-variogram*. The experimental semi-variogram, can provide an experimental (as in obtained

from data) covariance model. We will use this methodology to obtain an estimate of a Gaussian covariance model describing the noise. Please note that I may refer to the variogram or semi-variogram interchangeably, in all cases referring to the definition in Equation (4.3).

4.1.1 Semi-variogram models

The experimental semi-variogram is an estimate of the true variability in the data. However, a covariance matrix build on this estimate is likely not to have an inverse, which is a requirement in LSQ inversion. Therefore, in order to ensure we obtain a covariance matrix with an inverse, we choose to model the experimental semi-variogram using a number of pre-defined semi-variogram models, that are all positive definite. I list three of the most applied models below. The magnitude of **h** is denoted by h, also called the *lag*. I introduce the Spherical model:

$$\gamma^*(h) = \begin{cases} 0 & h = 0 \\ C_0 + C_1[\frac{3}{2}\frac{h}{R} - \frac{1}{2}\frac{h^3}{R^3}] & 0 < h < R \\ C_0 + C_1 & h \geq R, \end{cases} \tag{4.5}$$

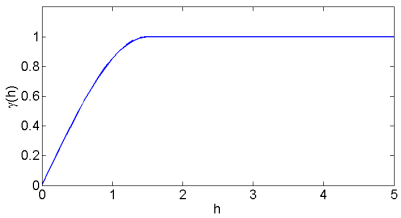


Figure 4.1: Spherical model with range = 1.5 and sill = 1, abbreviated: 1 Sph(1.5).

where C_0 is the *nugget effect*, R is the *range* and C_0+C_1 is the *sill* ($= \sigma^2$). Note, that C_0 is not equal to $C(\mathbf{0})$. The range is defined as the distance from where on samples become uncorrelated. This distance also marks maximum variance (the sill). Model variance is unchanged for higher lags. The nugget effect is also in itself a semi-variogram model, albeit rather boring as it is constant. This is not the case for the Exponential model:

$$\gamma^*(h) = \begin{cases} 0 & h = 0 \\ C_0 + C_1[1 - \exp(-\frac{3h}{R})] & h > 0, \end{cases} \tag{4.6}$$

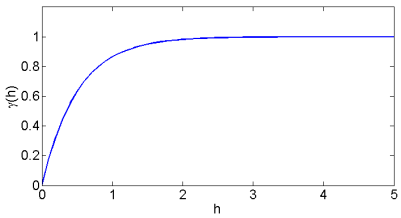


Figure 4.2: Exponential model with range = 1.5 and sill = 1, abbreviated: 1 Exp(1.5).

or the Gaussian model:

$$\gamma^*(h) = \begin{cases} 0 & h = 0 \\ C_0 + C_1[1 - \exp(-\frac{3h^2}{R^2})] & h > 0. \end{cases} \tag{4.7}$$

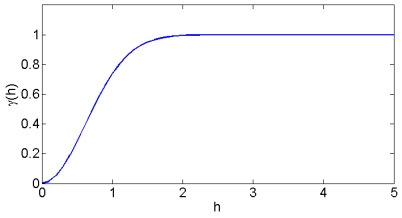


Figure 4.3: Gaussian model with range = 1.5 and sill = 1, abbreviated: 1 Gau(1.5).

The two latter models approach their sill asymptotically. The range is then usually defined as 95% of the sill. The models can be combined (nested) to obtain various shapes, without loosing positive definiteness (Goovaerts [1997]).

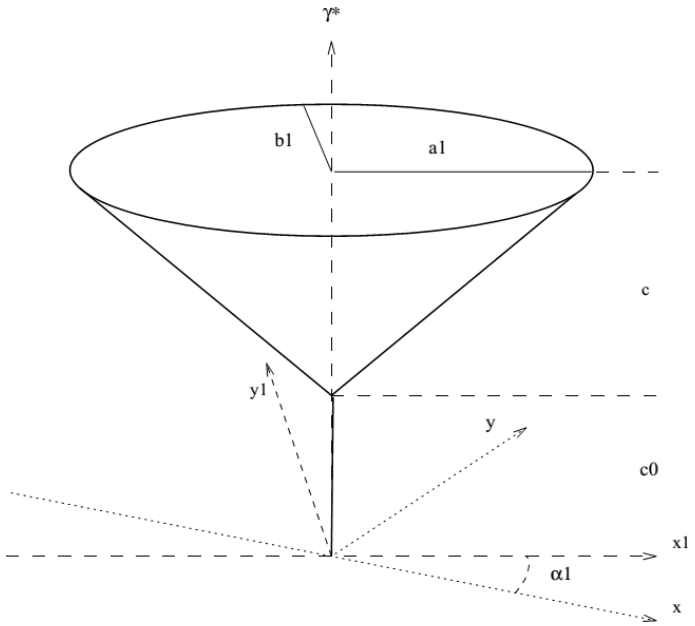


Figure 4.4: Sketch of elliptic cone 2-D semi-variogram model. From [Nielsen, 1994, p.12]

.

The above models are independent of orientation (isotropic). If anisotropy is desired, one has to introduce an angle dependent displacement vector. If the angle and coordinates are defined as in Figure 4.4, we have (in 2-D):

$$\begin{aligned} x &= x_1 \cos\alpha_1 - y_1 \sin\alpha_1 \\ y &= x_1 \sin\alpha_1 + y_1 \cos\alpha_1 \end{aligned} \tag{4.8}$$

$$\begin{aligned} x_1 &= x \cos\alpha_1 + y \sin\alpha_1 \\ y_1 &= -x \sin\alpha_1 + y \cos\alpha_1 . \end{aligned} \tag{4.9}$$

By rotating through the angle α_1 we can sample correlation in any direction. The magnitude of the displacement vector then becomes: $h = \sqrt{x^2 + y^2}$. Spatially correlated

acquisition noise is highly anisotropic and must of course be parameterized by such a model. Nielsen [1994].

4.1.2 Fast semi-variogram computation with FFT

From a practical point of view, it is necessary to take into account the time it takes to calculate a semi-variogram, i.e. the number of point pairs included in the calculation of the experimental variogram. In his paper from 1996 Marcotte details a way to efficiently calculate 2-D semi-variograms using the well known Fast Fourier Transform (FFT) algorithm. The method reduces the complexity of the computing the semi-variogram for a n_x x n_y matrix from $(n_x \cdot n_y) \cdot (n_x \cdot n_y - 1)/2$ to $(2n_y - 1)(2n_x - 1)log_2(2n_x - 1)$. Using the standard spatial approach one should not consider calculating a full variogram for matrices with more than 500^2 data points. With the FFT based approach semi-variograms for matrices with 1000^2 data points are calculable within a second. The backside is increased memory usage during calculation. The amount of RAM however is easily increased, whereas CPU speed is not. The method also introduces some variance issues at high lags (close to the edges of the variogram). However this is not important as statistics always get unreliable at the edges due to insufficient point pairs anyway. I will use this method to calculate the experimental semi-variogram, and from it, the covariance matrix.

4.1.3 Covariance matrix

When a fitting semi-variogram model has been found, the corresponding covariance matrix, $\mathbf{C}_m$ or $\mathbf{C}_d$, must be calculated. From Equation 4.3 we see, that $C(\mathbf{h}) = \sigma^2 - \gamma(\mathbf{h})$. Knowing the semi-variogram model we can then calculate the auto-covariance model. However, the LSQ inversion requires a covariance *matrix*: A $n_x n_y \cdot n_x n_y$ large matrix including explicit calculation of correlation coefficients from all points to all points. This is done by setting up a mesh grid with a constant cell size Δh and applying the chosen covariance model. Ex: $n_x = n_y = 4$. $\Delta h = 1$. Mesh grid matrices:

$$\hat{x} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix} \qquad \hat{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \\ 4 & 4 & 4 & 4 \end{bmatrix} \tag{4.10}$$

Then one can calculate the euclidean distance (difference) from all points to all points. With an isotropic Gaussian covariance model with sill=1 and range=2, this results in the covariance matrix in Figure 4.5.

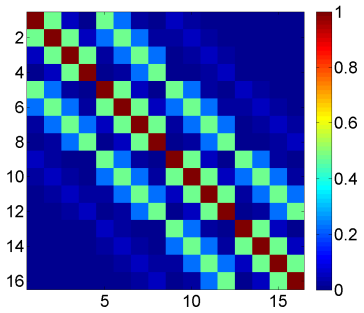


Figure 4.5: Isotropic Gaussian covariance matrix with sill=1 and range=2.

4.1.4 Drawing Gaussian realizations

When a (stationary) covariance matrix is obtained, a random realization of the Gaussian process is fairly simple to compute by Cholesky decomposition. The Cholesky algorithm expresses the covariance matrix $\mathbf{C}$ as a product of upper and lower triangular matrices:

$$\mathbf{C} = \mathbf{L}\mathbf{L}^t \tag{4.11}$$

Solving for $\mathbf{L}$ and multiplying by a random realization $\mathbf{w}$ of Gaussian process with zero mean and unit variance yields a Gaussian realization $\mathbf{x}$ with said mean and variance:

$$\mathbf{x} = \mathbf{L}\mathbf{w}. \tag{4.12}$$

Other, more efficient algorithms exist for computing random Gaussian realizations. For instance, the FFT Moving Average (FFT-MA) generator is an efficient numerical method for generating and conditioning large Gaussian simulations, as it is based on convolution in the Fourier domain. Figure 4.6 shows three such random Gaussian realizations. See Tarantola [2005] for a discussion and derivation of the Cholesky method and Ravalec et al. [2000] for a description of the FFT-MA algorithm.

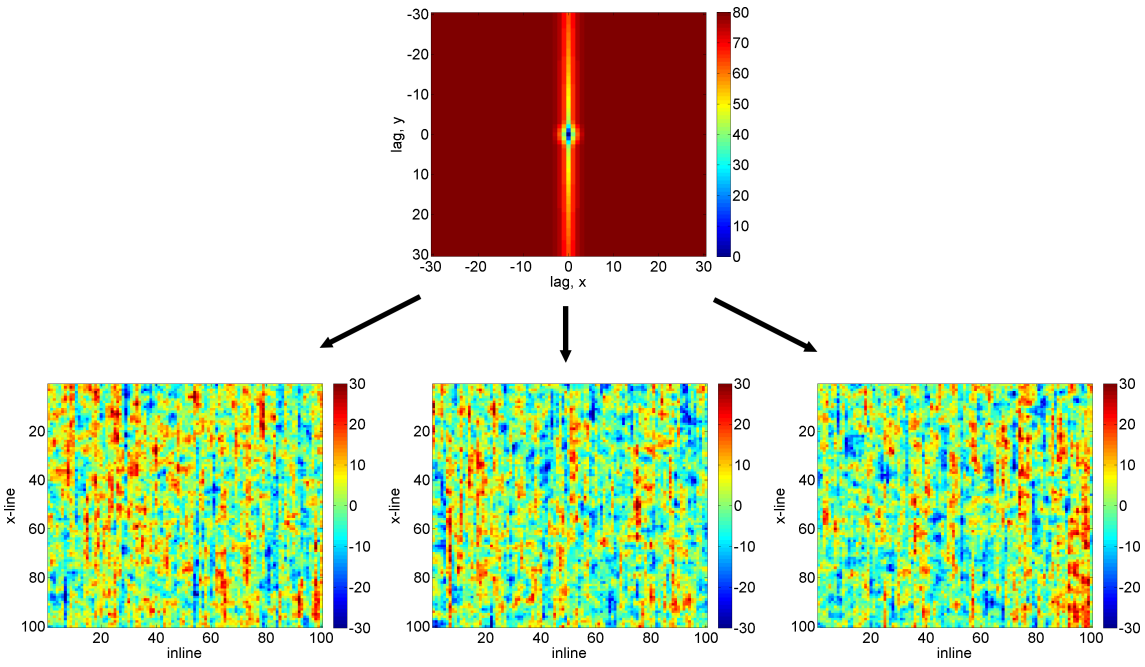


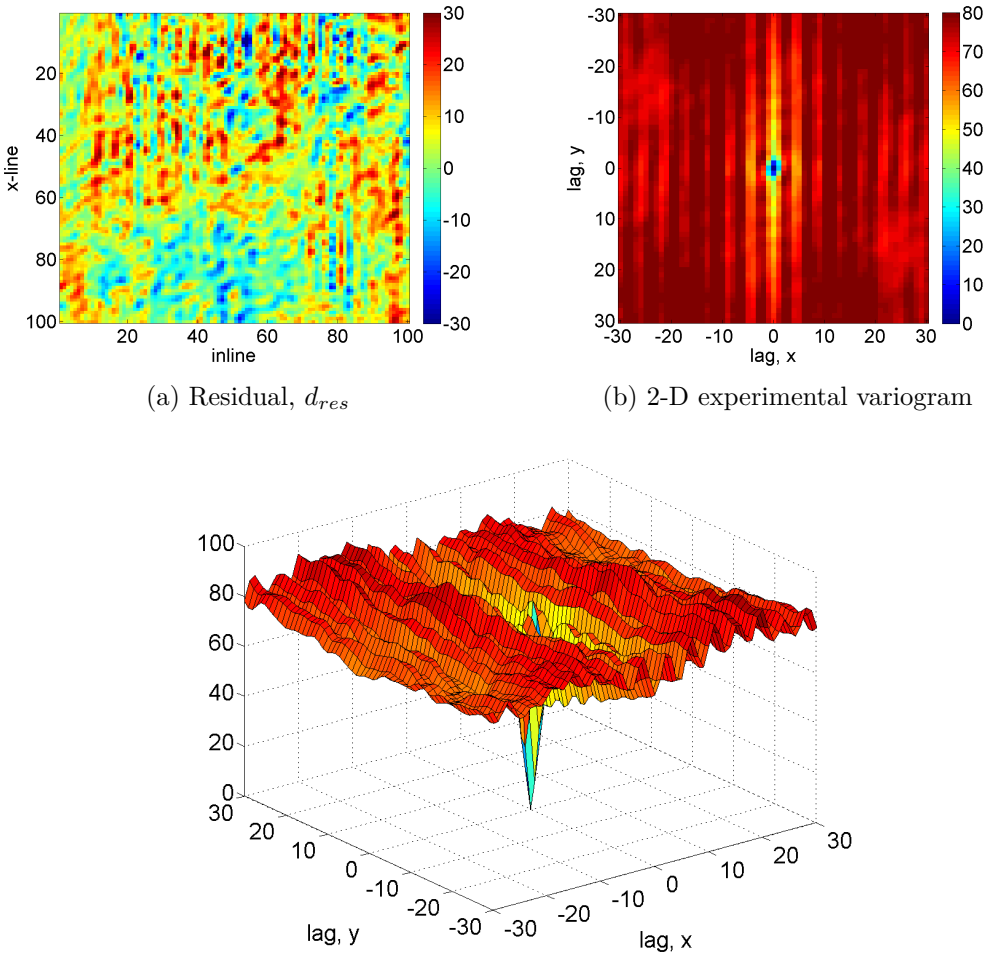
Figure 4.6: Random Gaussian realizations (bottom) of the covariance model (top) generated by FFT-MA.

4.2 Modeling

The success of the noise characterization and removal technique relies on the degree to which the noise can be estimated correctly. However, it also relies on the existence of the inverse of the noise covariance matrix $\mathbf{C}_d$. We found that the Residual, d_{res}, is a good description of the noise in the Observations, d_{obs}. Hence, we fit a positive definite semi-variogram model to the experimental semi-variogram, as represented by the Residual data, and convert that into a noise covariance matrix with an inverse now guaranteed to exist. When a model has been chosen, random Gaussian realizations of this model

should be compared to the Residual data for confirmation of its resemblance. I will begin by investigating the Residual of a real time slice of the Halfdan data and compute the experimental variogram on which we shall model $\mathbf{C}_d$.

4.2.1 The noise



(a) Residual, d_{res}

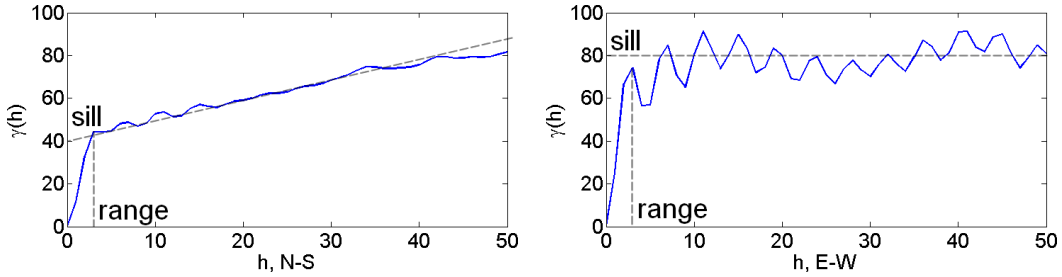
(b) 2-D experimental variogram

(c) 3-D experimental variogram. Same as (b) but shown in 3-D.

Figure 4.7: Residual data and 2-D variogram of layer 140 in the Halfdan data. The left figure shows part of layer 140 (the Residual). The right and lower its corresponding experimental 2-D variogram in 2 and 3-D. The mean has not been subtracted from the data.

Figure 4.7 shows data and corresponding 2-D variogram in 2 and 3-D of layer 140. Inline correlation (N-S) is strong and dominates the variogram giving it anisotropic characteristics. The data also have a smooth look. This is seen as the small, circular (isotropic) blue patch in the center (low lag) of the variogram. A small part of the signal also seems to oscillate, like a wave pattern, especially in the cross-line (E-W) direction. This is due to the repeating N-S pattern. Thus, we have identified three distinct characteristics of the variogram: A high correlation part, a low correlation part and an oscillating part. In the ideal case then, we need at least three different models to accurately describe the Residual.

A 1-D version of Figure 4.7 with more lags better illustrates the shape and magnitude of the variogram (see Figure 4.8).



(a) d_{res}. Inline direction. (b) d_{res}. X-line direction.

Figure 4.8: Experimental semi-variograms in 1-D.

The direction of anisotropy is obviously in the inline direction (this is why the 1-D semi-variogram (a) was depicted in this direction). The sill and range can also be estimated from this plot, as a combination of three distinct models: One isotropic model with range around 3, seen in both (a) and (b). One Linear, anisotropic model with range at least 50, best seen in (a). And a *hole model*[a] (describing oscillatory motions), best seen in (b). Unfortunately, only certain hole models belong to the group of positive definite models, and then only in certain dimensions (Deutsch and Journel [1998]). As a result I choose not to model the oscillations to ensure the covariance model stays positive definite in all cases. The isotropic, short range model could be either Gaussian, Spherical or Exponential. The Linear model might also be approximated by a long range Exponential model.

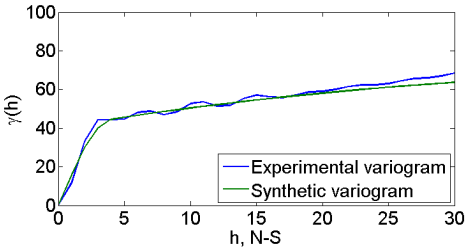
The sill of the first model is estimated from (a). A Linear model does not have a sill (unless it is constant). Hence, the *gradient* of a Linear model will then replace the sill parameter. However, assuming the Linear model can be approximated by a slowly increasing (long range) Exponential, the sill of the *Exponential* can approximated by elimination: By linear regression the Linear/Exponential model is estimated to be shifted +40 in variance (see (a)), assuming the isotropic model has variance = 40. If the (unmodeled) hole model has a constant mean variance of zero, the maximum sill (80) (see (b)) minus the isotropic model sill (40) equals the sill of the remaining model(s) - in this case the Exponential model.

I have listed two possible nested models in Table 4.1 and plotted them along with the experimental semi-variograms in Figure 4.9. From these models I calculate their corresponding covariance matrices $\mathbf{C}_{d,1}$ and $\mathbf{C}_{d,2}$. Making realizations of the different model combinations will reveal which combination is best and how to improve it.

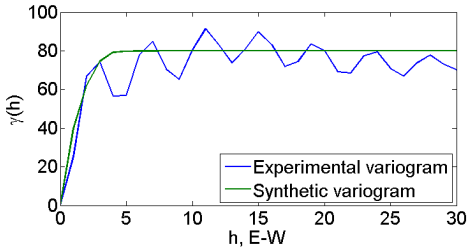
	Model	Sill	Range	Angle	Aniso. (b_1/a_1)
Model 1, $\mathbf{C}_{d,1}$	Exponential	40	100	180	.03
	Spherical	40	4	0	1
Model 2, $\mathbf{C}_{d,2}$	Exponential	40	100	180	.03
	Gaussian	40	3	0	1

Table 4.1: Two noise models based on the Residual (Figure 4.9). Angle is defined as degrees from North.

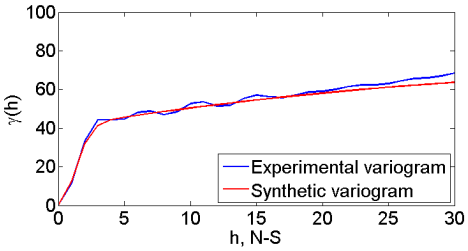
[a] E.g. the cosine hole model: $\gamma(h) = C_0[1 - \cos(\frac{h}{R}\pi)]$



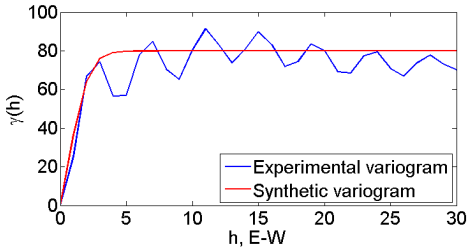
(a) Model 1. d_{res}. Inline direction.



(b) Model 1. d_{res}. X-line direction.



(c) Model 2. d_{res}. Inline direction.



(d) Model 2. d_{res}. X-line direction.

Figure 4.9: Experimental semi-variograms of the Residual in 1-D, with synthetic noise models from Table 4.1. The maximum lag calculated is reduced to 30 to increase detail in the plot.

Both proposed models have the same function and parameters modeling the long range part of the signal (an Exponential). The short range correlations however, have a greater impact on realizations of the model (short range correlations have a higher probability). Model 1 uses a Spherical short range model, whereas Model 2 uses a Gaussian. This should be evident in the realizations.

I present the first model (Model 1), three Gaussian realizations of it and their respective experimental semi-variograms, see Figure 4.10. The Gaussian realizations should be compared to the Residual data on which their Gaussian random process are modeled. We have already made sure to choose parameters close to those of the data (including the mean which is equal or close to zero in both cases). Thus, it is up to the interpreter to ensure that realizations and data look similar. It is possible in more or less simple cases to automatize this procedure. However, I prefer to keep some measure of human influence on the modeling, even though building an advanced algorithm for model fitting would be interesting.

Each realization is accompanied by its experimental semi-variogram to measure how well one can expect to retrieve the synthetic semi-variogram parameters (top figure). It is also these experimental semi-variograms which should be compared to the experimental semi-variogram of the Residual. Once again, the human eye should be applied to ensure likeliness.

