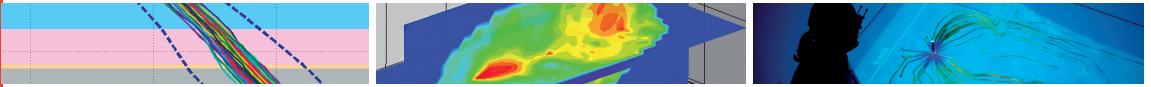




E.ON Energy Research Center



E.ON Energy Research Center

OPTIMIZATION OF GEOTHERMAL ENERGY RESERVOIR MODELING USING ADVANCED NUMERICAL TOOLS FOR STOCHASTIC PARAMETER ESTIMATION AND QUANTIFYING UNCERTAINTIES

Christian Vogt

GGE | Applied Geophysics and Geothermal Energy

**RWTHAACHEN
UNIVERSITY**

**Optimization of Geothermal Energy Reservoir Modeling using
Advanced Numerical Tools for Stochastic Parameter Estimation
and Quantifying Uncertainties**

Von der Fakultät für Georessourcen und Materialtechnik
der Rheinisch-Westfälischen Technischen Hochschule Aachen

zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigte Dissertation

vorgelegt von

Diplom-Physiker Christian Vogt

aus Mechernich

Berichter: Univ.-Prof. Dr.rer.nat. Christoph Clauser
Univ.-Prof. Dr.rer.nat. Thomas Kohl

Tag der mündlichen Prüfung: 18. Februar 2013

Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online verfügbar

Bibliographische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb-nb.de> abrufbar.

D 82 (Diss. RWTH Aachen University, 2013)

Herausgeber:

Univ.-Prof. Dr.ir. Dr.h.c. Rik W. De Doncker
Direktor E.ON Energy Research Center

Univ.-Prof. Dr. rer. nat. Christoph Clauser
Institute for Applied Geophysics and Geothermal Energy
E.ON Energy Research Center
Mathieustraße 10
52074 Aachen

Copyright Christian Vogt

Alle Rechte, auch das des auszugsweisen Nachdrucks, der auszugsweisen oder vollständigen Wiedergabe, der Speicherung in Datenverarbeitungsanlagen und der Übersetzung, vorbehalten.

Printed in Germany

ISBN: 978-3-942789-12-7

Verlag:

E.ON Energy Research Center, RWTH Aachen University
Mathieustraße 10
52074 Aachen

Internet: www.eonerc.rwth-aachen.de

E-Mail: post_erc@eonerc.rwth-aachen.de

Herstellung:

Druckservice Zillekens
Am Bachpütz 4
52224 Stolberg

Is suarach uisge teth a shireadh fo chloich fhuair (?)

A toolmaker succeeds as, and only as, the users of his tool succeed with his aid. However shining the blade, however jeweled the hilt, however perfect the heft, a sword is tested only by cutting. That swordsmith is successful whose clients die of old age.

– Frederick P. Brooks Jr.

ABSTRACT

Geothermal energy is an option for low carbon production of heat or electric energy. For further developments of this resource, a major obstacle is the risk of project failure due to uncertain estimates of flow rate and temperature (and, hence, produced power) of geothermal installations. In this work, I develop and apply stochastic methods and modeling strategies for predicting the variation of pressure, temperature, and their uncertainty with time within geothermal reservoirs based on observed thermal and hydraulic rock property distributions. This comprises stochastic forward and inverse modeling approaches for simulating heat and tracer transport as well as fluid flow numerically. The approaches reduce the corresponding *a priori* uncertainties of perturbed parameters and states drastically by 50 % – 67 % in case of temperature at a depth of 2000 m, depending on the target location. Furthermore, I estimate the spatial distribution of permeability as well as its uncertainty by applying the stochastic assimilation technique of Ensemble Kalman Filtering on production data for sedimentary rocks and fractured hard rocks. This addresses structure and parameter heterogeneity within the reservoir.

I study different geothermal reservoirs, such as (i) numerous synthetic reservoirs to test the tools of Sequential Gaussian Simulation combined with geostatistical post-processing and Ensemble Kalman Filter. (ii) Further, I quantify temperature uncertainties of a doublet system in a sedimentary reservoir in The Hague, The Netherlands. (iii) In addition to temperature uncertainties, I study pressure uncertainties at a reservoir in the north-eastern German basin. Here, also a single-well design for exploitation of geothermal energy along a fault zone proofs to represent an alternative to doublet layouts. By gradient-based deterministic Bayesian inversion, basal specific heat flow is revealed. (iv) Finally, I investigate the hard rock reservoir of the Enhanced Geothermal System at Soultz-sous-Forêts, France, using Sequential Gaussian Simulation and Ensemble Kalman Filtering in an equivalent porous medium approach. A tracer circulation test performed in 2005 provides data for the inversion. Applying the two different stochastic methods allows for identifying best estimates for the heterogeneously distributed hydraulic parameters, studying

their non-uniqueness, and comparing the results from stochastic (massive Monte Carlo, Ensemble Kalman Filter) and deterministic (gradient-based Bayesian inversion) estimation techniques. Based on the Ensemble Kalman Filter estimation results, I perform a long-term performance prediction with regard to transient temperature variation including corresponding uncertainties. The presented work flows constitute a method for creating calibrated reservoir models based on data which will allow the operators of a geothermal installation to compute production scenarios optimized with respect to profit or sustainability.

ZUSAMMENFASSUNG

Die Geothermie bietet eine vielversprechende Möglichkeit der CO₂-armen Energieversorgung in Form von Wärme oder elektrischem Strom. Große Unsicherheiten von erwarteter Temperatur und Fließrate verhindern bisher jedoch eine genaue Prognose der Entzugsleistung geothermischer Anlagen und stellen damit ein massives Hindernis für die Nutzung der Georessource dar. Ziel der vorliegenden Arbeit ist die Entwicklung stochastischer Methoden zur Prognose des zeitlichen Druck- und Temperaturverlaufs von geothermischen Reservoiren samt einer Quantifizierung der zugehörigen Unsicherheit auf Basis thermischer und hydraulischer Gesteinsparameter. Dazu werden stochastischen Vorwärtsmodellierungs- und Inversionsansätze zur Simulation von Wärme- und Stofftransport verwendet. Dieses Vorgehen reduziert Unsicherheiten der Parameter und Systemzustände drastisch, im Fall der Temperatur in etwa 2000 m Tiefe um 50 % bis 67 % je nach Ziellokation. Stochastisches Assimilieren von Produktionsdaten, beispielsweise mit Hilfe des Ensemble-Kalman-Filters, erlaubt zudem die Schätzung der räumlichen Verteilung der Permeabilität in sedimentären und geklüfteten Reservoiren sowie deren Unsicherheit.

Verschiedene Fallbeispiele geothermischer Reservoire werden untersucht. i) Synthetische Reservoir, um die numerischen Werkzeuge Sequentielle Gauß'sche Simulation — kombiniert mit einer geostatistischen Analyse — sowie Ensemble-Kalman-Fitler zu testen. (ii) Es werden Temperaturunsicherheiten einer geothermischen Dublette in einem sedimentären Reservoir in Den Haag in den Niederlanden quantifiziert. (iii) Außerdem werden zusätzlich zur Temperatur Unsicherheiten des Produktionsdrucks im Nordostdeutschen Becken betrachtet. Ein alternatives Konzept zur Dublette ermöglicht hier zudem die Ausbeute geothermischer Energie entlang einer Verwerfung unter Nutzung nur einer Bohrung. Gradientenbasierte deterministische Bayes'sche Inversion erlaubt die Schätzung der basalen Wärmestromdichte. (iv) Schließlich wird ein Enhanced Geothermal System in einem kristallinen Reservoir in Soultz-sous-Forêts, Frankreich, mit Hilfe der Sequentialen Gauß'schen Simulation und des Ensemble-Kalman-Filters betrachtet. Zu diesem Zweck liegen Daten aus einem Markierungsversuch aus dem Jahr

2005 vor. Die beiden verschiedenen stochastischen Methoden ermöglichen die Identifizierung bester Schätzwerte der hydraulischen Parameter, zeigen Uneindeutigkeit der Lösungen auf und erlauben einen Vergleich stochastischer (massives Monte Carlo, Ensemble-Kalman-Filter) und deterministischer (gradientenbasierte Bayes'sche Inversion) Schätzverfahren. Basierend auf dieser Schätzung wird auch eine Langzeitprognose in Hinblick auf die Temperaturentwicklung samt der zugehörigen Unsicherheiten angegeben.

Mit Hilfe der vorgestellten Verfahren können Betreiber einer geothermischen Anlage eine in Hinblick auf Nachhaltigkeit oder Profit optimierte Förderstrategie für jedes Reservoir individuell auf Basis kalibrierter Untergrundmodelle entwickeln.

Contents

1. Introduction	1
2. Uncertainty and Heterogeneity	7
2.1. A note about simplification, parameterization, and heterogeneity	9
3. Tools for Characterizing Geothermal Reservoirs and Quantifying Uncertainties	11
3.1. State of the Art of Characterizing Geothermal Reservoirs and Quantifying Uncertainties	11
3.1.1. Monte Carlo	12
3.1.2. Ensemble Kalman Filter	13
3.2. Forward Code	16
3.3. Gradient-based Bayesian Deterministic Inversion	18
3.4. Sequential Gaussian Simulation and Constraining Post-Processing	18
3.4.1. Constraining Post-Processing	20
3.5. Estimation of State Variables and parameters using the Ensemble Kalman Filter	21
3.5.1. Formalism	22
3.5.1.1. The Analysis Step	22
3.5.1.2. A closer inspection of the Kalman gain	25
3.6. Ensemble Sizes	25
4. Synthetic Test Scenarios	27
4.1. Test of massive Monte Carlo Inversion on a Synthetic 3D Model	27
4.1.1. Model Reconstruction with Varying Number of Wells	28
4.1.2. Constraining Post-Processing	29
4.2. 2D Synthetic Test Case for the Ensemble Kalman Filter	32
4.2.1. Results	32
4.2.1.1. Iterative Filtering	34
4.2.2. Comparison with SGSim	34
4.3. 3D Synthetic Test Case for EnKF in a Hard Rock Reservoir	36
4.3.1. Reference Model	36

4.3.2.	Data from 15 Observation Points	38
4.3.3.	Data from three Observation Locations	41
4.4.	Tracer vs. Temperature Data	41
5.	Assessment of Probabilistic Predictions for Porous Rocks	45
5.1.	Quantifying Uncertainties for a Given Reservoir Layout in The Hague, The Netherlands	46
5.1.1.	Temperature Uncertainty Related to Thermal Properties	46
5.1.1.1.	Thermal Model	48
5.1.1.1.1.	Quantifying and Reducing Uncertainty	49
5.1.1.1.2.	Constraining Post-Processing	50
5.1.1.1.3.	Results	52
5.1.1.2.	Additional Stochastic Simulation of Basal Specific Heat Flow	54
5.1.1.2.1.	Results	54
5.1.2.	Temperature Uncertainty Related to Hydraulic Properties	56
5.1.2.1.	Reservoir Model	57
5.1.2.1.1.	Propagation of the Cooling Front	57
5.1.2.2.	Quantification of Uncertainties in Temperature Variation with Time	59
5.1.2.2.0.1.	Target Histogram and Correlation Length	61
5.1.2.2.0.2.	Variability of Thermal Conductivity	62
5.1.2.2.1.	Results and Discussion	62
5.2.	Quantifying Uncertainties for a Free Reservoir Layout in the Northeastern German Sedimentary Basin	66
5.2.1.	Influence of Salt Domes on Temperature?	68
5.2.2.	The Geometric Model	68
5.2.3.	The Thermal Model	72
5.2.3.1.	Rock Properties for the Thermal Model	73
5.2.3.2.	Determining Specific Heat Flow by Inversion	73
5.2.3.3.	Quantifying Uncertainties of the Temperature Field	74
5.2.4.	Doublet Reservoir Model	75
5.2.4.1.	Hydraulic Rock Properties	76
5.2.4.2.	Results and Discussion for the Doublet Layout	77
5.2.5.	Single-Well Reservoir Model	82
5.2.5.1.	Results and Discussion for the Single-Well Layout	84
5.3.	Discussion	85

6. Appraisal of Methods for a Stochastic Parameter Estimation in a Fractured Crystalline Reservoir	87
6.1. The Geological Framework at Soultz-sous-Forêts	88
6.2. The Circulation Experiment of 2005	89
6.2.1. Motivation for Applying Massive Monte Carlo and EnKF Approaches .	92
6.3. Modeling Approach	92
6.4. Massive Monte Carlo Study of the Fractured Reservoir	94
6.4.1. Identification of Fitting Realizations	95
6.4.1.1. Root Mean Square Error	96
6.4.1.2. Kolmogorow-Smirnow Test	96
6.4.1.3. Normalized Kolmogorow-Smirnow test	97
6.4.2. Modeling Results	97
6.4.2.1. Characteristics of possible flow paths	99
6.4.2.2. Comparing Results with Discrete Fracture Network Approach	102
6.4.2.3. The Influence of Porosity	102
6.4.3. Discussion	106
6.5. Ensemble Kalman Filter Study of the Fractured Reservoir	108
6.5.1. Results and Discussion	109
6.5.2. Updating a Non-Gaussian Distribution	109
6.5.3. Long-Term Performance Prediction	113
6.6. Comparing Bayesian Inversion, Massive Monte Carlo, and Ensemble Kalman Filter	115
7. Summary and Conclusions	118
7.1. Outlook	122
A. Tool Implementation and Parallelization	139
A.1. SGSim	139
A.2. Ensemble Kalman Filter	140
B. Virtual Reality Visualization	141

Chapter 1.

Introduction

Counteracting the increasing global warming is one the greatest challenges of humanity in our generation. Since the reduction of emitted greenhouse gases, such as CO₂, is vital to achieve this goal, sustainable types of energy generation, such as solar or wind, are promising alternatives to fossil based energies or nuclear energy. The contribution of energy from sustainable sources to the energy consumption in Germany was recently increased from 10 % in 2009 to 20 % in 2011 (BMW, 2012). However, their power supply varies with weather and day time. Therefore, storage installations are required to use these sources for base load supply. In contrast, geothermal energy plants yield constant heat or electrical power supply over time. Theoretically, the huge geothermal resource (1200 EJ in Germany as estimated by Paschen et al. (2003)) alone can cover the energy consumption in Germany (13 411 PJ in 2009, BMW (2012)) for the next 90 years. Therefore, geothermal energy as a sustainable source may help to decrease CO₂ emissions (e. g. Rybach, 2010). However, geothermal energy does not yet contribute significantly to the thermal and electric energy production in the world in general or in Germany in particular (Lund et al., 2011; Bertani, 2012).

The risk of project failure is a major obstacle for the future development of geothermal energy. This includes (i) seismic risk during hydraulic injection, (ii) technological risk during drilling and operation, and (iii) the uncertainty of subsurface properties. While the first one has an impact mainly on the acceptance of society for geothermal operations, the latter two may result in damaged equipment (case ii) or insufficient energy exploitation (case iii). This may result from a certain targeted temperature, that is not reached inside a target horizon, or from a required fluid circulation, which can not be established at a sufficient rate. As a result, the investment for the project may be lost partly or totally. In this thesis, I will focus on the risk resulting from the uncertainty in subsurface properties. Accurately quantifying this risk will set the stage for making reliable financial decisions and for designing an optimal geothermal installation. In addition, the operator can avoid oversizing the installation.

As 95 % of the geothermal resource is located in reservoirs at depths below 3000 m (Paschen et al., 2003) — so called medium- (100 °C – 150 °C) and high- (> 150 °C) enthalpy reservoirs

(e. g. Clauser, 2006) — my work focuses on deep geothermal installations. Here, information on subsurface properties is particularly sparse due to the great depth, and parameter estimation and quantifying of uncertainty of special importance (Rybach, 2010). There is a small number of deep installations in Germany. They are located e. g. in Landau (3170 m – 3340 m, 160 °C, Frey & Milles (2007); Bertani (2012)) and in Unterhaching (3350 m – 3590 m, 122 °C, Berger et al. (2009); Bertani (2012)), or projected in Saxony (Wolf & Felix, 2009). This small number shows that the high potential of geothermal energy is not used by now. One reason for this is the uncertain temperature and flow rate before drilling the wells for the installation.

For deep installations, closed systems are available, such as deep engineered heat exchangers. But they are usually not as efficient as open systems because of their limited heat exchange surface. Therefore, I focus on open systems, which, however, require sufficient subsurface fluid pathways. This study considers mostly doublet or triplet systems where water is circulated continuously between two or three wells delivering heat from a geological reservoir.

Large geothermal power plants with significant production of electrical energy are still limited to high-enthalpy systems in countries with natural steam systems and elevated surface heat flow, such as Iceland, Turkey, or Indonesia (Bertani, 2012). In order to make geothermal energy available in regions without natural steam systems, engineered subsurface hot water systems — Hot Dry Rock (HDR) or Enhanced Geothermal System (EGS) — have been tested and installed at a few locations. Actual projects are e. g. the EGS at Soultz-sous-Forêts, France (Gérard et al., 2006), and the Habanero EGS in the Cooper Basin, Australia (Xu et al., 2012).

In an EGS reservoir, the host rock is used as a heat exchanger where pathways are created by hydraulic or chemical stimulation. Hydraulic stimulation or hydraulic fracturing consists of injecting water under high pressure into the subsurface in order to increase the pore pressure within the rock mass. This results in fracturing or opening of pre-existing fractures. If the rock is under shear stress, one face of the fracture will dilate, preventing the fracture to close again after pressure shut-down. Fracturing is indicated by concurrent microseismic events (Tarrahi & Jafarpour, 2012). As a consequence, either a fracture network or a single coin-shaped fracture zone develops, depending on history and stress field of the host rock. This yields an enhanced heat exchange surface which also supports circulation between suitably placed wells. In addition, chemical stimulation can be performed by injecting acids into the fracture system for dissolving minerals, thus increasing hydraulic permeability further.

The use of geothermal energy for an economic electrical power generation requires flow rates of 50 L s^{-1} – 100 L s^{-1} and temperatures of 150 °C – 200 °C (e. g. Clauser, 2006). Therefore, in regions with, for instance, a geothermal gradient of 30 K km^{-1} , EGS need to be implemented at depths below 4000 m for producing the sufficient thermal energy content. Jain et al. (2012) simulated EGS performance numerically and found optimal geometrical layouts for multiple

Chapter 1. Introduction

doublets and triplets assuming a reservoir comprising penny-shaped fractures. Covering a resource base for 500 years and an operation time of 31 years and concerning available areas with crystalline rock, they estimated that the EGS technique could deliver 258 TWh of electric energy and has the potential to deliver 42 % of German power production at any given time.

