

# IDEA League

MASTER OF SCIENCE IN APPLIED GEOPHYSICS  
RESEARCH THESIS

---

**Test**  
Subtitle

**Fabian Antonio Stamm**

---

September 28, 2017



# Test

## Subtitle

MASTER OF SCIENCE THESIS

for the degree of Master of Science in Applied Geophysics at  
Delft University of Technology  
ETH Zürich  
RWTH Aachen University  
by

Fabian Antonio Stamm

September 28, 2017

Department of Geoscience & Engineering	·	Delft University of Technology
Department of Earth Sciences	·	ETH Zürich
Faculty of Georesources and Material Engineering	·	RWTH Aachen University



**Delft University of Technology**

Copyright © 2013 by IDEA League Joint Master's in Applied Geophysics:

Delft University of Technology, ETH Zürich, RWTH Aachen University

All rights reserved.

No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying or by any information storage and retrieval system, without permission from this publisher.

Printed in The Netherlands, Switzerland, Germany

IDEA LEAGUE  
JOINT MASTER'S IN APPLIED GEOPHYSICS

Delft University of Technology, The Netherlands  
ETH Zürich, Switzerland  
RWTH Aachen, Germany

Dated: *September 28, 2017*

Supervisor(s):

---

Prof. Florian Wellmann, Ph.D.

---

Prof. Dr. Janos Urai

Committee Members:

---

Prof. Florian Wellmann, Ph.D.

---

Prof. Dr. Janos Urai

---

Miguel de la Varga, M.Sc.



---

# Abstract

Please pay particular attention to the preparation of your abstract; use this text as a guide. Every master thesis report must be accompanied by an informative abstract of no more than one paragraph (max 300 words). The abstract should be self-contained. No references, figures, tables, or equations are allowed in an abstract. Do not use new terminology in an abstract unless it is defined or is well-known from the literature. The abstract must not simply list the topics covered in the paper but should (1) state the scope and principal objectives of the research, (2) describe the methods used, (3) summarize the results, and (4) state the principal conclusions. Do not refer to the master thesis report itself in the abstract. For example, do not say, "In this thesis we will discuss". Furthermore the abstract must stand alone as a very short version of the master thesis report rather than as a description of the contents. Remember that the abstract will be the first and most widely read portion of the master thesis report. Readers will be influenced by the abstract to the point that they decide to read the master thesis report or not.





---

# Acknowledgements

First of all I want to thank all the people who have participated in this project .. Remember, often more people have (in some way) contributed to your final thesis than you would initially think of....

RWTH Aachen University  
September 28, 2017

Fabian Antonio Stamm



---

# Table of Contents

<b>Abstract</b>	<b>v</b>
<b>Acknowledgements</b>	<b>vii</b>
<b>Nomenclature</b>	<b>xv</b>
<b>Acronyms</b>	<b>xv</b>
<b>1 Methods</b>	<b>1</b>
1-1 Bayesian analysis and decision theory . . . . .	1
1-1-1 Basic elements . . . . .	1
1-1-2 Bayesian inference . . . . .	2
1-1-3 Estimation . . . . .	2
1-1-4 Expected loss and loss functions . . . . .	3
1-2 Application in structural geological modeling . . . . .	4
1-2-1 Markov chain Monte Carlo sampling (MCMC) . . . . .	6
1-2-2 Structural geological forward modeling . . . . .	7
1-2-3 Economic significance and reservoir valuation . . . . .	8
Original oil-in-place and recoverable volumes as value measures . . . . .	9
Net present value . . . . .	10
Abstract valuation using scores in the one-dimensional model . . . . .	10
1-2-4 Numerical computational implementation via Python, GemPy and PyMC . . . . .	10
1-3 Designing a case-specific loss function . . . . .	11
1-3-1 Including different risk-affinities in the loss function . . . . .	15
1-4 Constructing a 1D geological reservoir model . . . . .	16
1-5 Constructing a full 3D geological reservoir model . . . . .	16

<b>2</b>	<b>Results</b>	<b>17</b>
2-1	Results using the 1D geological model . . . . .	17
2-2	Results using the 3D geological model . . . . .	17
<b>3</b>	<b>Discussion</b>	<b>19</b>
	<b>Bibliography</b>	<b>21</b>
<b>A</b>	<b>The back of the thesis</b>	<b>23</b>
A-1	An appendix section . . . . .	23
A-1-1	An appendix subsection with C++ Lisitng . . . . .	23
A-1-2	A MATLAB Listing . . . . .	23
<b>B</b>	<b>Yet another appendix</b>	<b>25</b>
B-1	Another test section . . . . .	25

---

## List of Figures

1-1	woo a potential field! from Calcagno et al. (2008) . . . . .	8
1-2	The single steps of customizing the loss function are depicted in plots I - IV. . .	13
1-3	Loss based on the customized loss function (IV) for determined true scores of -750, 0 and 750. . . . .	14
1-4	Plotting of expected loss realizations after including the risk factor $r$ in the loss function (IV) for actors with risk-affinities ranging from risk-averse ( $r = 0.5$ and $0.75$ ), over risk-neutral ( $r = 1$ ), to risk-friendly ( $r = 1.25$ and $r = 1.5$ ). . . . .	16



---

## List of Tables





---

# Acronyms

**RWTH** Aachen University



---

# Chapter 1

---

## Methods

The methods utilized in this work are presented in the following chapter. Bayesian analysis and decision theory are introduced. Focus is laid on Bayesian inference, estimation of uncertain values and the use of loss functions in this context. These methods are incorporated in computational modeling of structural geological settings by programming in a python 3 environment. Central tools for model construction and conduction of the statistical evaluation are GeMpy and PyMC 2 in particular. These are also presented in this chapter.

### 1-1 Bayesian analysis and decision theory

As implied by the name, the problems and reasoning behind decision-making are examined in the field of decision theory ([Berger, 2013](#)). Such decision problems are commonly influenced by parameters that are uncertain. In statistical decision theory, available statistical knowledge is used to gain information on the nature of these uncertainties. (Such uncertain parameters can be considered as numerical quantities.) In order to find the best decision to a problem, it is possible to combine sample information with other aspects such as the possible consequences of decision-making and the availability of prior information on our uncertainties. Decision consequences are expressed as gains in economic decision theory and as losses, which equal negative gains, in statistics. Prior information might be given for example based on experience from previous similar problems or from expert knowledge (see Batvold and Begg). The approach of utilizing priors is known as Bayesian analysis, which is explained in the following ([Berger, 2013](#)). (It goes well with decision theory.)

#### 1-1-1 Basic elements

First, some basic elements are to be defined. The unknown (uncertain) quantity influencing decision-making is usually denominated as the state of nature  $\theta$  ([Berger, 2013](#)). Given statistical information on  $\theta$  in the form of probability distributions,  $\theta$  is called the parameter.

Decisions are also referred to as actions  $a$ . The outcome of statistical tests in form of information or statistical evidence is denoted as  $X$ . Loss is defined as  $L(\theta, a)$ , so  $L(\theta_1, a_1)$  is the actual loss incurred when action  $a_1$  is taken while the true state of nature is  $\theta_1$  (Berger, 2013). Loss, expected loss and loss functions are explained in detail further below.