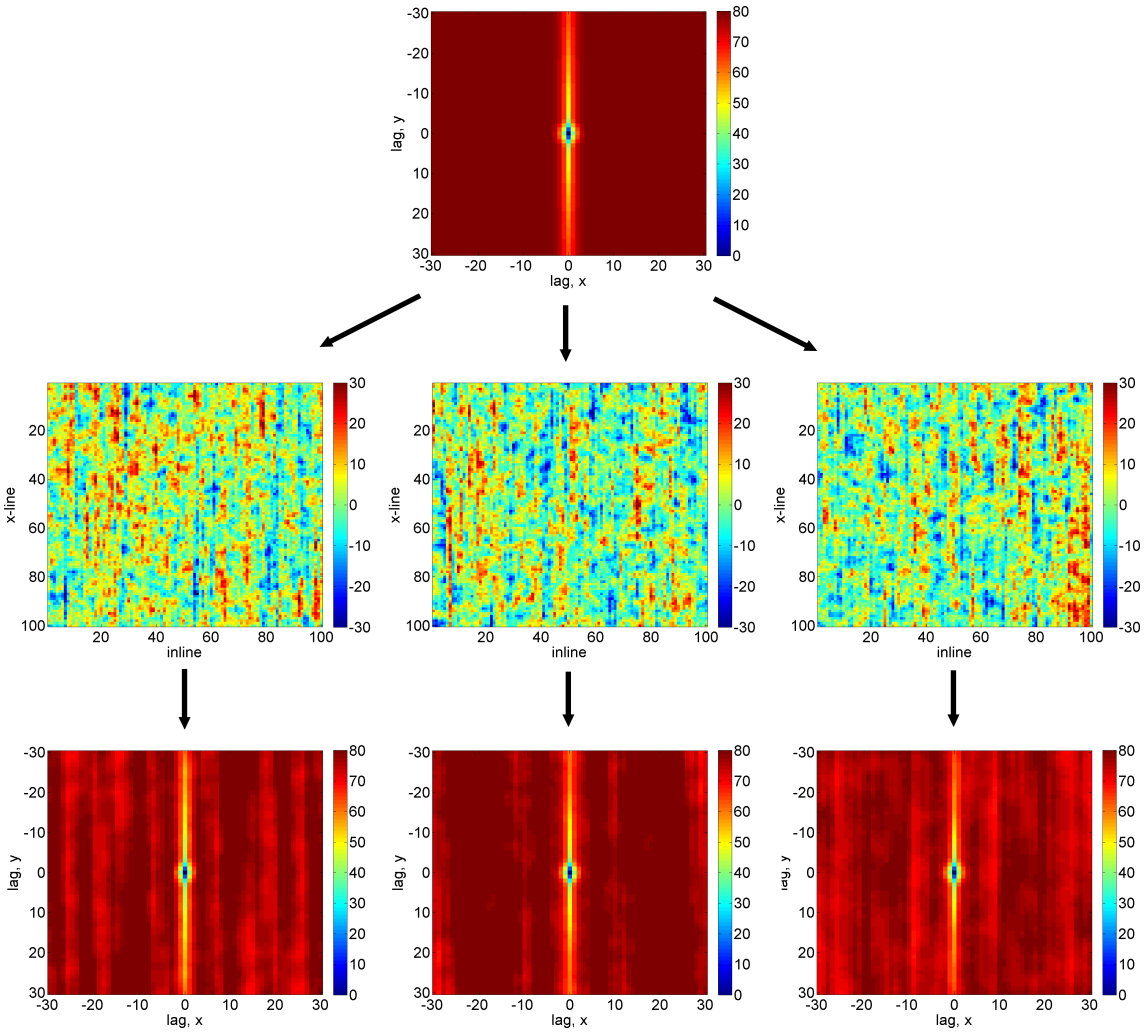
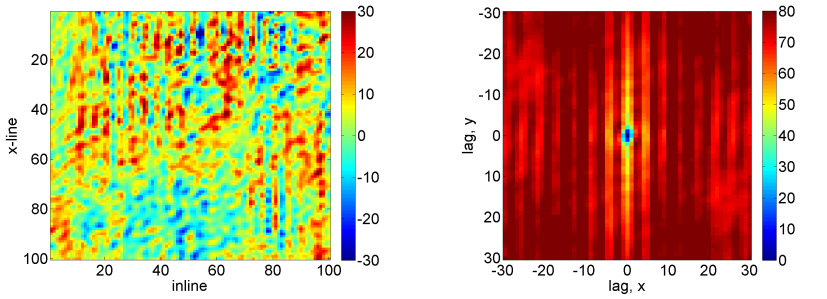


Figure 4.10: Model 1 from Table 4.1: 40 Exp(100,180,.03) + 40 Sph(4,180,1). Top: Synthetic semi-variogram. Middle: Gaussian realizations. Bottom: Experimental semi-variograms.

COMPARE TO:



(a) Residual (b) 2-D experimental variogram

Figure 4.11: Residual data and experimental semi-variogram.

The second model is presented in Figure 4.12.

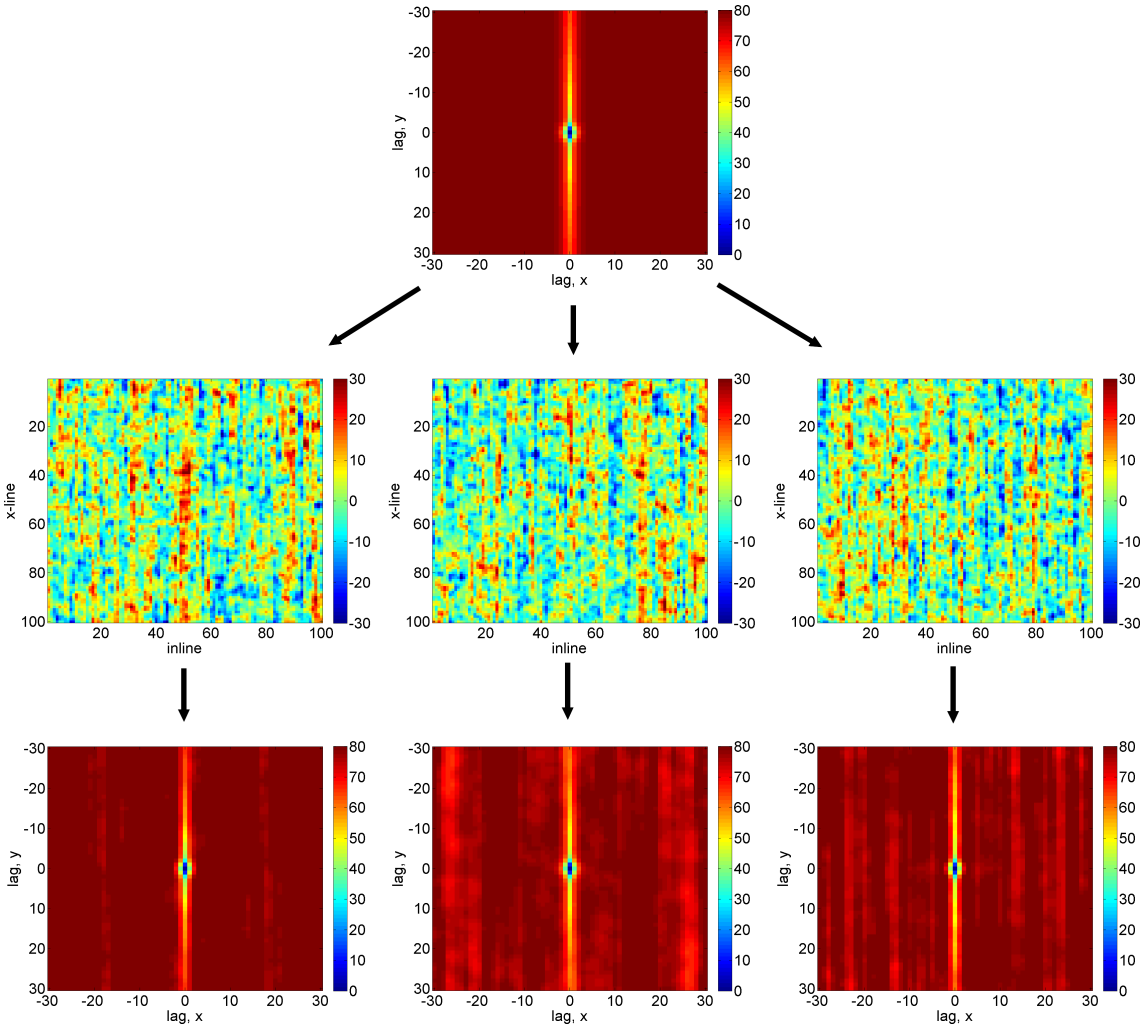
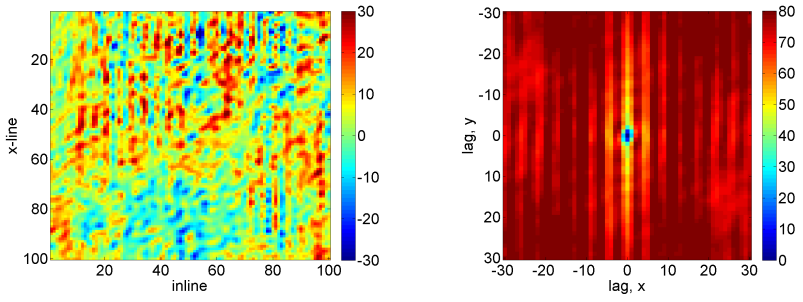


Figure 4.12: Model 2 from Table 4.1: 40 Exp(100,180,.03) + 40 Gau(3,180,1). Top: Synthetic semi-variogram. Middle: Gaussian realizations. Bottom: Experimental semi-variograms.

COMPARE TO:



(a) Residual (b) 2-D experimental variogram

Figure 4.13: Residual data and experimental semi-variogram.

The best semi-variogram model is then chosen, based on visual fit to the experimental semi-variogram and realizations of the model. In this case, the best resemblance is found in Model 2. Both models reproduce long range correlations well. However, Model 2 featuring the Gaussian short range model, ensures smooth features on the small scale found in the data but not reproduced in Model 1 with the Spherical model. It should be mentioned, that none of the models are perfect, in part due to the possibility, that the data might not be completely stationary. We will investigate this in Chapter 6. There are also some internal modeling errors introduced, due to the finite number of data points and the FFT based experimental semi-variogram calculation technique.

4.2.2 The geology

As explained in the Methodology section, knowing $Cov(s_1 + s_2)$ and having modeled $Cov(n_1 + n_2)$, we can find $Cov(2g)$. However, it must also be modeled to ensure positive definiteness of the geology covariance matrix $\mathbf{C}_m$. The procedure is the same for the geology as for the noise, only with the Residual replaced by the Observations. A 1-D modeling of the geology summed with the chosen noise model (Model 2) is drawn in Figure 4.14. Semi-variogram parameters are given in Table 4.2.

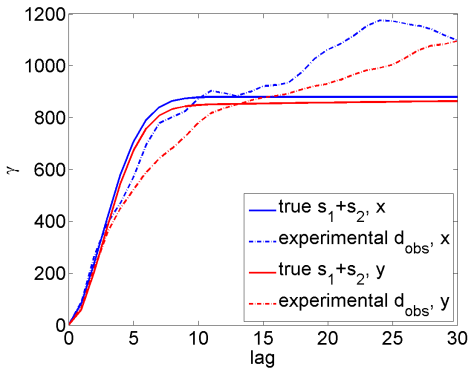


Figure 4.14: Experimental semi-variograms of the Observations in 1-D. Solid lines are the sum of geology and noise models.

	Model	Sill	Range	Angle	Aniso. (b_1/a_1)
Geology, $\mathbf{C}_m$	Gaussian	800	7	0	1
Noise (Model 2), $\mathbf{C}_d$	Exponential	40	100	180	.03
	Gaussian	40	3	0	1

Table 4.2: Geology and noise models based on the Observations and Residual. Angle is defined as degrees from North.

*

With the noise covariance matrix $\mathbf{C}_d$ and geology covariance matrix $\mathbf{C}_m$ now guaranteed to have an inverse, we can calculate the posterior of the linear least squares inverse problem in Equation (3.8). In the next chapter, we will see how this works with a synthetic example.

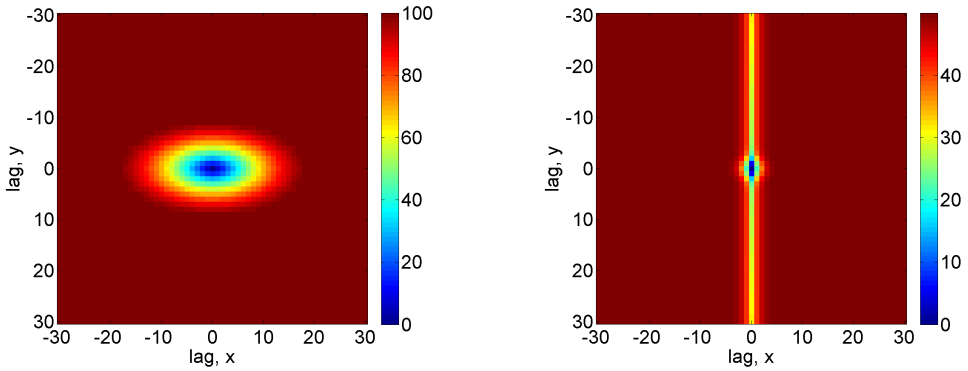
Chapter 5

Synthetic example

In this chapter, I conduct a synthetic experiment to evaluate the performance of the method. Hence, I will generate some data, model the noise, calculate the LSQ posterior (posterior noise-free geology) and compare to the known true noise-free geology. In this synthetic case I will have complete knowledge of all parameters making evaluation easy. The experiment is scaled down to one time slice, 90 pixels on each side, due to memory issues[a].

5.1 Generating the data

Two random data sets with zero mean are generated from Gaussian realizations of the semi-variogram models shown in Table 5.1. Each data set has the same geology, g, but different noise realizations, n_1 or n_2 (both noise realizations are generated from the same variogram model though). The noise model is a combination of an Exponential model with high correlation length in the N-S direction, and an isotropic Gaussian model for short correlation lengths. It is designed to mimick the acquisition imprint seen on the Halfdan data. The geology is an anisotropic Spherical model. Anisotropy in geology is vastly exaggerated compared to the Halfdan data. The models are illustrated in Figure 5.1.



(a) Geology variogram

(b) Noise variogram

Figure 5.1: 2-D variogram models used to generate the synthetic data in Figure 5.2.

[a] An experiment comparable in size to a full horizontal slice of the Halfdan data, would require about 14 terabytes of RAM (per layer).

	Model	Sill	Range	Angle	Aniso. (b_1/a_1)
Geology (g)	Spherical	100	20	90	.5
Noise (n_1 and n_2)	Exponential	25	300	180	.01
	Gaussian	25	3.5	0	1

Table 5.1: The reference models. The synthetic data are generated from the models listed above. Angle is defined as degrees from North.

As mentioned earlier, the first survey acquired is referred to as Base. The second survey is referred to as Monitor. Both are a combination of geology g and noise n_1 or n_2.

$$\text{Base: } s_1 = g + n_1 \qquad \text{Monitor: } s_2 = g + n_2.$$

g, n_1 and n_2 are generated from the models in Table 5.1 by an FFT based moving average algorithm (see section 4.1.4).

As explained in the Methodology section, the noise filtering must be applied to the summed signal, i.e. s_1+s_2 (the Observations). The noise is characterized from the Residual, s_1-s_2:

$$\text{Observations: } d_{obs} = s_1 + s_2 = 2g + n_1 + n_2 \qquad \text{Residual: } d_{res} = s_1 - s_2 = n_1 - n_2$$

All four data sets are presented in Figure 5.2.

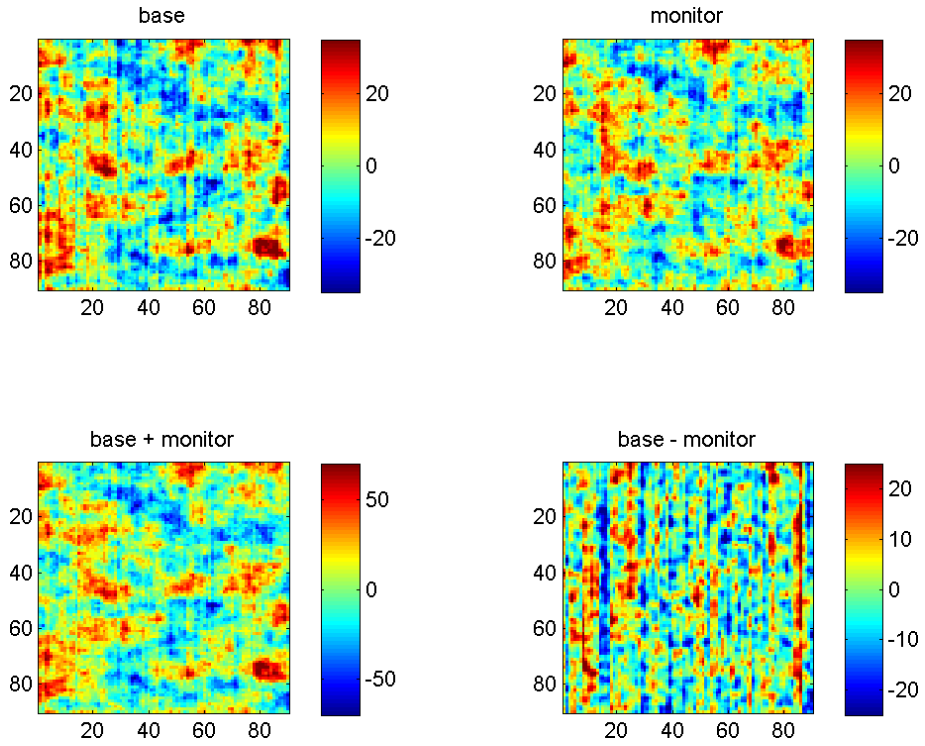


Figure 5.2: Gaussian synthetic data generated by FFT based moving average algorithm. d_{obs} is used as input observations for the inversion. d_{res} is used to characterize the noise.

5.2 Inversion results with true covariance models

The initial inversion results are designed to illustrate the best possible result you can expect from the defined setup. I.e. neither geology nor noise is modeled. All estimated

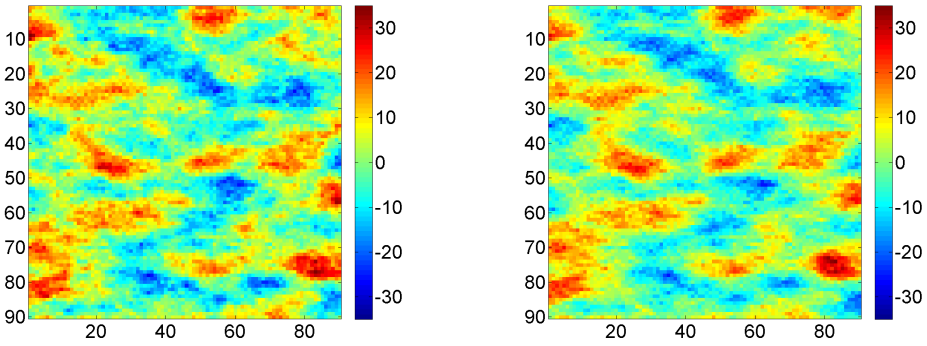
models used for this inversion are identical to the models used for generating the data (except the variance, which is increased in the inversion priors, since the summed and differenced signals have higher variance than the original data s_1 and s_2). Thus, the only information lost, is due to the finite number of data points. I give the geology and noise models in Table 5.2 from which $\mathbf{C}_d$ and $\mathbf{C}_m$ can be constructed and hence the least squares inversion solved by solving Equation (3.8).

	Model	Sill	Range	Angle	Aniso. (b_1/a_1)
Prior geology, $\mathbf{C}_m$ (true)	Spherical	400	20	0	.5
Noise, $\mathbf{C}_d$ (true)	Exponential	50	300	180	.01
	Gaussian	50	3.5	0	1

Table 5.2: "Estimated" (true) geology and noise models used in the first synthetic example (ideal case). Angle is defined as degrees from North.

I will investigate a number of properties of the inversion result. Comments will be made along the way, but I will keep the final conclusion to the last section in this chapter. I begin by comparing the posterior mean solution with the true geology g shown in Figure 5.3. Note, that through out the rest of the report, when referring to the posterior distribution I refer to $\tilde{\mathbf{m}}/2$ and $\tilde{\mathbf{C}}_m/4$. The inversion is applied to the *summed* result of two signals (Base and Monitor). Hence, as explained in the Methodology section, to obtain comparable reflection amplitude and variance, one has to divide $\tilde{\mathbf{m}}$ by 2 and $\tilde{\mathbf{C}}_m$ by 4.

5.2.1 Posterior mean solution



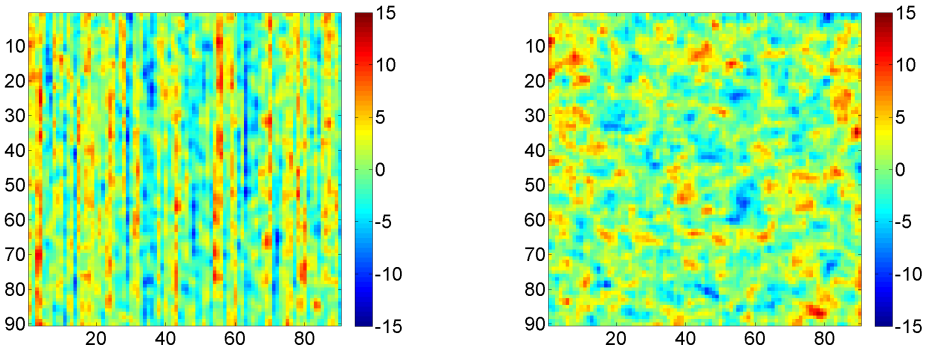
(a) True signal (g). (b) Posterior mean ($\tilde{\mathbf{m}}/2$).

Figure 5.3: Noise filter simulation with known solution. All geology and noise models used for the inversion are identical to the models used for generating the data.

This initial inversion is expected to perform extraordinarily well, since all assumptions about Gaussianity, linearity, additivity, etc. are true and valid. Hence, it is no surprise, that all major and minor features of the original geology g have been resolved, with no indication of striping visible. There is a slight decrease in peak amplitude ($max(|g|)$). This is expected as it is only the *mean* of the posterior solution. The true variability of the solution is better illustrated with a series of random Gaussian realizations (section 5.2.3).

5.2.2 Removed noise and residual error

Figure 5.4 illustrates removed (estimated) noise and the residual error. The removed part of the signal clearly shows high correlation in the N/S direction with some smoothness, as dictated by the prior.



(a) Removed noise, $d_{obs}/2 - \tilde{\mathbf{m}}/2$.

(b) Residual error, $g - \tilde{\mathbf{m}}/2$.

Figure 5.4: Left: Difference between the noisy Observations and the mean posterior solution. Right: Difference between the original noise-free observations and the mean posterior solution.

The residual error, $g - \tilde{\mathbf{m}}/2$, illustrates the difference between the mean posterior solution and the true solution. In this case, where we know the prior geology is identical to the true geology, the "error" is due to the smoothness of the mean posterior and can be thought of as the information lost in the inversion, if one uses the mean solution as the only estimate of the posterior. A deterministic solution of that kind fails to capture the full variability of the solution.

5.2.3 Gaussian realizations

The posterior solution in Figure 5.3 is the model with the Maximum A Posteriori probability, the MAP solution (see section 3.2). Maximum a posteriori probability also means minimum variance in the solution. Hence, the MAP solution tends to be too smooth. Instead, a way to illustrate the true variability of the solution, is to generate random Gaussian realizations of the posterior distribution $\mathcal{N}(\tilde{\mathbf{m}}/2, \tilde{\mathbf{C}}_m/4)$. 9 such realizations are shown in Figure 5.5.

The variance of the realizations is higher than the mean solution (on average), and small scale variability is now seen. The ability to generate a sample of posterior solutions to the problem, all with the same a posteriori probability, is one of the leading advantages of a probabilistic Gaussian inversion scheme.