As micro-seismic events due to hydraulic stimulation may cause damages on buildings at the surface, lacking acceptance of society may stop a geothermal project as was the case in Basel, Switzerland (Häring et al., 2008). Furthermore, insufficient post-stimulation permeability will prevent sufficient fluid circulation. Therefore, reliable information on the response of the host rock to hydraulic stimulation is vital in particular for the development and acceptance of the EGS technique. This response is of interest in particular with respect to the engineered permeability field and its uncertainty. Consequently, I will discuss parameter estimation and quantifying uncertainty also for EGS.

The particular research target for the various studies presented in this thesis is understanding the impact of uncertainty of reservoir properties on geothermal reservoir modeling as well as providing calibrated numerical reservoir models. Traditional ways for quantifying uncertainty, such as accounting for minimum and maximum values of parameters, do not provide useful predictions due to the large ranges of possible predictions. The widely practiced habit of providing only best and worst estimates or even no information on uncertainty at all is little useful for investors or operators of geothermal installations. In contrast, I discuss methods for quantifying the risk of failure due to unknown subsurface properties in a stochastic way. Thus, project failure may be avoided and, in addition, unnecessary monetary cost due to oversized installations or expensive insurances may be reduced significantly.

Calibrated numerical reservoir models obtained by forward and inverse modeling provide powerful tools supporting the reservoir development and operation. The position of the boreholes within a target horizon, its depth and natural temperature, and the circulation rate all influence the variation of reservoir pressure and temperature with time. For example, slow cooling and maintaining reservoir pressure are contradicting operation targets. This results e.g. from decreasing production pressure and simultaneously decreasing arrival time of the cooling front with increasing borehole distance. However, a calibrated numerical reservoir model helps preventing a thermal short circuit between the production and injection wells. In addition, it allows developing strategies for cooling a sufficiently large reservoir volume. Finally, a calibrated reservoir model comprising also information on uncertainty allows for computing production scenarios optimized with respect to the targets profit or sustainability.

This thesis is organized as follows: First, I discuss the involved uncertainties in geothermal reservoir modeling in Chapter 2. The numerical methods and program tools used for the forward and inverse modeling are described in Chapter 3. Testing of the stochastic estimation tools on

Chapter 1. Introduction

synthetic examples is reported in Chapter 4 for sedimentary as well as for hard rocks. Considering uncertainties of rock properties from borehole logging and laboratory measurements on cores and cutting, Monte Carlo techniques can be applied to quantify accurately both, rock property and state variable uncertainties, thus providing best estimates. The transient variations of reservoir states are also propagated in time in a probabilistic way. I present examples of natural porous reservoirs in The Hague, Netherlands, and the Northeastern German Basin in Chapter 5. Both, doublet and single-well concepts are studied here. In Chapter 6, an engineered geothermal system in fractured rock is studied for the EGS test site at Soultz-sous-Forêts based on an equivalent porous medium approach. In this approach, the fractured rock is treated as a porous medium. It assumes laminar flow in the damage zone along a fracture or fault. This is done in contrast to discrete fracture network approaches (see e. g. Gentier et al. (2010a)), where laminar or turbulent fluid flow is simulated in single discrete objects (fractures) within the rock matrix. Data from a tracer circulation test performed in this reservoir in 2005 are used to identify fluid pathways as well as their permeability in order to compare different parameter estimation methods. This comprises (i) gradient-based Bayesian deterministic inversion, (ii) massive Monte Carlo, and (iii) Ensemble Kalman Filter with respect to long-term prediction of the production temperature. Chapter 7 provides the conclusion of this thesis.

In this work, I apply the most suitable numerical method or modeling sequence for each individual problem, in my point of view. The use of virtual reality visualization techniques (Weiskopf & Erlebacher, 2004; Wolter, 2010) allows to explore the characteristics of different 3D realizations of the numerical models, see Appendix B.

This study presents a strategy for an iterative improvement of the information required by a geothermal reservoir model in porous rocks or in engineered fractured systems. Such a model may be updated as different kinds of observations become available, such as logging data, well tests, or even tracer experiments. Additionally, this study may contribute to an improved understanding of hydraulic connectivity resulting from hydraulic stimulation processes in EGS. Thus, future developments of the EGS technique may benefit from the applied inverse approach.

Since methods from geophysical inverse modeling and sequential filtering are presented, the terms *assimilation* and *calibration* are used synonymously in literature. I will use the first term for the process of sequential updating the model. I call the final model after all updates *calibrated*. In petroleum reservoir engineering the search for model parameters is called *history matching*. I.e., finding a suitable set of input parameters which fit reservoir production data of the past. This set can be used to provide a prediction of future reservoir performance. Therefore, *history matching* is an inverse problem. It is applied in this thesis on geothermal reservoirs.

As in most scientific research done today, the work presented was performed by me mostly in cooperation with those researchers shown in the author's list of the papers corresponding to the

Chapter 1. Introduction

individual chapters of this dissertation.

Publications

This dissertation is based on following manuscripts, which are submitted to or already published by scientific journals:

- Vogt, C., Mottaghy, D., Wolf, A., Rath, V., Pechnig, R., and Clauser, C., 2010. Reducing temperature uncertainties by stochastic geothermal reservoir modeling, *Geophysical Journal International*, **181**(1), 321–333. Contribution: synthetic tests of stochastic inverse modeling; quantifying uncertainty based on thermal rock properties; constraining post-processing.
- Mottaghy, D., Pechnig, R., and Vogt, C., 2011. The geothermal project Den Haag: 3D numerical models for temperature prediction and reservoir simulation, *Geothermics*, **40**(3), 199–210. Contribution: quantifying uncertainty based on hydraulic rock properties; constraining post-processing.
- Vogt, C., Kosack, C., and Marquart, G., 2012. Stochastic inversion of the tracer experiment of the EGS demonstration reservoir in Soultz-sous-Forêts — revealing pathways and estimating permeability distribution, *Geothermics*, **42**, 1–12. Contribution: stochastic parameter model; selecting best-fit realizations; identifying model characteristics.
- Vogt, C., Marquart, G., Kosack, C., Wolf, A., and Clauser, C., 2012. Estimating the permeability distribution and its uncertainty at the EGS demonstration reservoir Soultz-sous-Forêts using the ensemble Kalman filter, *Water Resources Research*, **48**, W08517. Contribution: synthetic tests; inverse modeling; interpretation of results
- Vogt, C., Iwanowski-Strahser, K., Marquart, G., Arnold, J., Mottaghy, D., Pechnig, R., Gnjezda, D., and Clauser, C., 2013. Modeling contribution to risk assessment of thermal production power for geothermal reservoirs, *Renewable Energy*, **53**, 230–241. Contribution: creating numerical models; doublet and single-well layouts; quantifying uncertainty based on hydraulic and thermal rock properties.
- Marquart, G., Wolf, A., Rath, V., and Vogt, C., 2012. Stochastic estimate of permeability in geothermal reservoir simulation using the ensemble Kalman filter method, *Geothermics*, submitted. Contribution: models design

Chapter 2.

Uncertainty and Heterogeneity

Uncertainty of subsurface properties is of major importance for the predictions of reservoir performance. Therefore, I consider the thermal power of a geothermal installation which controls the energy delivered over time and, hence, the financial profit (Manzella, 2010). Neglecting any operational losses, the thermal power P_{th} (W) of a geothermal installation is given by

$$P_{th} = (\rho c)_f \cdot \frac{dV}{dt} \cdot \Delta T, \quad (2.1)$$

where $(\rho c)_f$ is the volumetric heat capacity of the fluid ($\text{J m}^{-3} \text{K}^{-1}$), $\frac{dV}{dt}$ is the fluid circulation rate ($\text{m}^3 \text{s}^{-1}$) and ΔT is the fluid temperature drop (K). Here, $\frac{dV}{dt}$ and T are considered uncertain reservoir state variables. Since they depend on reservoir rock properties, e. g. thermal and hydraulic conductivity, it is necessary to quantify these rock properties, too.

Thus, the uncertainties of reservoir state variables are controlled by (i) uncertain geometry and (ii) uncertain rock properties. Addressing (i), I will discuss briefly geometric uncertainties in Chapters 6.4 and 5.2. the reservoir geometry is commonly defined by reflexion-seismics. When identifying geological layers and boundaries, the seismic vertical resolution (above one quarter of the dominant wavelength) is limited typically to 10 m – 60 m (Yilmaz, 1991). Therefore, this source of uncertainty has to be taken into account.

Faults as possible fluid pathways are of special interest in geothermal prospecting. Detecting faults is possible using wavenumber-wavenumber filtering of seismic data based on a Fourier transform of depth-distance profiles (Yilmaz, 1991). Faults and fissures can be identified also using Log-Gabor filters, band filtering for different kinds of image processing (Field, 1987), or a coherence-based algorithm described by Gersztenkorn & Marfurt (1999). In all cases, results can be verified by comparing their fractal distributions with the ones typically found for geologically mapped fault systems (Turcotte, 1997). For geothermal applications, these methods were applied for a sedimentary reservoir by Iwanowski-Strahser et al. (2011), discussed later in this thesis in detail, and for a fault zone in crystalline rock by Szalaiovà (2012).

Papadopoulos et al. (2011) addressed geometric uncertainties by inverting positions of differ-

ent layer boundaries from temperature data using a level-set function. For granite, which is also a target rock in this study, seismic identification of layered structures is particularly ineffective due to the absence of reflecting layers. In this case, inversion techniques are of particular interest for identifying fluid pathways.

However, for the sandstone reservoirs, the geometric model, i.e. the layer boundaries defined by geology and the discussed seismic imaging, is not varied by stochastic modeling and inversions applied in this work.

In this thesis, however, I focus mainly on uncertainties resulting from rock properties. Information on these are usually limited to secondary observations, e. g. from seismics, and very sparse direct measurements in few calibration boreholes. Therefore, observations of reservoir state variables are required, such as temperature, pressure, or concentration of a chemical tracer. Numerical forward modeling predicts the crucial reservoir state variables pressure and temperature before and during the operation of a geothermal installation. By stochastic forward and inverse modeling, I obtain also the uncertainties of pressure and temperature. From measurements of tracer concentration, pressure, or temperature, rock properties such as thermal and hydraulic conductivities, and their heterogeneous spatial distribution as well as their corresponding uncertainties can be estimated using deterministic or stochastic inverse modeling techniques.

To this end, estimates are obtained in a probabilistic framework based on Bayes' theorem (Tarantola, 2004). Here, an *a posteriori* probability density function (pdf) $Posterior(x, t)$ is obtained with respect to location in space x and time t . The function describes a stochastic distribution of rock properties, e. g. the permeability field, based on an *a priori* pdf $Prior(x, t)$ and the likelihood $L(x, t)$, provided by observed data:

$$Posterior(x, t) \propto L(x, t) \times Prior(x, t). \quad (2.2)$$

This relation shows that uncertainties result from measurement errors on the one hand, and from a lack of information on reservoir parameters on the other hand. The first can be addressed easily by the tools applied in the following. However, the latter has by far a larger impact on the modeling results because usually data are sparsely distributed. Measurement errors enter the data, whereas lack of information enters the *a priori* pdf.

Rock property and state variable uncertainties differ in space and time. In this study, uncertainty may be also correlated in space, e. g. due to the nature of a sedimentation process, yielding a smaller correlation length in the vertical than in the horizontal direction.

Transient uncertainties are mainly observed in the reservoir state variables pressure and temperature. For instance, the variation of production water temperature will be increasingly uncertain with time. In this work, the rock properties are assumed to be invariant in time and

hence, the uncertainty of rock properties has no transient dimension. However, engineering of a reservoir, natural seismic disturbance, or chemical reactions due to cold injected water into the reservoir all may affect rock properties, e. g. by reducing or increasing porosity by precipitation or thermal contraction, respectively. In this cases, the transient dimension becomes important and has to be investigated in future studies.

Often, uncertainty is assumed to follow a Gaussian probability distribution. This applies also to this study. However, a bi-modal probability distribution of permeability may be more appropriate, in particular for fluvial sediments or fracture controlled hard rock reservoirs (Zhou et al., 2011). Since the most inverse tools used in this work, such as the Ensemble Kalman Filter, presume Gaussian uncertainty, I study also the effect of a bi-model probability distribution on the estimation in Section 6.5.2.

2.1. A note about simplification, parameterization, and heterogeneity

A general guideline for building physical models, including reservoir models, is to keep the model as simple as possible, but not simpler. This paraphrases well the principle known as Occam's razor (Gauch, 2003). However, over-simplification of models must be avoided as well. Moore & Doherty (2006) demonstrated that simplifications made to obtain a unique model calibration in history matching may yield spurious results. While heterogeneity needs to be taken into account on the one hand, it is essential for all model predictions to be complemented by an estimate of their errors on the other hand. A single history-matched model may be useful, but unlikely sufficient for planning as long as it does not specify the uncertainties. Information on those would allow to estimate the planning risk. The complete solution to a history matching problem therefore requires an assessment of uncertainty in both, reservoir properties and, consequently, predictions of reservoir's performance (Oliver & Chen, 2011).

This approach is recommended e. g. by Gavalas et al. (1976) and Hunt et al. (2007) who stated that inversions based on a large number of parameters yield a better data fit than inversions based on a number of homogeneous zones selected by the modeler. Jorand et al. (2012)¹ stated that taking single values from a lithology or even appropriate mixing laws often yields significant errors. In general, assuming only a small number of parameters underestimates uncertainty, both in model properties and in predicted state variables (e. g. Celaya & Wahr, 1996; Oliver & Chen, 2011).

¹Jorand, R., Vogt, C., Marquart, G., and Clauser, C., 2012. Effective thermal conductivity for heterogeneous rocks based on lab experiments and numerical modeling, in preparation

Chapter 2. Uncertainty and Heterogeneity

Following this philosophy, I try to reach the appropriate level of simplicity by mostly simulating simplified geometric models with heterogeneous stochastic parameter distributions within these geometries.

Chapter 3.

Tools for Characterizing Geothermal Reservoirs and Quantifying Uncertainties

In this chapter, I present basically three different approaches (and related work) for parameter estimation and uncertainty quantification: (i) deterministic Bayesian inversion, (ii) massive Monte Carlo, and (iii) Ensemble Kalman Filter. The latter combines the advantages of the first two approaches to some degree. A comparison of the three inverse modeling approaches is provided at the end of this thesis. All three tools for inverse numerical modeling are implemented together with a forward modeling code which calculates system states from system parameters by simulating fluid flow, species transport, and heat transport in one simulator (SHEMAT-Suite) as a toolbox. SHEMAT-Suite (Rath et al., 2006) was developed from the SHEMAT code (Clauser, 2003).

3.1. State of the Art

Various codes can simulate heat, mass, and species transport in porous media such as subsurface rocks, e. g. SHEMAT (Clauser, 2003), TOUGH (Xu et al., 2006), ECLIPSE (Schlumberger, 1999), or COMSOL Multiphysics[®] (Pryor, 2011). The studies presented are based on the parallized code developed by Rath et al. (2006).

Inverse methods are powerful tools to calibrate numerical models on observed data (Hill & Tiedeman, 2006). In addition to best estimates, the methods applied will also provide information on the uncertainty of the derived parameters. Uncertainties can be quantified and best estimates can be obtained using deterministic gradient-based methods or stochastic methods.

Several numerical methods from both families exist to estimate the properties of the subsurface using data from direct rock property measurements at individual wells, such as stochastic Markov Chain Monte Carlo (e. g. Elfeki & Dekking, 2001) or deterministic Kriging (e. g. Caers, 2005). Other approaches estimate rock properties from observations of state variables by solving the inverse problem (e. g. Mottaghy, 2007; Oliver & Chen, 2011).

Deterministic approaches (e. g. Rath et al., 2006; Tarantola, 2004) have the advantage of converging fast towards a sufficiently close numerical solution. However, if the initial values for the involved parameters are too different from the true values or the solution found represents not the global but only a local optimum, a deterministic approach will yield no or an insufficient solution (Sambridge & Mosegaard, 2002). Stochastic Monte Carlo algorithms can be applied to overcome this obstacle. Additionally, even when applying automatic differentiation (Rall, 1981), the computing power available today limits the gradient-based deterministic approach to a small number of estimated parameters. This may result in an over-simplification of the problem. In contrast, Monte Carlo methods are more suitable to handle heterogeneity because model parameters can be estimated for each cell of the computational grid. Therefore, in this work I focus mainly on the stochastic family of inverse modeling tools.

Wolf (2011) described implementation and parallelization of the deterministic tool in SHEMAT-Suite in detail. As this thesis focuses mainly on Monte Carlo methods, their implementation is described in the Appendix A. Marquart et al. (2012) reported on optimal input parameters for the EnKF for 2D and 3D cases. For a detailed overview of related work concerning deterministic methods, see Rath et al. (2006).

3.1.1. Monte Carlo

Even though structure and properties of the Earth's crust are strongly deterministic, our knowledge on them is limited to indirect (e. g. seismic) measurements or sparse direct measurements, based on mostly very few boreholes. In light of this limitation, the subsurface parameters are treated as randomly distributed according to an observed or assumed distribution. The parameter variances are related to their uncertainties.

Applying the Monte Carlo approach means generating and processing an ensemble of realizations of the subsurface model. All of these realizations must be equally likely in order to reflect reality represented by data, taking into account the available prior geological and petrophysical information. An overview on different Monte Carlo approaches for the inverse problem is given by Mosegaard & Sambridge (2002). A Monte Carlo approach may become the only feasible method when the problem is highly dimensional. All other numerical methods have an absolute error that decreases no faster than $N^{-1/M}$ for N samples in M -dimensional space, whereas the absolute error of the Monte Carlo method decreases with $N^{-1/2}$, i.e. independently of the dimension of the space of possible samples (e. g. Fishman, 1996).

Monte Carlo methods can be used for estimating parameters and their uncertainty (represented by the spread of samples) or to identify an optimal set of model parameters. The latter may be viewed as special case of the more general sampling. This study applies a version of the Monte Carlo method called Sequential Gaussian Simulation (SGSim). Its significant advantage is that it

can estimate a large number of parameters. Therefore, estimation is not only possible for a mean parameter value of a geological unit but for the entire heterogeneous parameter field. Moreover, applying this approach inherently quantifies the parameter uncertainty.