### 1-1-2 Bayesian inference

Bayesian inference is most importantly characterized by its preservation of uncertainty, in contrast to standard statistical inference (Davidson-Pilon, 2015). Probability is seen as a measure of belief for an event to occur. It has been argued by Davidson-Pilon (2015) that this Bayesian approach is intuitive and inherent in the natural human perspective. These beliefs can be assigned to individuals (Davidson-Pilon, 2015). Thus, different and even contradicting beliefs about the probability of an event might be held by different individuals, based on variations and disparities in the information available to each one individual (Davidson-Pilon, 2015).

The initial belief or guess about an event  $A$  can be denoted as  $P(A)$  (Davidson-Pilon, 2015). This is used as the so-called prior probability on which Bayesian updating is based. The beliefs about the occurrence of an event are revalued in the presence of additional information, i.e. the observation of new evidence  $X$ . These observations are included as likelihoods  $P(X|A)$ . This process of updating results in a posterior probability  $P(A|X)$  (Davidson-Pilon, 2015). It is important to note that the prior is not simply discarded but re-weighted by Bayesian updating. It was also pointed out by Davidson-Pilon (2015) that by utilizing an uncertain prior, the potential for wrongfulness of the initial guess is already included. This means that Bayesian updating is about reducing uncertainty in a belief and reaching a guess that is less wrong (Davidson-Pilon, 2015). Bayesian updating is defined by and conducted via the following equation, called the Bayes' Theorem:

$$P(A|X) = \frac{P(X|A)P(A)}{P(X)} \propto P(X|A)P(A) \quad (1-1)$$

-Law of Large Numbers!

### 1-1-3 Estimation

The resulting posterior distribution can be used to acquire point estimates for the true state of nature  $\theta$ . Common and simple examples for such estimators are the mode (i.e. the generalized maximum likelihood estimate), the mean and the median of a distribution (Berger, 2013). The presentation of a point estimate should usually come with a measure for its estimation error. According to Berger (2013), the posterior variance is most commonly used as an indication for estimate accuracy. However, it is argued by Davidson-Pilon (2015) that by using pure accuracy metrics, while this technique is objective, it ignores the original intention of conducting the statistical inference in cases, in which payoffs of decisions are valued more than their accuracies. A more appropriate approach can be seen in the introduction of loss and the use of loss functions (Davidson-Pilon, 2015).

### 1-1-4 Expected loss and loss functions

Loss is a statistical measure of how bad an estimate is, i.e. how much is lost by making a certain decision. Gains are considered by statisticians as negative losses (Davidson-Pilon, 2015). The magnitude of an estimate's loss is defined by a loss function, which is a function of the estimate of the parameter and the true value of the parameter (Davidson-Pilon, 2015):

$$L(\theta, \hat{\theta}) = f(\theta, \hat{\theta}) \quad (1-2)$$

So, "how bad" a current estimate is, depends on the way a loss function weighs accuracy errors and returns respective losses. Two standard loss functions are the absolute-error and the squared-error loss function. Both are simple to understand and commonly used (Davidson-Pilon, 2015).

As implied by its name, the absolute-error loss function returns loss as the absolute error, i.e. the difference between the estimate and the true parameter (Davidson-Pilon, 2015):

$$L(\theta, \hat{\theta}) = |\theta - \hat{\theta}| \quad (1-3)$$

Accordingly, losses increasing linearly with the distance to the true value are returned for respective estimates. This means that all differences between relative errors are weighed equally, no matter whether they are found in the realm of relatively small or relatively large errors (Hennig and Kutlukaya, 2007).

Using the squared-error loss function returns losses that increase quadratically with distance of the estimator to the true parameter value (Davidson-Pilon, 2015):

$$L(\theta, \hat{\theta}) = |\theta - \hat{\theta}|^2 \quad (1-4)$$

This exponential growth of loss also means that large errors are weighed much stronger than small errors. This might come with over-valuation of distant outliers and misrepresentation of magnitudes in distance. Regarding this, the absolute-error loss function can be seen as more robust (Davidson-Pilon, 2015).

Both of these standard loss functions are symmetric and can be described as objectively aiming at a high precision in estimating the true parameter value. Davidson-Pilon (2015) and Hennig and Kutlukaya (2007) propose that it might be useful to move away from these type of objective loss functions to the design of customized loss functions that specifically reflect an individual's (i.e. the decision maker's) objectives, preferences and outcomes. Hennig and Kutlukaya (2007) argue that choosing and designing a loss function involves the translation of informal aims and interests into mathematical terms. This process naturally implies the integration of subjective decisions and subjective elements. According to Hennig and Kutlukaya (2007), this is not necessarily unfavorable or "less objective", as it may better reflect an expert's perspective on the situation and contribute to a productive scientific discussion. The standard loss functions defined above are symmetric, but can easily be adapted to be asymmetric, for example by weighing errors on the negative side stronger than those on the positive side. Preference over estimates larger than the true value (i.e. overestimation) is thus incorporated in an uncomplicated way (Davidson-Pilon, 2015; Hennig and Kutlukaya,

2007). Much more complicated designs of loss functions are possible, depending on purpose, objective and application (Davidson-Pilon, 2015). Some case-specific loss functions are designed in the following chapters of this work.

The presence of uncertainty during decision making implies that the true parameter is unknown and thus the truly incurred loss  $L(\theta, a)$  cannot be known at the time of making the decision (Berger, 2013; Davidson-Pilon, 2015). The Bayesian perspective considers unknown parameters as random variables and samples that are drawn from the posterior distribution as possible realizations of the unknown parameter, i.e. all possible true values are represented by this distribution (Davidson-Pilon, 2015). A suitable alternative to the actual loss is to consider a decision's expected loss and to make a decision that is optimal in relation to this expected loss (Berger, 2013).

Given a posterior distribution  $P(\theta|X)$ , the expected loss of choosing an estimate  $\hat{\theta}$  over the true parameter  $\theta$  (after evidence  $X$  has been observed) is defined by the function below (Davidson-Pilon, 2015):

$$l(\hat{\theta}) = E_{\theta}[L(\theta, \hat{\theta})] \quad (1-5)$$

The expectation symbol  $E$  is subscripted with  $\theta$ , by which it is indicated that  $\theta$  is the respective unknown variable. This expected loss as defined above, is also referred to as the risk of estimate  $\hat{\theta}$  (Davidson-Pilon, 2015).