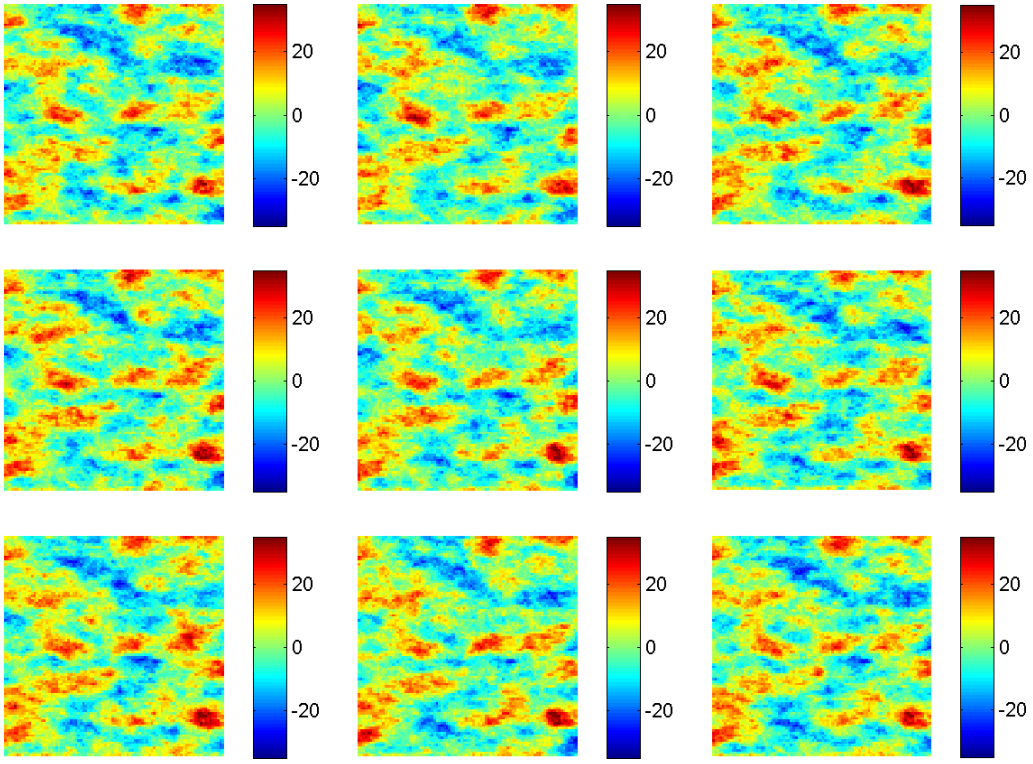
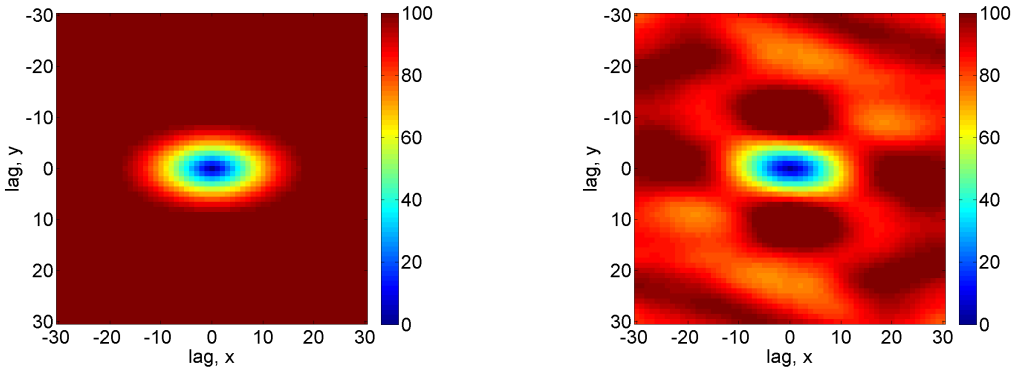


Figure 5.5: Realizations of posterior distribution $\mathcal{N}(\tilde{\mathbf{m}}/2, \tilde{\mathbf{C}}_m/4)$. All realizations have the statistical properties of the least squares solution to the inversion. All carry a clear resemblance with the noise-free geology in Fig. 5.3.

5.2.4 Variograms

Figure 5.6 compares the original variogram model used to generate the synthetic geology, with the experimental variogram of the mean posterior solution. The shape and amplitude of the noise-free geology variogram is well preserved. The slight decrease in variance and slight increase in range is another sign of the inherent smoothness of the mean solution.



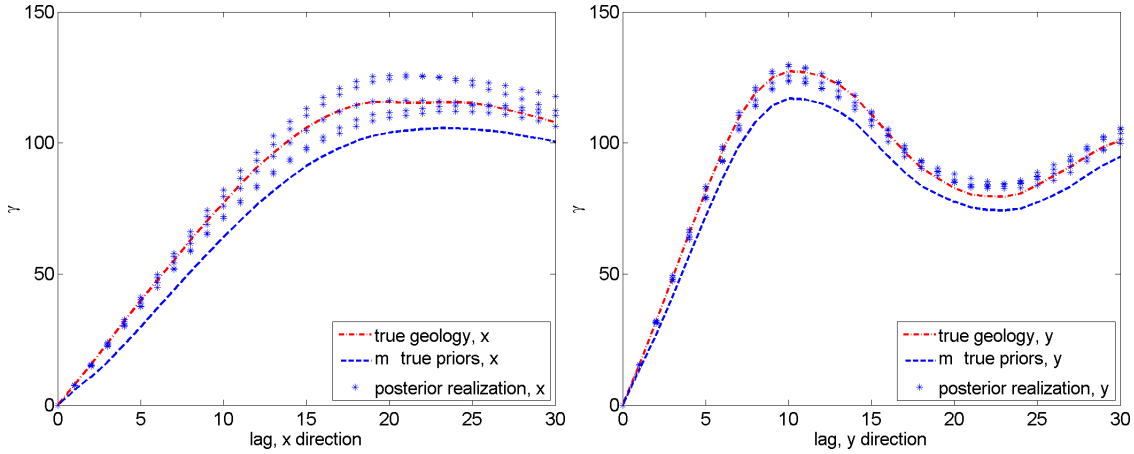
(a) Synthetic variogram of prior geology model.

(b) Experimental variogram of posterior mean.

Figure 5.6: Before and after variogram comparison.

The success of the inversion is better illustrated by directional 1-D variograms. Figure 5.7 shows how close the posterior solution and in particular its realizations come to the real

solution, as a function of lag distance. The red line marks the true geology, the blue line is the mean posterior solution, and the stars are realizations of the posterior. The posterior realizations are all scattered around the true geology, albeit with some uncertainty (spread). The too low variance of the mean solution is evident here.



(a) Major principal direction (180 degrees). (b) Minor principal direction (90 degrees).

Figure 5.7: Comparing semi-variograms of; the true geology, the mean posterior solution, and five posterior realizations.

*

It has been established, that, in this ideal case where both noise and geology are known, the characteristics of the posterior pdf follow those of the reference model closely. I will now advance to the central question of this thesis, and replace the true noise model with a model estimated from the residual, d_{res}.

5.3 Inversion results with noise modeled on the Residual

The following inversion will aim to remove the noise from d_{obs} by modeling the noise covariance on d_{res}. To investigate how well the noise is characterized and modeled under otherwise ideal conditions, the geology is not modeled, i.e. is assumed perfectly modeled. The previous "true model", or reference model, inversion will serve as a benchmark. I will compare the solutions from the "true model" inversion, $\tilde{\mathbf{m}}$ with the "estimated model" inversion, $\tilde{\mathbf{m}}^{mod}$. I will also evaluate how wrong or vague priors, which arguably are sometimes used in the industry, impact the result.

5.3.1 Modeling the noise

As described in the Methodology section, I model the noise variogram on the experimental variogram of the Residual. The result of fitting is seen in 1-D and 2-D in Figure 5.8 and 5.9.

The solid blue line in Figure 5.8 is the modeled noise. Have in mind that, in real life, the true distribution (green dashed line) is not known. However, it is straight forward to fit a model well to the Residual (blue dash-dot line), as it is well behaved, in the sense, that range and sill are relatively easy to determine. Moving to 2-D, I check and model anisotropy not observed in x, y directions. And finally 3D for a thorough visualization (not shown here).

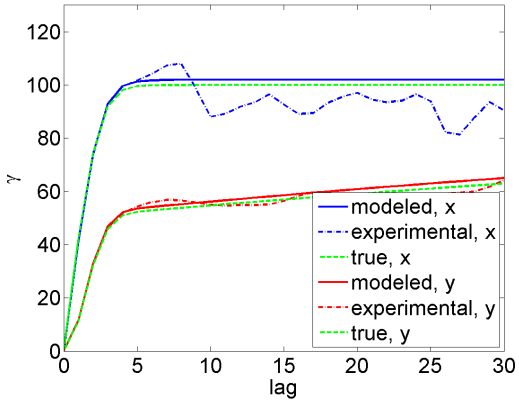


Figure 5.8: Modeling the noise variogram from the experimental variogram of the Residual. x is East/West, y is North/South. The experimental variogram is well behaved and easy to model.

Random Gaussian realizations are generated and compared to the true Residual. If model and Residual appear similar, I proceed with the inversion. Figure 5.9 illustrates the modeling procedure.

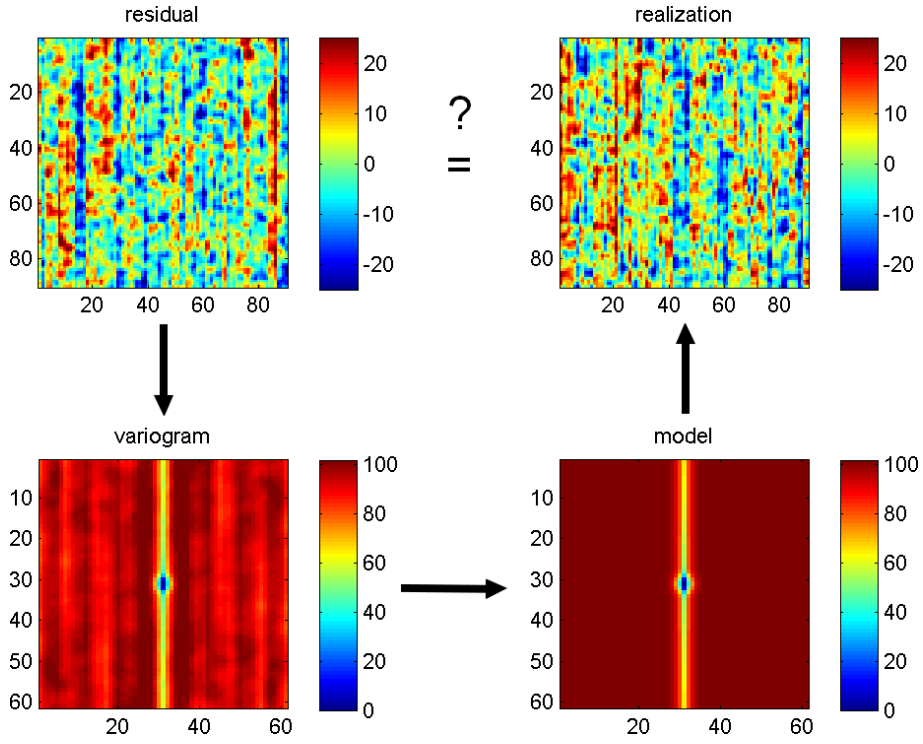


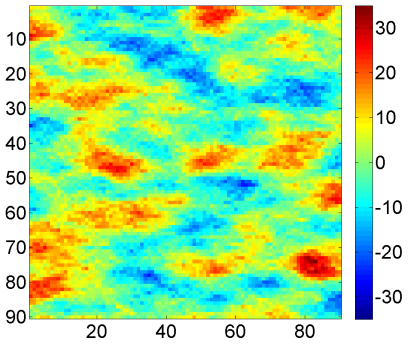
Figure 5.9: Noise variogram modeling. The noise is modeled on the experimental variogram of the Residual d_{res}. Random Gaussian realizations are generated to ensure likeliness to the true residual.

The resulting noise model is included in an updated version of Table 5.2 in Table 5.3.

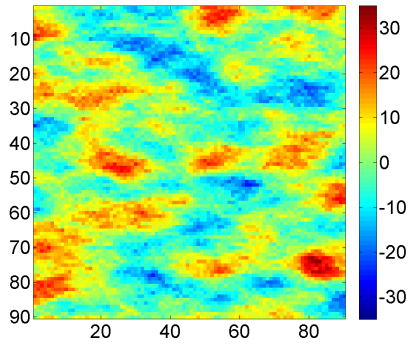
	Model	Sill	Range	Angle	Aniso. (b_1/a_1)
Prior geology, $\mathbf{C}_m$ (true)	Spherical	400	20	0	.5
Noise, $\mathbf{C}_d$ (true)	Exponential	50	300	180	.01
	Gaussian	50	3.5	0	1
Prior geology $\mathbf{C}_m^{mod}$ (modeled)	Same as true geology				
Noise $\mathbf{C}_d^{mod}$ (modeled)	Exponential	51	280	180	.012
	Gaussian	51	3.5	0	1

Table 5.3: The newly fitted noise model is tabled, along with the true geology and noise models.

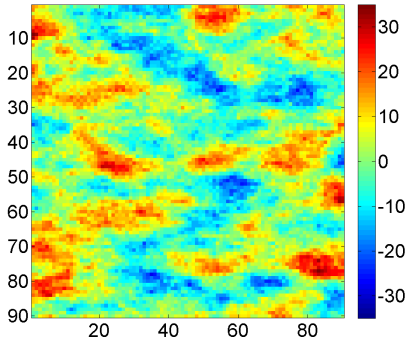
5.3.2 Posterior mean solution



(a) Mean posterior with true priors, $\tilde{\mathbf{m}}/2$.



(b) Mean posterior with modeled noise, $\tilde{\mathbf{m}}^{mod}/2$.



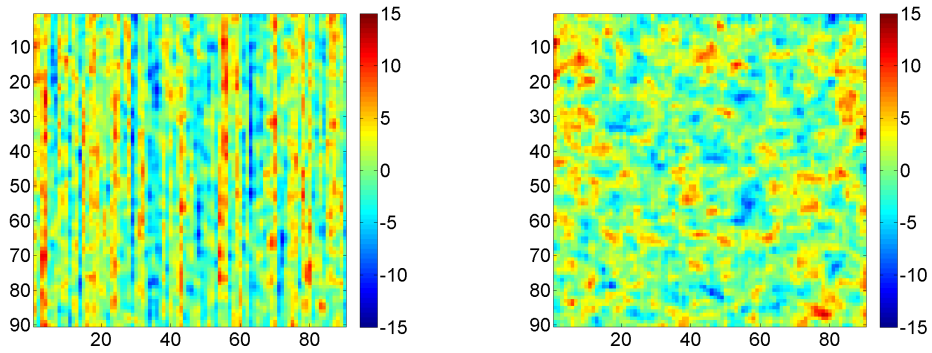
(c) True signal (g).

Figure 5.10: Noise filter simulation with known solution. The noise covariance $\mathbf{C}_d^{mod}$ was modeled on the covariance of the Residual. The geology covariance $\mathbf{C}_m^{mod}$ was identical to that of the true geology.

Figure 5.10 (b) shows (half) the mean posterior $\tilde{\mathbf{m}}^{mod}$ obtained using $\mathbf{C}_d^{mod}$ as noise covariance. As was the case in the former inversion (a), striping is completely removed, and little difference is seen between $\tilde{\mathbf{m}}^{mod}/2$ and the true solution g (c). Only a slight smoothing and drop in variance is noticeable.

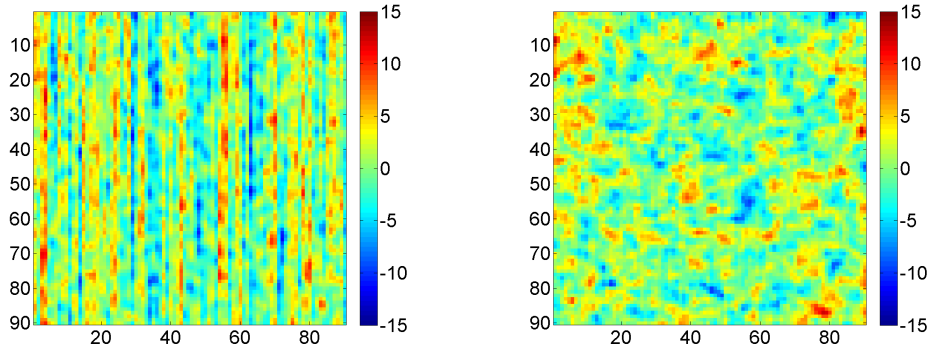
5.3.3 Removed noise and residual error

Figure 5.11a and 5.11b illustrate removed (estimated) noise ($d_{obs}/2 - \tilde{\mathbf{m}}^{mod}/2$) and residual error ($g - \tilde{\mathbf{m}}^{mod}/2$). The removed noise and residual error from the previous "true noise" inversion (Figure 5.4) is repeated in (c) and (d) for comparison.



(a) Removed noise, $d_{obs}/2 - \tilde{\mathbf{m}}^{mod}/2$ (modeled noise).

(b) Residual error, $g - \tilde{\mathbf{m}}^{mod}/2$ (modeled noise).



(c) Removed noise, $d_{obs}/2 - \tilde{\mathbf{m}}/2$ (true noise model).

(d) Residual error, $g - \tilde{\mathbf{m}}/2$ (true noise model).

Figure 5.11: Left: Difference between the noisy Observations and the mean posterior solution. Right: Difference between the original noise-free observations g and the mean posterior solution.

No significant difference between (a), (c) and (b), (d) is seen. The modeled noise and the true noise essentially produce the same result.

5.3.4 Gaussian realizations

Gaussian realizations of the posterior distribution $\mathcal{N}(\tilde{\mathbf{m}}^{mod}/2, \tilde{\mathbf{C}}_m^{mod}/4)$, seen in Figure 5.12, are also similar to the previous realizations of the posterior reference model $\mathcal{N}(\tilde{\mathbf{m}}/2, \tilde{\mathbf{C}}_m/4)$.

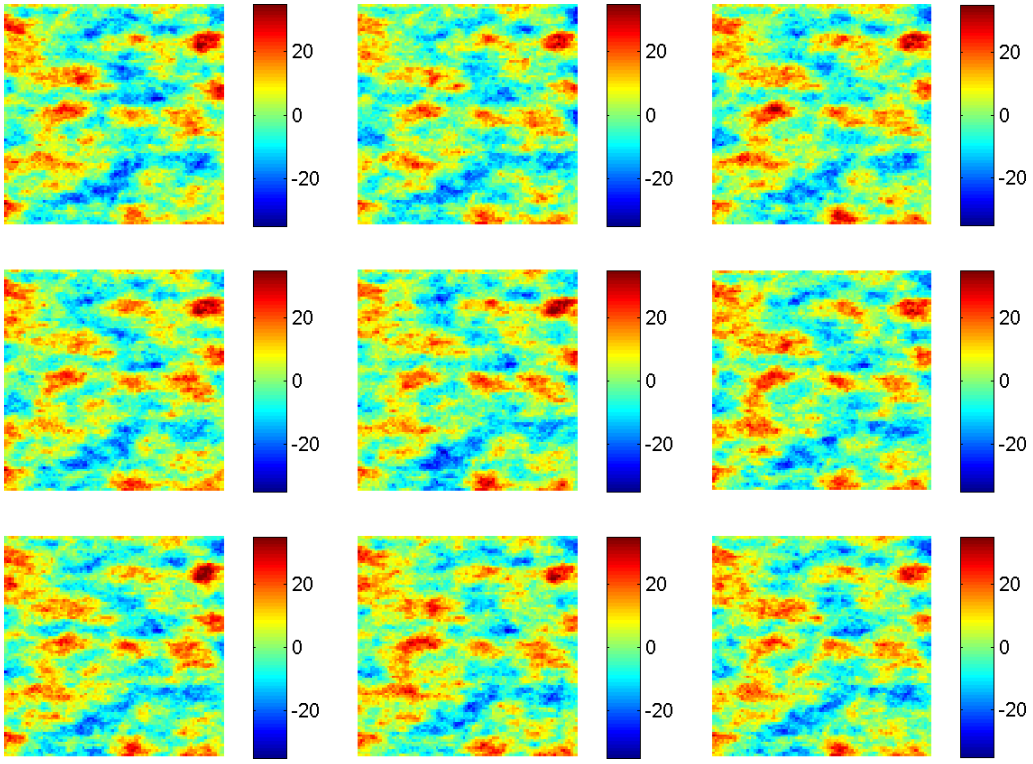
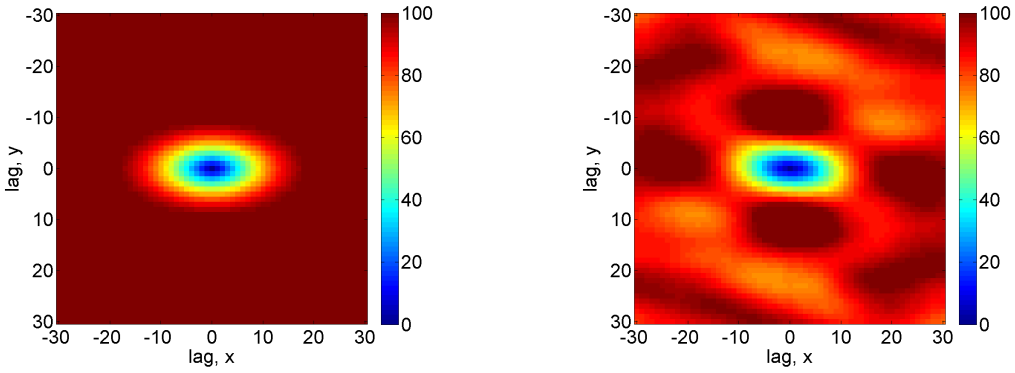


Figure 5.12: Random realizations of the posterior distribution $\mathcal{N}(\tilde{\mathbf{m}}^{mod}/2, \tilde{\mathbf{C}}_m^{mod}/4)$.

5.3.5 Variograms

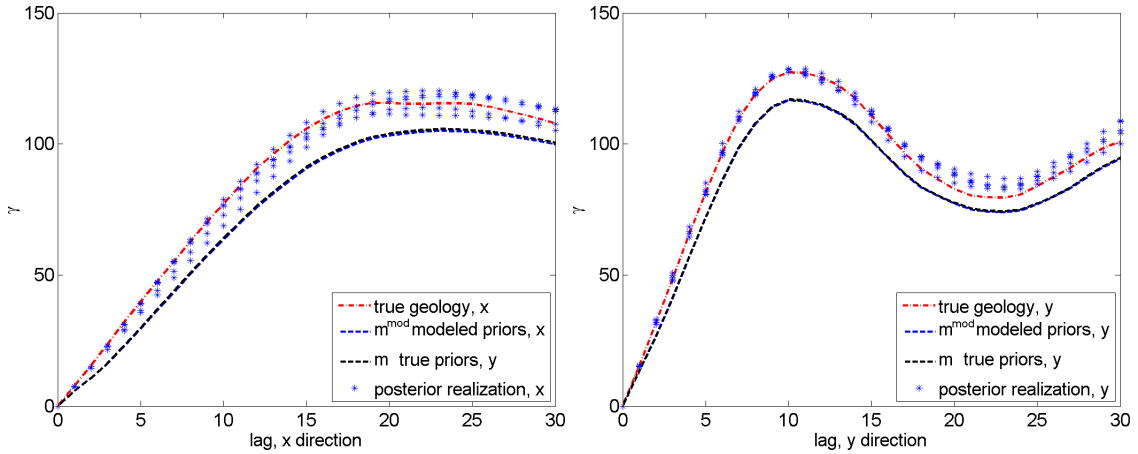


(a) Synthetic variogram of prior geology model.