Stochastic numerical techniques for estimating heterogeneous rock properties in the whole reservoir from data samples, so-called geostatistical algorithms, are presented e. g. by Caers (2005) or Chiles & Delfiner (1999). Various software tools for geostatistical estimation of rock properties are available (Remy, 2005; Paradigm, 2011). They are based on the widely-used geostatistical software library GSLIB (Deutsch & Journel, 1998). Geostatistical algorithms have been applied successfully in research on hydrogeology (e. g. Nowak, 2005; Cooley, 2004; Kitanidis, 1997), nuclear waste repositories (Neuman & Wierenga, 2002), and hydrocarbon reservoir characterization (e. g. Li et al., 2008; Campos, 2002).

Uncertainty estimation based on Monte Carlo techniques is described by (e. g. Robert & Casella, 2004). On overview Monte Carlo simulation for solving geophysical inverse problems is reported by Sambridge & Mosegaard (2002). More references in this field are given in the following chapter. In geothermal reservoir characterization however, to the best of my knowledge, geostatistical algorithms have been used rarely although they are promising for uncertainty quantification and hence risk estimation.

3.1.2. Ensemble Kalman Filter

Similar to a massive Monte Carlo approach, the Ensemble Kalman Filter (EnKF) is a stochastic technique based on a forward propagation of an ensemble of realizations. At successive instants in time, different kinds of data measured in various boreholes are collected in one data vector and used to update the system variables in a least-squares sense. This way, the match between observation and simulation is improved. The statistics of the ensemble yield the optimal parameter estimate (ensemble mean) and its uncertainty (ensemble standard deviation). As a stochastic data assimilation method, the EnKF combines the advantages of classical Monte Carlo techniques, such as simplicity, inherent measure of uncertainty, and the suitability for highly dimensional problems and parallelization, with fast convergence. As a drawback, the EnKF is only useful for transient problems. In these cases, however, the EnKF is a promising and sophisticated tool for geothermal reservoir parameter estimation.

The EnKF was used widely for predicting the behavior of complex physical non-linear systems and their uncertainty. It is based on the original Kalman Filter (Kalman, 1960), which is the optimal data assimilation method for linear dynamics with additive model errors. It turns out to be the best Bayesian estimator if the model as well as the observations are assumed to be perturbed with Gaussian error statistics (Cohn, 1997). The extension of the Kalman Filter method towards non-linear dynamics is called Extended Kalman Filter. As a drawback, it suf-

fers from huge computational effort. To overcome this problem, the Ensemble Kalman Filter was introduced by Evensen (1994). It turns out that the EnKF can be applied to large, non-linear numerical problems with Gaussian error statistics. The formalism is introduced in Section 3.5.

In the past, the EnKF method has been developed particularly in meteorology and oceanography (Evensen, 2003; Houtekamer & Mitchell, 1998). It has been successfully applied also in hydrogeology (e. g. Herrera, 1998; Reichle et al., 2002; Hendricks Franssen & Kinzelbach, 2008) and hydrocarbon reservoir simulation (e. g. Chen & Zhang, 2006; Gu & Oliver, 2006; Dong et al., 2006; Lawniczak et al., 2008; Krymskaya et al., 2009) mostly in generic studies on synthetic data. Seiler et al. (2011) used the EnKF to estimate the permeability field of a real oil reservoir beneath the North Sea. Jointly with reservoir porosity and permeability, they estimate a sealing factor between 0 and 1, indicating whether faults inside their model are open, closed, or partially sealed. An overview on recent development of the EnKF in history matching is given by Aanonsen et al. (2009) and Oliver & Chen (2011). Recently, Tarrahi & Jafarpour (2012) used the EnKF to characterize EGS permeability based on microseismic events. However, to the best of my knowledge, neither the EnKF nor similar sophisticated data assimilation methods have been used yet for calibrating non-synthetic reservoir models for geothermal energy production.

Moradkhani et al. (2005b) used the EnKF for sequentially estimating model parameters and state variables. Chen & Zhang (2006) reported that state variable data are more effective for estimating formation properties and transport characteristics than rock property data from limited locations. Vrugt et al. (2005) applied a combined global optimization for parameters with sequential data assimilation to estimate model parameters and state variables simultaneously. Nowak (2009) presented a purely parameter-based implementation of the EnKF called Kalman Ensemble Generator in contrast to the previous parameter- and state variable-based implementations. He used his code to investigate tracer propagation in a synthetic test case. This approach is not followed as I wish to update the hydraulic head, tracer concentration, and temperature in the models simultaneously.

For joint Gaussian state variable and observation statistics, the EnKF ensemble update approximates the best estimate of mean and covariance of the Gaussian posterior probability distribution in the Bayesian sense. Non-Gaussian error statistics can be handled by the more general particle filters (e. g. Ristic et al., 2004) which account for the full probability density function in the parameter space. Therefore, these filters are promising alternative methods to the standard EnKF. Where the EnKF updates the ensemble members themselves, particle filters update the weights of the single ensemble members (particles) without changing them. Particle filters have been already applied successfully in groundwater modeling (Moradkhani et al., 2005a; Rings et al., 2010; Montzka et al., 2011). Note that the EnKF represents a special particle filter for Gaussian error distributions. However, the application of particle filters is usually limited

to estimating only a few unknown parameters at the cost of integrating a very large ensemble (typically more than 10^4 realizations, Evensen (2009)) in simulation time.

Jafarpour & Tarrahi (2011) studied the impact of wrong prior knowledge on the EnKF update based on uncertain variogram parameters and, hence, based on a uncertain spatial correlation. They found that over-estimating the uncertainty in a reservoir avoids convergence of the EnKF towards a non-optimal solution. They realized this by accounting for a number of different correlation lengths instead of just one when creating the initial ensemble. This approach results in better estimates. To overcome the obstacles resulting from poor prior knowledge, Krymskaya et al. (2009) proposed the Iterative Ensemble Kalman Filter. It shows no advantage over the EnKF when the initially guessed model parameters are close to the real ones. But it is significantly better in case of insufficient or wrong prior information. Here, the filter is repeated globally after the data assimilation, replacing the initial guess with the estimated model state in each iteration, until the estimated model state is not changed any more significantly. Gu & Oliver (2006) also obtained improved results with an iterative approach for a multiphase-flow system. Hendricks Franssen & Kinzelbach (2008) tested and discussed an iterative EnKF approach for single-phase saturated groundwater flow. In contrast, they did not find any advantage for the iterative method in this case, possibly due to weaker non-linearities in their application. I do obtain significantly better results with an iterative approach which are presented in the following chapters.

In contrast to the Monte Carlo-based EnKF, deterministic ensemble-based data assimilation approaches were introduced under the name Ensemble Square Root Filters (Tippett et al., 2003). They show a significantly better performance over the EnKF for small ensembles, but lack the EnKFs simplicity and its capability to perform realistic applications where the system dimensions exceeds the ensemble size. Sakov & Oke (2008) proposed a simple modification to the EnKF called Deterministic Ensemble Kalman Filter. It showed to be a linear approximation of the Ensemble Square Root Filter, which combines the simplicity, numerical effectiveness, and flexibility of the EnKF with the performance of the Ensemble Square Root Filters. However, both approaches assume unperturbed observations. Therefore, the EnKF must be used in order to account for measurement errors and benefit from lowest computational effort of all filter methods mentioned. This applies in particular when a filter inbreeding is avoided using a large number of realizations. Filter inbreeding is a possible filter divergence caused by a loss of ensemble variability. As a consequence, I apply the original EnKF in this study. The EnKF algorithm is implemented as a module into the forward modeling tool SHEMAT-Suite (Rath et al., 2006).

With a sufficiently large number of stochastic realizations, the EnKF can handle uncertainties deriving from different sources simultaneously (Hendricks Franssen & Kinzelbach, 2008). In a

massive Monte Carlo approach, in contrast, and if the influence of the uncertain parameter on the forward simulation is sufficiently strong, the number of realizations required for exploring the probabilistic space increases dramatically with the number of uncertainty sources. Nevertheless, realizations generated using Sequential Gaussian Simulation form an appropriate initial ensemble for the EnKF algorithm.

Compared to other methods which can address heterogeneity, such as pilot-point or self-calibration (see Hendricks Franssen et al., 2009), the EnKF is supposed to be simpler, to require less computing time, to avoid re-calibration with new data available, and to handle easily multiple sources of uncertainty (Zhou et al., 2011).

3.2. Forward Code

Fluid flow through a porous medium is described by Darcy's law, (Darcy, 1856):

$$\mathbf{v} = -\frac{\mathbf{k}}{\mu_f}(\nabla p + \rho_f g \nabla z), \quad (3.1)$$

where \mathbf{v} is specific discharge (or Darcy velocity) ($\text{m}^3 \text{m}^{-2} \text{s}^{-1}$), \mathbf{k} the tensor of hydraulic permeability (m^2), μ_f fluid dynamic viscosity (Pa s), ρ_f fluid density (kg m^{-3}), g gravity (m s^{-2}), p fluid pressure (Pa), and z (m) the vertical space coordinate (positive upwards).

In the following, hydraulic head h (m) is used instead of pressure. The corresponding fluid pressure can be calculated from the variation of fluid density and head with depth as (e.g. Marsily, 1986; Clauser, 2003):

$$p(z, h) = p_0 + \int_z^0 \rho_f(\tilde{z}) g(h - \tilde{z}) d\tilde{z}, \quad (3.2)$$

where $p_0(z = 0) \approx 10^5 \text{ Pa}$ is the atmospheric pressure at the Earth's surface $z = 0$.

The equation for fluid flow implemented here is derived from equations (3.1) and (3.2), and the equation of continuity, using an Oberbeck-Boussinesq approximation (e.g. Diersch & Kolditz, 2002; Clauser, 2003):

$$\rho_f g (\alpha + \phi \beta) \frac{\partial h}{\partial t} = \nabla \cdot \left[\frac{\rho_f g \mathbf{k}}{\mu_f} (\nabla h + \rho_r \nabla z) \right] + W. \quad (3.3)$$

Here, ϕ is porosity while α and β denote the compressibility (Pa^{-1}) of the rock and the fluid phase, respectively. W corresponds to a mass source term (s^{-1}). Relative density ρ_r is defined by $\rho_r = \frac{\rho_f - \rho_0}{\rho_0}$, where ρ_0 is density at reference conditions (i.e. temperature and pressure).

The heat transport equation follows in an analog way from Fourier's equation $\mathbf{q} = \lambda_e \nabla T$ and

the conservation of energy (e. g. Clauser, 2003; Beardsmore & Cull, 2001):

$$(\rho c)_e \frac{\partial T}{\partial t} = \nabla \cdot [\boldsymbol{\lambda}_e \nabla T] - (\rho c)_f \mathbf{v} \cdot \nabla T + H. \quad (3.4)$$

\mathbf{q} is specific heat flow (W m^{-2}). T is temperature ($^{\circ}\text{C}$), $(\rho c)_e$ and $(\rho c)_f$ are the effective volumetric heat capacities of the fluid-saturated porous medium and the fluid ($\text{J m}^{-3} \text{K}^{-1}$), respectively. $\boldsymbol{\lambda}_e$ is the tensor of effective thermal conductivity of the fluid saturated porous medium ($\text{W m}^{-1} \text{K}^{-1}$), and H a heat generation rate source term (W m^{-3}). Effective values are appropriate averages defined for the fluid-rock system, for instance the arithmetic mean for specific heat capacity, and the geometric or square-root means for thermal conductivity (e. g. Clauser, 2011).

The transport of an inert tracer species is given by the transient species transport equation (e. g. Clauser, 2003):

$$\phi \frac{\partial C}{\partial t} = \nabla \cdot [\mathbf{D} \nabla C] - \mathbf{v} \cdot \nabla C. \quad (3.5)$$

Here, C denotes the concentration of the dissolved species (mol m^{-3}) and \mathbf{D} the hydrodynamic dispersion tensor ($\text{m}^2 \text{s}^{-1}$), defined by a combination of transversal and longitudinal dispersion lengths and a molecular diffusion coefficient. To keep the approach simple, \mathbf{D} is regarded as a scalar. No discrimination is made between transversal and longitudinal dispersion lengths, and molecular diffusion is neglected as it is usually orders of magnitude smaller than hydrodynamic dispersion. Thus, dispersion is represented by $D = \alpha_d |\mathbf{v}|$, where α_d is the isotropic dispersion length.

The physical properties of rock matrix and fluid (and thus the Darcy velocity, equation 3.1) vary with temperature and pressure. These functions provide a non-linear coupling between the flow equation (3.3) and the heat transport equation (3.4). For the studies described in this thesis, both, thermal conductivity and hydraulic permeability are assumed to be isotropic within one grid cell, that is, they are treated as scalars λ and k , respectively. Non-linearities resulting from fluid and rock properties and their pressure and temperature dependences are implemented according to Rath et al. (2006), where further details on the forward code SHEMAT-Suite can be found.

Note that the linear relationship in equation (3.1) is valid for laminar flow only. In close vicinity to a production or injection borehole, turbulent flow may occur, resulting in spurious pressure or filtration velocities. This has to be taken into account when comparing observed pressure in real reservoirs and simulated ones. In this thesis, no reality pressure data are used for inversions.

3.3. Gradient-based Bayesian Deterministic Inversion

A straightforward Bayesian inverse approach (Tarantola, 2004) was chosen here, which is based on a parameter set minimizing the nonlinear functional:

$$\Theta = [\mathbb{D} - \Psi^f(\mathbb{P})]^T C_{\mathbb{D}}^{-1} [\mathbb{D} - \Psi^f(\mathbb{P})] + (\mathbb{P} - \mathbb{P}^0)^T C_{\mathbb{P}}^{-1} (\mathbb{P} - \mathbb{P}^0). \quad (3.6)$$

Here, $\Psi^f(\mathbb{P})$ is the vector of data simulated by the forward model based on the parameter vector \mathbb{P} , and \mathbb{D} is the observed data vector. *A priori* values for parameters \mathbb{P}^0 and inverse covariance matrices $C_{\mathbb{D}}^{-1}$ and $C_{\mathbb{P}}^{-1}$ need to be specified beforehand. A minimum of the functional in equation 3.6 is called the maximum *a posteriori* estimate and can be found by a quasi-Newton iteration:

$$\mathbb{P}^{k+1} = \mathbb{P}^k + \alpha \cdot (J^T C_{\mathbb{D}}^{-1} J + C_{\mathbb{P}}^{-1})^{-1} \cdot \{J^T C_{\mathbb{D}}^{-1} [\mathbb{D} - \Psi^f(\mathbb{P}^k)] - C_{\mathbb{P}}^{-1} [\mathbb{P}^k - \mathbb{P}^0]\}, \quad (3.7)$$

where the Jacobian $J = \frac{\partial \Psi^f}{\partial \mathbb{P}}$ refers to the matrix of partial derivatives of p and T with respect to the parameters and α is a damping parameter. The implementation includes data and parameter space formulations.

The model Jacobian J required for the inversion is obtained by automatic differentiation of the forward code (e. g. Rall, 1981; Griewank, 2000), using the tool ADIFOR (Bischof et al., 1992, 1996).

3.4. Sequential Gaussian Simulation and Constraining Post-Processing

The ensemble of realizations of a rock property distribution is synthesized using the Monte Carlo algorithm of Sequential Gaussian Simulation (SGSim) (Deutsch & Journel, 1998). In this approach, multiple parameter realizations of the same geometrical reservoir model are generated. Each realization corresponds equally likely to the real situation defined by data.

To this end, SGSim uses a Kriging interpolation technique for spatially distributed data (e. g. Deutsch & Journel, 1998). Kriging stands for a family of least-square regression methods yielding optimal estimates of the target parameters. The estimator $\hat{\zeta}_0$ for a parameter at a certain point x_0 is written as linear combination of the N related values ζ at the measuring points x_i :

$$\hat{\zeta}_0(x_0) = \sum_{i=1}^N \nu_i \zeta(x_i). \quad (3.8)$$

The weights ν_i are calculated by minimizing — in a least-squares sense — the residual R_0 between estimator $\hat{\zeta}_0$ and unknown true value ζ_0 , i.e., $R_0 = (\hat{\zeta}_0 - \zeta_0)^2$.

For each node, Kriging returns a mean value, the *Kriging mean*, and an average squared error called *Kriging variance*. Approximating the variogram by suitable weighting of the available data, *Kriging* accounts for the underlying spatial data correlation. The variogram represents a measure for the dissimilarity of spatial data as a function of the separation of pairs of data (Deutsch & Journel, 1998).

I assume spherical variograms (Deutsch & Journel, 1998) as the simplest Kriging variogram type given by:

$$\gamma(r) = S \cdot \left(\frac{3}{2} \frac{r}{a} - \frac{1}{2} \frac{r^3}{a^3} \right) \quad \text{for } r \leq a \quad (3.9)$$

$$\gamma(r) = S \quad \text{for } r > a. \quad (3.10)$$

Here, S denotes the so-called sill, a threshold value, r the distance, and a the range when S is reached, the correlation length.

The SGSim algorithm assigns a rock property to each grid point of a single realization according to a probability density distribution as follows: (i) the geometrical structure of a reservoir model is discretized on a discrete grid; (ii) the rock property distribution is transformed into a Gaussian shape of zero mean and unit variance; (iii) the algorithm follows a random path through the model; (iv) for each grid node, neighbouring data and already simulated nodes are used for a Kriging interpolation (Caers, 2005) of the target property; (v) a property value drawn randomly from the distribution defined by the *Kriging mean* and *variance* is assigned to the node. Therefore, these values account for (a) assumed or observed property distributions, (b) correlation length, (c) primary data, such as borehole measurements, (d) secondary data, such as seismic data; (vi) a realization is completed after property values have been assigned to all nodes of the model; (vii) the rock property distribution assigned to the model is transformed back from a Gaussian into the original distribution; (viii) more realizations are created by following other random paths. Each of these realizations is equally likely with respect to the real situation defined by the data (Figure 3.1).

The state variables p , T , and C are obtained for each realization of a rock property distribution from a forward simulation of fluid flow as well as heat and tracer transport. This does not provide only average values and error estimates of a target rock property or state variable at any location in the geothermal reservoir, but also their local probability distribution. Additionally, the approach allows the stochastic simulation of boundary conditions.