By the Law of Large Numbers, the expected loss of  $\hat{\theta}$  can be approximated drawing a large sample size  $N$  from the posterior distribution, respectively applying a loss function  $L$  and averaging of the number of samples (Davidson-Pilon, 2015):

$$\frac{1}{N} \sum_{i=1}^N L(\theta_i, \hat{\theta}) \approx E_{\theta}[L(\theta, \hat{\theta})] = l(\hat{\theta}) \quad (1-6)$$

Minimization of a loss function returns a Bayesian point estimate known as Bayes action  $\delta^P(X)$ , which is the estimate, action or decision with the least expected loss according to the loss function (Berger, 2013). For a unimodal and symmetric absolute-error loss function, the Bayes action is simply the median of the posterior distribution, while using squared-error loss it is the mean (Davidson-Pilon, 2015; Berger, 2013). The MAP (maximum a posteriori) estimate is the minimizing solution for the posterior using zero-one loss (Davidson-Pilon, 2015). The possibility of more than one minimum also implies that several Bayes actions can exist for one problem (Berger, 2013).

Davidson-Pilon (2015) implemented different risk affinities by simply introducing a risk parameter into the loss function. By using different values for this parameter, it can be represented how comfortable an individual is with being wrong and furthermore which "side of wrong" is preferred by this decision maker (Davidson-Pilon, 2015). This approach to expressing risk-affinities is used the design of loss functions in this work.

## 1-2 Application in structural geological modeling

In this work, these methods of Bayesian analysis and decision theory are applied in the field of geological structural modeling. The fundamental approach follows closely the research

conducted by [de la Varga and Wellmann \(2016\)](#) and builds upon their findings.

According to them, structural geological modeling can be regarded as a forward problem and the elements of Bayesian inference can be specified in this context as follows:

1. **Mathematical forward model ( $M$ ):** The connections between parameters  $\theta$  and observed data  $y$  are defined in this mathematical model. ...
2. **Model parameters ( $\theta$ ):** These model-defining parameters, such as the position or dip of layers and faults in geological settings, can be deterministic or stochastic. In the latter case, they are uncertain parameters to which a probability distribution is assigned.
3. **Observed data ( $y$ ):** This is any type of additional information that can be related to the forward modeling results and might possibly be used to reduce uncertainty. Such data can be gained by measurements and observations, for example by core sampling, well-log analysis or seismic acquisition.
4. **Likelihood functions ( $p(y|\theta)$ ):** Links between the previous parameters  $\theta$  and the additional data  $y$  are established by these functions in way that they reflect the likelihood of the parameter states given the observations ([de la Varga and Wellmann, 2016](#)).

A fundamental sequence of the inference process was proposed by [Gelman et al. \(2014\)](#), adapted by [de la Varga and Wellmann \(2016\)](#) and is subsequently adjusted for the application in this work as follows:

1. **Setting up a full probability model:** A multi-dimensional joint probability space is to be generated, taking into account the probability distributions of every model parameter  $\theta$ . (- here, this not only results in the creation of geological models, but is subsequently also coupled with algorithms which deduct economic realizations from these models.)
2. **Conditioning on observed data:** Subsequently, an appropriate posterior distribution  $p(\theta|y)$  is to be calculated by conditioning the parameters  $\theta$  on the observed data  $y$  given the likelihood  $p(y|\theta)$ . This is the step of Bayesian updating of the belief about the parameter uncertainty given new information. In a chosen model ( $M$ ), this is achieved by linking parameters and data through deterministic operations which are additionally compared to the likelihood functions. It is pointed out by [de la Varga and Wellmann \(2016\)](#) that any combination of parameter-observation connections is allowed, i.e. not all parameters need to be necessarily connected to all observed data. After all conditional probabilities have been set up, the Bayes Theorem (Equation 1-1) is applied to attain the posterior. However, due to the multi-dimensionality given in geological problems, the use of Markov chain Monte Carlo methods is advised to achieve this as described in Section 1-2-1 below.
3. **Evaluation of the posterior model:** Depending on the aim of the study, a post-processing analysis can be conducted accordingly. [de la Varga and Wellmann \(2016\)](#) focused on the examination of the posterior distributions of the parameters  $\theta$  and the generated models, particularly regarding information entropy within the model space.

In this work, the geological models are assigned an economic meaning by declaring them potential petroleum reservoirs and introducing customized loss functions to reflect the economic interest of decision makers in developing respective resource extraction projects (see Section 1-1-4). Shifts in Bayes actions are considered a measure for the influence of Bayesian inference on decision-making and the significance of additional observations for different decision makers.

### 1-2-1 Markov chain Monte Carlo sampling (MCMC)

Despite the apparent simplicity of the Bayes Theorem, a direct analytical calculation and exact inference of the posterior distribution  $P(A|X)$  is rarely possible in non-idealized cases, due to intractability in multi-dimensional spaces (Hoffman and Gelman, 2014; de la Varga and Wellmann, 2016). Thereby arises the necessity to resort to methods of statistical inference approximation. Markov chain Monte Carlo (MCMC) sampling has proven to be a generally applicable and reliable method for exploring multi-dimensional parameter spaces in an intelligent way (Hoffman and Gelman, 2014; Davidson-Pilon, 2015). Gilks (2005) has emphasized the significance of MCMC for the application of Bayesian statistics.

In the ordinary Monte Carlo approach, random independent samples are drawn from a target distribution in order to approximate its shape (Gilks, 2005; de la Varga and Wellmann, 2016). High-dimensional parameter spaces as found in Bayesian applications lead to complex shapes and often make independent sampling infeasible (Gilks, 2005). This can be solved by extending the Monte Carlo principle with a Markov chain, in which every sample iteration of the parameter  $\theta^{(t+1)}$  is dependent uniquely on the previous value  $\theta^{(t)}$  (Gilks, 2005; de la Varga and Wellmann, 2016).

The general principle of MCMC can be described as follows: Drawing representative samples from an target distribution of unknown shape is based on the conduction of a so-called random walk on the parameter distribution space. T sampling steps are to be performed. The first sampling location is chosen at random. With each subsequent step, a new sampling location is proposed. The new sample value is then related to the previous step. According to a weight defined by the scaled up candidate density of the value, the proposed step is then accepted or rejected. In the case of acceptance, the value is added to the sample trace and the process is continued from the current location. In the case of rejection, sampling is reverted to the previous accepted step (Schaaf, 2017; de la Varga and Wellmann, 2016). The intention behind this concept it to achieve convergence of the sampling algorithm towards areas of high probability (Davidson-Pilon, 2015).

Variations in the way of how new sample steps are proposed and in the acceptance-rejection condition result in different single MCMC sampling methods (Schaaf, 2017; de la Varga and Wellmann, 2016). Various algorithms for random MCMC walks have been developed for over more than six decades and advancements have still been made in recent years. Common examples for such algorithms are the Metropolis-Hastings samplers as devised by Metropolis et al. (1953) and generalized by Hastings (1970). The Gibbs sampler (Geman and Geman, 1984) is another well-known method. For the purpose of this work, an adaptive Metropolis-Hastings sampler is used.