(b) Experimental variogram of posterior mean.

Figure 5.13: Variogram comparison between the synthetic (true) model and the experimental variogram of the modeled mean posterior.

Figure 5.13 compares the original variogram model used to generate the synthetic geology, with the experimental variogram of the mean posterior solution. The similarity to the true prior solution continues, as the shape and amplitude of the geology variogram is well preserved, with a notable decrease in variance and increase in range. The 1-D plot illustrates the shortcomings better (see Figure 5.14). However, once again, the Gaussian realizations turn out evenly spread out around the true solution, with a tendency to increase in uncertainty with increasing lag.



(a) Major principal direction (180 degrees). (b) Minor principal direction (90 degrees).

Figure 5.14: Comparing semi-variograms of; the true geology, the posterior mean with true and modeled noise, and five realizations hereof (with estimated noise).

5.4 Inversion results with badly estimated geology and noise models

As mentioned earlier, in conventional noise processing the noise is often estimated from expert knowledge or other information which uncertainty is not necessarily well known, leading to uncertain and/or unreliable results. Despite this, the resulting error or bias in the posterior solution is almost never investigated in practice. In the synthetic case, knowing the true solution, it is possible to confirm, both visually and probabilistically, just how good or bad a de-noising scheme has performed. In this case, I have modeled the noise on the Residual, and confirmed, that the noise modeling technique successfully removes striping. But is it actually necessary to know that much about the noise to remove it?

In this section, I will investigate what happens when noise and geology is not well known, beginning with a common choice of noise: Uncorrelated noise, or *nugget* noise. I will also investigate the effect of a wrong geology prior, with too long range, and too high variance. The last iteration will include both wrong models at the same time. Table 5.4 updates the previous model tables.

	Model	Sill	Range	Angle	Aniso. (b_1/a_1)
Prior geology, $\mathbf{C}_m$ (true)	Spherical	400	20	0	.5
Noise, $\mathbf{C}_d$ (true)	Exponential	50	300	180	.01
	Gaussian	50	3.5	0	1
Prior geology $\mathbf{C}_m^{mod}$ (modeled)	Same as true geology				
Noise $\mathbf{C}_d^{mod}$ (modeled)	Exponential	51	280	180	.012
	Gaussian	51	3.5	0	1
Prior geology, $\mathbf{C}_m^{bad}$ (bad)	Gaussian	500	30	0	.7
Noise, $\mathbf{C}_d^{bad}$ (bad)	Nugget	1	N/A	N/A	N/A

Table 5.4: Variations of noise or geology models estimated wrongly are tabled, along with the modeled noise, true noise and true geology.

5.4.1 Posterior mean solution

We already saw, that inverting with an estimated noise model yields less than optimal results. The noise, however, was modeled on the Residual, which gave it a good chance of being close to the true noise. If no or only vague information exists on the data, the geology and noise models must be chosen with corresponding vagueness. I have chosen two such models, listed in Table 5.4 and performed the inversion. Figure 5.15 shows the posterior mean solutions of all four combinations of good and bad priors along with the true geology and noisy Observations. Only the mean posteriors are shown.

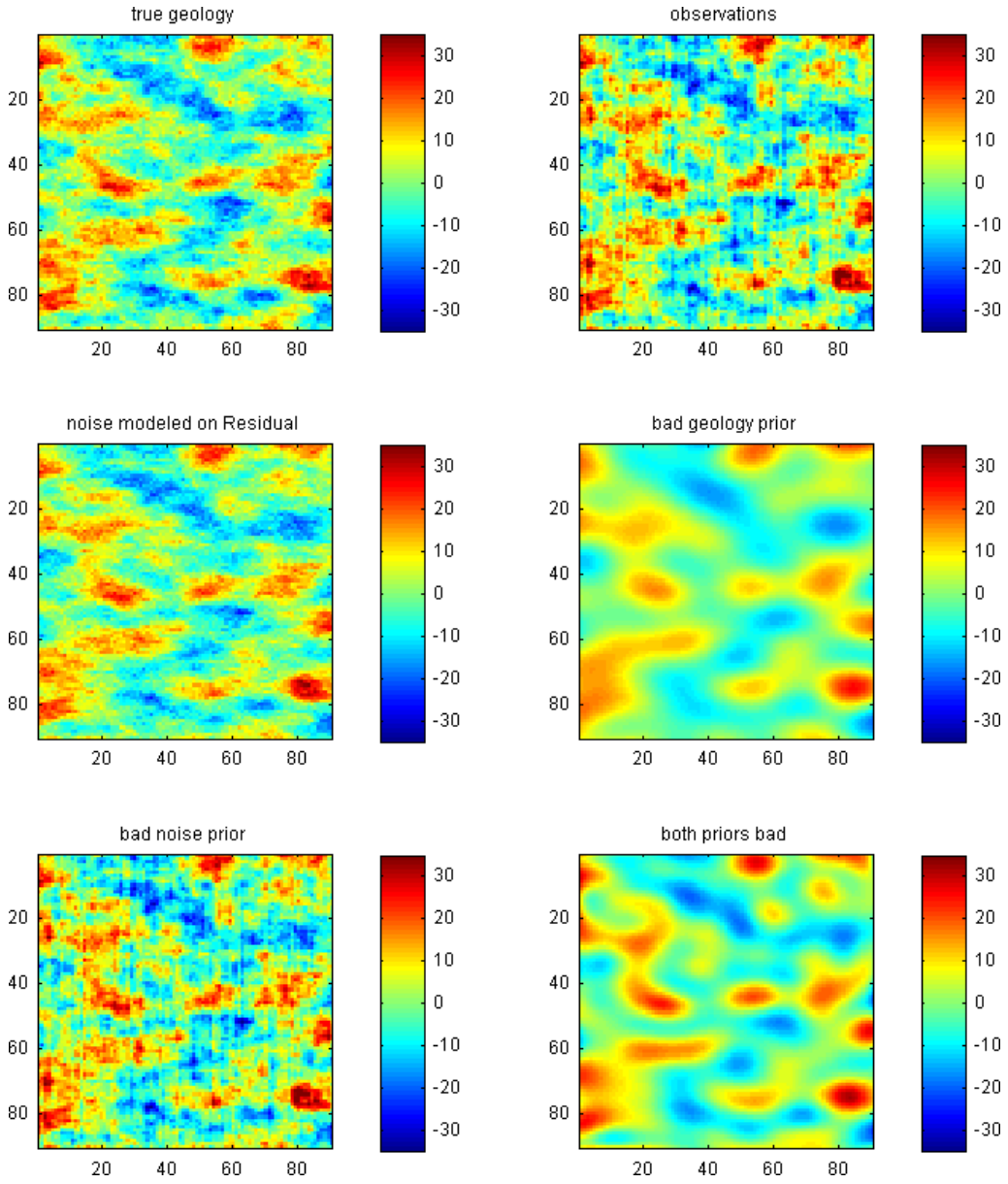
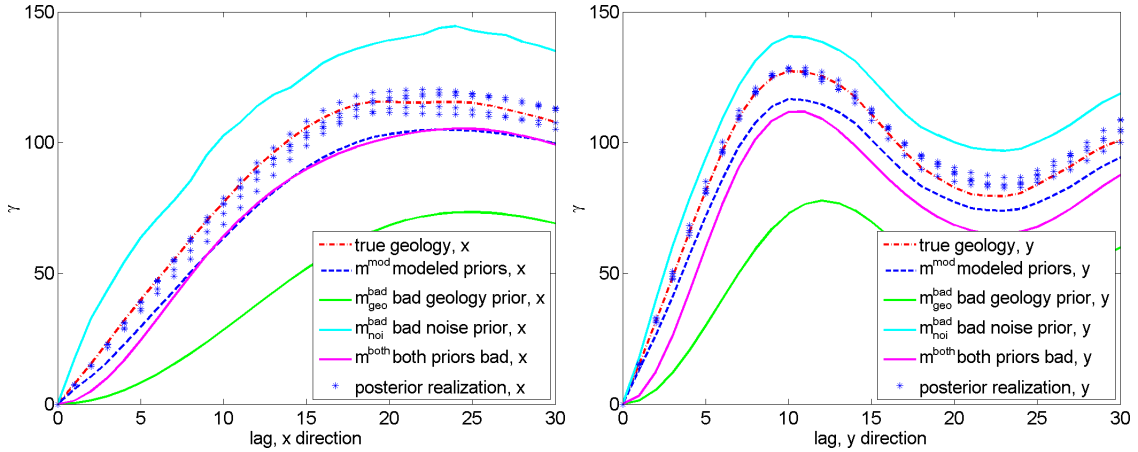


Figure 5.15: Posterior mean solutions of all four combinations of good and bad priors, as well as the true noise-free geology, and the noisy Observations.

It is obvious from simply looking at the different posterior means, that the bad priors have either destroyed the solution, or not affected it at all. Starting with the $\tilde{\mathbf{m}}_{geo}^{bad}$ (bad geology), the long range and high variance of the prior has smoothed the solution severely. Noise is, however, efficiently removed. Looking at the $\tilde{\mathbf{m}}_{noi}^{bad}$ (bad noise), the situation is reversed, as a nugget model can never remove the correlated noise present in the data. Hence, the solution is almost identical to the Observations, i.e. noise has not been removed. Combining the two bad priors $\mathbf{C}_d^{bad}$ and $\mathbf{C}_m^{bad}$ we get $\tilde{\mathbf{m}}^{both}$ (both priors bad) which of course is a cross between $\tilde{\mathbf{m}}_{geo}^{bad}$ and $\tilde{\mathbf{m}}_{noi}^{bad}$. The nugget prior nudges the solution towards shorter range, which is an improvement over the long range $\tilde{\mathbf{m}}_{geo}^{bad}$ solution, but still too smooth to be useful to any geologist or petroleum engineer. The difference in variance and range is better illustrated in a variogram which is discussed in the next section.

5.4.2 Variograms



(a) Major principal direction (180 degrees). (b) Minor principal direction (90 degrees).

Figure 5.16: Comparing semi-variograms of posterior solutions based on different combinations of true, modeled or wrong priors.

Figure 5.16 shows semi-variograms in the two principal directions of the mean posterior solutions shown in Figure 5.15. The three bad solutions are colored cyan, magenta, and green. The red and blue lines are the covariance of the geology and $\tilde{\mathbf{m}}^{mod}$ repeated from earlier semi-variograms. The smoothness of $\tilde{\mathbf{m}}_{geo}^{bad}$ and $\tilde{\mathbf{m}}^{both}$ is seen here as too long range in the y direction and too small gradient (initially). $\tilde{\mathbf{m}}_{noi}^{bad}$ on the other hand, overestimates the variance, but otherwise follows the true geology well, since it basically looks like the Observations, which, in a statistical sense, is close the true geology. The increase in variance is due to the nature of uncorrelated noise: No restraint is put upon variations near or far from any point. All can vary randomly, giving rise to high variance. The $\tilde{\mathbf{m}}_{noi}^{bad}$ solution's immediate close relationship with the true geology hence, is not a sign of success. Rather a sign of the inability to move towards the true solution. $\tilde{\mathbf{m}}^{mod}$ is clearly the best estimate of the noise-free geology, in particular if one looks at the random realizations.

5.4.3 Assumptions

No assumptions made during this synthetic example; stationary, additive noise, geological stability, alignment of data sets, normal distributed data, are violated. Only the number of data points included in the modeling and inversion can alter each outcome slightly.

It would, however, be interesting to investigate to what degree any assumption could be weakened. However, due to the limited time set for this project, these questions were not investigated.

5.5 Conclusion

The synthetic example has provided an insight into the inversion and modeling procedure, not possible with real data. By knowing the true solution to the problem, i.e. the noise-free geology, the posterior solutions can be graded and the success of the de-noising procedure quantified. Several posterior distributions based on different noise and geology models were compared to a posterior reference distribution, $\tilde{\mathbf{m}}$, where the noise and geology models were identical to the models used to generate the synthetic data. No solution can be expected to surpass the performance of the reference distribution. The posterior estimate with modeled noise, $\tilde{\mathbf{m}}^{mod}$, was very close to $\tilde{\mathbf{m}}$, which lends credibility to the noise characterization technique. Inversions based on wrong prior information, made to simulate the absence of good prior information, performed much worse. Hence, estimating the noise model from the Residual $s_1 - s_2$, and filtering the Observations $s_1 + s_2$, has proven highly successful in this synthetic event where all assumptions are valid. It remains to be seen, if the method works on real data.

Chapter 6

Preliminary studies of the data

The method has been tested successfully on synthetic data. However, before we attempt to apply it to real data, those data must pass some validation checks. I have been provided with two cubes of seismic reflection data from the upper part of the Halfdan field in the North Sea. In the following sections of this chapter, I will investigate if these data can be assumed to have some or all of the following properties:

- Normally distributed
- Zero mean
- First and second order stationarity

Besides these three statistical properties, the data should also comply with the following three properties:

- *Additive noise*
- *Correct geographical alignment of data sets*
- *Geological stability*

6.1 History of the data

Two data sets have been kindly provided by Maersk Oil. The data are in the form of marine reflection seismics (SEG-Y format), and were taken at the Halfdan field in the North Sea. The acquisition boat dragged several lines of hydrophones along what becomes the inline direction of the data. This created a strong inline noise acquisition imprint. The first data were taken in 2005. 7 years later, in 2012, the survey was repeated. Combined, the two 3-D data sets constitute one 4-D data set. The data have been modified to include only the overburden, i.e. above the reservoir, and are from a part of the section which does not exhibit changes in elastic properties between the two surveys. Hence, the geological component can be assumed constant. This would not be true for parts including the reservoir, which does show differences.

Each original data set measured 1751 by 800 by 450 cells, (12.5 m, 12.5m, 4 ms). This was cut down to 1751 by 651 by 31 (every fifth layer between layer 50 and 200) to resolve memory issues and the fact that the data were irregularly sized. The data have been through Krichhoff Prestack time migration (PSTM) and stacking, prior to my analysis. However, it is not known if the data are suitable for Bayesian inversion. The coming analysis will address this issue.

6.2 Gaussianity

A true Gaussian distribution is completely described by its mean and variance. Other interesting characteristics are skewness and kurtosis. These parameters are closely related to the *moments* of the distribution. The m'th moment about the mean μ is

$$\mu_m = E[(X - E[X])^m] = \int_{-\infty}^{+\infty} (x - \mu)^m f(x) dx, \tag{6.1}$$

where X is a stochastic variable, x is a realization of the stochastic process, and $f(x)$ is a probability density function. Notice that I am talking about the *central* moment here, which means the moment is centered around the mean μ, which is defined

$$\mu = E[X] = \int_{-\infty}^{+\infty} x f(x) dx. \tag{6.2}$$

As just mentioned, the mean and variance completely describes the true Gaussian distribution. Skewness and kurtosis, related to the third and fourth moment, play no role in characterizing the true normal distribution. If the distribution is not truly Gaussian however, skewness and kurtosis must be calculated to quantify how close to or far from Gaussian it is. Real data are usually not perfectly Gaussian. In the discrete case, we have the sample central moment

$$\mu_m = \frac{1}{n} \sum_{k=1}^{n} (x_k - \mu)^m, \tag{6.3}$$

where n is the sample size, and $\mu = \frac{1}{n} \sum_{k=1}^{n} x_k$. I will continue to investigate the mean, variance, skewness and kurtosis of the Halfdan data.

The data set stretches over large distances, both horizontally and vertically. Therefore, I will list the parameters for several parts of the data set, both in the form of Observations (Base and Monitor added) and Residual (Base and Monitor differenced). The results are presented in Table 6.1 and 6.2 and discussed in turn in the following sections. For comparison, I calculated the same parameters for the synthetic data generated in the previous chapter, which are known to be zero mean and Gaussian. All investigated data are shown in Figures 6.1 and 6.2, except the 3-D cube. Histograms are also provided. The investigated layer, layer 140, is chosen for its extremely noisy nature.

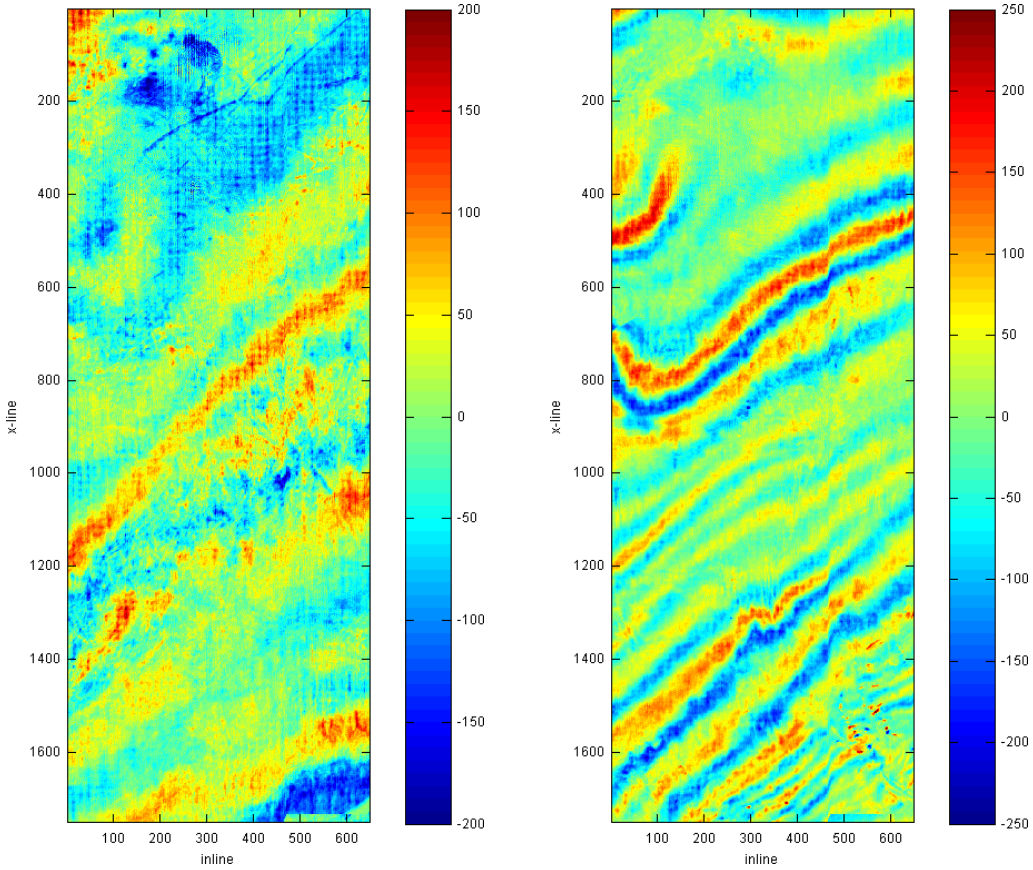
Observations ($s_1 + s_2$)	Mean	Variance	Skewness	Kurtosis
Full 3-D cube	2.765	3205	0.6148	10.460
Single horizontal layer, shallow	-7.244	2338	-0.2462	3.406
Single horizontal layer, deep	-2.540	3386	-0.1151	3.310
Patch of 200x200, shallow	-21.697	1074	-0.4378	3.951
Patch of 200x200, deep	-23.996	1865	0.5522	3.208
Noisy synthetic data, d_{obs}	-3.397	1950	0.0751	3.239

Table 6.1: Statistics on Observations. The mean, variance, skewness and kurtosis of different parts and sizes of the Halfdan data and synthetic data.

Residual ($s_1 - s_2$)	Mean	Variance	Skewness	Kurtosis
Residual full 3-D cube	0.970	198	0.7660	11.600
Residual single horizontal layer, shallow	-1.020	99	0.0246	3.395
Residual single horizontal layer, deep	0.533	45	-0.0132	3.345
Residual patch of 200x200, shallow	2.032	86	-0.1500	3.074
Residual patch of 200x200, deep	1.417	40	0.2731	3.117
Residual synthetic data, d_{obs}	0.460	181	0.0357	2.876

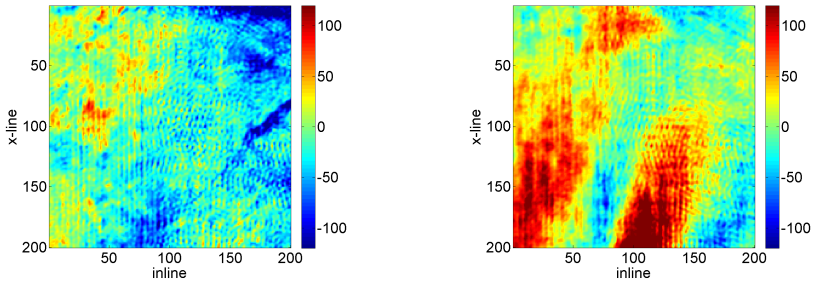
Table 6.2: Statistics on Residual. The mean, variance, skewness and kurtosis of the residual of different parts and sizes of the Halfdan data and synthetic data.

OBSERVATIONS



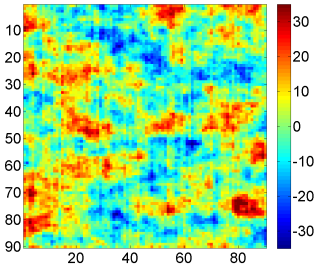
(a) Shallow layer.

(b) Deep layer.