The SGSim algorithm is integrated as a module into a fluid flow and heat transport simulator (Rath et al., 2006). This way, the generated realizations can be used directly as input for the fluid

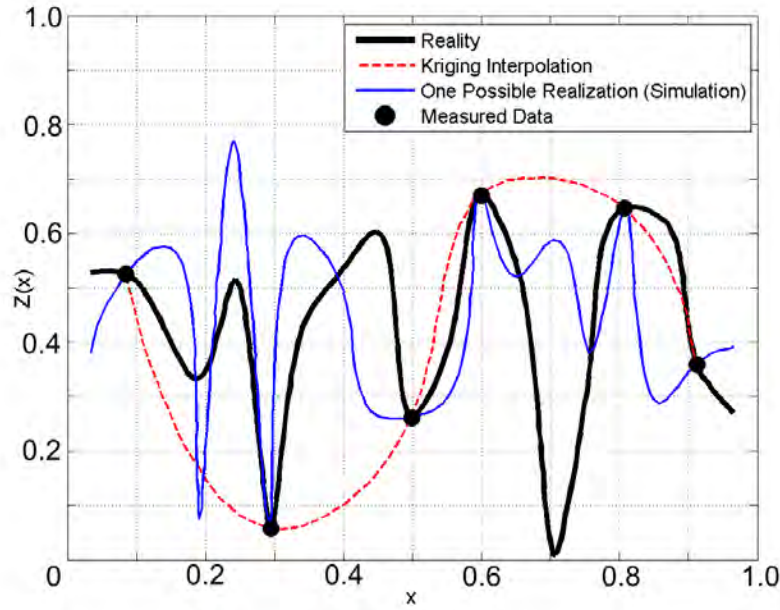


Figure 3.1.: 1D example for a geostatistical parameter estimation showing the *Kriging* interpolation technique and a realization generated with SGSim (after Schafmeister, 1999).

flow and heat transport simulations, without any format conversion. Further, already existing models implemented in the flow simulator can be updated easily with a stochastic simulation.

3.4.1. Constraining Post-Processing

Compared to hydrocarbon reservoirs, geothermal reservoirs are usually explored by only few boreholes. Therefore, additional information is used here in order for constraining the stochastic results and hence, minimizing the uncertainty further, or finding optimal solutions. To this end, observations of state variables such as temperature, hydraulic head or tracer concentration can be used to identify the best fitting realizations of a massive Monte Carlo approach. That is, the single realizations are ranked according to an objective function and realizations with large misfit are discarded from the ensemble in a constraining post-processing.

3.5. Estimation of State Variables and parameters using the Ensemble Kalman Filter

Like the original Kalman Filter, the Ensemble Kalman Filter in principle consists of three major parts:

1. A state and parameter vector is identified. For my applications this is the entire uncertain parameter and state variable space of a geothermal reservoir model.
2. A forward-in-time step is computed via simulation of fluid flow, heat and species transport in order to generate a model prediction.
3. An analysis step is performed comprising an update the state and parameter vector by minimizing in a least-square sense the difference of the forecast, obtained from the forward step, and available observation data describing the actual state. In the cases studied, the objective of the filter is finding optimal thermal or hydraulic parameters with regard to the observed data by varying the state variables and parameters in each grid element of a zone of interest inside all model realizations from the initial ensemble.

These parts are addressed for each forward-in-time step of the simulation. A new state and parameter vector, and hence, updated reservoir parameters, are obtained in each analysis step assimilating data when they come available. Therefore, the Kalman Filter is a sequential filter method. During this procedure, the problem's error covariance matrix needs to be calculated for performing the analysis step. Since the Kalman Filter is limited to linear problems, the Extended Kalman Filter was derived to address non-linear problems. It extends the original Kalman Filter using a tangent linearization (Jacobian) to calculate the error covariance matrix.

However, the Extended Kalman Filter has two fundamental disadvantages: (i) in order to identify a correlation matrix a Jacobian needs to be computed in the analysis part, at large computational cost; (ii) high-order statistic moments are not accounted for due to the performed linearization. In contrast, the EnKF uses a large ensemble of state and parameter vectors. In this study, the ensemble members are identified by different realizations of the subsurface model. In contrast to a single state and parameter vector propagated in time, here the state and parameter vector ensemble represents the probability density of the state variables and parameters. Therefore, the EnKF yields the error covariance matrix from the ensemble of state variables and parameters provided that the ensemble is sufficiently large. Hence, the Jacobian needs not be computed avoiding associated errors and time effort.

The full probability density function of the stochastic problem would be described by an infinite ensemble of model state variables and parameters. A finite ensemble size N therefore

yields an approximation of the probability density function. The corresponding sampling error decreases with $\frac{1}{\sqrt{N}}$ (Evensen, 2003). As a Monte Carlo technique, the EnKF can inherently avoid local minima in the optimization function to a certain degree. Moreover, the EnKF is very suitable for parallelization because the forward computations of the different realizations are independent of each other. Only during the analysis step when data are assimilated, information is required from all individual realizations of the entire ensemble. Then, all simulation progresses in the parallel computational threads need to wait until the last ensemble member is fully computed.

3.5.1. Formalism

I present here only a brief summary of the EnKF since it and its mathematical background for solving stochastic differential equations has been well documented in literature (e. g. Evensen, 2003). The probability density of the model states and parameters is represented by an ensemble of combined state and parameter vectors $\Psi = [\Psi_1, \dots, \Psi_N]$ generated with the SGSim algorithm (see Section 3.4). In contrast to Section 3.3, Ψ contains now not only states, but also rock properties. All of the N ensemble members reflect the true state variables and parameters equally likely according to all prior knowledge (defined by data and errors). The state and parameter vector contains the model state variables (hydraulic head, temperature and tracer concentration, in the forward equations denoted by h , T , and C , respectively) and the parameters (permeability denoted by k in the forward equations). Therefore, the parameters are updated according to the state variables. As before, k is assumed as isotropic in each grid cell and is, hence, a scalar.

$\Psi_{k,i}^f$ and $\Psi_{k,i}^a$ represent the forecast of the state variables (obtained by performing a forward step i in time), and the updated state variables and parameters after the analysis step (see equation 3.13 below), respectively. The index of an ensemble's realization is denoted by k . The two vectors are linked by the following equation:

$$\Psi_{k,i}^f = F(\Psi_{k,i-1}^a) + e_{k,i}^f. \quad (3.11)$$

Here, the forward operator F integrates the model from one time step to the next one. In the applications studied here, it is represented by the flow and transport equations described in Section 3.2. The uncorrelated Gaussian uncertainty of the prediction model is represented by $e_{k,i}^f$.

3.5.1.1. The Analysis Step

The analysis step requires a number of observations which characterize the true model state. Observations are assimilated by the EnKF in the time step when they become available. The vector of observations $\mathbb{D}_{k,i}$ is linked to the true model state Ψ_i^t by the measurement operator H .

This operator yields the grid cells of the model where data are observed. For the applications in this study, most elements of measurement operator H are equal to zero, elements equal to one are located on the diagonal and identify measurements. The measurement error is denoted by $e_{k,i}^{\mathbb{D}}$ individually for each realization. Here, $\mathbb{D}_{k,i}$ contains the hydraulic head or tracer concentration data or both data sets, taken from a synthetic reference realization or observed in the real reservoir. Perturbed data for a reference is used to evaluate the performance of the EnKf.

$$\mathbb{D}_{k,i} = H\Psi_i^t + e_{k,i}^{\mathbb{D}}. \quad (3.12)$$

Generating perturbed observations by adding error $e_{k,i}^{\mathbb{D}}$ with assumed or given statistics is required for obtaining not only the correct mean but also the correct variance of the updated ensemble. Burgers et al. (1998) reported that otherwise a variance too small would make the system diverge finally, in particular for a small ensemble size. This type of divergence is often referred to as filter inbreeding in the literature (e. g. Evensen, 2003).

Using the perturbed observations, the actual analysis step is performed by computing the Kalman Filter analysis equation for each ensemble member:

$$\Psi_{k,i}^a = \Psi_{k,i}^f + \alpha K_i (\mathbb{D}_{k,i} - H\Psi_{k,i}^f). \quad (3.13)$$

This filtering step may be damped by the factor α for obtaining stable convergence (Hendricks Franssen & Kinzelbach, 2008) with typical values of $0.1 \leq \alpha \leq 1.0$. The so-called Kalman gain K_i for the i -th time step follows from a minimization of the *a posteriori* error covariance in a least-square sense and is defined by

$$K_i = \frac{C_{\mathbb{S},\mathbb{P},i}^f H^T}{H C_{\mathbb{S},\mathbb{P},i}^f H^T + C_{\mathbb{D},i}}, \quad (3.14)$$

where $C_{\mathbb{D},i}$ is the error covariance matrix of the observations arising from the measurement errors $e_{k,i}^{\mathbb{D}}$. $C_{\mathbb{S},\mathbb{P},i}^f$ is the error covariance matrix of the model ensemble of state variables \mathbb{S} and parameters \mathbb{P} calculated from

$$C_{\mathbb{S},\mathbb{P},i}^f \approx \frac{1}{N-1} \sum_{k=1}^N [\Psi_{k,i}^f - \overline{\Psi}_i^f][\Psi_{k,i}^f - \overline{\Psi}_i^f]^T, \quad (3.15)$$

using the ensemble mean

$$\overline{\Psi}_i^f = \frac{1}{N} \sum_{k=1}^N \Psi_{k,i}^f, \quad (3.16)$$

which is considered the best estimate for the unknown true states and parameters. The error of

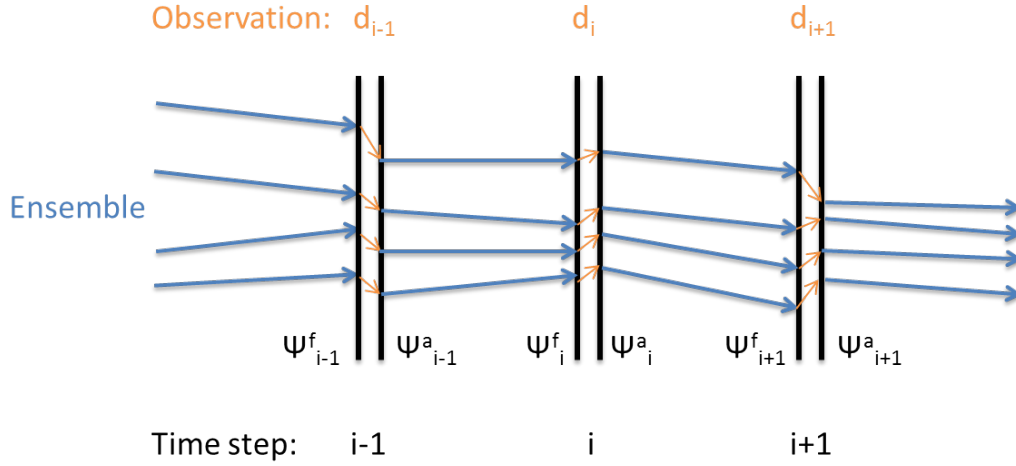


Figure 3.2.: Schematic representation of the EnKF data assimilation process during three analysis steps (after Moradkhani et al., 2005b).

this approximation will decrease with the ensemble size N as $\frac{1}{N}$. Therefore, the error statistics are represented adequately by this method if an appropriate ensemble size is used.

The posterior error covariance matrix of the analysis is now given by

$$C_{\mathbb{S},\mathbb{P},i}^a = [I - \alpha K_i H] C_{\mathbb{S},\mathbb{P},i}^f, \quad (3.17)$$

where I is the identity matrix. Thus, the ensemble spread, and hence, uncertainty is reduced by assimilating increased information from measurements — based on the Kalman gain (the “trust” in the data, see the following Section 3.5.1.2) and the grid cells with data available. Note that the EnKF update is optimal only for Gaussian model and measurement error statistics. However, this ideal situation is never realized in the given non-linear models. Therefore, the results obtained here are always smoothed (as discussed in Chapter 6.5.2). Figure 3.2 shows the data assimilation process schematically.

For my applications, the estimated ensemble mean of the tracer concentration curve is far away from the observed one. Therefore, I apply multiple iterations of the EnKF following the approach of Krymskaya et al. (2009). The underestimation of ensemble variance coming with the application of the iterative approach is compensated at least partially by the applied damping of the Kalman gain (here by a factor 0.2). In a synthetic case, global iterations are stopped when no further improvement of the estimate is obtained. However, there is no control about that in a realistic case. Therefore, for a real world reservoir, iterations are stopped when no significantly better match between simulated and observed data (here: tracer) is obtained.

3.5.1.2. A closer inspection of the Kalman gain

The Kalman gain from eq. 3.14 can be interpreted as a measure of "trust" in the observations with regard to the quality of the model prediction during the analysis step. Inspection of equation 3.14 allows to identify two special cases:

$$\lim_{C_{\mathbb{D},i} \rightarrow 0} K_i = \frac{1}{H} \quad (3.18)$$

and

$$\lim_{C_{\mathbb{S},\mathbb{P},i}^f \rightarrow 0} K_i = \frac{H^T - HH^T}{C_{\mathbb{D},i}} = 0^1. \quad (3.19)$$

The equations show that the Kalman gain weights the the residual within the parentheses in equation 3.13 stronger and hence "trusts" the actual observations more if the error covariance matrix of the observations approaches zero (the state and parameter vector is updated with the residual term). On the other hand, it weights the residual less heavily, which means it "trusts" the observations less if the *a priori* error covariance matrix of the model prediction $C_{\mathbb{S},\mathbb{P},i}^f$ approaches zero (the state and parameter vector is not updated).

3.6. Ensemble Sizes

Massive Monte Carlo studies in sedimentary reservoirs such as the ones presented here require prior information on rock parameters. Therefore, the probability density distribution for the studied rock property used in the SGSim algorithm is sampled from well constrained log and laboratory data. Hereby reasonable results are justified even though an only coarsely sampled *a posteriori* probability distribution is obtained from a relatively small number of realizations (400 - 10 000). This is supported by the reasonably Gaussian shapes of histograms of temperature at the sites studied, see e. g. Chapter 5.1.1.1.3.

Also for transient simulation runs, the ensemble size showed Gaussian shape during the whole simulation time.

For the Ensemble Kalman Filter applications, parameter space is explored more sophisticatedly. Here, the Monte Carlo ensemble is used for determining a state variable and parameter correlation matrix. Previous studies, e. g. Chen & Zhang (2006), reported that a smaller ensembles size (> 100) is sufficient for the EnKF compared to the massive Monte Carlo approach. In particular, they found only 200 realizations sufficient for a 3D two-phase petroleum reservoir study. Nevertheless, larger ensemble size avoids filter divergences. Therefore, a ensemble size of 300 – 880

¹Note that the equation is evaluated for a matrix H containing just zeros and ones on the diagonal

is used for large problems. This ensemble size appears to be sufficient to estimate the reservoir standard deviation of the permeability correctly; because the deviations of the ensemble mean from synthetic reference fields are in the same range as the ensemble variance. An ensemble of about 500 realizations was also found sufficient by Hendricks Franssen & Kinzelbach (2008) for a hydro-geological model. Ensembles even larger can not be computed feasibly given the computing power available (see Appendix A).

In both cases, massive Monte Carlo and EnKF, the number of free parameters of the problem ($> 20\,000$) exceeds the number of realizations by far. Hence, the parameter space cannot be explored rigorously. However, the fact that spatial correlation exists between the parameters allows suitable fits given the computed ensemble sizes. Spatial correlation results in similar values in a certain area defined by correlation length. Therefore, an inverse method does not need to estimate each single grid value and, hence, requires less ensemble members. An additional argument for relatively small ensembles is that the individual realizations are characterized by zones of high and low parameter values. This smooths out for the ensemble mean, but provides a sufficient ensemble variance for data fits.

Chapter 4.

Synthetic Test Scenarios

In this chapter, I evaluate with 2D and 3D test scenarios the value of the stochastic tools for estimating hydraulic or thermal rock properties and capturing their spatial heterogeneities as well as the associated uncertainties. As synthetic tests of the gradient-based Bayesian inversion were already reported in Rath et al. (2006), I focus on the stochastic methods, massive Monte Carlo and Ensemble Kalman Filter.

4.1. Test of massive Monte Carlo Inversion on a Synthetic 3D Model

To test the stochastic Monte Carlo approach, a synthetic model is created and later a constraining post-processing applied to select the realizations which fit temperature data best. Results are compared using, on the one hand, SGSim with varying number of wells and, on the other hand, unconditioned SGSim with constraining post-processing.

A small model (Table 4.1) is generated comprising $12 \times 12 \times 12$ cells and four geological units. This corresponds to a total model size of $1.2 \text{ km} \times 1.2 \text{ km} \times 1.2 \text{ km}$. Using SGSim, an individual Gaussian distribution of thermal conductivity λ is simulated in each unit. In one central permeable layer, a bimodal logarithmic distribution of permeability k is modeled with values between 10^{-16} m^2 and 10^{-12} m^2 . A flow corresponding to a hydraulic head difference of 30 m passes through the model in the y-direction. This high value is chosen to emphasize the effects of the hydraulic properties on the temperature field. This results in horizontal 1D-Péclet numbers $Pe_y > 40$ for each realization. For the vertical direction it is defined for as:

$$Pe_y = \frac{q_{adv}}{q_{cond}} = \frac{(\rho c)_f v_y Y}{\lambda_e}, \quad (4.1)$$

where q_{adv} and q_{cond} are the amount of heat transported by advection and conduction, respectively, v_y is the Darcy velocity in vertical direction (pointing to the surface), and Y defines a

Table 4.1.: Properties of the synthetic 3D model.

Parameter	Value
Mesh size	$12 \times 12 \times 12$
Resolution	$100 \text{ m} \times 100 \text{ m} \times 100 \text{ m}$
Permeability	$10^{-16} \text{ m}^2 - 10^{-12} \text{ m}^2$
Bulk thermal conductivity	$1 \text{ W m}^{-2} \text{ s}^{-1} - 5.5 \text{ W m}^{-2} \text{ s}^{-1}$
Head difference	30 m
Temperature at top	11°C
Basal specific heat flow	80 mW m^{-2}
Stochastic parameters	permeability, bulk thermal conductivity

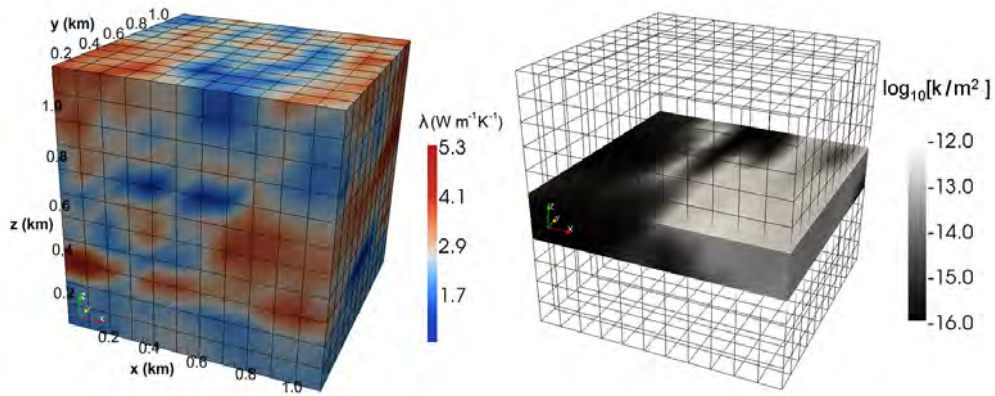


Figure 4.1.: The reference model: Single realization of thermal conductivity λ (left) and permeability k of a permeable layer (right) generated by the SGSim algorithm. The wire frame is used to emphasize the permeable layer, whose top is located at a depth of 600 m with a formation thickness of 360 m.

characteristic length, which is the reservoir length here. λ_e is the effective bulk thermal conductivity of the saturated rock. Therefore, the advective heat transport dominates in this problem.