In Metropolis-Hastings methods, each sampling step at iteration  $t$  is determined by a candidate probability distribution  $q(\theta, \theta')$ , from which a proposed sample  $\theta'$  is drawn (de la Varga and Wellmann, 2016). The acceptance-rejection condition is defined by the acceptance ratio



$a(\theta', \theta)$ :

$$a(\theta', \theta) = \frac{p(\theta')p(y|\theta')}{p(\theta)p(y|\theta)} \quad (1-7)$$

To ensure a thorough exploration of the probability space, the transition to higher probability densities should not be enforced in every case but selectively. This is assured by relating the acceptance ratio from Equation 1-7 to a random value  $u$  from a Uniform distribution  $U(0, 1)$  as follows:

$$\theta^{(t+1)} = \begin{cases} \theta' & \text{if } a(\theta', \theta) > U(0, 1) \\ \theta^t & \text{otherwise} \end{cases} \quad (1-8)$$

Thereby, the algorithm assigns high probabilities to high-density points and low probabilities to low-density points, so that the chain state is moved accordingly (de la Varga and Wellmann, 2016).

Metropolis methods are furthermore defined by the step size scale factor that is chosen. While large steps are good for exploration of the space and mixture in the chain, acceptance rates are low. Small steps have better acceptance rates, but lead to slower exploration and convergence of the algorithm (de la Varga and Wellmann, 2016).

The Adaptive Metropolis (AM) by Haario et al. (2001) is used in this work. It adapts the traditional Metropolis-Hastings by incorporating the ability of continuous step size tuning during convergence, by taking into account the full information saved along the process. This is achieved by generating a covariance matrix that is updated every iteration. The adaptive nature of the process enables fast convergence for non-linear distributions while maintaining ergodicity Haario et al. (2001); de la Varga and Wellmann (2016). Its suitability for multi-dimensional distribution spaces make it an excellent method for dealing with complex models such as the structural geological models in this work (Schaaf, 2017)

### 1-2-2 Structural geological forward modeling

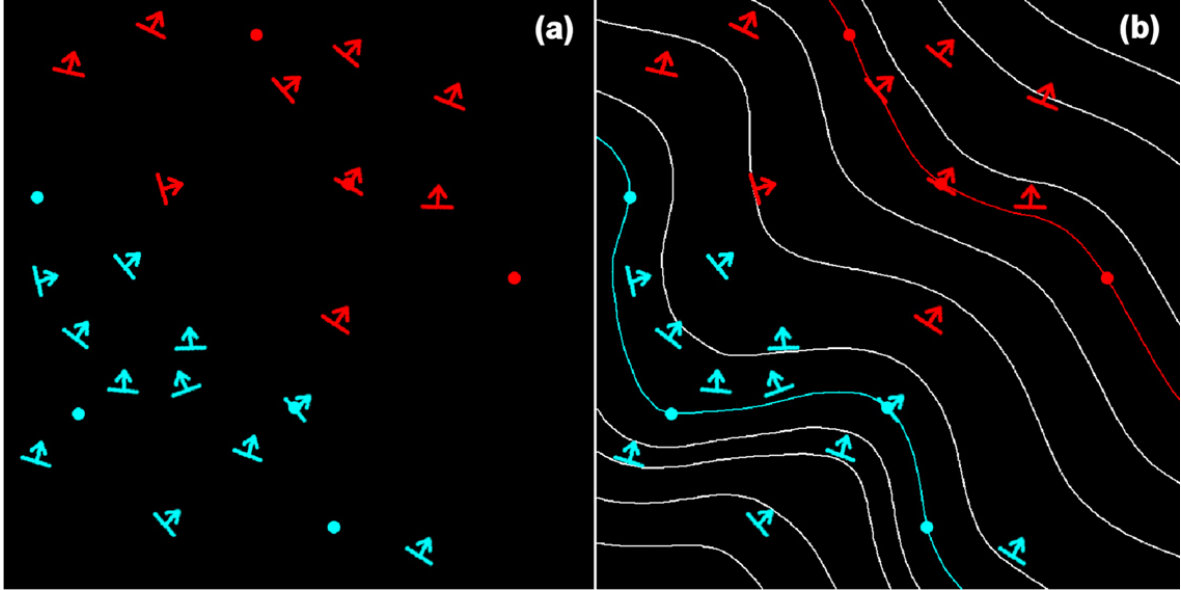
Performing forward modeling in the context of structural geology requires the use of a suitable modeling step. Additionally, for the application in a probabilistic setting, the method should enable fully automatic reconstruction of the model, when parameters are changed. The application in this work follows the example of de la Varga and Wellmann (2016) and relies on the use of implicit interpolation for geological modeling, a method developed and elaborated by Lajaunie et al. (1997) and Calcagno et al. (2008).

This implicit method relies on the interpolation of a potential field scalar function  $T(\vec{x})$  of any point  $(\vec{x})$  in a 3D space, and thus reflects the geometry of geological structures (Calcagno et al., 2008). Modeling  $T(\vec{x})$  is achieved by cokriging that regards two forms of data: (i) contact points on geological interfaces through increments of the potential field  $T(\vec{x}) - T(\vec{x}')$  and (ii) orientation data as gradients, i.e. partial derivatives of the potential field  $\delta T(\vec{x})/\delta u_\beta$  in each direction  $u$  (Calcagno et al., 2008). The respective estimator is defined as follows:

$$T(\vec{x}) - T(\vec{x}_0) = \sum_{\alpha=1}^M \mu_\alpha (T\vec{x}_\alpha - T(\vec{x}'_\alpha)) + \sum_{\beta=1}^N \nu_\beta \frac{\delta T}{\delta u_\beta}(\vec{x}_\beta), \quad (1-9)$$

where  $\vec{x}_0$  is an arbitrary origin,  $M$  and  $N$  are the total number of data points and partial derivatives respectively, and their relative contributions are weighed by the factors  $\mu_\alpha$  and  $\nu_\beta$ . Furthermore,  $T$  is assumed to be a random function defined by polynomial drift and a stationary covariance  $K(h)$  (Calcagno et al., 2008). The use of a cubic covariance model is suggested by Calcagno et al. (2008), based on the results from studies conducted by Aug (2004) and Chilès et al. (2004).

Resulting potential fields can be used to describe geological interfaces as iso-surfaces in any



**Figure 1-1:** woo a potential field! from Calcagno et al. (2008)

kind of 3D geometry (Calcagno et al., 2008). Fault geometries can be interpolated analogously. These can be infinite in the 3D space, interrelated in a fault network or finite. To account for the effect of faults on geological layers, discontinuous potential fields are created by applying discontinuous drift functions in the cokriging system. Additionally, geological rules allow for the representation of several types of interactions between sets of geological layers (Calcagno et al., 2008).

It is pointed out by Calcagno et al. (2008), that this method is particularly appropriate for cases in which knowledge about the geology is only given a sparse locations and is thus applicable for a wide variety of typical problem in geological settings. Its suitability for this work is furthermore emphasized by the possibility to modify the topology-defining geological pile to achieve different geometric realizations without altering the data. The model can thus be updated in the face of new data or interpretations (Calcagno et al., 2008).

### 1-2-3 Economic significance and reservoir valuation

In this work, the generated structural models are to be assigned an economic significance in order to approach a real setting, in which hypothetical actors could have an interest in taking a decision, and thus, to evaluate the influence of the Bayesian inference of the decision-making process. Changes in Bayes actions are considered as a measure for this effect.