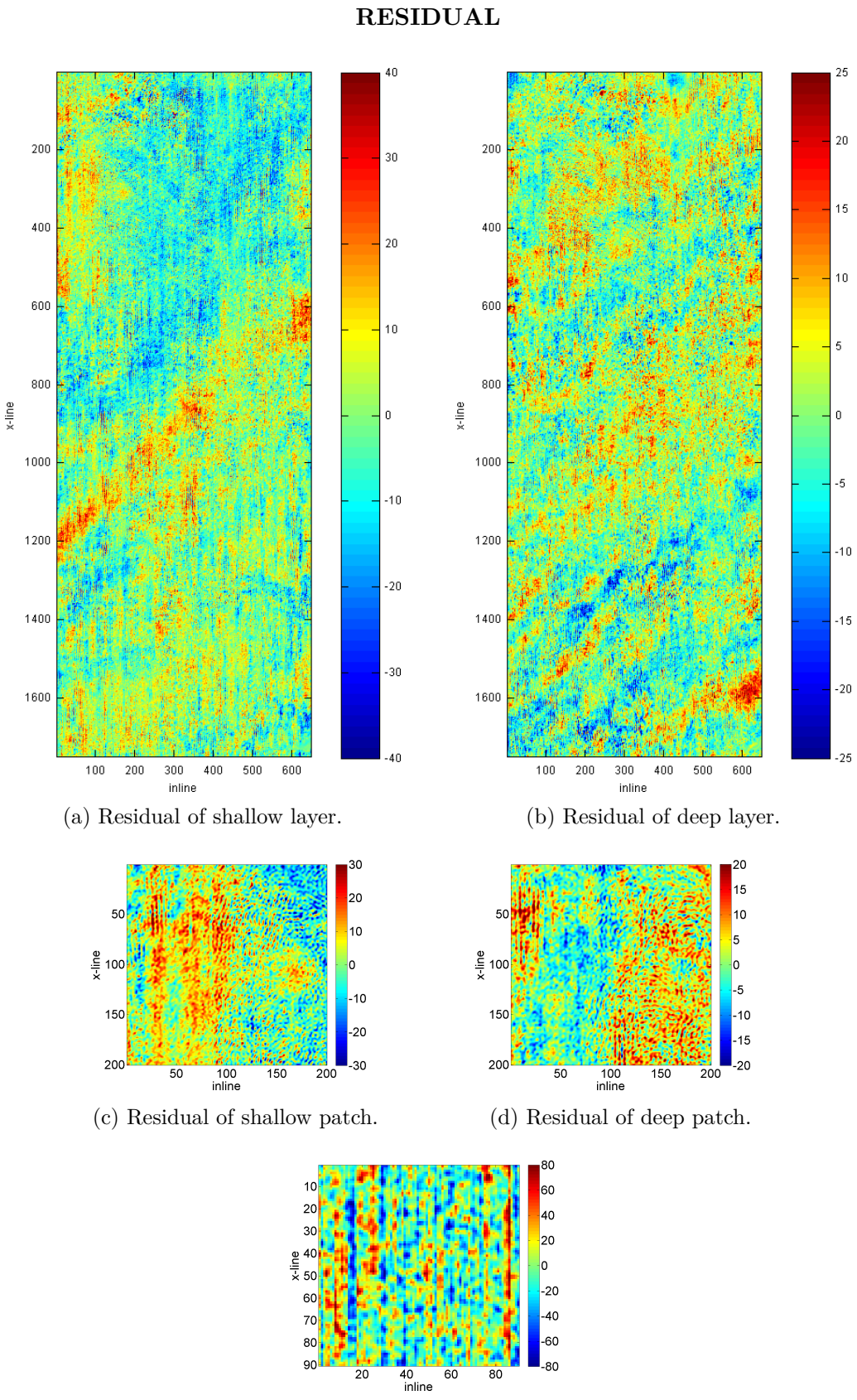
(c) Shallow patch.

(d) Deep patch.



(e) Synthetic d_{obs}.

Figure 6.1: Investigated data.



(a) Residual of shallow layer.

(b) Residual of deep layer.

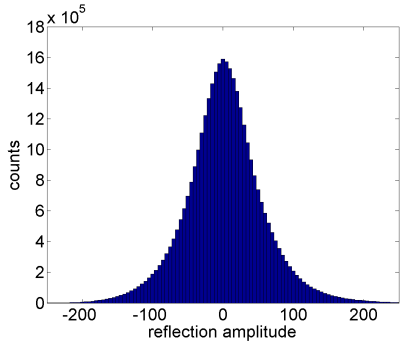
(c) Residual of shallow patch.

(d) Residual of deep patch.

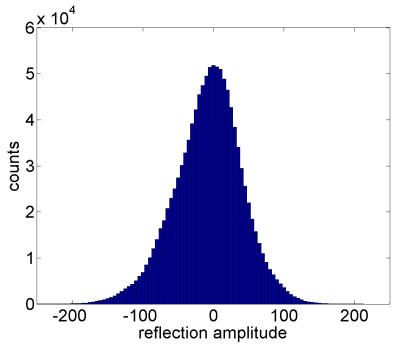
(e) Residual of synthetic d_{obs}.

Figure 6.2: Residual of investigated data.

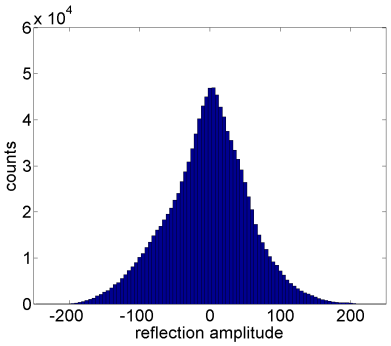
OBSERVATIONS



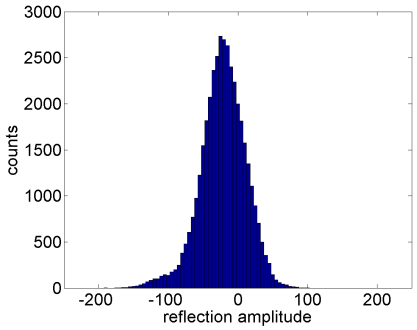
(a) Full 3-D cube.



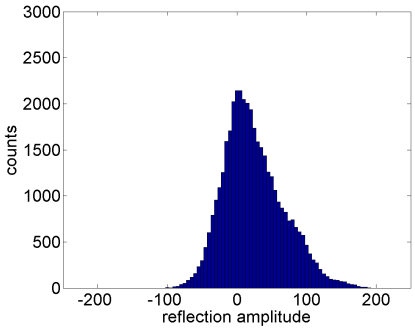
(b) Shallow layer.



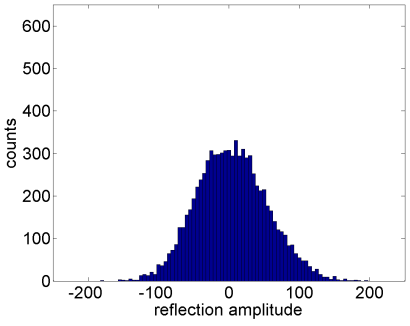
(c) Deep layer.



(d) Shallow patch.



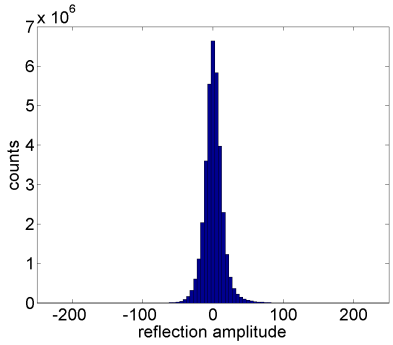
(e) Deep patch.



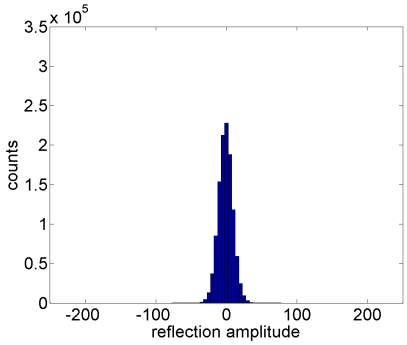
(f) Synthetic d_{obs}.

Figure 6.3: Histograms of investigated Observations data.

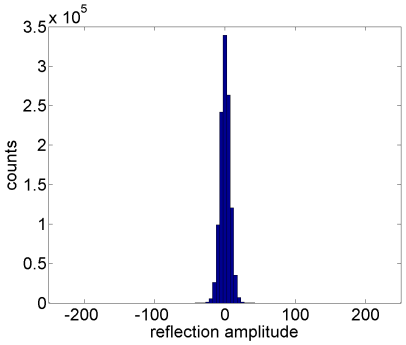
RESIDUAL



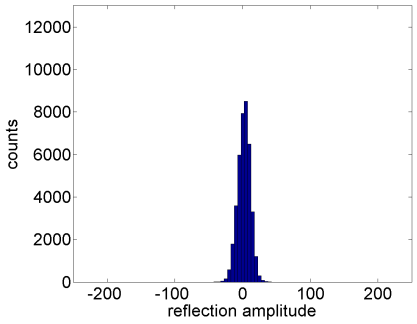
(a) Residual of full 3-D cube.



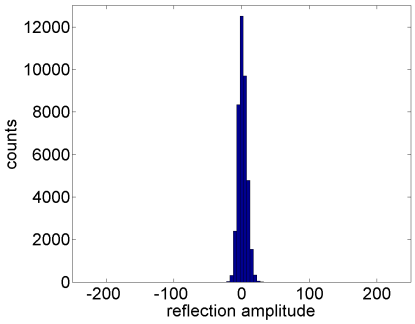
(b) Residual of shallow layer.



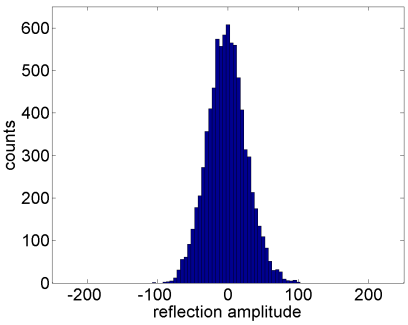
(c) Residual of deep layer.



(d) Residual of shallow patch.



(e) Residual of deep patch.



(f) Residual of synthetic d_{obs}.

Figure 6.4: Histograms of investigated Residual data.

The mean

The mean, or *expected value*, is an intuitive variable that one would expect to find, if a certain random variable process was repeated an infinite amount of times and multiplied by its probability density function. The sample mean is an approximation to the true mean that needs to converge for n approaching infinity. That is why having a large number of data points ensures the best possible estimate of the mean. As briefly mentioned, the sample mean is defined

$$\mu = E[X] = \sum_{k=1}^{n} x_k p_k = \frac{1}{n} \sum_{k=1}^{n} x_k, \tag{6.4}$$

where p_k is the probability of the k'th value of x. If X is a stochastic process, p_k becomes $1/n$, and we have the right hand side of Equation (6.4). The mean specifies the center of the Gaussian distribution and is assumed to be zero prior to data investigation. However, the sample mean of the Halfdan is not zero, or even constant. I repeat the means from Table 6.1 and 6.2 for convenience in Table 6.3. You will also need Table 6.4 (next page) as I refer briefly to variances.

	Mean of Observations	Mean of Residual
Full 3-D cube	2.765	0.970
Single horizontal layer, shallow	-7.244	-1.020
Single horizontal layer, deep	-2.540	0.533
Patch of 200x200, shallow	-21.697	2.032
Patch of 200x200, deep	-23.996	1.417
Noisy synthetic data, d_{obs}	-3.397	0.460

Table 6.3: The mean of different parts and sizes of the Halfdan data and synthetic data.

The effect of the non-zero mean is, if not taken into account, a bias in the inversion result. To prevent this, one can demean the data by subtracting the sample mean from all data points. Looking at Observations (the first column of Table 6.3), the smallest investigated data sets have a sample mean of -21.7 and -24.0. Hence, it deviates from zero, but stays within one standard deviation (square root of variance) equal to 32.8 and 43.2, respectively. The single horizontal layers come much closer to the anticipated zero-mean, with a sample mean of -7.2 and -2.5, respectively, and twice the variance, on the same order of magnitude as the synthetic data. The large difference between means in Single and Patch suggests the smallest data set only depict local behavior, and might be considered too small to adequately characterize the geology. The complete cube tips the mean to the positive side, but stays close to zero. With a variance of 3,205, and a sample mean of 2.8, any bias introduced by assuming zero mean would be small. The millions of data points involved could lead one to believe, that the sample mean is close to the true mean. And it would be, if the random process generating all the data was first order stationary. However, the vertical distance between the upper and lower layers is thousands of feet. It is not likely, that the upper and lower layers can be described by the same process and thus have the same mean.

The mean of the Residual, however, is close to zero in all cases. This suggests, that the process generating the noise is more stationary than the geology, and that a zero mean noise process is likely.

Variance

The second central moment describes the squared differences with the mean summed. Usually it is called the variance, and is carried by the symbol σ^2. It describes how much any data point is expected to vary from another.

$$\mu_2 = \sigma^2 = E[(X - \mu)^2] = \frac{1}{n}\sum_{k=1}^{n}(x_k - \mu)^2. \tag{6.5}$$

Variance is the second variable needed to describe the standard Gaussian distribution.

	Variance of Observations	Variance of Residual
Full 3-D cube	3205	198
Single horizontal layer, shallow	2338	99
Single horizontal layer, deep	3386	45
Patch of 200x200, shallow	1074	86
Patch of 200x200, deep	1865	40
Noisy synthetic data, d_{obs}	1950	181

Table 6.4: The second moment of different parts and sizes of the Halfdan data and synthetic data.

One of the assumptions made in the synthetic example, is second order stationarity, meaning the variance must be invariant to translation. Looking at Table 6.4 we see significantly different amount of variance in Single horizontal layer and Patch. This suggests second order stationarity does not hold. Only 1-D variance is presented in the table (data has been vectorized), but the assumption should still hold. They are of the same order of magnitude though. The difference in variance supports the earlier statement, that a 200 x 200 patch might be too small to accurately describe the mean and variation in the process behind the data.

In the Observational data, the shallow layers seem to have lower variance. It also so happens, that noise is most visible in the shallow layers. Random or correlated noise usually increases variance. That variance then is smaller in the shallowest layers, is unexpected, but supports the earlier statement, that different processes created the upper and lower layers. The Residual data however, have lowest variance in the deepest layers, opposite the Observations. This is expected due to the way the data have been processed (stacking) leading to better signal-to-noise ratio in the deeper layers.

Skewness

Skewness is a measure of the extent to which the distribution "leans" to one side or another relative to the mean. The definition of skewness, is

$$\mu_3 = E[(\frac{X - \mu}{\sigma})^3] = \frac{E[(X - \mu)^3]}{(E[(X - \mu)^2])^{3/2}}, \tag{6.6}$$

also called the third standardized moment. The sample skewness follows

$$\mu_3 = \frac{\frac{1}{n}\sum_{k=1}^{n}(x_k - \mu)^3}{(\frac{1}{n}\sum_{k=1}^{n}(x_k - \mu)^2)^{3/2}}. \tag{6.7}$$

A distribution leaning to the right thus has positive skewness, and negative to the left. A normal distribution is symmetric, and therefore has zero skewness. The data however, are weakly or moderatley non-symmetric and skewed. This is evident from Table 6.5 but also the histograms of the data in Figure 6.3.

	Skewness of Observations	Skewness of Residual
Full 3-D cube	0.6148	0.7660
Single horizontal layer, shallow	-0.2462	0.0246
Single horizontal layer, deep	-0.1151	-0.0132
Patch of 200x200, shallow	-0.4378	-0.1500
Patch of 200x200, deep	0.5522	0.2731
Noisy synthetic data, d_{obs}	0.075	0.0357

Table 6.5: The skewness of different parts and sizes of the Halfdan data and synthetic data.

Skewness is prominent in both Observations and Residual for Patch and 3-D cube data. For the Observations, the full data cube has the highest absolute skewness of 0.61, closely followed by the smallest data sets with 0.55 and 0.44. The order is repeated for the Residual, although the skewness is lessened in all but the Full 3-D cube. Only the Single layers come close to the low level skewness of the synthetic data. Once again, this suggests that the full data cube takes in too much data from too different distributions, and the small data set takes in too little.

Kurtosis

The last distribution variable I will be investigating is kurtosis. Kurtosis refers to the peakedness or curvature of a distribution and is related to the fourth central moment. It is often used as a measure of tail length. A long, or heavy tail, suggests more variance is due to infrequent but extreme deviations. The opposite means most events lie close to the mean, yielding steep side slopes and short tails. Kurtosis is defined as the standardized fourth moment

$$\mu_4 = E[(\frac{X-\mu}{\sigma})^4] = \frac{E[(X-\mu)^4]}{(E[(X-\mu)^2])^2}. \tag{6.8}$$

The sample kurtosis is

$$\mu_4 = \frac{\frac{1}{n}\sum_{k=1}^{n}(x_k-\mu)^4}{(\frac{1}{n}\sum_{k=1}^{n}(x_k-\mu)^2)^2}. \tag{6.9}$$

A Gaussian distribution should ideally have no excess kurtosis. Excess kurtosis is defined as kurtosis minus 3. The kurtosis of the data is repeated in Table 6.6.

	Kurtosis of Observations	Kurtosis of Residual
Full 3-D cube	10.460	11.600
Single horizontal layer, shallow	3.406	3.395
Single horizontal layer, deep	3.310	3.345
Patch of 200x200, shallow	3.951	3.074
Patch of 200x200, deep	3.208	3.117
Noisy synthetic data, d_{obs}	3.239	2.876

Table 6.6: The kurtosis of different parts and sizes of the Halfdan data and synthetic data.

All data sets have slight excess kurtosis, including the synthetic data sets. Only the Full 3-D cube is excessively kurtosious. This is hardly visible in Figure 6.3a or 6.4a, but clear in Table 6.6.

Conclusion

From initial investigation of the four variables mean, variance, skewness and kurtosis, in various segments of the data, we have learned that the Halfdan data are hard pressed to be considered Gaussian in anything than the Single horizontal layers. The smaller data sets introduced a bias due to non-zero mean and higher skewness, while the Full 3-D cube was excessively skewed and highly kurtosious. It was however noted, that the Residual in all but the Full 3-D cube, was much closer to a Gaussian distribution than the Observations. I will continue the investigation with some normality tests.

6.3 Kolmogorov-Smirnov test

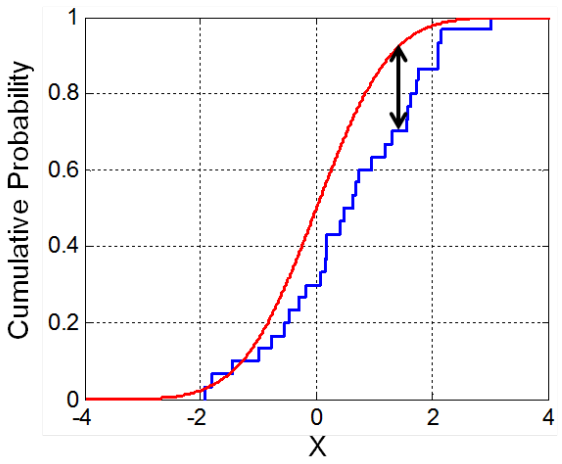


Figure 6.5: Illustration of the Kolmogorov-Smirnov statistic. Red line is CDF, blue line is an ECDF, and the black arrow is the K-S statistic. Source: `http://en.wikipedia.org/wiki/Kolmogorov%E2%80%93Smirnov_test`

It is possible to test for normality in more efficient ways than looking at standardized central moments. One common test is the Kolmogorov-Smirnov (K-S) test which compares the empirical distribution function of the sample with the cumulative distribution function

of a reference distribution, in our case the normal distribution. The probability that the sample distribution is a realization of the reference distribution, is calculated. A generic illustration is seen in Figure 6.5. Table 6.7 lists the calculated probabilities. Each data set was compared to two reference Gaussians; one with zero mean and variance equal to the variance of the data tested, and another with both mean and variance equal to that of the data. The reference Gaussians are my null hypothesis'.

	Gaussian probability for Observations Zero mean / Sample mean	Gaussian probability for Residual Zero mean / Sample mean
Full 3-D cube	0 / 0	0 / 0
Single horizontal layer, shallow	0 / 0	0 / 8.206e-81
Single horizontal layer, deep	0 / 0	0 / 3.723e-76
Patch of 200x200, shallow	0 / 2.267e-32	0 / 3.428e-12
Patch of 200x200, deep	0 / 7.780e-119	0 / 5.050e-23
Noisy synthetic data, d_{obs}	3.524e-11 / 3.352e-26	5.405e-17 / 0.953

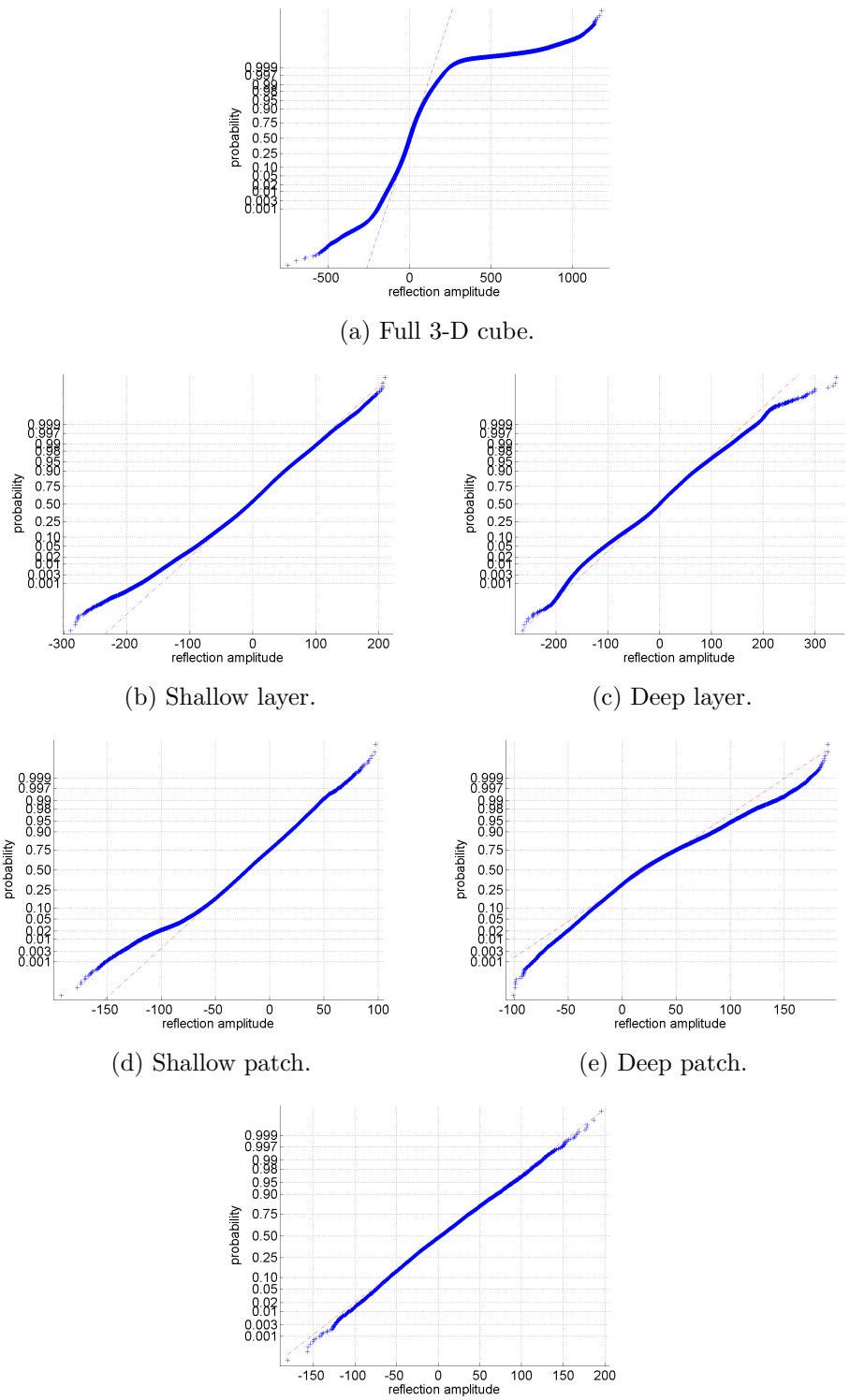
Table 6.7: Kolmogorov-Smirnov tests of different parts and sizes of the Halfdan data and synthetic data.

Clearly, the mean has a significant impact on the result, even for the synthetic data, where the sample mean is very close to zero. The zero mean hypothesis is completely rejected (0 % probability) in all cases but the synthetic data. The sample mean increases the probability, though not to a degree validating the use of the term Gaussian about the distribution.

From Table 6.7 it follows, that the Kolmogorov-Smirnov test rejects the null hypothesis at a significance level of more than 99.9% for all but the synthetic residual with sample mean. However, this should not discourage any further investigation or use of the data in a Gaussian framework. The test is just one way of checking Gaussianity and is not definitive. Also, if we loose the Gaussian assumption, the whole inversion scheme becomes invalid.