A transient fluid flow and heat transport simulation is performed for a period of 10 years. This realization is called the reference model and is shown in Figure 4.1.

4.1.1. Model Reconstruction with Varying Number of Wells

Assuming that boundary conditions, probability distributions, and correlation lengths are well known, the SGSim algorithm is run to reproduce the reference model. The simulation is conditioned by data: thermal conductivity and permeability from a varying number of synthetic wells inside the reference model. The positions of the wells are chosen randomly. A total of three

realizations using the same number of wells form one group. Target parameters and state for this test are thermal conductivity, permeability, and temperature. The latter is of special interest because it is a state variable which can be measured directly in wells. Additionally, temperature is a state variable of major significance with respect to geothermal energy.

Figure 4.2 illustrates the temperature differences in the temperature fields ΔT between the reference model and simulated realizations from a group using data from three wells. It shows a horizontal cross-section in the center of the model. The three patterns are different, but indicate a similar level of uncertainty. In total, the simulation runs yield comparable patterns within each group and very different ones between the groups. Not surprisingly, the differences decrease with increasing number of wells used. It is of interest that some simulated realizations, in this case model (a), reflect the reference model more successfully, although they are based on a relatively small number of wells.

Typical realizations for a different number of wells are shown in Figure 4.3. As expected, the quality of the fit to the reference model increases with the number of wells. Interestingly, simulations using only a small amount of wells (Figures 4.3(b) 4.3(c)) do not yield better results than an unconditioned simulation (Figure 4.3(a)). Therefore, an unconditioned simulation appears preferable in the case of only few available wells. This situation is typical for projects dealing with geothermal energy. However, some successful members of the ensemble, such as the one shown in Figure 4.3(d), fit the reference distribution quite well, even though they are constrained by relatively few wells. The spatial permeability distribution is supposed to be mainly responsible for the shape and magnitude of the temperature field because in operated reservoirs fluid flow usually affects temperature stronger than conductive heat transport (see Chapter 5.1.2). Thus, different realizations feature large ranges of temperature differences ΔT , varying from less than 1 K up to 20 K. Considering the absolute temperature mean of 30 °C in the analyzed layer of the reference model, the uncertainty in the temperature prediction is very large. Successful realizations, in terms of a successful fit of the temperature in the reference model, result mostly from a successful fit of the corresponding permeability distribution.

4.1.2. Constraining Post-Processing

A simple constraining technique is proposed for identifying the best fits of the reference model and discarding all unsuccessful ones. To this end, a larger ensemble of 50 realizations is generated using unconditioned SGSim. A large ensemble size is essential when applying constraining methods in order to obtain a sufficient number of successful members. Unfortunately, generating large ensembles is at the price of large computing time. However, for the small synthetic model in this case, 50 realizations can be generated in short time even on a single processor core. Again, information is used from three synthetic wells ($i = 1, 2, 3$) inside the reference

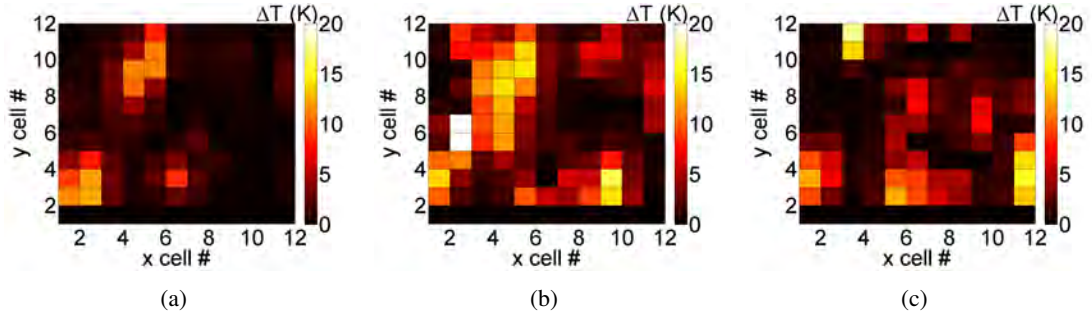


Figure 4.2.: Temperature differences ΔT between reference model and three simulated realizations (a-c) from the same group using data from three wells. The mean temperature in the displayed layer is 30°C . A uniform flow originates at $y = 0$.

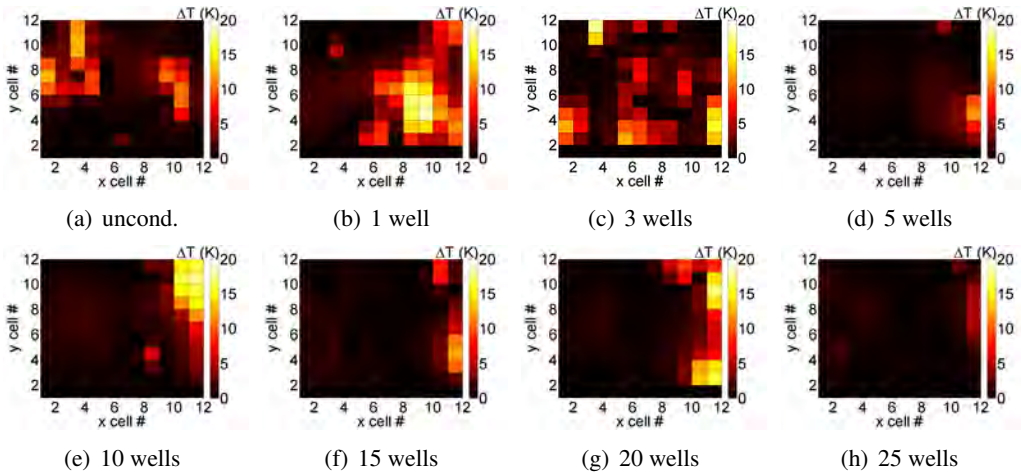


Figure 4.3.: Temperature differences ΔT between reference model and one simulated realization from different groups: (a) unconditioned; (b) - (h) conditioned to an increasing number of wells. A uniform flow originates at $y = 0$.

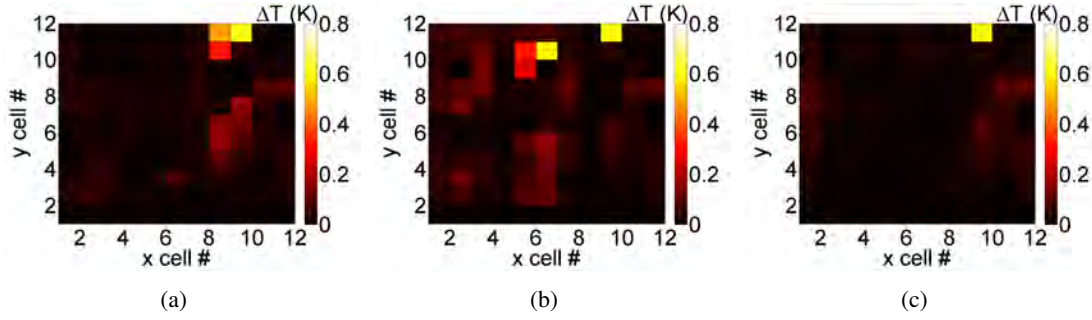


Figure 4.4.: Temperature differences ΔT between reference model and the three successful realizations after performing constraining post-processing for ensemble conditioning. An uniform flow originates at $y = 0$.

model. One single temperature value T_i^{ref} from the permeable layer is recorded for each well instead of the rock properties. This corresponds to measured bottom-hole temperatures (BHT) in boreholes. These temperatures are compared to the corresponding temperatures of a simulated realization T_i^{sim} by using three fitting parameters c_i :

$$c_i = \frac{|T_i^{sim} - T_i^{ref}|}{T_i^{ref}} \quad (4.2)$$

To select the three best fitting members of the ensemble, realizations are discarded which yield one of the fitting parameters above an appropriate threshold (here: 7.6×10^{-4}). This technique is repeated for each realization. The temperature difference ΔT of the three successful realizations is shown in Figure 4.4 to assess the quality of fit. The figure should be compared with Figure 4.2, which illustrates the results using three wells for conditioning SGSim without constraining post-processing.

Even though the SGSim algorithm is run unconditionally with respect to thermal conductivity and permeability, constraining post-processing based on temperature data reduces the maximum misfit between simulation and reference model from 20 K to 0.8 K. Obviously, constraining post-processing yields a major improvement in uncertainty minimization. Therefore it is performed during stochastic modeling of real geothermal reservoirs, as described in Chapter 5.1.1. As stated above, in the case of exploited reservoirs, hydraulic properties affect the temperature field stronger than thermal properties and their spatial distribution. Consequently, constraining techniques are particularly effective in selecting realizations after stochastic simulations of hydraulic properties in systems where fluid flow affects significantly the temperature. A more sophisticated alternative to constraining post-processing is the Ensemble Kalman Filter method.

Table 4.2.: Properties of the synthetic 2D model.

Parameter	Value
Mesh size	$20 \times 20 \times 1$
Resolution	$20 \text{ m} \times 20 \text{ m} \times 20 \text{ m}$
Log permeability / m^2	-14 ± 1
Circulation rate	$0.016 \text{ m}^3 \text{ s}^{-1}$
Concentration of injected tracer	0.5 mmol L^{-1}
Stochastic parameter	permeability

4.2. 2D Synthetic Test Case for the Ensemble Kalman Filter

A production scenario of a geothermal reservoir is a transient problem and therefore suitable to test the potential of the EnKF for inverse geothermal reservoir modeling. As permeability controls the production flow rate of hot water, it is a parameter of crucial importance for characterizing geothermal reservoirs. Hence, the EnKF is used in this test scenario to recover the spatial distribution of the reservoir permeability.

The synthetic reference model comprises a heterogeneous 2D permeability field (Figure 4.5, Table 4.1) denoted by k , generated with SGSim. It consists of 20×20 grid blocks and has the dimension of $1 \text{ km} \times 1 \text{ km}$. Water is injected in a well at the center with a rate of $0.016 \text{ m}^3 \text{ s}^{-1}$ and produced from four wells close to the corners (Figure 4.6). In addition, a chemical tracer is injected at the central well for ten days at a concentration of 0.5 mmol L^{-1} . The entire simulation time lasts 1 000 days, with one assimilation performed every 10 days. Hydraulic head and tracer concentration data at the well locations are used to calibrate the permeability field using the EnKF. Therefore, only sparse data from five locations are assumed available, as typical in geothermal production scenarios.

The borehole data are assumed to be measured once a week and used to update hydraulic head and the decimal logarithm of permeability. To that end, permeability k is varied at random according to a Gaussian distribution $\log_{10}[k/\text{m}^2] = -14 \pm 1$ with a correlation length of 300 m in an ensemble of 500 realizations generated by SGSim. In addition, a damping of the Kalman Gain is implemented to avoid filter convergence. An optimal damping of $\alpha = 0.2$ and an optimal system noise of 0.01 are found by manual optimization.

4.2.1. Results

Even though information from just five wells is available, the EnKF reproduces the major characteristics of the reference field, as illustrated in Figure 4.7(a) by the ensemble mean as the

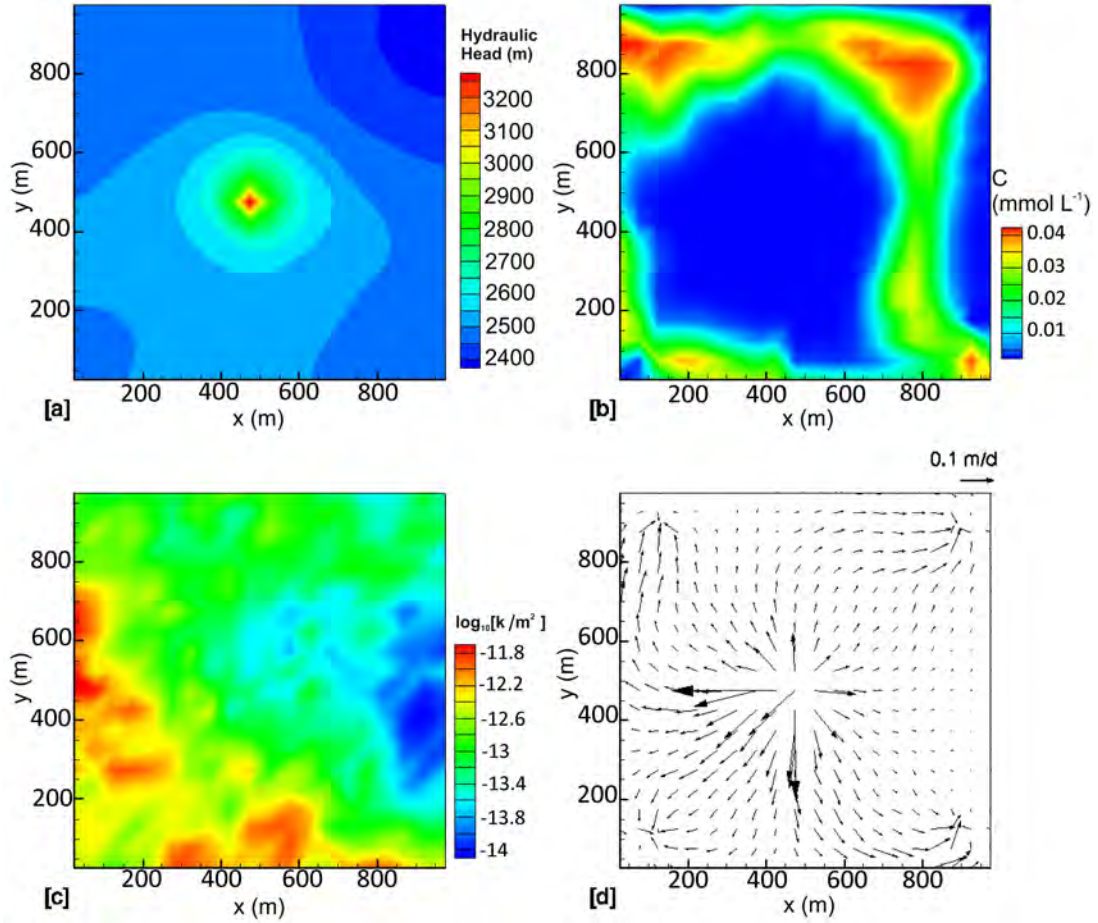


Figure 4.5.: Synthetic reference model: (a) shows the head in meters and (d) the Darcy velocity field (the black arrow indicated 0.1 m d^{-1}) after a simulation time of 1000 days; (b) shows the tracer concentration cloud after 150 days and (c) the (static) permeability field generated with SGSim.

best estimator for the permeability field. The ensemble standard deviation, illustrated in Figure 4.7(b), indicates that the field is best known along the tracer pathways and, not surprisingly, at the well locations.

Thus, the Ensemble Kalman Filter appears to be a powerful tool for history matching of production data from geothermal reservoirs. Similarly to Jafarpour & Tarrahi (2011), I obtain best results by a unconditioned SGSim generation of the ensemble members, i.e., by not fixing permeability values to the boreholes. Assigning permeability values results also in fixed permeability values in the surrounding grid cells according to the specified correlation length.

4.2.1.1. Iterative Filtering

Following the iterative approach of Krymskaya et al. (2009), the EnKF estimation is repeated globally with respect to the entire data assimilation time, replacing the initial guess with the estimated model state. The estimation improves by approximately 10 % of the *root mean square error* between reference and simulation when applying this technique for one single iteration in the given test case. More iterations did not yield any significant improvement.

4.2.2. Comparison with SGSim

To evaluate the performance of the EnKF, I compare it to a massive Monte Carlo approach including constraining post-processing similar to the study in Section 4.1. To this end, I generate a varying number of realizations of the log permeability field and rank them according to their quality in fitting the reference field. This quality ranking is obtained using the *root mean square error* E_i between simulated values d_i^{sim} for each realization i and n observations measured in the reference model d^{obs} :

$$E_i = \sqrt{\frac{1}{n} \sum_{j=1}^n (d_{i,j}^{sim} - d_j^{obs})^2}. \quad (4.3)$$

A small E_i corresponds to a high ranking and therefore a high quality of fit. For the fit, I use tracer and hydraulic head data from the synthetic reference at the production wells. Then, following the constraining post-processing approach, the ten best-fitting realizations are identified. Finally, a mean log permeability field is calculated as the result of the inversion.

The *root mean square error* between the log permeability field of the obtained mean and the reference field (designated by E without index i) is a measure to evaluate the quality of the fit. Here, n denotes the number of grid cells considered in the estimation process. Figure 4.8 illustrates the resulting ensemble means for ensemble sizes of 200 and 2000 realizations. As expected, a *root mean square error* of $E = 0.93$ for 200 and $E = 0.34$ for 2000 realizations, respectively, indicates an improved quality of fit with increasing ensemble size. Reproducing the

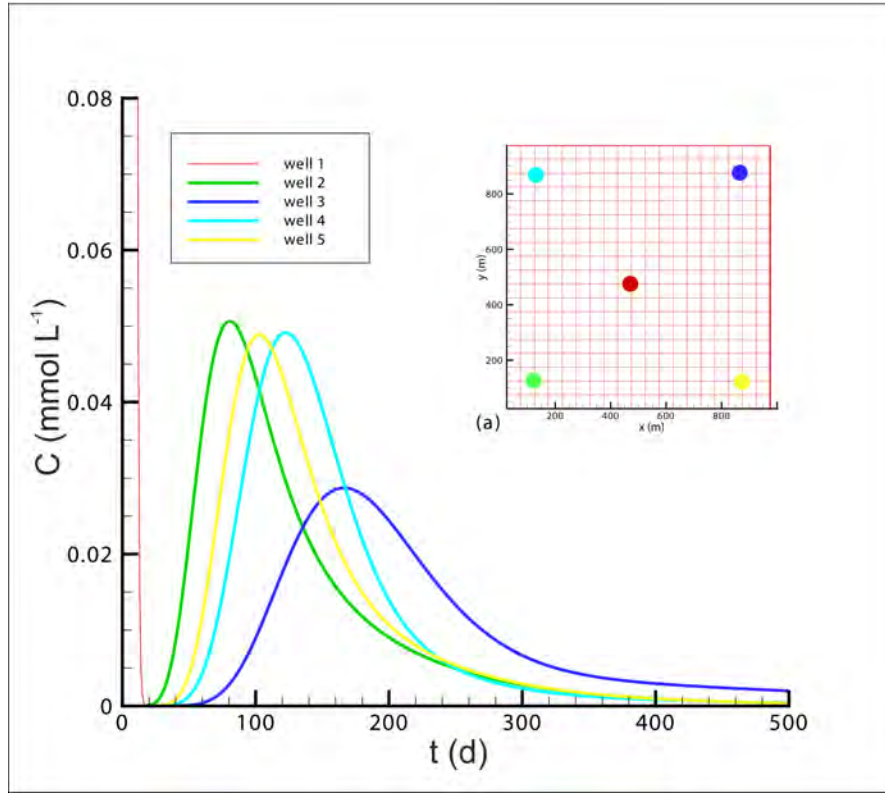


Figure 4.6.: Concentration C of a chemical tracer at four production wells with time t . The position of the wells is color marked in the inlay sketch (a).