For this purpose, it is assumed that a hydrocarbon reservoir system is represented by each geological model realization. These are typically comprised stratigraphically of at least one potential reservoir formation, succeeded by at least one sealing layer above. Additionally, a trapping mechanism is normally required (source?).

The resulting reservoirs are to be valued in a way that respective estimations on the basis of loss functions can be conducted, i.e. Bayesian decision theory is applicable. In the typical petroleum industrial setting, the main question is one of how much of the resource can be produced and how high the return on investment will be (Dean, 2007).

### Original oil-in-place and recoverable volumes as value measures

Before pressure and production tests have been conducted (i.e. before production has started), volumetric estimation is the only approach to assess the amount of hydrocarbons in place in a reservoir. From this value, recoverable reserves can be estimated based on an estimated recovery factor (Dean, 2007).

Oil-in-place and gas-in-place volumes are calculated based on:

1. Subsurface rock volume containing hydrocarbons. This is mainly defined by thickness and areal extend of the accumulation.
2. Weighted average effective porosity of the reservoir rock.
3. Water saturation in the reservoir rock.
4. Hydrocarbon fluid properties (Dean, 2007).

As the cases used in this work are purely artificial, a pure oil accumulation is assumed. Using the factors listed above, the respective equation for original oil-in-place (OOIP) is formulated as follows:

$$OOIP = A * h * \phi * (1 - S_W) * 1/FvF \quad (1-10)$$

Where OOIP is returned in  $m$ . The hydrocarbon-filled rock volume is defined by the drainage area  $A$  in  $m$  and the net pay thickness  $h$  in  $m$ . Porosity  $\phi$  and water saturation  $S_W$  (interstitial water) are given in fraction of the rock volume. A dimensionless factor for the change in oil volume between reservoir conditions and standard conditions at surface is represented by the formation volume factor  $FvF$ . Thus, shrinkage of the oil volume brought to the surface is determined by  $1/FvF$  (Dean, 2007).

Subsequently, the effectively recoverable oil volumes (ROV) can be calculated by multiplying the OOIP with a recovery factor  $RF$ :

$$ROV = OOIP * RF = A * h * \phi * (1 - S_W) * 1/FvF * RF \quad (1-11)$$

the recovery factor is influenced by a number of fluid properties, such as viscosity, density, solution oil/gas ratio and the formation volume factor. Thus, it is difficult to estimate (Dean, 2007). Globally, the ultimate average recovery factor for oil fields is about 35% (Labastie, 2011). For gas accumulations, however, recovery factors range typically between 70 and 90% (Dean, 2007).

The models in this work are completely hypothetical and not based on real data. The inputs to these equations can thus be chosen arbitrarily to test the applied methods. However, to come to significant conclusion, it makes sense to utilize values that approximately represent real possible scenarios. Furthermore, as 3D geological structures are modeled here, the hydrocarbon-filled rock volume ( $A * h$  in Equations 1-10 and 1-11) is defined by every realization of the uncertain geological model. The rock volume is thus an uncertain factor. Remaining factors can be implemented deterministically or as uncertain values, too. For most factors, it seems most suitable to define a normal distribution around a known typical value or global average, as it will return a high probability for common scenarios and occasionally produce cases of rare, more extreme values. (INSERT WHAT IS CHOSEN IN THE END - MAYBE MORE SENSE TO TAKE FIXED VALUES FOR THE FACTORS THEN RV, SO THAT ITS INFLUENCE IS CLEAR; THEN MAYBE INCLUDE POROSITY? OR COMPARE BOTH POSSIBILITIES?).

Consequently, OOIP and ROV calculations represent the step following and using the results from computing the geological model, so that the realizations are assigned values that can be interpreted from a perspective of economic interest. The resulting values from numerous model runs are used as a base probability distribution for the true value, on which loss function estimation can be applied as described in Section 1-1-4.

### Net present value

hmmm...

### Abstract valuation using scores in the one-dimensional model

For the simplified 1D model that is described further below, OOIP and recoverable volume calculations are not applicable, as these require a 3D setting. Given only layer positions in one vertical dimension, it is resorted to an abstract way of reservoir valuation by defining a scoring system. The reservoir score is made dependent on three uncertain parameters which can be deduced from the 1D layer positions: (1) reservoir thickness, (2) reservoir top depth and (3) seal thickness.

### 1-2-4 Numerical computational implementation via Python, GemPy and PyMC

Bayesian analysis can be conducted using probabilistic programming (Salvatier et al., 2016). The implicit method of forward geological modeling, described above, is to be embedded in such a framework. For doing so, the programming language of choice in this work is Python. The merits of Python have been pointed out by Behnel et al. (2010); Salvatier et al. (2016); Langtangen (2008). Development is facilitated by an expressive but concise and clean syntax that is easy to learn. Python is dynamic, compatible with multiple platforms and offers good support for numerical computing. Integration of other scientific libraries and extension via C, C++, Fortran or Cython are easily possible (Behnel et al., 2010; Salvatier et al., 2016; Langtangen, 2008). Python is thus a straightforward tool for the implementation of central components of Bayesian analysis, such as custom statistical distributions and samplers (Salvatier et al., 2016).

The 3D geological modeling step in this work is implemented using GeMpy, an open-source, Python-based software that is able to generate and visualize complex 3D structural geological models based on the potential field interpolation method elaborated in Section 1-2-2 De la Varga and Schaaf (2017). Its design allows for its application in a probabilistic setting, particularly by coupling it with PyMC (see below). At the time of writing this, GemPy is still under development (version 0.995), but is already functioning for the purpose of this work.

For conducting the Bayesian analysis, GemPy is combined with the Python library PyMC, which was developed for conducting Bayesian inference and prediction problems in an open-source probabilistic programming framework (Davidson-Pilon, 2015; Salvatier et al., 2016). Different model fitting techniques are provided in PyMC, such as the *maximum a posteriori* (MAP) method and several (MCMC) sampling methods, including the Adaptive Metropolis explained in Section 1-2-1 (Salvatier et al., 2016). The components which are used to construct a statistical model, are represented by *Deterministic* or *Stochastic* variables in PyMC. the values of *Deterministic* variables are completely dependent on its parents' values, as defined by a respective mathematical function (Salvatier et al., 2016). *Stochastic* variables are used to represent uncertain parameters  $\theta$  or observed stochastic variables as likelihood functions  $p(y|\theta)$  (Salvatier et al., 2016; de la Varga and Wellmann, 2016). Complex mathematical relations between *Stochastic* variables can be described through *Deterministic* variables (de la Varga and Wellmann, 2016). Furthermore, PyMC allows for the creation of own object definitions inheriting from the class descriptions of these two variable types.

As of writing this, the newer PyMC3 is not yet supported by GemPy. Thus, it is resorted to PyMC 2 for 3D geological settings in this work. However, the newer PyMC 3 is used in a simple 1D scenario. Salvatier et al. (2016) point out that the development of PyMC 3 is continuing, as the inclusion of further tools is planned for future updates.