6.4 Normal probability plots

Another way of testing normality is to look at normal probability plots. It is based on normal order statistic medians or means (look up rankits), which are superimposed on a line joining the first and third quartiles of the data, which is then extrapolated. Normality is evaluated visually, by noting how far the data (blue stars, see next figure) are from the empirical normal distribution (red dash). A Gaussian sample is expected to lie close to the straight line. The normal probability plots of the previously investigated data are presented in Figures 6.6 and 6.7.



(a) Full 3-D cube.

(b) Shallow layer.

(c) Deep layer.

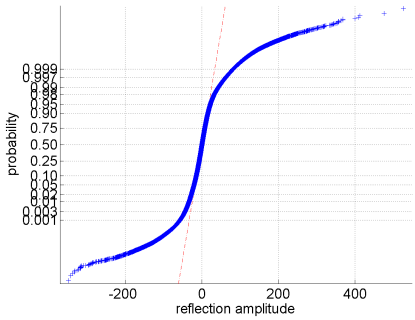
(d) Shallow patch.

(e) Deep patch.

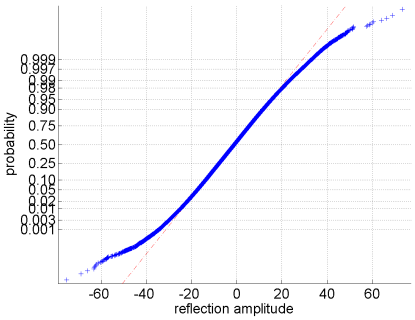
(f) Synthetic d_{obs}.

Figure 6.6: Normal probability plots of investigated data.

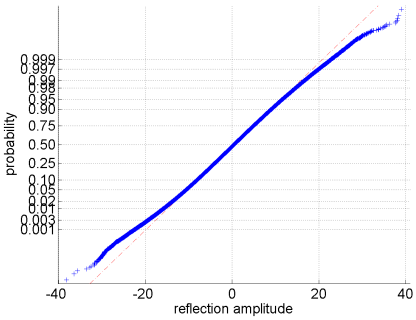
RESIDUAL



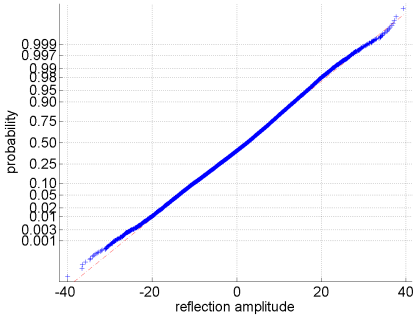
(a) Residual of full 3-D cube.



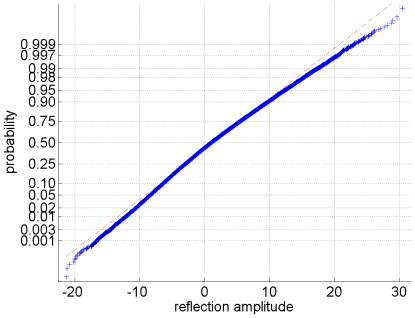
(b) Residual of shallow layer.



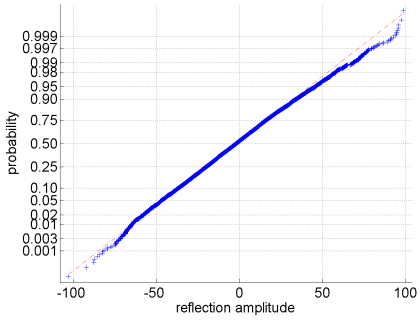
(c) Residual of deep layer.



(d) Residual of shallow patch.



(e) Residual of deep patch.



(f) Residual of synthetic d_{obs}.

Figure 6.7: Normal probability plots of investigated data.

The Gaussian or non-Gaussian properties of the data are clearly illustrated in the previous figures. In contrary to the previous K-S tests, all data sets seem fairly Gaussian, following the line well in most cases, except for the Full 3-D cube. Small deviations are expected, especially at extreme values where the number of data points is small. Excluding the Full cube, this supports the earlier statement, that we should not or need not abandon the Gaussian assumption lightly.

6.5 Stationarity

To avoid breaking the stationarity assumption, the stochastic process or processes describing the distribution, from which the Observations are a realization, must not change over time or space. Formally, the mean must be constant over the data space (first order stationary) and the auto-covariance function must be position independent (second order stationary). That the mean and variance is not constant with depth is of no concern. But we have also seen that mean and variance of Observations are not constant within one time slice in shallow or in deep layers, thus making the mean and covariance of Observations non-stationary.

The mean and variance of the Residual data vary much less. The mean is close to zero, and the variance changes only slightly. From this, the noise can be assumed to be first and second order stationary, at least if the noise description is limited to a single layer.

6.6 Other requirements

Additive noise
It is very difficult to assess whether the Halfdan data are contaminated only by additive noise. It is a popular simplification to assume additive (linear) noise, which first of all comes from the fact, that it is fast and easy to handle. Unless it seriously contradicts the prior knowledge about the noise process, all noise is assumed additive. Furthermore, it is often also assumed white and Gaussian[a] and hence uncorrelated. The assumption is well known, and widely accepted, because the studies required to prove or disprove additivity would be disproportionately demanding, compared to the seismic survey itself, in both acquisition and processing time. Also, it is questionable whether a possibly more realistic non-linear noise model will contribute significantly to the accuracy of the uncertainty estimates and interpretation of the data in general. I was able to produce synthetic data with additive noise, with roughly the same statistical properties as the Halfdan data, and subsequently generate realizations that convincingly emulated the true data (see Chapter 5). Hence, the additive noise assumption is believed to be valid.

Correct geographical alignment of data sets
This noise filtering technique relies on 4-D data from which we get the Observations and the Residual. To avoid blurring when summing or differencing the data, the data sets must be aligned perfectly so coordinate $(x, y)_{Base}$ lines up with coordinate $(x, y)_{Monitor}$. Misalignment will cause blurriness and reduce the effective resolution of the combined signals. The alignment of the two data sets were done by professionals and assumed related with negligible error compared to the size of an individual grid cell.
Geological stability
Vertical alignment of the data is also important and relies, as the horizontal alignment,

[a] Identically distributed and statistically independent (iid).

on the GPS and the data processing. Assuming the GPS is correct and both surveys undergo the same processing sequence, geological instability of the region becomes the largest source of error. The Halfdan field is known to be a production field and thus risks various production related structural deformations, e.g. compaction. If deformation has occurred, velocity stretching, among other things, will distort the combined signal just like horizontal displacement, only in the vertical direction. However, the investigated data all originate from above the reservoir, far above the production zone. It is presumed, that deformation is negligible in this relatively shallow part of the formation, and thus the assumption of geological stability holds. Also, more than 10 random trace pairs (Base and Monitor) have been examined, by matching peaks and troughs down the trace. All matched up well, especially in the deeper, less noisy layers suggesting no deformation. Two of the inspected trace pairs are shown below.

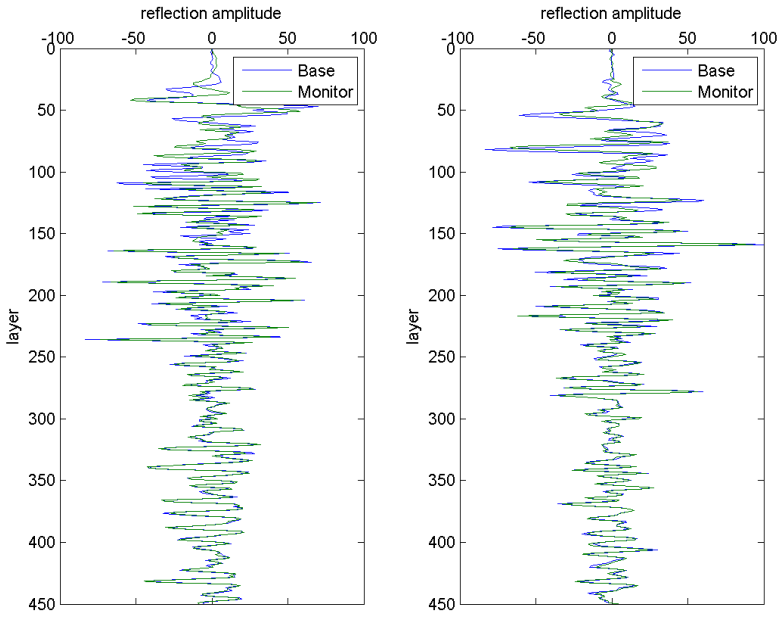


Figure 6.8: Two random trace pairs from the Halfdan data. Peaks and troughs line up well, indicating no deformation in the formation.

6.7 Conclusion

It has been established, that the data are far from perfectly Gaussian. In particular the Full 3-D cube has shown significant deviation from Gaussian properties. Of course, since I aim to remove horizontal noise, the Gaussianity of the Full 3-D cube is not strictly important. I will not attempt to invert the Full cube. It does however reveal the fact, that no one process can be expected to describe all the data. Only regional, horizontal stationarity can be assumed. As a consequence, a new covariance model must be chosen for each inverted section. The size of the inverted region also showed its importance in the former investigation. Local trends are a problem, and it should be considered, if one should sequentially invert smaller regions, while correcting for the trend, or invert a whole layer at a time, assuming zero mean. The Observations especially were shown to be excessively skewed and not zero-mean, while the Residuals exhibited more Gaussian values.

It is concluded, that the Residual (the noise estimator) can be considered first and second order stationary and Gaussian with zero mean, and its covariance matrix be estimated from large or small patches of data. The Observations (i.e. the data set for inversion) are significantly more complicated and may not be Gaussian of nature. K-S tests reject the Gaussian hypothesis, but normal probability plots show less reason to worry. Hence, the Observations are not strictly Gaussian, but retains the likeliness of a normal distribution and likely does not belong to any other analytical distribution. Since the Gaussian assumption must be accepted to proceed with the inversion, and the data are not extremely non-Gaussian (e.g. bimodal), I move on to the real inversion with confidence.

Chapter 7

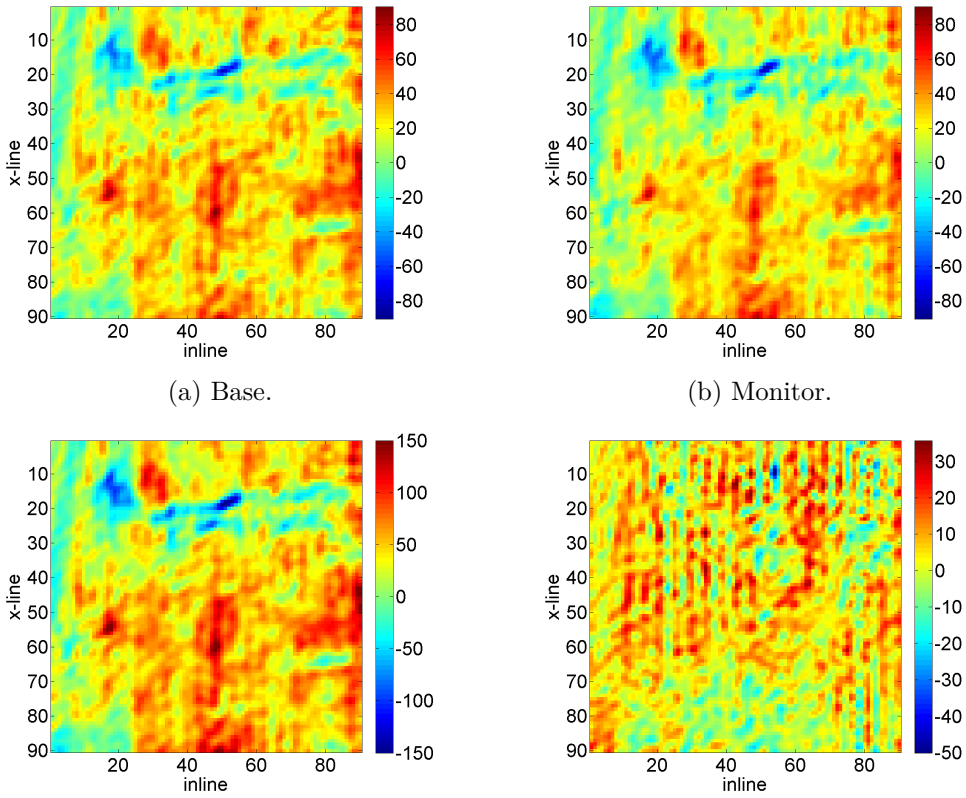
Noise filtering of Halfdan data

With the Halfdan data thoroughly analyzed in the previous chapter, it is time to apply the developed method on real data for the first time. The approach is identical to Chapter 5 where I tested the method on synthetic data. The data analysis in Chapter 6 raised questions about the statistical properties of the data. The most severe was non-stationarity, seen as a change in variance with depth. As the noise filtering will only be horizontal, this will not be an issue, if each layer is modeled and processed independently. The analysis also showed the data should be de-meaned if inverting smaller patches of data. I present the results based on one of the most noisy layers in the shallow part of the data, layer 140. However, the method is applicable to any layer, and in fact all 4-D data sets, provided they comply with the assumptions discussed in the former chapter.

I invert a large data set and a small data set. The small data set is inverted in the spatial domain, restricting the sample size to no more than 90^2 data points due to memory issues. The large data set is inverted in the Fourier domain where the size of a covariance matrix of a N x N matrix is reduced from N^4 to N^2. CPU time is also reduced significantly. However, some approximations are needed to make the method work in the frequency domain. The impact of these approximations have not been investigated. However, I present the results to show that the noise characterization is also valid for larger areas. I begin in the spatial domain with a small subset of layer 140.

7.1 Small subset inversion

As first described in the Methodology section, the two data sets, Base and Monitor, are summed and differenced, creating two new data sets; Observations, d_{obs}, and Residual, d_{res}. The following inversion will aim to remove the noise from d_{obs} by modeling the noise on d_{res}.



(a) Base.

(b) Monitor.

(c) Observations, d_{obs}.

(d) Residual, d_{res}.

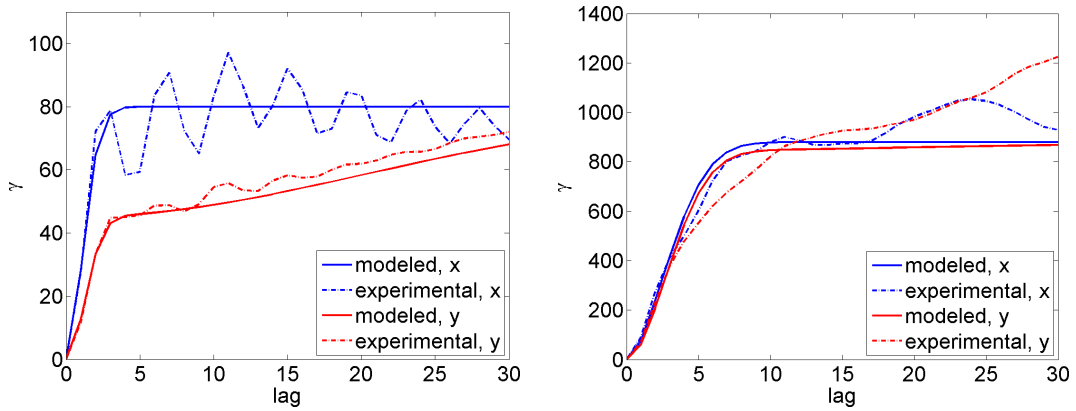
Figure 7.1: Data chosen for inversion.

7.1.1 Modeling the noise

With risk of repeating myself, I model the noise semi-variogram on the experimental semi-variogram of the Residual, exactly as described in the Methodology section and in the synthetic example. Figure 7.2 and 7.3 illustrate the modeling procedure. The resulting noise model and prior geology are shown in Table 7.1.

	Model	Sill	Range	Angle	Aniso. (b_1/a_1)
Prior geology, $\mathbf{C}_m$	Gaussian	800	7	0	1
Noise, $\mathbf{C}_d$	Gaussian	35	50	180	.045
	Gaussian	45	3	0	1

Table 7.1: The geology and noise model estimated from the semi-variogram of the Observations and Residual is tabled.



(a) d_{res} and noise model.

(b) d_{obs} with combined noise and geology model.

Figure 7.2: 1-D experimental semi-variogram of Residual and Observations data with modeled noise and geology.

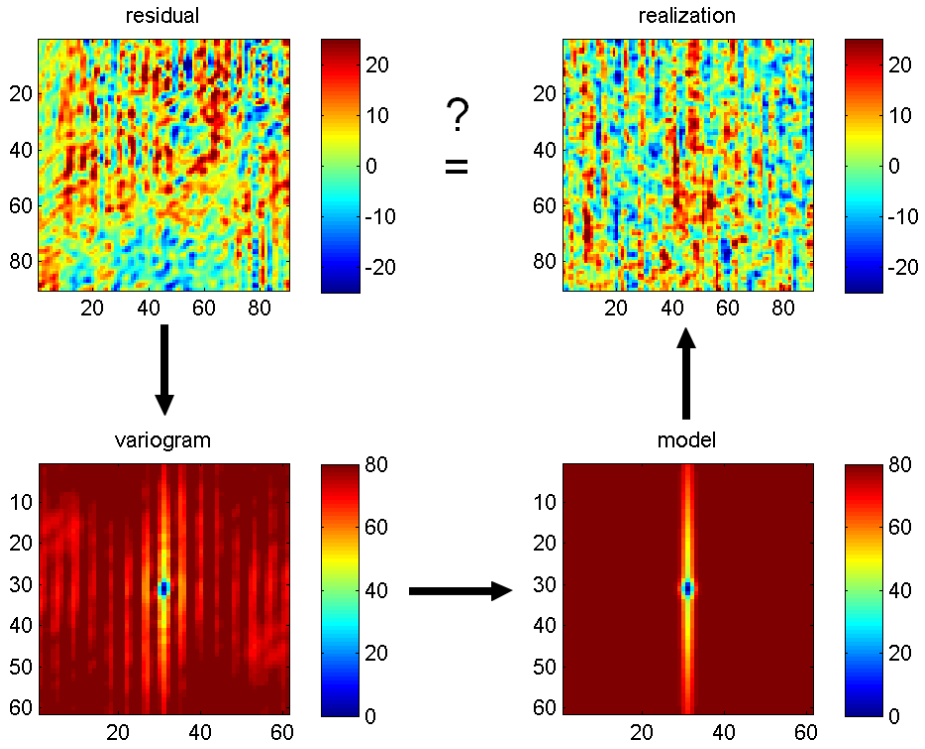
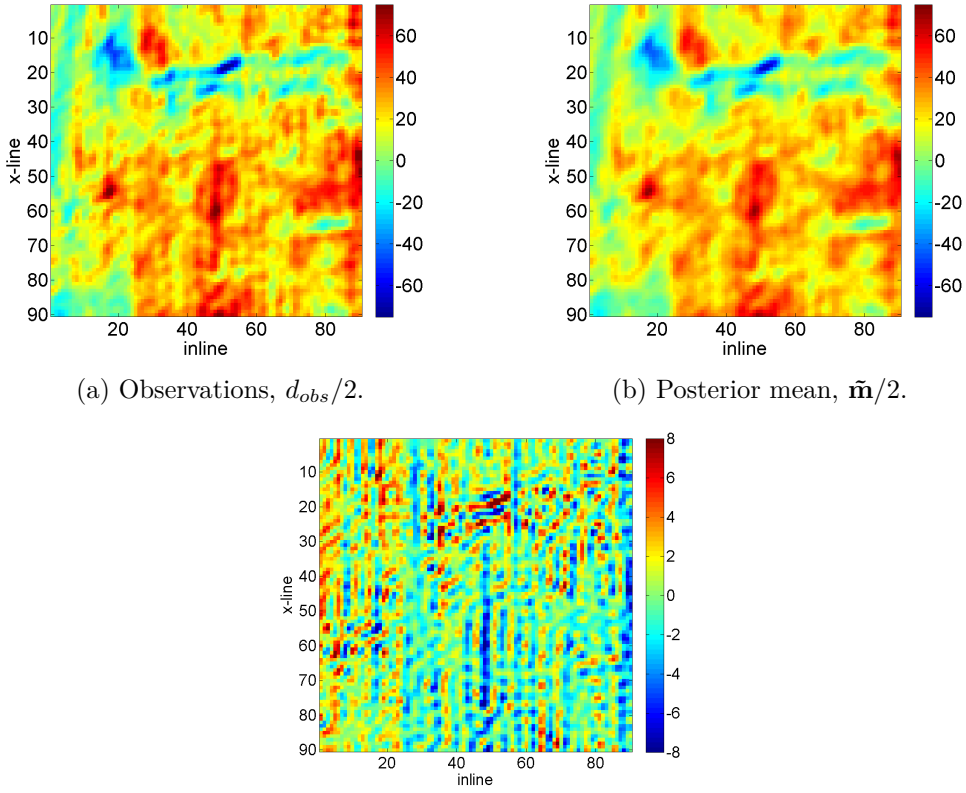


Figure 7.3: Noise semi-variogram modeling. The noise is modeled on the experimental semi-variogram of the Residual d_{res}. Random Gaussian realizations are generated to ensure likeliness to the true residual.

With the best models found, I can construct $\mathbf{C}_m$ and $\mathbf{C}_d$ and perform the least-squares inversion. The Observations, which had a mean of 37.1, are de-meaned.

7.1.2 Posterior mean solution



(a) Observations, $d_{obs}/2$.

(b) Posterior mean, $\tilde{\mathbf{m}}/2$.

(c) Removed noise $d_{obs}/2 - \tilde{\mathbf{m}}/2$.

Figure 7.4: Posterior mean solution to the inverse problem (b) with removed noise (c) and original Observations (a).

Figure 7.4 shows the posterior mean solution in comparison to the original Observations d_{obs} (the original mean has been restored). The inversion has removed the subtle inline striping, without removing large features not characterized by the noise covariance model. The removed noise, Figure 7.4c, convincingly illustrates the removed striping. Some less stripy features were also removed. This may or may not be actual noise. The short range, isotropic part of the noise semi-variogram model allows the removal of small scale features, even though they do not look like striping.

A set of 9 random Gaussian realizations of the posterior is presented in Figure 7.5 and experimental 1-D semi-variograms in Figure 7.6. Variance is obviously increased in the random realizations and some remnant striping is visible. It is likely that some striping will remain, as we did not model the oscillations seen in Figure 7.2a.