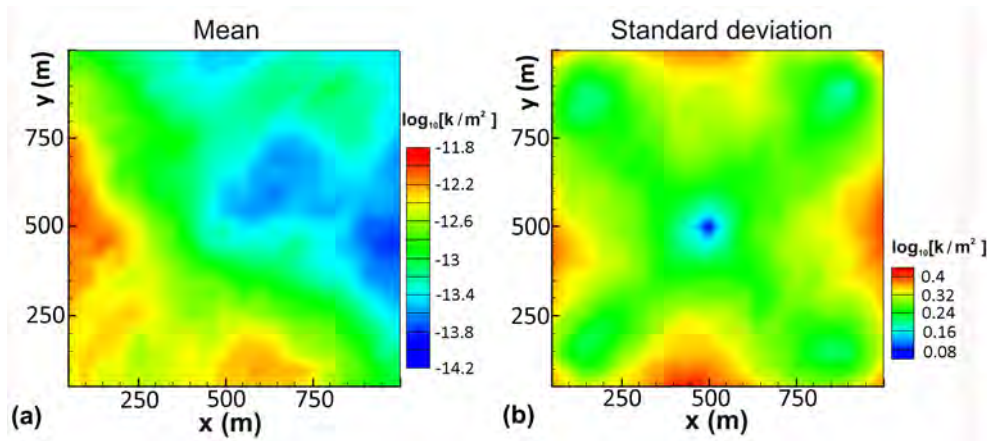


Figure 4.7.: Result of EnKF estimation: ensemble mean (a) and standard deviation (b).

basic characteristics of the reference field requires at least an ensemble size of 2 000 realizations. Increasing the ensemble size from 200 to 2 000 members improves the *root mean square error* by a factor of about three. However, increasing ensemble size further to 16 000 realizations yields no further improvement.

Comparison of Figures 4.7(a) and 4.8 indicates a better fit by EnKF than by massive Monte Carlo. This is supported by the *root mean square error* of $E = 0.26$ which is smaller by a factor of about 0.77 than the best result obtained by the massive Monte Carlo approach. This holds even for an EnKF ensemble size of 150 realizations, yielding $E = 0.29$.

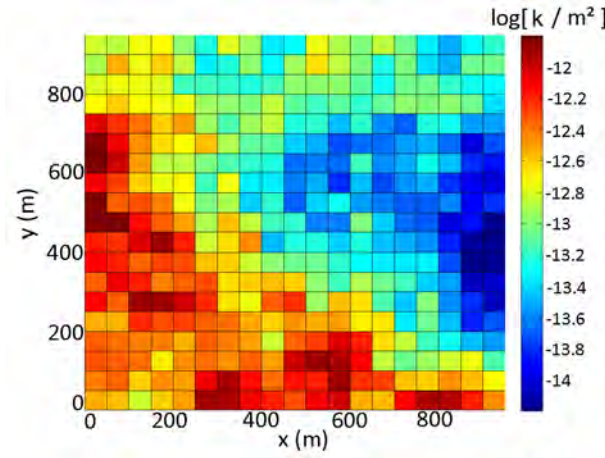
In summary, the EnKF yields better data fits in these synthetic test cases than a massive Monte Carlo approach with constraining post-processing. In addition, it requires much smaller ensemble sizes resulting in a significantly shorter computation time. As a further advantage, the EnKF is following rigorously a Bayesian scheme, provides an inherently correct probability density function (pdf) and, hence, correct uncertainty. In contrast, constraining post-processing yields only an approximation of the probability density function. However, constraining post-processing is still an appropriate tool for steady-state problems lacking transient data which could be used for sequential assimilation.

4.3. 3D Synthetic Test Case for EnKF in a Hard Rock Reservoir

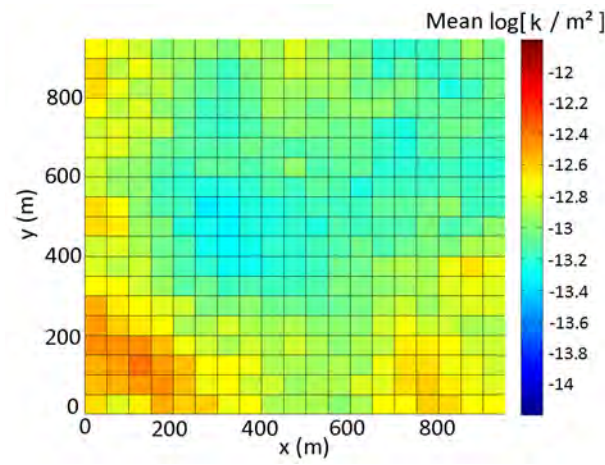
A real reservoir test case will be described in Chapter 6.5. There, the EnKF is applied to an estimation of permeability of an Enhanced Geothermal System (EGS) at the European test site Soultz-sous-Forêts. There are data available from just three wells for estimating permeability in a fractured reservoir. Here, I study the capability of the EnKF for this task on a Soultz-like synthetic reference model. It reflects the characteristics of all known conditions with regard to geometry, permeability, injected concentrations, and pumping rates of the Soultz reservoir during the tracer circulation test of 2005. One stochastically generated realization of the permeability field is assumed as the known reference. Note that this reference model is not part of the ensemble updated during the subsequent EnKF estimation. The model geometry, discretization, and boundary conditions are identical to the real reservoir model and are described in Chapter 6.5.

4.3.1. Reference Model

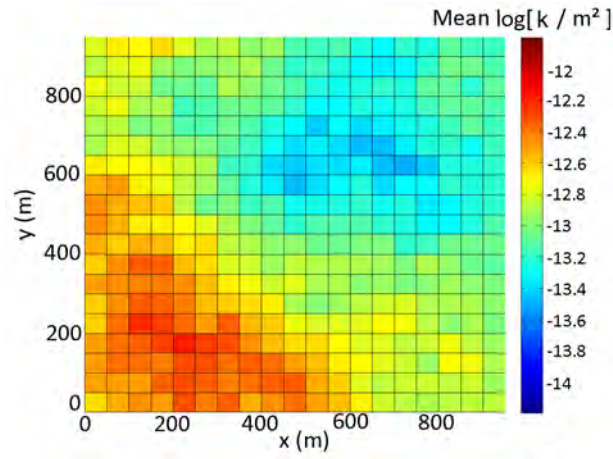
The reference realization (Figures 4.9 and 4.10(a)) is characterized by a low hydraulic connectivity between GPK3 and GPK4 and a well-connected although curved zone between GPK3 and GPK4. In the following, this reference permeability field is estimated using the EnKF based on hydraulic head and tracer concentration data computed from the reference model. The duration



(a) Reference



(b) 200 realizations



(c) 2000 realizations

Figure 4.8.: Reference field (a) and ensemble means of permeability k from massive Monte Carlo inversion using constraining post-processing and different ensemble sizes (b,c).

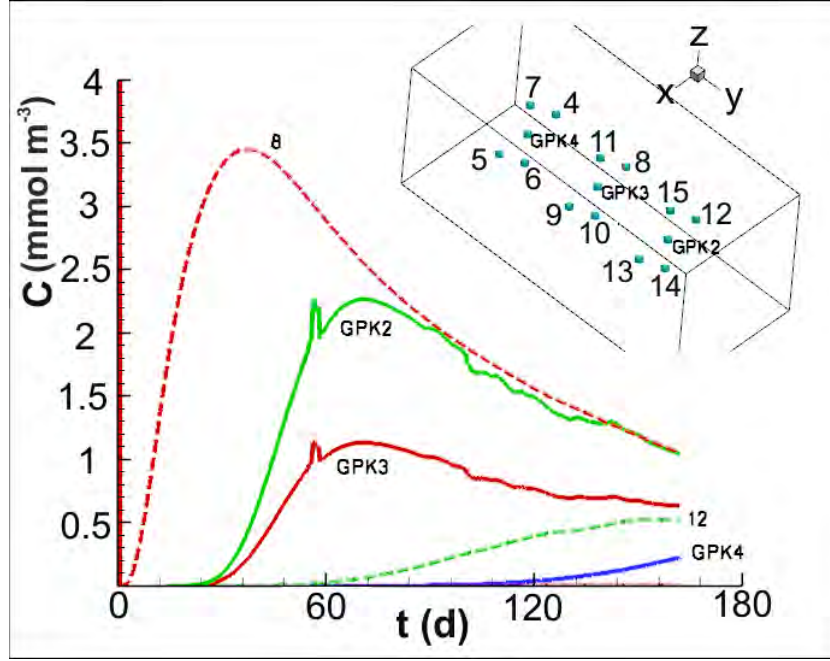
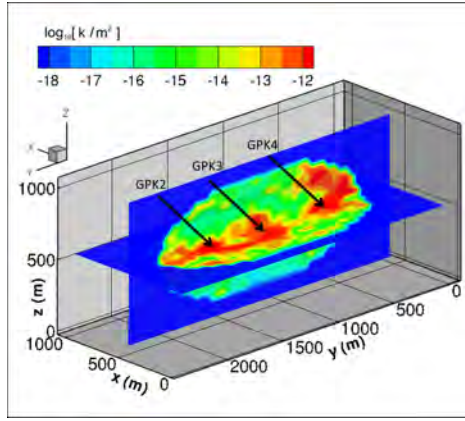


Figure 4.9.: Variation of tracer concentration C with time t at selected from a total of 15 observation points of the synthetic reference model for a hard-rock EGS. A significant concentration is observed only at 5 of these 15 points during the synthetic experiment.

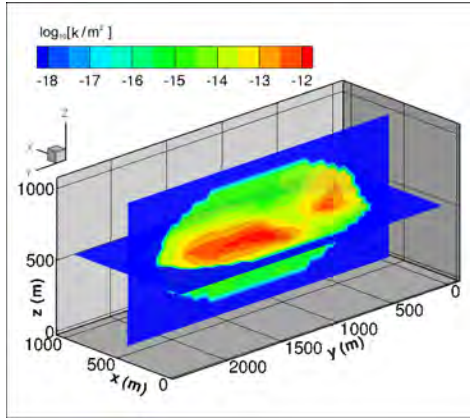
of the synthetic tracer experiment (300 days) is twice the duration of the original tracer test at Soultz in order to allow a tracer breakthrough (or at least arrival) at the observation locations in this synthetic test case. The simulation time is discretized into 6 000 equal time steps of 12 hour duration each.

4.3.2. Data from 15 Observation Points

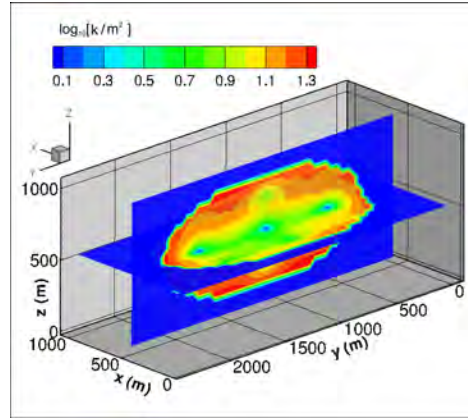
First, I estimate the permeability field using tracer concentration data from three active wells (GPK2, GPK3, GPK4) and twelve additional passive monitoring points (1 – 12, shown in Figure 4.9). A significant concentration is observed only at five of these points during the numerical experiment. Nevertheless, observations from all 15 points are used to update the ensemble. For the EnKF estimation, I use this concentration data at 48 times steps distributed evenly over the simulation time. The data are perturbed by small errors of $10^{-6} \text{ mol m}^{-3}$ to account for the measurement errors in a real-world experiment. Due to the strong heterogeneity and the wide parameter distribution (see Figure 6.3(a)), an ensemble size of 300 realizations is necessary to prevent filter inbreeding and a large underestimation of uncertainties. The initial mean permeability is homogeneous at $k = 10^{-14} \text{ m}^2$ except for the injection and production cells.



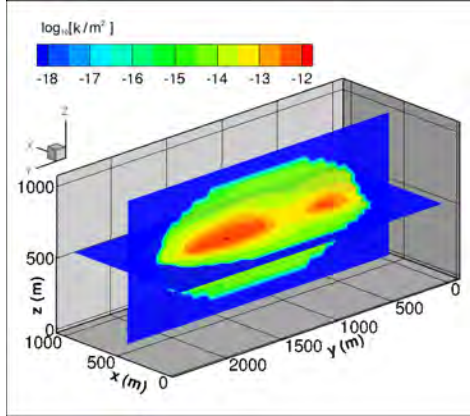
(a) Reference



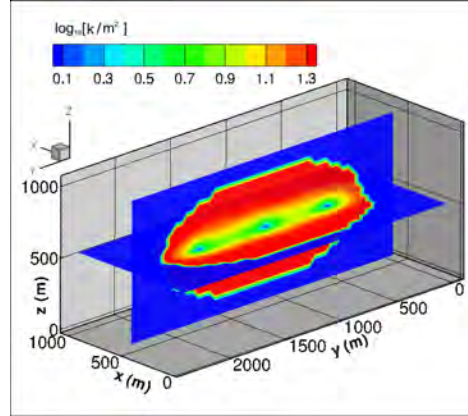
(b) μ , 15 observation points



(c) σ , 15 observation points



(d) Ensemble μ , 3 observation points



(e) Ensemble σ , 3 observation points

Figure 4.10.: Permeability field k of the synthetic reference field (a) and estimation results: ensemble mean μ (b,d) and standard deviation σ (c,e) based on data from 15 (a,b) and 3 (c,d) observation points, respectively.

The Kalman Gain in the assimilation process was dampened by a factor of $\alpha = 0.2$, also in order to avoid filter inbreeding (see Section 3.5). According to initially performed calibrations of the filter and to the results of Marquart et al. (2012), this value is supposed to be optimal for this problem. No model error was implemented. Following Krymskaya et al. (2009), the EnKF is reinitialized with the latest assimilated ensemble as new *a priori* information in three iterations to improve the estimates. A fourth iteration does not improve the result any further.

The estimated ensemble mean μ is illustrated in Figure 4.10(b). Compared to the reference condition, the estimation reflects the basic characteristics of the permeability field. Even though the estimated field is much smoother than the reference, it corresponds very well to the curved pathways between GPK3 and GPK2 and the low connectivity between GPK3 and GPK4. However, small-scale variations within the permeability field, particularly if they are located far away from the flow path (for example above GPK4), are not identified in the estimation.

In Figure 4.10(c), the ensemble standard deviation σ is shown as a measure for the uncertainty of the estimation. Along the flow paths between the boreholes with high tracer concentration, the uncertainty is lowest at about half an order of magnitude. It becomes larger further away with more than one order of magnitude. The location of the observation points, in particular of the three active boreholes is reflected also by the standard deviation.

In order to quantify the quality of the estimation and the reliability of the uncertainty measure, I consider again in each grid cell the ensemble *root mean square error* (RMSE), and the ensemble *spread*. These are considered measures for the error of the estimates and the ensemble variability. For a real-world reservoir, the *spread* is the only available measure of uncertainty, as the *root mean square error* is only defined if the true reference for the estimate is known.

The *root mean square error* E is already defined Section 4.2.2. Again, a small E corresponds to a good fit. In contrast, the *spread* S is defined using the ensemble mean μ_j^{sim} in each estimated grid cell j :

$$S = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (d_j^{sim} - \mu_j^{sim})^2}. \quad (4.4)$$

Figures 4.11(a) and (b) show the variation of these two quantities with time. They illustrate a clear convergence of the estimation result with time as well as the improvement obtained by a second and third iteration.

If E and S are approximately identical, the ensemble variance captures the estimation error correctly. This is not the case here. Comparing two error measures shows that the ensemble *spread* is lower by 10 % than the *root mean square error*. Therefore, I conclude that the standard deviation (Figure 4.10(c)) underestimates the actual deviation from reference model. According to this finding, the ensemble loses too much variance during the assimilation process. For

compensating this, I suppose an uncertainty increased by a factor of 0.1 (corresponding to the difference of E and S) as more likely than the error given in Figure 4.10(c). As the test scenario is similar to the Soultz reservoir, the uncertainties resulting for the application using real data are supposed to be larger by this factor, too.

4.3.3. Data from three Observation Locations

There are only three observation boreholes in the real Soultz reservoir. Therefore, I study additionally the performance of the EnKF for this low data density. Figure 4.10(d) illustrates the result. Here, the ensemble mean does not show the curved shape of the path between GPK3 and GPK2 in the reference field. This is due to the lacking information from this region. However, the magnitude of the permeability field and its very basic characteristics are still reproduced. Therefore, the EnKF should be capable of capturing these basic features at Soultz.

The corresponding uncertainty is shown in Figure 4.10(e). It is much larger than for 15 observation points (Figure 4.10(c)). The increased uncertainty indicates the reduced quality of fit due to the reduced number of data sampling points. Still, the underestimation of uncertainty for three data sampling points illustrated by E (Figure 4.11(c)) and S (Figure 4.11(d)) is not larger than the one for 15 data sampling points.

As a results, an inversion of the basic characteristics of the flow paths inside the Soultz reservoir appears feasible based on EnKF and data from the three boreholes.