### 1-3 Designing a case-specific loss function

As explained before in Section 1-1-4, the standard symmetric loss functions provide objectively good estimators minimizing expected loss by returning the median or mean respectively. However, assigning an economic notion to our model and assuming the case of an actor or decision maker in any field, naturally necessitates the consideration of preferences, interests and the overall subjective perspective such an individual or for example a company might have. Further constraints, properties and factors can also be specific to the field, industry or generally to the problem at hand. Consequently, the design of a more specific non-standard and possibly asymmetric loss function might be required, so that an adapted Bayesian estimator can be found. One that includes subjective aspects and difference in weighting of particular gains or losses, arising from an actor's inherent preferences and the environment in which the actor has to estimate or make a decision. In the face of several uncertain parameters, a perfectly true estimate is virtually unattainable. However, an attempt can be made to design a customized loss function that returns a Bayesian estimator involving the least bad consequences for an individual in a specific environment. Assuming the reservoir setting case and valuation methods introduced in Section 1-2-3), such a customization attempt is made and explained step by step in the following. (SOURCES?)

For the purpose of estimation, it makes sense that one of the standard loss functions is chosen as a basis and a customized loss function is developed from there. The absolute-error loss

seems most appropriate for this case of hydrocarbon reservoir value estimation. Ideally, an actor would like to know the exact true value of interest, say the OOIP, so that investments or resources can be allocated appropriately in order to acquire economic gains. This allocation is the decision to be made or action to be taken. Deviations from the unknown true value in the form of over- and underestimation bring about an error and loss accordingly. In this case, it is assumed that investments increase linearly with linear growth in the value of the resource. For this reason, the absolute-error loss function is favored here over the squared-error loss function. It is chosen as the base function to which further development steps refer, based on mostly logical case-specific assumptions:

1. Step I: The standard symmetrical absolute-error loss function is chosen as a starting point for further customization steps:

$$L(\theta, \hat{\theta}) = |\theta - \hat{\theta}| \quad (1-12)$$

2. Step II: Considering the development of a hydrocarbon reservoir, it can be assumed that over-investing is worse than under-investing. Overestimating the size of an accumulation might for example lead to the installation of equipment or facilities that are actually redundant or unnecessary. This would come with additional unrecoverable expenditures. Consequences from underestimating, however, may presumably be easier to resolve. Additional equipment can often be installed later on. Hence, overestimation is weighted stronger in this loss function by multiplying the error with an overestimation factor  $a$  ( $= 1.25$ ):

$$L(\theta, \hat{\theta}) = |(\theta - \hat{\theta})| * a \quad (1-13)$$

3. Step III: The worst case for any project would be that its development is set into motion, expecting a gain, only to discover later that the value in the reservoir does not cover the costs of realizing the project, resulting in an overall loss. A petroleum system might also turn out to be a complete failure, containing no value at all, although the actor's estimate indicated the opposite. Here, this is referred to as worst case or fatal overestimation. A positive value is estimated, but the true value is zero or negative. This is worse than the "normal" non-fatal overestimation, where both values are positive and a net gain is still achieved, which is only smaller than the best possible gain of expecting the true value. Fatal overestimation is included in the loss function by using another weighting factor  $b$  ( $= 2$ ) that replaces  $a$ :

$$L(\theta, \hat{\theta}) = |(\theta - \hat{\theta})| * b \quad (1-14)$$

(In other words: Fatal overestimation is twice as bad as simple underestimation.)

4. Step IV: A worst case or fatal underestimation can also be derived from the idea of estimating a zero or negative value, when the true value is actually positive. This is assumed to be worse than non-fatal overestimation, but clearly better than fatal overestimation. No already owned resources are wasted, it is only the potential value that is lost, i.e. opportunity costs that arise from completely discarding a reservoir with

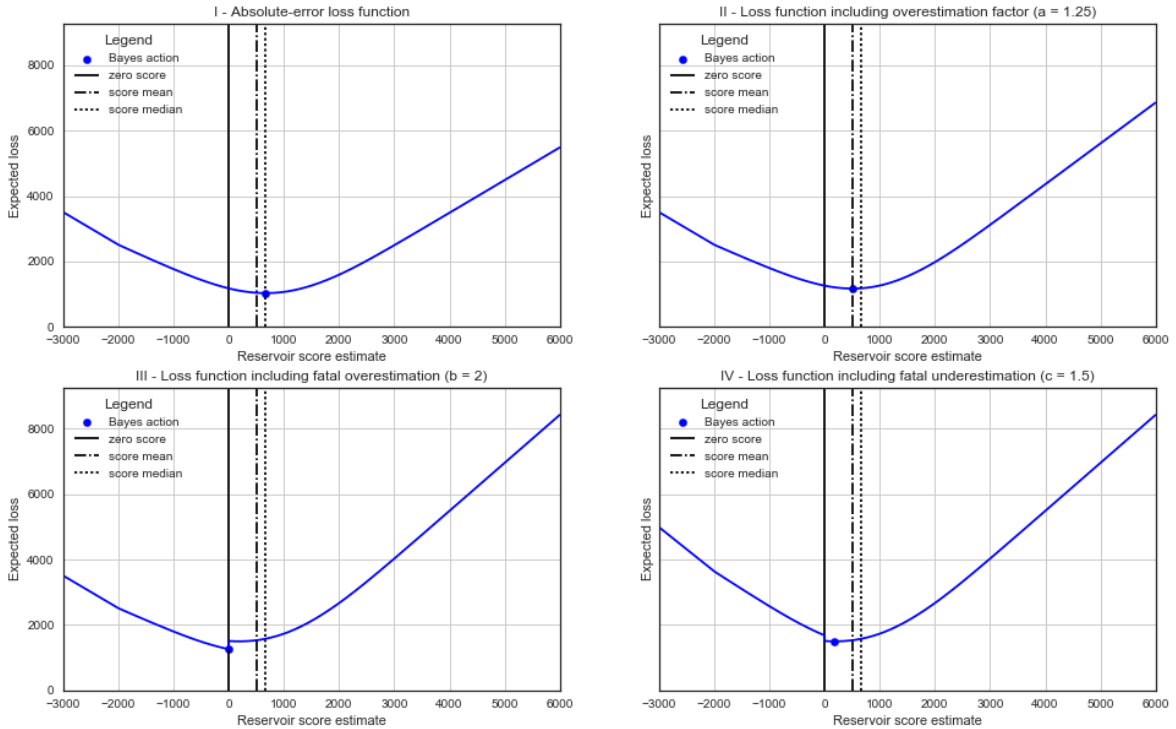
a potential gain equal to the positive true value. Fatal underestimation is weighted using a third factor  $c$ :

$$L(\theta, \hat{\theta}) = |(\theta - \hat{\theta})| * c \quad (1-15)$$

Combining these adaption steps and the conditions defined in them, results in the following customized loss function:

$$L(\theta, \hat{\theta}) = \begin{cases} |\theta - \hat{\theta}|, & \text{for } 0 < \hat{\theta} < \theta \\ |\theta - \hat{\theta}| * a, & \text{for } 0 < \theta < \hat{\theta} \\ |\theta - \hat{\theta}| * b, & \text{for } \theta \leq 0 < \hat{\theta} \\ |\theta - \hat{\theta}| * c, & \text{for } \hat{\theta} \leq 0 < \theta \end{cases}, \text{ with } a, b, c \in \mathbb{Q} \quad (1-16)$$

Realizations of the four adaption steps are depicted in the plots in Figure 1-2. It is important to note that the weighting factors defined above can take basically any numerical values but should be chosen in a way that they appropriately represent the framework conditions of the problem. Here, for example, it is assumed that normal overestimation is 25% ( $a = 1.25$ ), fatal overestimation 100% ( $b = 2$ ) and fatal underestimation 50% ( $c = 1.5$ ) worse than normal underestimation.