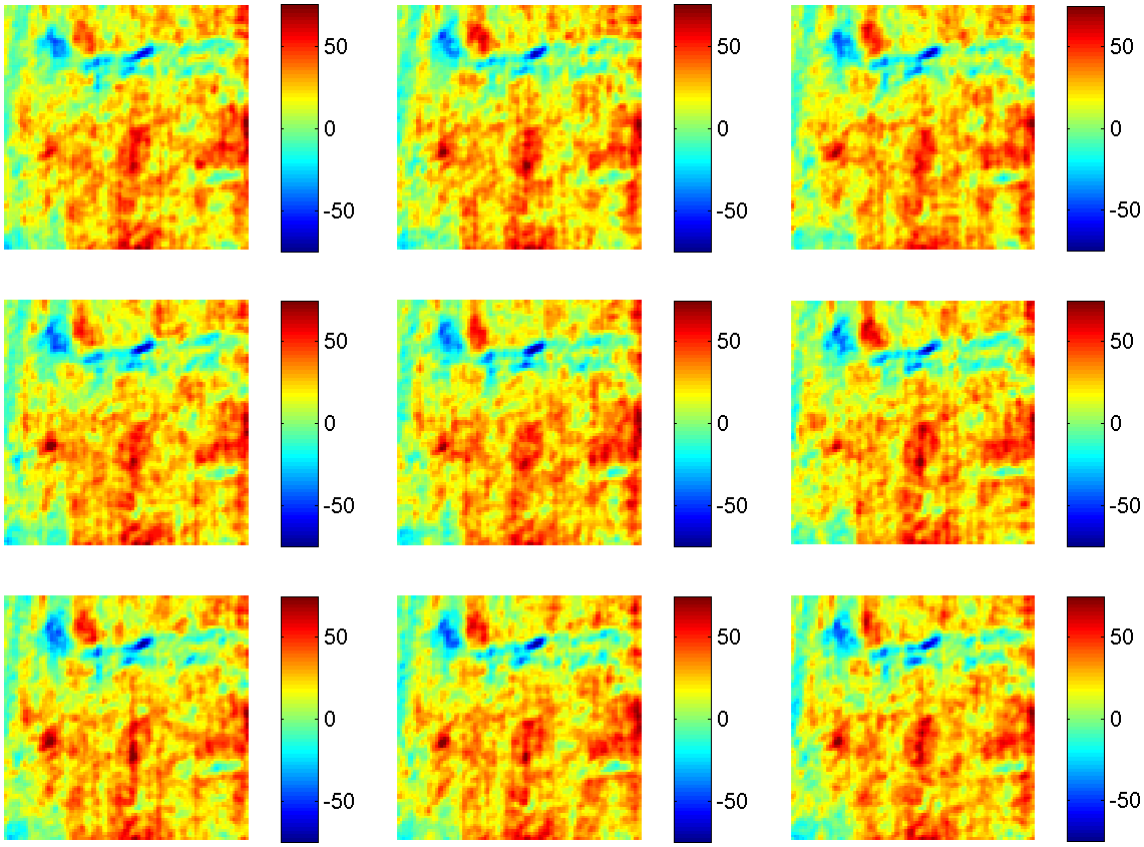
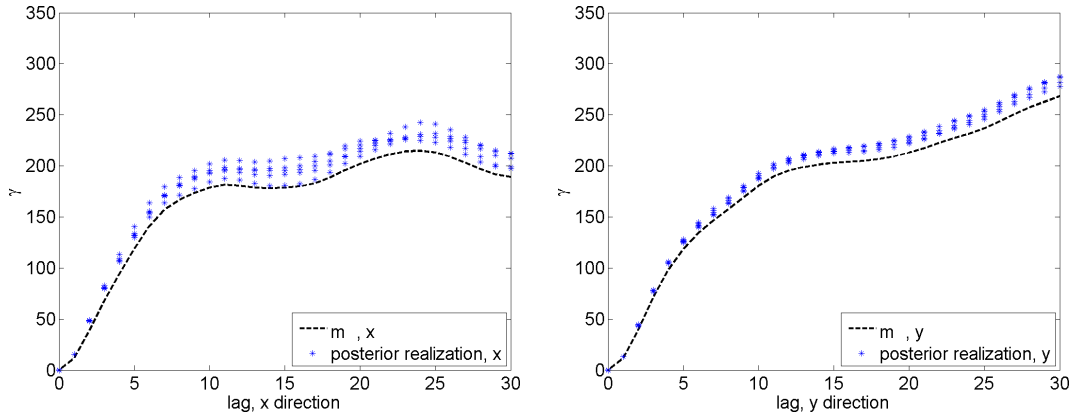


Figure 7.5: Gaussian realizations of the posterior distribution $\mathcal{N}(\tilde{\mathbf{m}}/2, \tilde{\mathbf{C}}_m/4)$.



(a) Minor principal direction (E-W). (b) Major principal direction (N-S).

Figure 7.6: 1-D experimental semi-variograms of the posterior mean, including five realizations.

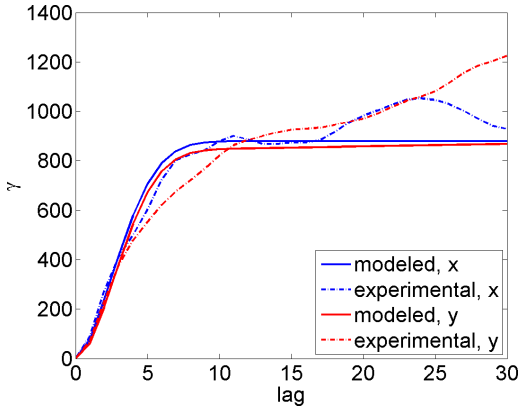
7.1.3 Spherical geology prior

I would like to present an alternate solution than the one just presented, to illustrate how small changes in the prior model can impact the result. I introduce a new geology prior based on a Spherical model with slightly longer range. A comparison between the former and latter models is shown in Figure 7.7. The fit of the Spherical model is worse than the Gaussian at lags 0-3. At longer distances the two models are equally good.

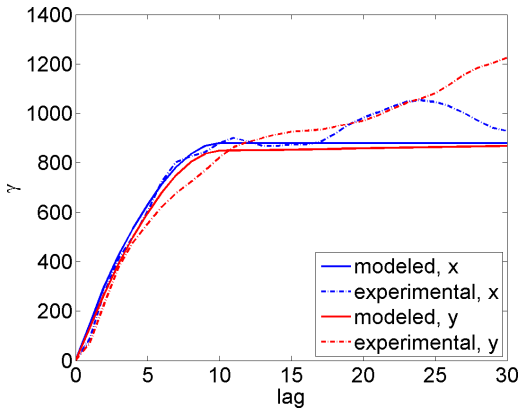
Figure 7.8 compares the solution from the two models: The Spherical model solution $\tilde{\mathbf{m}}^{Sph}$ seen in (b) is less smooth, more irregular on a small scale, than the Gaussian $\tilde{\mathbf{m}}^{Gau}$ seen in (a). The bumpy character of the Spherical solution is better seen from the removed part of the signal in Figure 7.9. The Spherical model has clearly removed some striping. However, small scale observations (which are possibly noise) are not efficiently attenuated.

	Model	Sill	Range	Angle	Aniso. (b_1/a_1)
Prior geology, $\mathbf{C}_m^{Gau}$	Gaussian	800	7	0	1
Alternate prior geology, $\mathbf{C}_m^{Sph}$	Spherical	800	10	0	1
Noise, $\mathbf{C}_d$	Gaussian	35	50	180	.045
	Gaussian	45	3	0	1

Table 7.2: An alternate prior geology model is introduced. The noise model is unchanged.



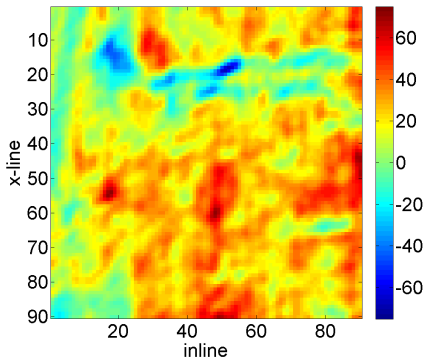
(a) d_{obs} with combined Gaussian geology model and noise.



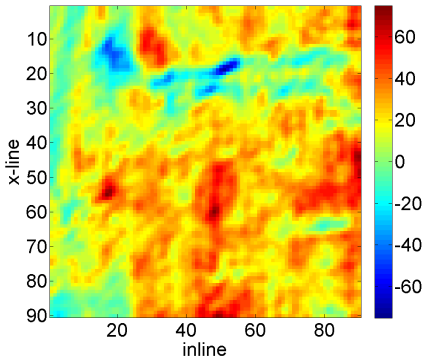
(b) Spherical geology model + noise.

(c) d_{obs} with combined Spherical geology model and noise.

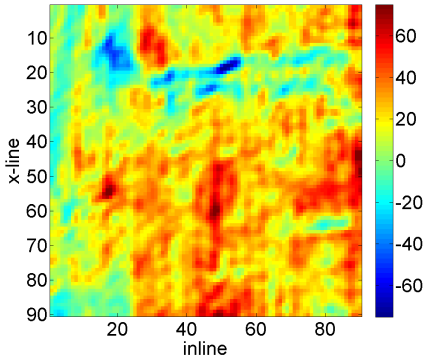
Figure 7.7: 1-D experimental semi-variograms of the two different models.



(a) Mean posterior $\tilde{\mathbf{m}}^{Gau}/2$ with Gaussian geology prior

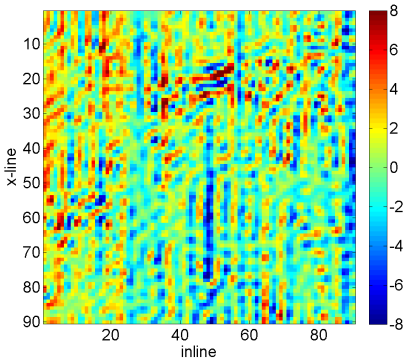


(b) Mean posterior $\tilde{\mathbf{m}}^{Sph}/2$ with Spherical geology prior.

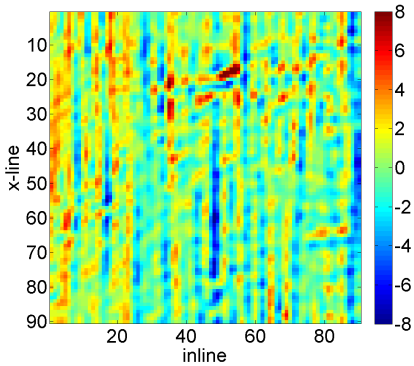


(c) Observations, $d_{obs}/2$.

Figure 7.8: Two mean posterior solutions to the inverse problem compared to the Observations.



(a) Removed noise, $d_{obs}/2 - \tilde{\mathbf{m}}^{Gau}/2$.



(b) Removed noise, $d_{obs}/2 - \tilde{\mathbf{m}}^{Sph}/2$.

Figure 7.9: Removed (estimated) noise.

Gaussian realizations of the posterior from the Spherical model is presented in Figure 7.10 and experimental 1-D semi-variograms in Figure 7.11. Realizations of the posterior based on a Spherical prior model does not yield terrible results. However, one should always choose the covariance model that best describes data.

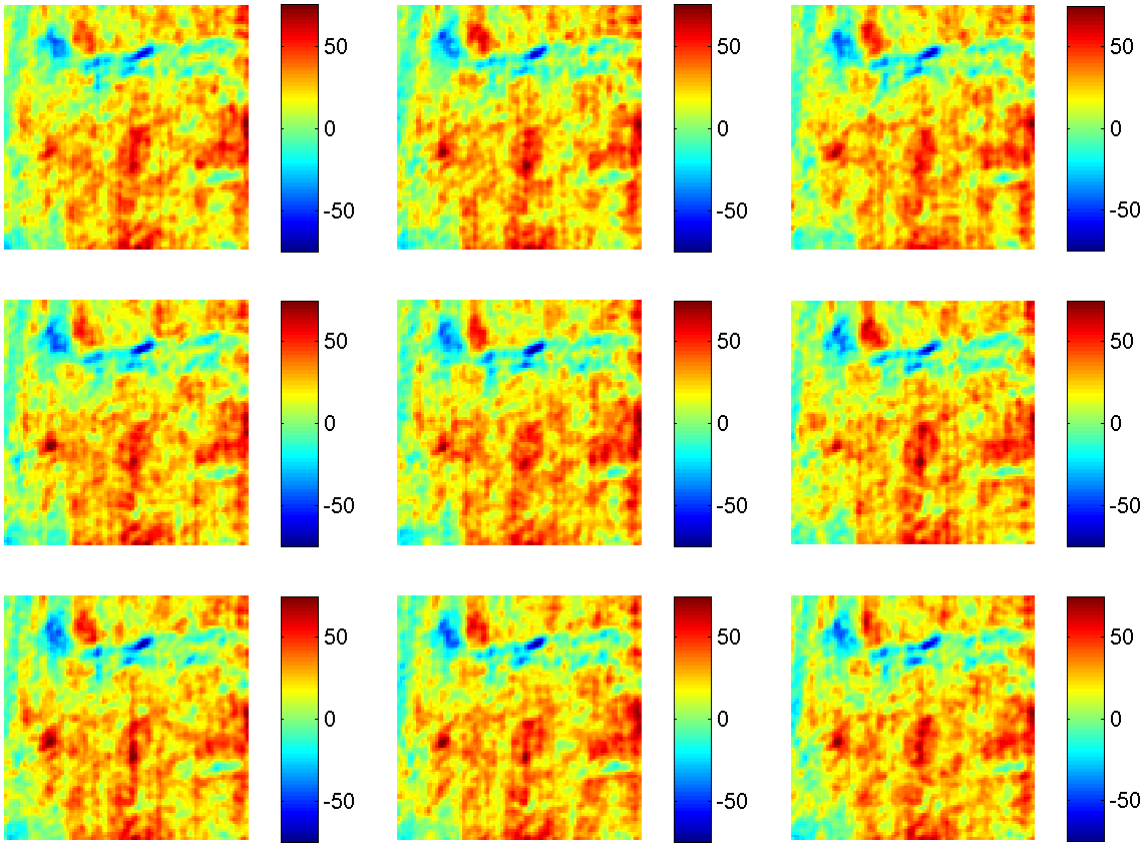
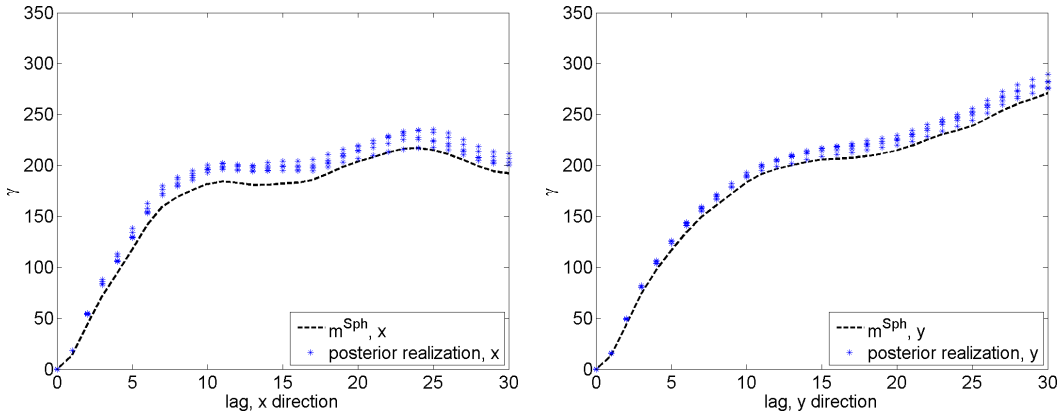


Figure 7.10: Gaussian realizations of the posterior distribution $\mathcal{N}(\tilde{\mathbf{m}}^{Sph}/2, \tilde{\mathbf{C}}_m^{Sph}/4)$ (Spherical geology model).

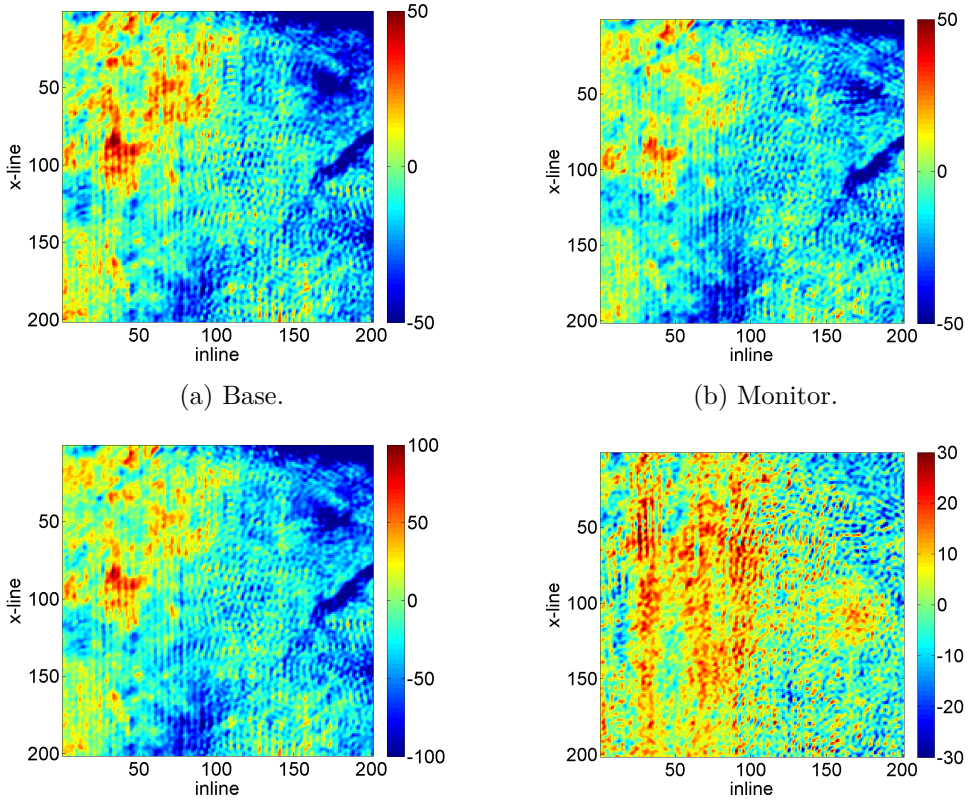


(a) Minor principal direction (E-W). (b) Major principal direction (N-S).

Figure 7.11: 1-D experimental semi-variograms of the posterior mean, including five realizations.

7.2 Larger subset inversion (FFT)

To show that the modeled noise covariance can also be removed on larger data sets, I model the noise on a 200 x 200 patch of data in layer 140. So much information (remember, the covariance matrix is $n_x n_y \cdot n_x n_y = 40,000 \cdot 40,000 = 1,600,000,000$ data points) is impossible to store and invert on normal computers. Hence, we must solve the inverse problem in the Fourier domain, where the covariance matrix is reduced to $n_x \cdot n_y = 200 \cdot 200 = 40,000$), and the inversion is solved as a deconvolution problem. The problem can now be solved in a matter of seconds. There are however issues of both technical and theoretical nature when converting a non-periodic signal to a periodic signal. Issues which are yet to receive proper attention. Those concerns aside, I present data set and the inversion result below. The data are shown in Figure 7.12.



(a) Base.

(b) Monitor.

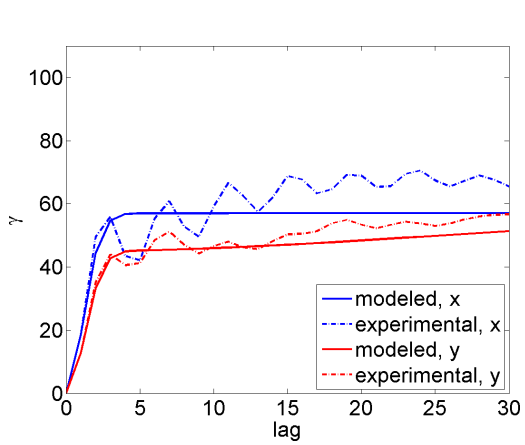
(c) Observations, d_{obs}.

(d) Residual, d_{res}.

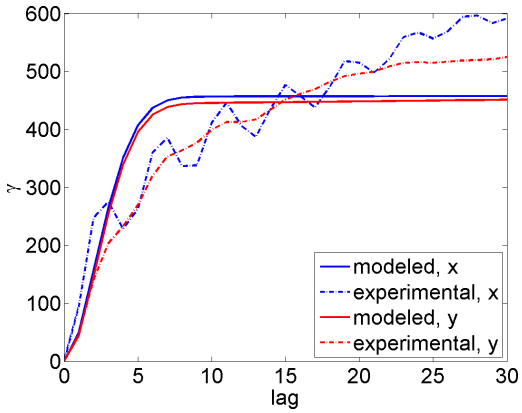
Figure 7.12: Data chosen for inversion.

	Model	Sill	Range	Angle	Aniso. (b_1/a_1)
Prior geology, $\mathbf{C}_m$	Gaussian	400	6	0	1
Noise, $\mathbf{C}_d$	Gaussian	12	60	180	.035
	Gaussian	45	3	0	1

Table 7.3: The noise model estimated from the semi-variogram of the Residual is tabled, along with a geology model based on the Observations.

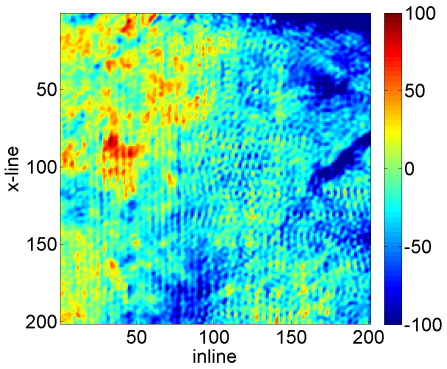


(a) d_{res} and noise model.

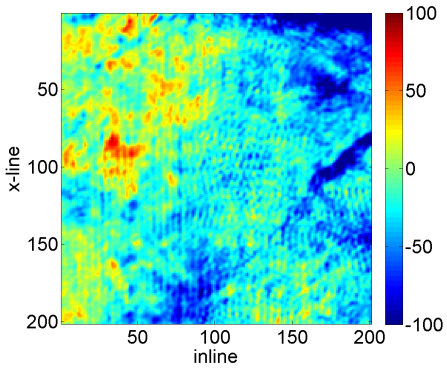


(b) d_{obs} with combined noise and geology model.

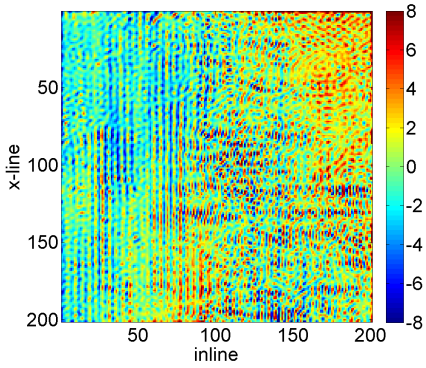
Figure 7.13: 1-D experimental semi-variogram of Residual and Observations data with modeled noise and geology.



(a) Observations, d_{obs}.



(b) Posterior mean, $\tilde{\mathbf{m}}$.



(c) Removed noise, $d_{obs} - \tilde{\mathbf{m}}$.

Figure 7.14: Posterior mean solution to the inverse problem (b) with removed noise (c) and original Observations (a).

The mean posterior solution is shown in Figure 7.14b along with the original Observations. 7.14c shows the difference between the least-squares estimator (LSE) and the Observations. What has been removed from the Observations then, is mainly N-S correlated, while the lower frequency part of the observations perceived as true geology, is

retained. The bad fit in Figure 7.13b is possibly due to some trend in the data.

7.3 Conclusion

The noise characterization technique was applied to real 4-D seismic reflection data from the Halfdan field in the North Sea. The noise was modeled by fitting a positive definite semi-variogram model to the experimental semi-variogram of the Residual data set. A LSQ inversion technique was then used to filter the modeled noise. Several sizes of data sets were filtered using spatial as well as frequency domain techniques. Modeling the noise on the Residual proved easy when the sill and range were well defined, and noise was partly removed. However, some remnant noise was visible in the posterior solution to both the smaller and larger data set. This is possibly due to the inability to model the oscillations in the Residual and unattended trends in the data (non-stationarity).

When applied to the smaller data set, two solutions with different prior models were investigated. The best data-model fit was obtained with Gaussian noise and Gaussian geology. This also yielded the most pleasing inversion results. In the synthetic chapter, the true geological model was known, and hence could be compared to the posterior solution. When working with real data, this is not possible. However, the success of the synthetic example ensues a high degree of trust in the method, and hence also in the real data inversion result.