4.4. Tracer vs. Temperature Data

As an alternative or in addition to tracer data, temperature data can be used for estimating reservoir permeability (as done in the study to evaluate the performance of constraining post-processing based on a different number of wells). Both types of data are compared by an EnKF parameter estimation of permeability following the synthetic test model in Section 4.2. The simulation is run in one global iteration for 40 years. Temperature data are assimilated during the entire simulation time with one ensemble update per year. Tracer data are assimilated for only 5 years with five ensemble updates per year. The resulting *root mean square error* is illustrated in Figure 4.12. It shows that the assimilation by tracer data outperforms the one by temperature data. In addition, the tracer propagates on different time scales compared to temperature. Therefore, the estimation is available much earlier. Nevertheless, assimilating temperature data in combination with tracer data yields an improved estimation comparable to the one of an additional global EnKF iteration and may still be of use during the long-term operation of a geothermal system.

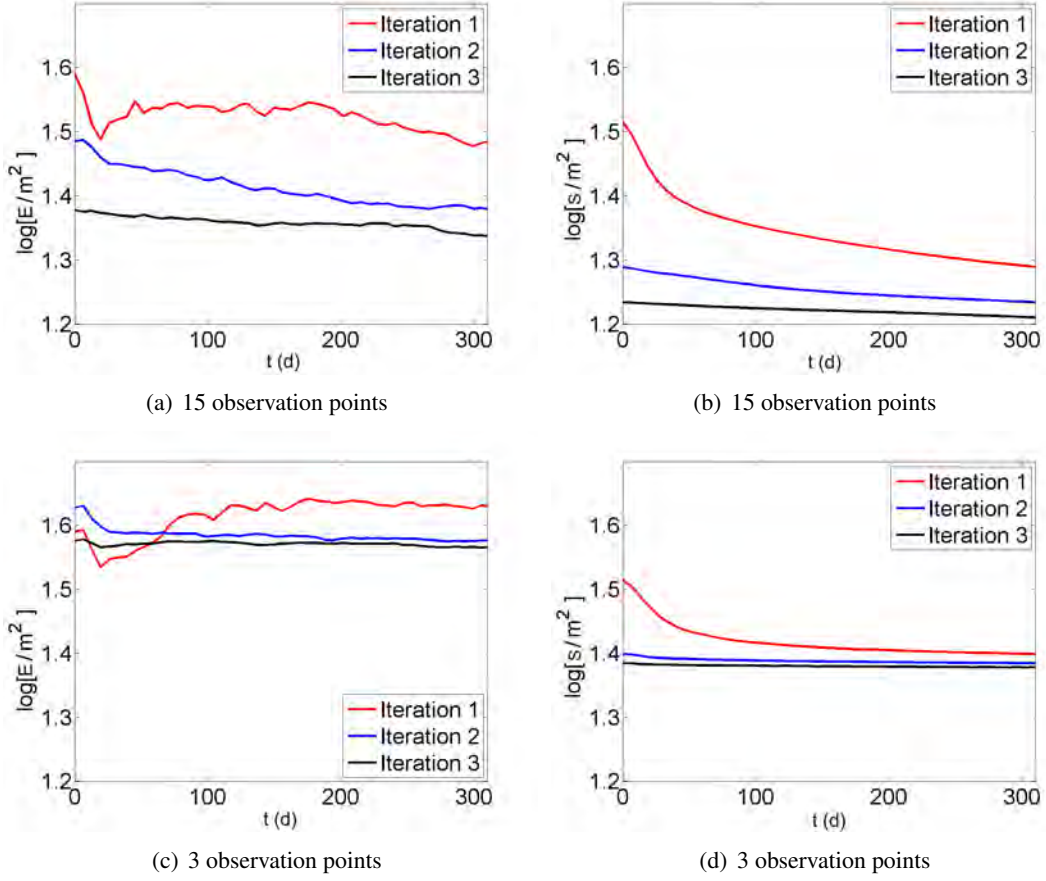


Figure 4.11.: Ensemble *root mean square error* E (a,c) and *spread* S (b,d) of log permeability for all ensemble nodes during the assimilation process based on data from 15 and three synthetic observation points, respectively.

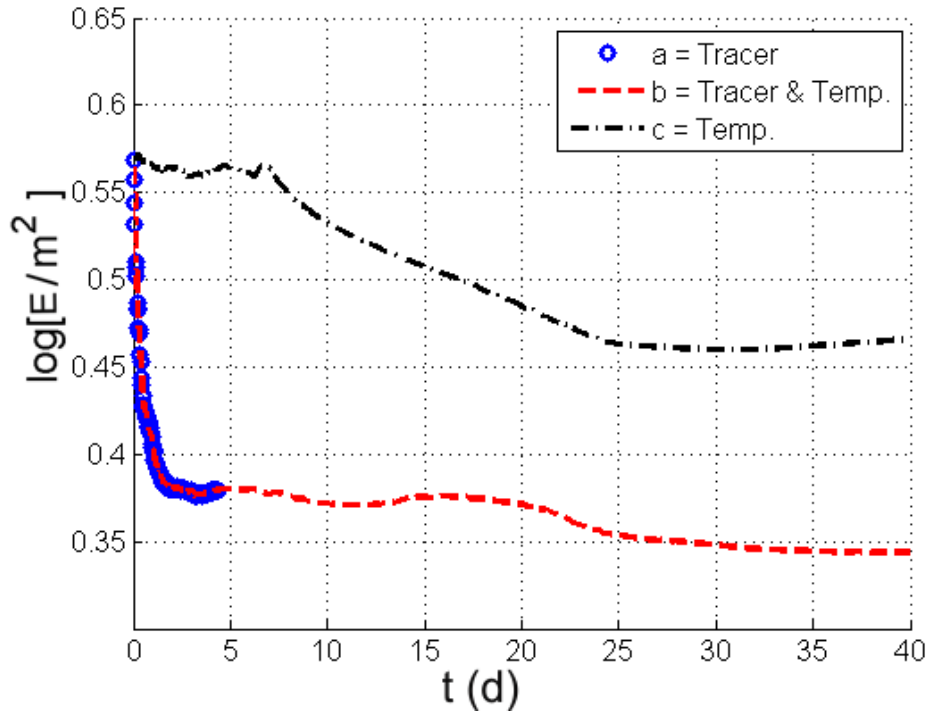


Figure 4.12.: *Root mean square error of EnKF estimation of log permeability based on hydraulic head combined with tracer only (a, dots), both, tracer as well as temperature (b, dashed line), and temperature only (c, dashed and dotted line). Modified after Vogt & Widera (2012).*

Chapter 4. Synthetic Test Scenarios

Having evaluated the performance of the stochastic tools for synthetic test cases with satisfying results, I apply the tools during modeling sequences for different real world reservoirs in the following chapters.

Chapter 5.

Assessment of Probabilistic Predictions for Porous Rocks

As stated above, quantifying and minimizing uncertainty as well as capturing heterogeneity are vital for simulating technically and economically successful geothermal reservoirs. To this end, I apply stochastic modeling sequences, Monte Carlo studies, based on (i) creating an ensemble of possible realizations of a reservoir model, (ii) forward simulation of fluid flow and heat transport, and — if possible — (iii) constraining post-processing using observed state variables.

In this chapter, I study two scenarios for quantifying uncertainties of geothermal reservoir properties and their transient variation. First, I model a geothermal doublet in The Hague, The Netherlands, which started operation recently. In this case, doublet layout parameters such as the well distance are fixed. The decision for actual placement of the boreholes, however, was based on the study of Mottaghy et al. (2011b) and on this study. Second, I investigate a reservoir in the Northeastern German basin of confidential location. In this academic study of a hypothetical installation, the well placement is not fixed. The main differences of first and second study are that the second one (i) combines stochastic Monte Carlo forecast with Bayesian inversion instead of constraining post-processing, (ii) uses more detailed seismic information, (iii) includes pressure predictions, and (iv) investigates a target layer with lower permeability and smaller thickness (by a factor a four), which yields a smaller distance between the wells (0.5 km instead of 2 km), and presents a single-well concept.

The presented approaches based on comprehensive statistical data on thermal rock properties are of special interest in particular when no temperature logs, but only bottom-hole temperature data are available for estimating the subsurface temperature field.

5.1. Quantifying Uncertainties for a Given Reservoir Layout in The Hague, The Netherlands

In order to demonstrate the stochastic approach and its advantages, an exploration scenario is simulated for a current geothermal district heating project in The Hague, Netherlands. The first steady-state simulations focus on thermal rock properties in Section 5.1.1. Additionally, the stochastic modeling of boundary conditions is also applied to this field study. Hydraulic properties are in the focus in Section 5.1.2 where also transient variations of the reservoir states are modeled. I show that the risk within a geothermal project can be estimated more accurately using the modeling sequence.

5.1.1. Temperature Uncertainty Related to Thermal Properties

In this initial step, the temperature in the target region is predicted and its uncertainty is accurately quantified.

A geothermal doublet will be installed at a depth of 2300 m to supply up to 4000 households with geothermal heat. The Delft sandstone (van Balen et al., 2000) at the transition between Jurassic and Cretaceous formations is identified as the target reservoir. Its permeability varies between 10^{-12} m^2 and 10^{-13} m^2 . Technical requirements for formation temperature and flow rate are 75°C and 42 L s^{-1} , respectively. Model parameters such as reservoir geometry are reported by Simmelink & Vandeweyer (2008). A map of the studied area is shown in Figure 5.1.

Thermal rock properties (porosity, thermal conductivity, heat generation rate) are assigned to each layer using a statistical data set. Using standard log interpretation procedures (Doveton & Cable, 1979), the lithological composition and the rock porosity can be computed. In this study, the gamma-ray log (GR) measured in a calibration borehole is used to separate the clay and the sandstone content plus carbonate volume fractions, while density log (RHOB) is used to calculate the formation porosity. From the calculated volume fractions and porosity, continuous profiles of thermal conductivity (Figure 5.2) are generated using the geometrical mixing law in Hartmann et al. (2007). The composition of the matrix (sandstone or carbonate) is known from the analysis of the cuttings. Based on these thermal conductivity profiles, mean and standard deviation of effective thermal conductivity are calculated for all stratigraphic units in the 3D model. In addition, profiles of radiogenetic heat generation rate are calculated from gamma-ray logs following the empirical approach of B  cker & Rybach (1996).



Figure 5.1.: Map of the studied area containing the proposed target location and borehole labels, as well as the dimensions of the model. Adapted from Mottaghy et al. (2011a).

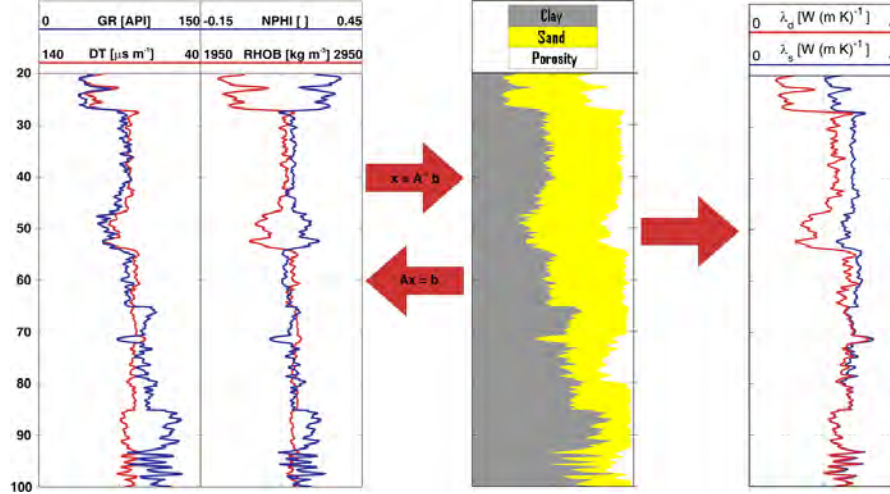


Figure 5.2.: Calculating dry and saturated thermal conductivity (λ_d and λ_s) profiles from logs b (GR: gamma-ray; DT: slowness; RHOB: density; NPHI= neutron porosity). The matrix A contains the log response with regard to the different rock types. The correlation between shale content and log responses is used to calculate the volume fractions from log data. Based on these and on the thermal conductivity of the single rock components from laboratory measurements, thermal conductivity in the subsurface can be determined using an appropriate mixing law. Adapted from Hartmann et al. (2005).

5.1.1.1. Thermal Model

Consistent with available geological information (Simmelink & Vandeweyer, 2008), there is neither significant thermal free convection nor topographically driven fluid flow. Therefore, heat conduction is dominant. The temperature prediction is obtained from a steady-state simulation of heat transport. For comparison of model and reality, ten corrected bottom-hole temperatures (BHT) are available distributed evenly in the study area. The modeling process aims for a temperature prediction at the proposed target location near the center of the model. This target location may change during the further planning progress of the project.

The 3D geometric model (Mottaghy et al., 2011a) represents a volume of $22.5 km \times 24.3 km \times 5 km$. It is discretized into $150 \times 162 \times 100$ grid cells. The model parameters are summarized in Table 5.1. It comprises nine geological layers identified and implemented as separate units. Each unit is characterized by a different saturated thermal conductivity and heat production rate. The mean and the standard deviation of the thermal conductivity in each layer characterize a unit.

A first estimate for the possible range in the predicted temperature is obtained from the simulations using consistently either minimum values (mean – standard deviation) or maximum values (mean + standard deviation) of thermal conductivity in each layer.

Table 5.1.: Properties of the 3D thermal model.

Parameter	Value
Mesh size	$150 \times 162 \times 100$
Resolution	$50 \text{ m} \times 50 \text{ m} \times 50 \text{ m}$
Bulk thermal conductivity	$0.8 \text{ W m}^{-2} \text{ s}^{-1} - 6 \text{ W m}^{-2} \text{ s}^{-1}$
Temperature at top	$11 \text{ }^{\circ}\text{C}$
Basal specific heat flow	$(63 \pm 3) \text{ mW m}^{-2}$
Stochastic parameters	bulk thermal conductivity, basal specific heat flow

The result of the prediction at the locations of the BHT data is illustrated in Figure 5.3. Obviously, the means are sufficiently consistent with the BHT, but the uncertainty appears large, if not over-estimated.

5.1.1.1. Quantifying and Reducing Uncertainty

The stochastic modeling sequence explained in Section 4.1 is applied to obtain better uncertainty estimates. Instead of considering only minimum and maximum values defined by means and standard deviations, the original probability distribution (in depth) of the saturated thermal conductivity in the calibration borehole is simulated in six of the nine geological units using the SGSim algorithm (Deutsch & Journel, 1998). As an example for the different input histograms to be reproduced using SGSim, the measured distributions of Layer 3 (Lower Cretaceous Supergroup) and Layer 4 (Jurassic Supergroup) are illustrated in Figure 5.5. The original histograms are calculated using different logs, as stated above. (Figure 5.2).

Although more than 50 exploration boreholes exist in the studied region, no temperature logs are available, only ten BHT measurements shown as red dots in the eight profiles in Figure 5.3. Calculated thermal conductivity data was available for just one calibration borehole with high vertical data density. With this single borehole in an area of about 550 km^2 , the borehole density is low, so that unconditioned SGSim has to be performed. This can be done without restricting the results as shown before in Section 4.1.1. However, the logs in the other boreholes indicate similar lithological conditions within the stratigraphic layers.

Vertical correlation lengths are calculated using variograms (see Section 3.4). The data required for the variograms are obtained from the calibration borehole. The variograms yield correlation lengths between 50 m and 450 m for the individual geological layers.

Due to the nature of sedimentation there is an anisotropy between vertical and horizontal correlation lengths (Caers, 2005). Vertical correlation lengths are usually much shorter than horizontal ones. For the horizontal directions, the borehole distribution with just one borehole is

insufficient to calculate reliable variograms. Fortunately, the horizontal direction is less important for predominantly vertical conductive heat flow. However, to obtain a maximum fluctuation in the ensemble of realizations, large correlation lengths are implemented in the horizontal directions by multiplying the corresponding vertical correlation length in each unit. Deutsch & Journel (1998) use a maximum factor of 16 between vertical and horizontal correlation lengths which is adopted in this study, too. This yields horizontal correlation lengths between 800 m and 7200 m. However, these values can be considered more or less arbitrary due to the dominantly vertical direction of heat flow. Nevertheless, simulating fluid flow in the reservoir, e.g. for performance prediction during the operation of the doublet, will require reasonable values in the future.

Using these correlation lengths, 1000 temperature fields are simulated based on 1000 realizations of thermal conductivity in the model. One example is shown in Figure 5.4. The ensemble is simulated in parallel on two quadcore processors with a clock rate of 3 GHz each. Thus, a computing time of 127 hours is needed to generate the full ensemble of 1000 realizations. Constraining post-processing is performed in order to reduce the uncertainty further by discarding unsuitable ensemble realizations.

5.1.1.1.2. Constraining Post-Processing

Constraining methods are useful to further reduce the uncertainty significantly. Therefore, constraining post-processing is performed in this case, too, although constraining is less effective in systems lacking significant fluid flow, as stated above in Section 4.1.2. In this post-processing, the temperatures of all 1000 realizations of the ensemble are compared to the measured BHT. Several studies discuss the reliability of corrected BHT data and their errors (e. g. Förster, 2001; Hermanrud et al., 1990; Deming, 1989). Förster (2001) and Hermanrud et al. (1990) reported that BHT values corrected with less advanced methods, such as the Horner plot method (Horner, 1951), underestimate the formation temperature by $8 \text{ K} \pm 8 \text{ K}$ in their studies. This underestimation is not confirmed by simulations (Figure 5.3). However, according to Hermanrud et al. (1990), advanced methods for correcting BHT usually yield values close to formation temperature with a standard deviation of $\pm 9 \text{ K}$. This standard deviation of $\pm 9 \text{ K}$ is used as a constraining parameter for post-processing. Thus, a realization is discarded from the ensemble, if the difference between the temperature T_i^{sim} simulated at the position of the i th BHT measurement and the corrected BHT value BHT_i exceeds $\pm 9 \text{ K}$ for at least one BHT. Thus, a realization is discarded if (for any i):

$$|T_i^{sim} - BHT_i| > 9 \text{ K} . \quad (5.1)$$

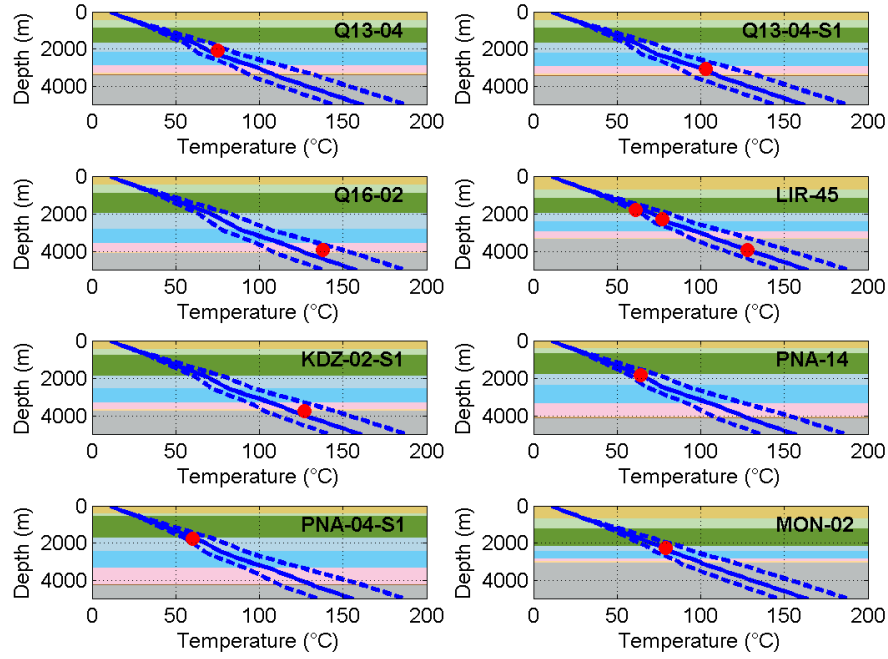


Figure 5.3.: Temperature prediction (blue line) for the location of BHT observations, BHT values (red dots), and uncertainty estimate (blue dashed lines). Labels refer to boreholes shown in Figure 5.1.