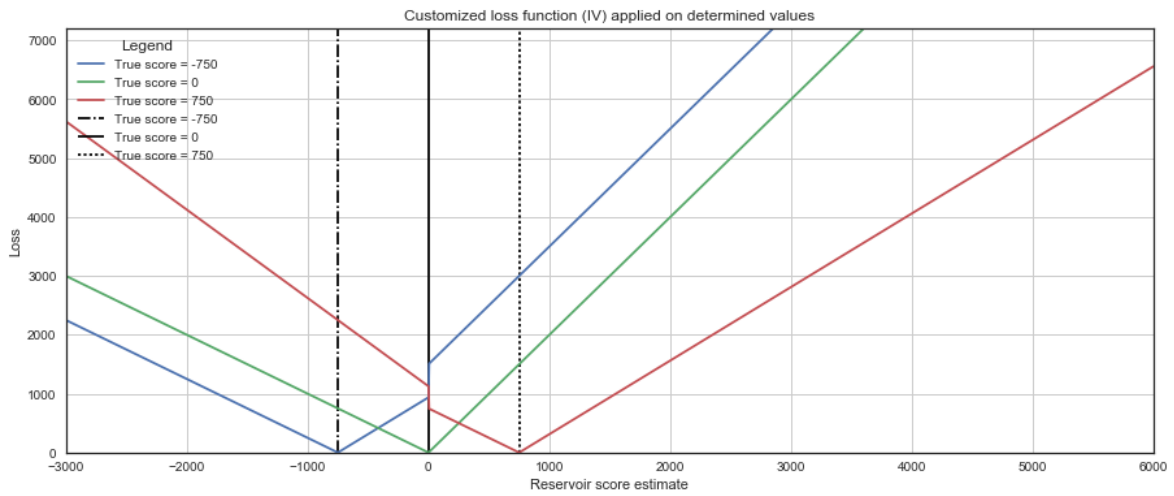


**Figure 1-2:** The single steps of customizing the loss function are depicted in plots I - IV.

In Figure 1-2 it can be observed, that assigning a stronger weight on overestimation (plot II) steepens the curve on the right hand side and shifts the minimum to the left, i.e. to a lower estimate. Using  $a = 1.25$ , the Bayes action changes from the median, to a value close



to the mean of the distribution. The shift and steepening are significantly reinforced by the introduction of fatal overestimation (plot III). With  $b = 2$ , the Bayes action drops to the zero score estimate. It can also be noted, that by defining a condition dependent on the algebraic sign of the values, according to which only losses for positive estimates are multiplied by  $b$ , a distinct jump appears at the zero score boundary. Due to a similar condition, the same effect is observed on the negative side of estimate values, where the curve has also been steepened, after including fatal underestimation (plot IV). This comes with a shift of the minimum towards positive values. It is also to be noted that with every customization step, the overall expected loss is increased.



**Figure 1-3:** Loss based on the customized loss function (IV) for determined true scores of -750, 0 and 750.

The implementation of this customized loss function (IV) using single determined values for the true score is plotted in Figure 1-3. This helps to clarify the way real losses result for each guess, relative to given true score values. The expected loss is acquired by arithmetically averaging such loss realizations based on the true score probability distribution by using formula 1-6.

It has to be emphasized that this is just one possible proposal for loss function customization. There exists not one perfect design for such a case. Slight to strong changes can already be implemented by simply varying the values of the weighting factors  $a$ ,  $b$  and  $c$ . Fundamentally different loss functions can also be based on a significantly different mathematical structure. Loss functions are customized regarding the problem environment and according to the subjective needs and objectives of an actor. Thus, they are mostly defined by the actor expressing his perspective. Changes in the individual's perception and viewpoint might lead to further customization needs even later on. (Especially considering individual persons as actors, psychological aspects may play a significant role.)

(Estimate or prediction? True value only known if a "yes"-action (development) is taken.)



### 1-3-1 Including different risk-affinities in the loss function

One can assume that several actors in one sector or decision environment may have the same general loss function, but different affinities concerning risks. This might be based for example on the different psychological factors or economic philosophies followed by companies. It might also be based on budgets and options such actors have available. An intuitive example is the comparison of a small and a large company. A certain false estimate or error might have a significantly stronger impact on a company which has a generally lower market share and only few projects, than on a larger company which might possess a higher financial flexibility and for which one project is only one of many development options in a portfolio.

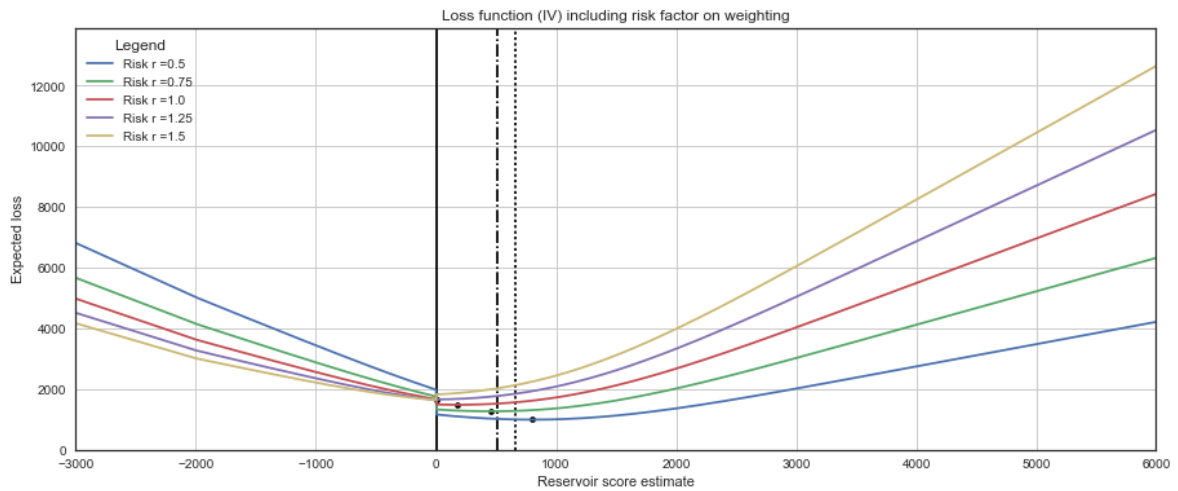
In the following, the loss function is further adapted to consider different risk-affinities of several actors. Representing risk behavior in a loss function can also be done in different ways and regarding different types of risks. Here, bidding lower is considered the cautious, risk-averse option, as smaller losses can be expected from underestimating. Guessing higher is deemed riskier. Losses from overestimation are greater. However, bidding correctly on a higher value, will also return a greater gain. It is assumed that risk-friendly actors care less about fatal underestimation, i.e. they will rather develop a project than discard it. In the loss function, risk is simply included using a risk factor  $r$  which alters the weighting factors  $a$ ,  $b$  and  $c$  respectively:

$$L(\theta, \hat{\theta}) = \begin{cases} |\theta - \hat{\theta}|, & \text{for } 0 < \hat{\theta} < \theta \\ |\theta - \hat{\theta}| * (a * r), & \text{for } 0 < \theta < \hat{\theta} \\ |\theta - \hat{\theta}| * (b * r), & \text{for } \theta \leq 0 < \hat{\theta} \\ |\theta - \hat{\theta}| * (c * (r^{-0.5})), & \text{for } \hat{\theta} \leq 0 < \theta \end{cases}, \text{ with } a, b, c, r \in \mathbb{Q} \quad (1-17)$$

According to this, for  $r = 1$  the risk-neutral loss function is returned, since  $a$ ,  $b$  and  $c$  are not altered. For  $r < 1$ , the weight on overestimating is reduced and increased for fatal underestimation. This represents a more risk-friendly actor that is willing to bid on a higher estimate to attain a greater gain. For  $r > 1$ , the overestimation weight is increased, the underestimation weight decreased in the loss function and respectively more risk-averse actors are prompted to bid on lower estimates.

The factor  $r$  can take basically any positive values. However, since risk-neutrality is expressed by  $r = 1$ , values  $0 < r < 2$  are considered to be the most appropriate choices to represent both sides of risk-affinity equally.

Implementing it as a factor  $r$  leads to different steepnesses of the plotted curves depicting loss and expected loss. The effect on the latter is visualized in Figure 1-4, where 0.5, 0.75, 1, 1.25 and 1.5 were chosen as values for  $r$ . It can be observed that the minima for expected loss, i.e. the Bayesian estimators for differently risk-affine actors, are located at different estimates. Mean and median are clearly surpassed by the best estimate of the most risk-friendly actor ( $r = 0.5$ ), while for the most risk-averse actors ( $r = 0.5$  and  $r = 0.75$ ), the Bayes action equals a zero score estimate (and thus the decision to take on action). In Figure 1-4 it can also be recognized that the expected loss is generally lower for risk-friendlier actors (on the side of positive estimates).



**Figure 1-4:** Plotting of expected loss realizations after including the risk factor  $r$  in the loss function (IV) for actors with risk-affinities ranging from risk-averse ( $r = 0.5$  and  $0.75$ ), over risk-neutral ( $r = 1$ ), to risk-friendly ( $r = 1.25$  and  $r = 1.5$ ).

## 1-4 Constructing a 1D geological reservoir model

## 1-5 Constructing a full 3D geological reservoir model

---

## Chapter 2

---

# Results

Text

**2-1 Results using the 1D geological model**

**2-2 Results using the 3D geological model**



---

## Chapter 3

---

# Discussion



---

## Chapter 4

---

# Conclusion





---

# Bibliography

- Aug, C. (2004). *Modélisation géologique 3D et caractérisation des incertitudes par la méthode du champ de potentiel*. PhD thesis, École Nationale Supérieure des Mines de Paris.
- Behnel, S., Bradshaw, R., Citro, C., Dalcin, L., Seljebotn, D. S., and K, S. (2010). Cython: The best of both worlds. *Computing in Science Engineering*, 13(2).
- Berger, J. O. (2013). *Statistical decision theory and Bayesian analysis*. Springer Science & Business Media.
- Calcagno, P., Chilès, J.-P., Courrioux, G., and Guillen, A. (2008). Geological modelling from field data and geological knowledge: Part i. modelling method coupling 3d potential-field interpolation and geological rules. *Physics of the Earth and Planetary Interiors*, 171(1):147–157.
- Chilès, J.-P., Aug, C., Guillen, A., and Lees, T. (2004). Modelling the geometry of geological units and its uncertainty in 3d from structural data: the potential-field method. In *Proceedings of international symposium on orebody modelling and strategic mine planning, Perth, Australia*, volume 22, page 24.
- Davidson-Pilon, C. (2015). Bayesian methods for hackers: Probabilistic programming and bayesian inference.
- De la Varga, M. and Schaaf, A. (2017). Gempy. <https://github.com/cgre-aachen/gempy/>.
- de la Varga, M. and Wellmann, J. F. (2016). Structural geologic modeling as an inference problem: A bayesian perspective. *Interpretation*, 4(3):SM1–SM16.
- Dean, L. (2007). Reservoir engineering for geologists - volumetric estimation. *The Monthly Magazine of the Canadian Society of Petroleum Geologists*, pages 11–14.
- Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2014). *Bayesian data analysis*, volume 2. Taylor & Francis.

- Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):721–741.
- Gilks, W. R. (2005). *Markov chain monte carlo*. Wiley Online Library.
- Haario, H., Saksman, E., and Tamminen, J. (2001). An adaptive metropolis algorithm. bernoulli 7 223–242. *Mathematical Reviews (MathSciNet): MR1828504 Digital Object Identifier: doi*, 10:3318737.
- Hastings, W. (1970). Monte carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109.
- Hennig, C. and Kutlukaya, M. (2007). Some thoughts about the design of loss functions. *REVSTAT - Statistical Journal*, 5(1):19–39.
- Hoffman, M. D. and Gelman, A. (2014). The no-u-turn sampler: Adaptively setting path lengths in hamiltonian monte carlo. *The Journal of Machine Learning Research*, 15(1):1593–1623.
- Labastie, A. (2011). En route: Increasing recovery factors: A necessity.
- Lajaunie, C., Courrioux, G., and Manuel, L. (1997). Foliation fields and 3D cartography in geology: Principles of a method based on potential interpolation. *Mathematical Geology*, 29(4):571–584.
- Langtangen, P. H. (2008). *Python scripting for computational science*. Springer Verlag, New York.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092.
- Salvatier, J., Wiecki, T. V., and Fonnesbeck, C. (2016). Probabilistic programming in python using pymc3. *PeerJ Computer Science*, 55(2).
- Schaaf, A. (2017). Geological inference based on kinematic structural models. Master’s thesis, RWTH AACHEN UNIVERSITY, Aachen, Germany.

---

# Appendix A

---

## The back of the thesis

### A-1 An appendix section

#### A-1-1 An appendix subsection with C++ Listing

```
// 1
// C++ Listing Test
//
#include <stdio.h>
for(int i=0;i<10;i++)
{
    cout << "Ok\n";
}
6
```

#### A-1-2 A Matlab Listing

```
%
% Comment
%
n=10;
for i=1:n
    disp('Ok');
end
5
```



---

## Appendix B

---

# Yet another appendix

### **B-1 Another test section**

Ok, all is well.