Chapter 8

Discussion and conclusion

8.1 Method

A method for characterizing and modeling noise in 4-D seismic reflection data has been developed. The purpose of the method is to minimize errors introduced by priors based on intuition. A more objective prior is obtained by ignoring external information such as "expert knowledge". Instead, priors are estimated straight from the data. The noise in particular is estimated from the experimental semi-variogram of the differenced data set (the Residual). The prior information takes the form of Gaussian pdfs and can be used in any probabilistic inversion.

This investigation has focused specifically on characterizing and removing horizontal 2-D spatially correlated acquisition noise in time slices of 4-D data. The method can without difficulty be applied to other types of noise, horizontal or vertical. Or expanded to work in 3 dimensions, provided positive definite 3-D geostatistical models are available, and that assumptions such as stationarity still holds.

A special form of linear least squares inversion was used to demonstrate how well this noise estimation method works. By letting the mapping operator map into itself, rather than between model and data space, the characteristics of the (data derived) noise covariance matrix are filtered from the data.

An important note is, that the noise model found by this method may be used in already existing seismic inversion methods relying on Gaussian noise models, allowing inversion directly on noisy data, without fitting the noise in the solution.

8.2 Results

The method has been tested on synthetic and real 4-D data. The synthetic case showed, that the method, under optimal conditions, can remove a combination of high amplitude correlated and uncorrelated noise (S/N = 2) with only a slight decrease in accuracy. The noise proved easy to model, even with severe noise. All striping was removed. It was also shown that, if badly estimated priors were used, the inversion results deteriorated severely.

When applied to real 4-D data, assumptions such as first and second order stationarity were challenged. The data were also mildly non-Gaussian. However, non of the assumptions were deemed so severe, that the method could not be safely applied. The noise was relatively easy to model with a combination of isotropic and anisotropic Gaussian models,

and the inversion results were pleasing in both smaller and larger data sets. Most N-S correlated noise was removed. However, some noise persisted, possibly due to unmodeled oscillations in the experimental noise covariance. The method proved less successful in estimating the prior model (the noise-free geology) of the real data, possibly due to non-stationarity in the Observations.

8.3 Discussion

It was unexpected to find how relatively well behaved the experimental semi-variogram of the Residual was, and consequently, how easy it was to model the noise. The preliminary data analysis also showed, that the Residual data, were much more Gaussian than the Observations. This implies, that the nature of the noise is indeed a Gaussian process. It was not possible though to model the oscillating hole effect in the noise (also to a lesser extent present in the geology), due to the non-positive definiteness of the hole model. It is believed, that the success of the inversion could improve considerably, if this oscillating behavior was included in the model.

It would also be interesting to consider what might happen if we had some data in a part of the section where there are production-related changes (which are the true targets of 4-D surveys). The work flow would then be to build a noise model using the difference from somewhere with no geological difference (as was done in this thesis), then use this noise model on the difference between Base and Monitor (not the sum), in a region where there are changes to the geology. One would then be estimating the 4-D signal. Knowing how the geology has changed as a consequence of production is very interesting when predicting the elastic properties of the reservoir.

8.4 Areas of interesting future research

As the variance of the data has been shown to vary with depth, and to a lesser extent, with position, different noise models must be estimated for different areas of the data. As of yet, the task of fitting noise and geology models by hand is tedious. If the method is to be implemented in commercial software, some form of automation is needed. To make the method fully autonomous, the noise and geology priors should be modeled automatically. However, it is doubtful if a fully automated approach can ever be trusted. The properties of the noise seem to be well characterized by the developed method, but the properties of the prior geology is not. It is believed, that some subjective information, e.g. a geologist, should be introduced when choosing this prior.

Bibliography

E. Oran Brigham. *The Fast Fourier Transform.* Prentice-Hall, Inc., 1974.

Arild Buland and Henning Omre. Bayesian linearized avo inversion. *Geophysics*, 68(1), January-February 2003.

Clayton V. Deutsch and André G. Journel. *Geostatistical Software Library and User's Guide.* Oxford University Press, 2. edition, 1998.

Adri Duijndam. *Detailed Bayesian Inversion of seismic data.* PhD thesis, Delft, 1987.

Pierre Goovaerts. *Geostatistics for Natural Resources Evaluation.* Oxford University Press, 1997.

Edward H. Isaaks and R. Mohan Srivastava. *An introduction to Applied Geostatistics.* Oxford University Press, 1989.

R. Lynn Kirlin, William J. Done, and Stephen J. Hill. *Covariance analysis for seismic signal processing.* SEG Books, 1999.

D. Marcotte. Fast variogram computation with fft. *Computers & Geosciences*, 22(10): 1175–186, 1996.

William Menke. *Geophysical Data Analysis - Discrete Inverse Theory.* Academic Press, Amsterdam, Boston, rev. edition, 1989. ISBN 0-12-490921-3.

Klaus Mosegaard and Albert Tarantola. Monte carlo sampling of solutions to inverse problems. *Journal of Geophysical Research*, 100(B7):12,431–12,447, 1995.

Klaus Mosegaard and Albert Tarantola. *Probabilistic Approach to Inverse Problems*, pages 237–265. Academic Press, 2002.

Allan A. Nielsen. Geostatistics and analysis of spatial data. 2009.

Allan Aasbjerg Nielsen. *Analysis of Regularly and Irregularly Sampled Spatial, Multivariate, and Multi-temporal Data.* PhD thesis, Technical University of Denmark, 1994. Institute of Mathematical Modelling.

OGP and IAGC. *An overview of marine seismic operations*, April 2011. Report No: 448.

Mickaele Le Ravalec, Benoit Noetinger, and Lin Y. Hu. The fft moving average (fft-ma) generator: An efficient numerical method for generating and conditioning gaussian simulations. *Mathematical Geology*, 32(6), 2000.

John A. Scales, Martin L. SMith, and Sven Treitel. *Introductory Geophysical Inverse Theory.* Samizdat Press, 2001.

Albert Tarantola. *Inver Problem Theory*. SIAM, 2005.

Albert Tarantola. Popper, bayes and the inverse problem. *Nature Physics*, 2:492–494, August 2006. Commentary.

Albert Tarantola and Bernard Valette. Inverse problems = quest for information. *Journal of Geophysical Research*, 50:159–170, 1982.

O. Yilmaz. *Seismic Data Processing*, volume 2 of *Investigations in Geophysics*. Society of Exploration Geophysics, 1987.

Appendix A

Sum of variances

The following proof uses the fact that expected values are additive:
$E(X \pm Y) = E(X) \pm E(Y)$.

Lemma A.1. *Let X be a real valued stochastic variable with second moment. And let $E(X) = \mu$. Then*

$$Var(X) = E(X^2) - \mu^2.$$

Proof.

$$\begin{aligned} Var(X) &= E[(X-\mu)^2] \\ &= E(X^2 - 2X\mu + \mu^2) \\ &= E(X^2) - 2\mu E(X) + E(\mu^2) \\ &= E(X^2) - 2\mu \cdot \mu + \mu^2 \\ &= E(X^2) - \mu^2. \end{aligned}$$

□

Theorem A.2. *If X and Y are independent, real valued random variables with second moments, then*

$$Var(X \pm Y) = Var(X) + Var(Y).$$

Proof.

$$\begin{aligned} Var(X \pm Y) &= E[(X \pm Y)^2] - \mu_{X\pm Y}^2 \\ &= E(X^2 \pm 2XY + Y^2) - (\mu_X \pm \mu_Y)^2 \\ &= E(X^2) \pm 2E(XY) + E(Y^2) - (\mu_X^2 \pm 2\mu_X\mu_Y + \mu_Y^2) \\ &= E(X^2) \pm 2E(XY) + E(Y^2) - \mu_X^2 \mp 2\mu_X\mu_Y - \mu_Y^2 \\ &= E(X^2) - \mu_X^2 \pm 2[E(XY) - \mu_X\mu_Y] + E(Y^2) - \mu_Y^2 \end{aligned}$$

Rewriting the terms in the square brackets, and using that the covariance of two independent stochastic variables is zero, we see that the two terms cancel:

$$E(XY) - \mu_X\mu_Y = E(XY) - E(X)E(Y) = Cov(X,Y) = 0.$$

It then follows, that

$$\begin{aligned} Var(X \pm Y) &= E(X^2) - \mu_X^2 + E(Y^2) - \mu_Y^2 \\ &= Var(X) + Var(Y). \end{aligned}$$

□

Appendix B

Bayes' theorem

The simplicity of Bayes' theorem, and the fact that it allows the inclusion of prior probability, has made it popular in many different areas where inversion techniques are applied. Why? An inverse problem is, in a nutshell, nothing more than solving a set of equations. However, all real problems suffer from under- or overdeterminacy, multidimensionality and uncertainty in the form of noise. Bayes' theorem offers a way to make the problem well-posed by combining a priori knowledge with information obtained from observations. Building on previous definitions, we assume that the data space is linear (and thus $\mu_d(\mathbf{d}) = const.$) and that modeling uncertainties are negligible (compared to observations). The relationship between model and data can then be written

$$\theta(\mathbf{d}|\mathbf{m}) = \delta(\mathbf{d} - g(\mathbf{m})), \tag{B.1}$$

where δ is the delta function. Hence, we are absolutely certain, that our forward model g is correct, and we have an exact relation between model parameters and data:

$$\mathbf{d} = g(\mathbf{m}). \tag{B.2}$$

The likelihood function then simplifies to

$$L(\mathbf{m}) = \rho_d(g(\mathbf{m})), \tag{B.3}$$

and equation (3.2) becomes

$$\sigma_m(\mathbf{m}) = k\rho_m(\mathbf{m})\rho_d(g(\mathbf{m})). \tag{B.4}$$

By identifying $f(\mathbf{m}|\mathbf{d}_{obs}) \equiv \sigma_m(\mathbf{m})$, $f(\mathbf{d}_{obs}|\mathbf{m}) \equiv \rho_d(g(\mathbf{m}))$, $f(\mathbf{m}) \equiv \rho(\mathbf{m})$, and $f(\mathbf{d}_{obs}) \equiv 1/k$, we obtain the well known *Bayes theorem*:

$$f(\mathbf{m}|\mathbf{d}_{obs}) = \frac{f(\mathbf{m})f(\mathbf{d}_{obs}|\mathbf{m})}{f(\mathbf{d}_{obs})}. \tag{B.5}$$

Bayes theorem is thus a special case of Tarantola and Valette's more general theory relating posterior probability with prior probability scaled by the likelihood function and the constant vector $f(\mathbf{d}_{obs})$. Bayes theorem can also be derived with the assumption that observational uncertainty are negligible (compared to modeling uncertainties), or with a combination of both, as long as they are independent.

Appendix C

Learning objectives

To ensure that my thesis complies with the DTU requirements to professionalism and documentation, I have formulated a set of learning objectives, including goals for the process, result, presentation and dissemination.

During the course of this project, I will

- Describe noise in relation to signal-to-noise optimization.
- Describe the noise removal technique.
- Describe and characterize the data made available by Maersk Oil (contact: Adam Cherret).
- **Main objective: Design a specific noise characterization and modeling method which can be applied to the before mentioned data.**
- Discuss the choice of method and theory.
- Apply the method to the data.
- Explore possible expansions to the method, e.g. automation.
- Have a critical theoretical and empirical discussion and conclusion on the results.
- Assess whether the method can be realized in commercial practice.
- Write the assignment in a structured, adequate, concise and clear manner, and moreover in compliance with good practice in written geophysics so the method can be reproduced by others.
- Outline my work in a weekly journal.
- Present and adjust the thesis in accordance with the achieved results.

Appendix D

MATLAB scripts

The code responsible for generating the synthetic data, calculating the semi-variograms, running the inversion and so on, accumulate to several thousand lines of code. I find it disrespectful to the environment to include the code in any work that the reader might want to have a printed copy of. Hence, I have uploaded all my scripts to my homepage `http://clundmand.wix.com/christianljensen#!projects/ctzx` for public download.

Appendix E

Journal

Week 1, 1^{st}-3^{rd} of March
Had my first meeting with Adam Cherrett from Maersk Oil. He provided the data in the form of two surveys geographically identical but temporally spaced, and told me about the acquisition method and what type of noise I should expect and look for. The trick is to subtract the two surveys and look at the residual. This is assumed to be a good representation of the noise. Each data block is 800 by 1750 by 450 giving a total of 630,000,000 data points per survey. The vast size makes it impossible to read it all at once in MATLAB and must be extracted trace by trace. Thomas provided me with a reader that subsamples the data set.

I spend most of my time looking through the data at various angles and researching existing noise characterization techniques. I found it difficult to find books or papers about structured noise in seismics.

Week 2, 4^{th}-10^{th} of March
Met up with Klaus this week to discuss basic ideas for the project: What is noise really? What type of noise id present in my data? How should I remove the noise; by overfitting perhaps? I also met with Thomas. He liked the idea about overfitting. We also talked about characterizing different types of noise independently, i.e. the striping noise as one type, the overfitting residual noise as another, and more.
Looked into Gaussian probability theory, spectral subtraction and f-k filtering as well as general FFT based filtering methods. Extracted a 99 by 99 by 450 data cube from each data set. Inspected the spatial and temporal cross correlation of the data by FFT (wrote code myself). Noticed noisyness/striping and variance decreases with depth. Made a 3D tapering function so reduce ringyness when using FFT.

Week 3, 11^{th}-17^{th} of March
Read paper about automatic factorial co-kriging (AFACK), by T. Coleou. They reduce striping from acquisition imprint by covariance modeling. Made a random function that produces 3D matrices with random entries. Read about the subtle differences between correlograms and variograms, and came up with an idea to use PCA to make better semi variograms. The highest principal components (PCs) contain the parts of the image with highest variance (through the 450 time steps). By making a semi variogram from the highest PC it will be easier to fit a model to it since it contains only variance in the principal direction (I think). I can calculate more semi variograms, fit models, and then combine them into a more advanced model if needed. Now I only need to figure out how to make the 2D semi variograms. Wrote my first entries in LaTeX about my progress so far. Read about Maximum Autocorrelation Factor, or MAF analysis. This supposedly is better than PCA. Started reading Seismic Data Processing by Yilmaz [1987]. At meeting

with Klaus we agree that I should characterize the noise by its mean and covariance. This stochastic description will enable me to use Gaussian statistics. I read the first 100 pages of Covariance Analysis for Seismic Signal Processing by Kirlin et al. [1999].

Week 4, 18^{th}-24^{th} of March

Met with Professor Allan Nielsen to discuss PCA, MAF, MAD, MNF and other data analysis methods. I also met with Thomas and got an action plan down: Model variogram and see if it can reproduce your data. Started reading An Introduction to Applied Geostatistics, by Isaaks and Srivastava [1989] and calculated basic summary statistics and q-q plots etc. for Base, Monitor and residual. Realized that the data have a bias. They do not have zero mean, and only fairly fits a Gaussian model. Produced synthesized semi-variograms and calculated realizations hereof, i.e. simulated data from variogram models. Tried detecting striping by correlation coefficient in a moving window. It didn't work, but the idea should be developed.

Week 5, 25^{th}-31^{st} of March

2^{nd} meeting with Adam. Looked at various variograms and histograms. Agreed I should make a 2D variogram simulator so I could test how different size data sets, tapering and detrending affect the variogram. Got 1D synthetic semi-variogram simulator code from Thomas. Realized most graduate programs had reached their application deadlines. Updated CV and wrote motivational letter. Contacted Maersk and Dong Energy. Prepared variogram simulation document in LaTeX. Investigated variogram variance development as function of number of data points - variance decreases proportional to data square side length. Investigated edge effects in Fourier domain - striping in real data thought to be due to edge effects is not evident in simulated data, hence the edge effects are real. Went to the DSE carrier expo. Fell ill at the end of the week.

Week 6, 1^{st}-7^{th} of April

Still feeling ill. Attended meeting for young SPE members. Applied for three positions at Schlumberger. Produced 31 complete horizontal slabs of the data between layer 50 and layer 200. Learned I must taper and zeropad (in that order) before using FFT. Found that the variogram sill and nugget (exhaustive data set) decreases with depth. Range decreases with depth. Tried out new approach: Maximum Noise Fraction, or MNF. Didn't learn much - needs several bands of data. I tried substituting the frequency bands with layers from the residual survey, but I wasn't able to separate the noise. Also tried calculating semi-variogram of five high variance PCA components of residual. Maybe this can be used to automate the variogram modeling. Read five chapters about Fast Fourier Transform in Brigham [1974]. Implemented and tested 2D exponential variogram model. Modified synthetic semi-variogram simulator code from Thomas to handle 2D. Added cosine function to mattaper.m. Designed flowcharts. Begun writing sections "The data" and "Noise characterization". Updated profile on Jobindex.dk. Applied for job at MAN Turbo & Diesel.

Week 7, 8^{th}-14^{th} of April

Wrote of few more lines in "The data". Assessed Gaussian properties of the data. Both Lilliefors and Kolmogorov-Smirnov tests rejected the hypothesis of Gaussianity in all the data as well as for individual layers. The test statistics were very bad! If tested on smaller parts of the data (99x99) several layers are accepted as Gaussian with 5 per cent significance interval... -> Nonstationarity? Started working on demeaning in windows of data to determined if some parts of the data are more Gaussian than others (stationarity). Finally started working a way to remove the noise (least squares inversion). Got the first working synthetic noise removal method to work!! :D Started working on noise removal on real data. Was successful in removing striped noise by modeling the covariance for geology and

noise. Attended carrier seminar. Attended personal carrier counseling meeting. Worked on poster (just for fun).

Week 8, 15^{th}-21^{st} of April

Rewrote my CV. Went to a doctor's appointment. Investigated if the same noise and geology model could remove noise other data than they were modeled on. This worked very well, although it got less convincing in the deep layers. Removing noise in this not-very-noisy environment was almost futile. It only decreased the contrast. Weekly seminar. Job seeking seminar. Found that variance amplitude was a plausible measure of noise filtering effectiveness. My noise filter reduces variance by 25-50 per cent for all layers. Even those temporally very far from the model. Started subdividing a layer into smaller squares to produce a complete noise reduced image. Completed code for removing noise in windows. PROBLEM: Each window now has different variance and mean, meaning the complete image looks like a quilted rug. Effect decreases with increasing window size. IDEA: Let windows overlap, then remove overlap at the end.

Week 9, 22^{nd}-28^{th} of April

Wrote an overview of stuff I had so far, and what I wanted to do, that I wanted to put in the report. Also the order of things. Shifted things around in my report and added many lines several places. Took a closer look at how the LSQ estimator distribution looked compared to the true distribution and how it depended on number of data points. Weekly seminar. Made 2 multi-distribution cumulative histograms instead of 6 single-distribution regular histograms. Meeting with Thomas: Results look great! Size of covariance matrix an issue. Can be circumvented by calculating cov.matrix from larger area than output model, but still not bigger than 50x50. That way boundaries will be much less prominent. Problems are then to verify that the small area used can truthfully represent the covariance of the observed data. Also, many more (smaller) covariance matrices has to be calculated due to overlap. We also talked about calculating the likelihood of my solution and compare it to the likelihood of the true solution (in the synthetic case), and also calculate bad solutions (from bad priors) to verify using the prior I have modeled. Meeting with Adam: He didn't like that I remove noise on single surveys with noise estimation taken from residual survey (both surveys). He suggested I take my synthetic example further, by allowing for different noise models in different locations: Calculate Gaussian probability that two blocks come from the same variogram model. Looked up proof for sum of independent Gaussians.

Week 10, 29^{th}-5^{th} of April/May

Felt sick the first few days. Implemented double-signal idea in synthetics. Worked fine. Implemented double-signal idea in real data. No significant change in noise reduction. Read paper for weekly seminar (Bosch et al. 2010). Found references to papers I need to read. Had quick talk with Thomas about meeting Adam. Weekly seminar. Started writing an application for a job at Qeye Labs.

Week 11, 6^{th}-12^{th} of May

Finished and sent the job application to Qeye Labs. Got a reply from Qeye Labs the day after. I was invited for a quick interview (screening) on Monday next week. Used most of the week preparing for the interview.

Week 12, 13^{th}-19^{th} of May

Job interview on Monday. Studied the properties of circulant and symmetric Toeplitz matrices when used with FFT. Made synthetic data set. Wrote sequential lsq filtering with overlap. Studied effects of wrong geology prior and noise prior. Added virtual box to geology to investigate how well it got resolved.

Week 13, 20^{th}-26^{th} of May

Worked on making the inversion faster by turning the inversion into a convolution in the

Fourier domain. Didn't manage to make it work in 2D. Wrote three pages about the inverse problem, the approach of Tarantola and Valette, Bayes' theorem, and LSQ inversion. Applied for job at DONG Energy.

Week 14, 27^{th}-2^{nd} of May/June

Worked yet again on making the FFT approach (inversion) work, without success. Computer broke down! It took two days to get it up to speed again. Wrote some more on the inverse theory part. Switched around the order of some chapters.

Week 15, 3^{rd}-9^{th} of June

Status meeting with Thomas. Walked through the essential parts of the project. Thomas was pleased with my progress. I should quantify my results with a 1D cloud variogram. In the synthetic case I know how the variogram should look like, hence realizations of my posterior solution should lie on top of the true model. Demonstrate with good and bad priors. Met with Adam and Hanno from Maersk to sort out the FFT inversion/deconvolution approach. Sorted out some issues, but still not entirely sure. Also because the FFT introduces more assumptions and "hacks" to the covariance matrix.

Week 16, 10^{th}-16^{th} of June

Generated and compared solutions based on good and bad priors via 1D semi-variograms. This 1D comparison of solutions is supported by realizations of the different solutions. I am planning to quantify the "goodness" by calculating the probability that posterior covariance "A" is equal to posterior covariance "B". I also had a meeting with Klaus Wednesday. I presented my work so far, and he suggested I should start writing the thesis, i.e. not attempt any new stuff. If I finish up before time, I could always pursue one of many minor issues. Worked on presentation for QI job interview.

Week 17, 17^{th}-23^{rd} of June

Job interview with QI Labs on Monday. Finished first draft of the Methodology section. Started working on writing the Synthetic example section.

Week 18, 24^{th}-30^{th} of June

Finished first draft of Synthetic example section. Started Preliminary studies of the data section.

Week 19, 1^{st}-7^{th} of July

Working on Preliminary data analysis. The data (Observations) are not really as Gaussian as I'd hoped. But the Residual seems fairly Gaussian, at least. Finished data analysis. Proceeding with noise filtering on real data.

Week 20, 8^{th}-14^{th} of July

Finished a rough version of the real noise filtering section. The method seems to work extraordinarily well. I read the whole manuscript through and sent the draft to Thomas for a look-through. This can be considered a pre-alpha version of the final product. Got feedback from Thomas and edited a lot.

Week 20, 15^{th}-21^{st} of July

Made a complete rewrite of the variogram section. Now with more theory. I also wrote the abstract. A red thread is forming in the report. I finally seem to be getting things together. However, I continue to find new information on stuff I'd like to do or had liked to do differently. I must be strong and feel comfortable with what I have accomplished.

Week 20, 22^{nd}-28^{th} of July

All chapters are more or less finished. Wrote discussion and conclusion. Sent it off for review. Took Tuesday off after working for 14 days straight. I clear up some possible misunderstandings unavoidable when mixing inverse problem theory and geostatistics. Run the simulations again to get the right color scale.

Week 20, 29th of July - 1st of August

I make a last effort to express myself as clear and concise as possible. It is taking longer than expected, and I get new ideas every day. The 1st of August has arrived. 5 long months have passed. But I am done and pleased with the result. Now , I just need to worry about the defense.