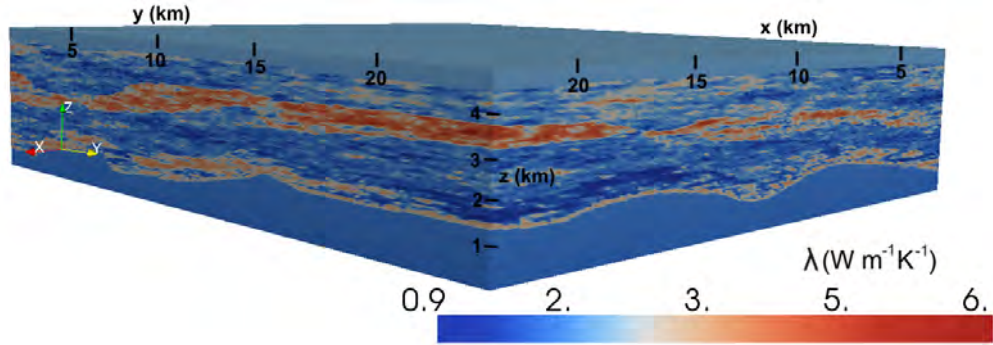


Figure 5.4.: One realization of a thermal conductivity distribution in the model.

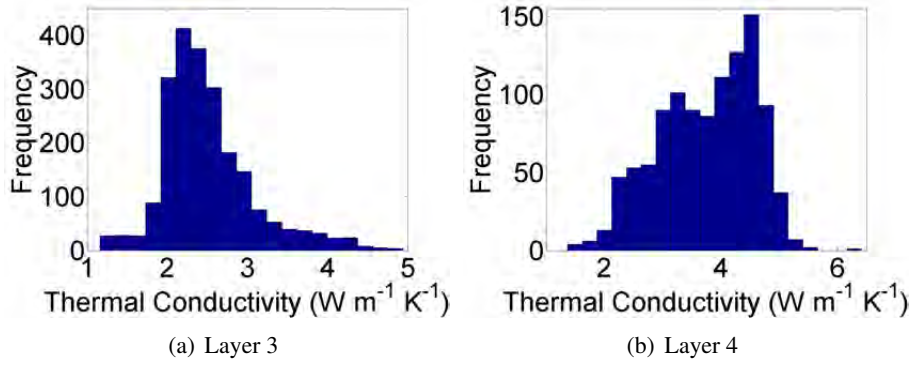


Figure 5.5.: Original distributions of the thermal conductivity in the Lower Cretaceous Supergroup (a) and the Jurassic Supergroup (b) obtained from the logging data. The average thickness of these layers is approximately 1100 m and 550 m, respectively.

5.1.1.1.3. Results

The probability distribution of the temperature at the proposed target location is shown for 1000 original realizations in Figure 5.6(a) and for those 488 realizations which survived the constraining post-processing in Figure 5.6(b). The full distribution width of the temperature is reduced from 16 K to 12 K as a result of the post-processing. Due to the Gaussian shape of the distribution, the standard deviation σ is reduced from 2.0 K to 1.8 K. In particular, realizations with very high or very low temperature predictions are removed from the ensemble. This eliminates too optimistic or pessimistic assumptions suggested by the stochastic modeling sequence without constraining the ensemble. The distribution mean $\mu = 79^\circ\text{C}$ stays constant at the proposed target position.

As in Figure 5.3, the result of the temperature predictions at the locations of the BHT is illustrated in Figure 5.7(a) for 488 realizations. The BHT are reproduced by the ensemble with an appropriate uncertainty.

In total, the final temperature prediction obtained from the ensemble of 488 surviving realizations is illustrated in Figure 5.8(a) in comparison with the initial uncertainty. Obviously, the uncertainty is minimized significantly by applying the stochastic modeling sequence. As further positive result in terms of project feasibility, according to the uncertainty quantification applied in this study, the likelihood for the formation temperature exceeding the required temperature of 75°C is 97 % and 99 % without or with constraining post-processing, respectively.

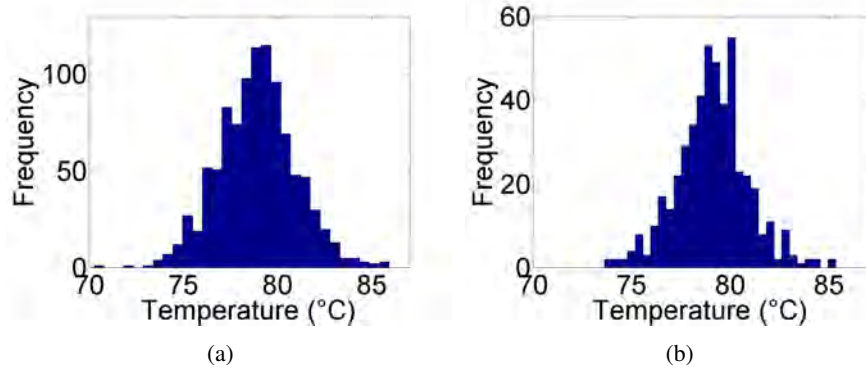


Figure 5.6.: Temperature at proposed target location (a) for the original ensemble of 1000 realizations and (b) for the ensemble of 488 realizations after constraining post-processing, further reducing uncertainty.

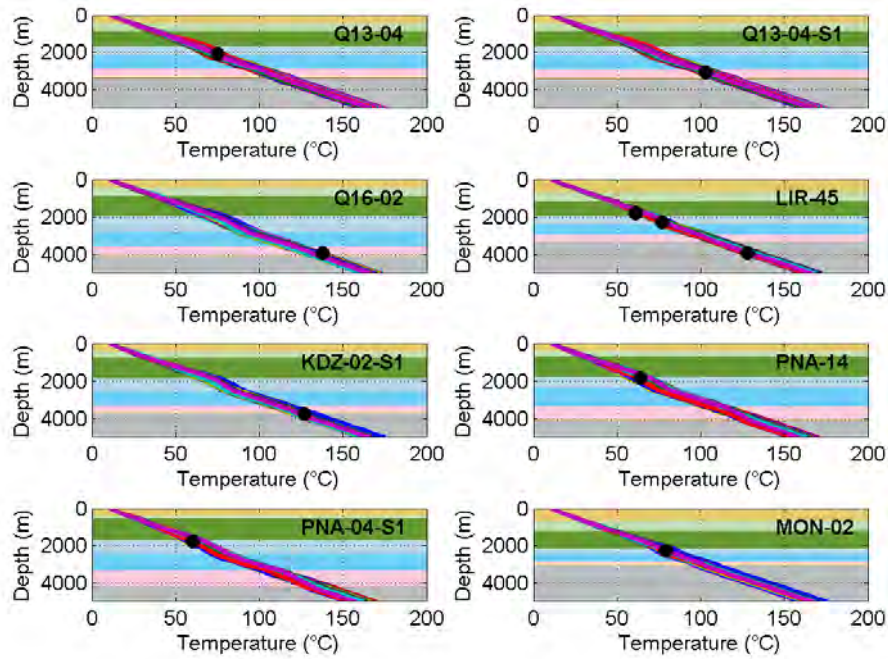


Figure 5.7.: Temperature prediction (colored lines) at the locations of BHT observations and BHT values (black dots) after applying constraining post-processing.

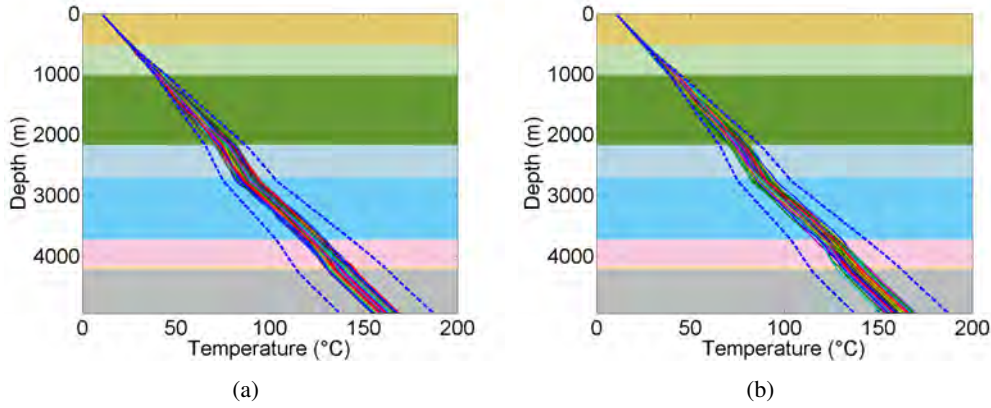


Figure 5.8.: Temperature prediction (colored lines) at proposed target location based on stochastic simulation of thermal conductivity compared to the original uncertainty estimate (blue dashed lines) (a). Temperature prediction (colored lines) based on additional stochastic simulation of basal specific heat flow and corresponding original uncertainty estimate (blue dashed lines) (b). The original uncertainty estimate (dashed lines) is based on an average value for the basal specific heat flow.

5.1.1.2. Additional Stochastic Simulation of Basal Specific Heat Flow

Taking advantage of the possibility for stochastic modeling of boundary conditions, an additional stochastic simulation of the basal specific heat flow is integrated in the modeling sequence. The objective is to consider this uncertain parameter of major significance in the modeling process. The mean of $q = 63 \text{ mW m}^{-2}$ is varied with a standard deviation of $\pm 3 \text{ mW m}^{-2}$. This small fluctuation of the basal specific heat flow is chosen due to the good agreement of the final corrected BHT with the predicted temperatures obtained from the simulated realizations illustrated in Figure 5.3.

The simulated value is assigned to each grid node at the bottom boundary of the model. Again, an ensemble of 1000 realizations is generated. Now, with an expected larger distribution width due to basal specific heat flow fluctuation, constraining post-processing is essential to identify reliable realizations.

5.1.1.2.1. Results

The results of the complete stochastic modeling sequence including boundary simulation are illustrated in Figure 5.9 and Figure 5.8(b). As expected, the full distribution width without constraining post-processing increases significantly from 16 K to 24 K at target depth compared to the distribution discussed in Section 5.1.1.1.1. The mean of 79°C does not change, but the standard deviation increases to 4.3 K.

Table 5.2.: Results of the stochastic modeling sequence: temperature mean μ , standard deviation σ and full distribution width.

Method	μ ($^{\circ}\text{C}$)	σ (K)	Full width (K)
Original estimate	77	–	25
Before post-processing without variation of basal specific heat flow	79	2.0	12
After post-processing without variation of basal specific heat flow	79	1.8	16
Before post-processing with variation of basal specific heat flow	79	4.3	24
After post-processing with variation of basal specific heat flow	79	2.1	11

However, performing the constraining post-processing yields results very similar to the results in Section 5.1.1.1.3 obtained without specific heat flow variation. The standard deviation of the distribution is $\sigma = 2.1$ K. The fraction of realizations which survived the constraining post-processing decreases from 488 / 1000 to 245 / 1000 because of the larger distribution width. Results are summarized in Table 5.2.

The uncertainty of the basal specific heat flow before performing constraining post-processing yields a lower likelihood of 80 % for the formation temperature exceeding the required temperature of 75 $^{\circ}\text{C}$. Here, this likelihood is 80 % without constraining post-processing. But interestingly, this likelihood is increased to 99 % again by applying the constraining post-processing, even when the basal specific heat flow is varied. This is a result from constraining post-processing, which discards realizations with large data misfit. Therefore, uncertainty estimates obtained with and without stochastic basal specific heat flow simulation yield similar results. The prize to pay is a reduction in ensemble size and, thus, lower statistical relevance of the results. The Ensemble Kalman Filter (discussed e. g. in Chapter 4) overcomes this problem.

In summary, the presented modeling sequence is very effective in reducing uncertainties significantly compared to prior estimates and thus helping to reduce risk of failure and cost, even though no hydraulic properties are involved like in Chapter 4.1. Additionally, the stochastic simulation of boundary conditions allows a more comprehensive approach to quantify uncertainty in geothermal reservoirs.

A first observation of the temperature T_{prod} of the produced brine measured at the surface has become available recently. Most important, the results are supported by the finding $T_{prod} =$

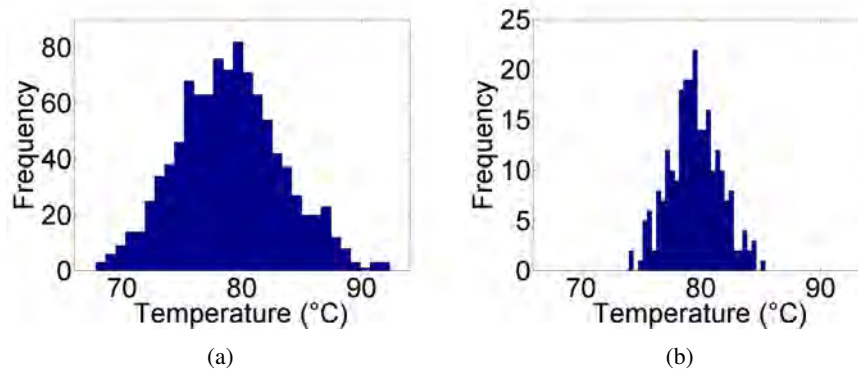


Figure 5.9.: Temperature at proposed target location (a) for the original ensemble of 1 000 realizations and (b) for the ensemble of 245 realizations after constraining post-processing including a stochastically simulated basal specific heat flow.

76 °C¹. Due to cooling on the way up to the surface, this value is assumed to be even higher at depth. Very rough analytical estimates assuming heat losses in pipe flow indicate a temperature drop of about 1.0 K in this case. Considering the standard deviation of my result of about 2 K, the measured temperature corresponds to the estimated temperature of $T_{est} = 79$ °C.

When interpreting the results, certain modeling constraints have to be considered such as neglecting mechanical clogging and stress field. They are discussed in detail for the reservoir in the northeastern German basin in the following Section 5.2.

5.1.2. Temperature Uncertainty Related to Hydraulic Properties

In the second phase of the study, a detailed reservoir model is set up. The new model is characterized by a finer discretization for resolving the target layer better for fluid circulation. This model is of smaller extension to keep the computing demand in a reasonable range. The models simulate the temperature evolution at the location of the production well and the propagation of the cold water front from the injection towards the production well. The main questions are (i) whether the production temperatures can be maintained throughout the years, and (ii) how far the cooling extends around the injector. This includes again quantifying of uncertainties. Whereas I focused on thermal properties before, now, hydraulic properties and their heterogeneities are investigated additionally.

¹pers. comm.: F. Schoof, Aardwarmte Den Haag, Stadzijde 5, 2543 VZ Den Haag, The Netherlands

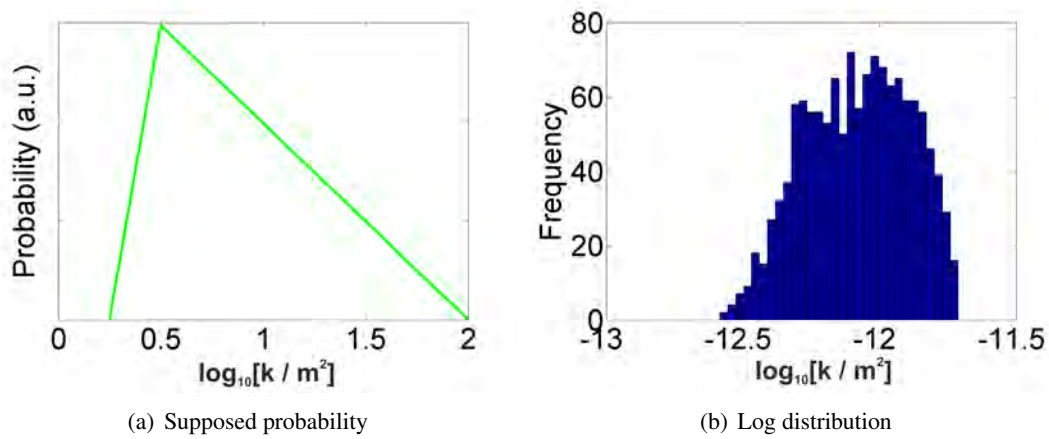


Figure 5.10.: Assumed permeability distribution k of the Delft sandstone according to Simmelink & Vandeweyer (2008) (a). This distribution is discretized and transformed logarithmically, obtaining a histogram for SGSim simulation (b).

5.1.2.1. Reservoir Model

The model itself is based, just like the regional model, on seismic interpretations by TNO² (Simmelink et al., 2007; Simmelink & Vandeweyer, 2008). In addition to the thermal parameters, the simulations now require hydraulic properties. These are much less known than the thermal ones, and show much larger variations than these. The crucial parameter permeability was re-evaluated by Simmelink & Vandeweyer (2008) for the Delft sandstone (Figure 5.10). The mean reservoir thickness, the second crucial parameter for determining reservoir transmissivity, is 55 m.

The general properties of the reservoir model are shown in Table 5.3. For initial sensitivity studies, mean values for the hydraulic parameters are used. At the locations of the injection and production wells, the grid is refined in order to deal with higher fluid velocities at these points. Figure 5.11 shows this model and the refinement. The depths for the injection and production borehole is about 1800 m and 2230 m, respectively. Boundary conditions and parameters of the reservoir model are adapted for compatibility with the results of the temperature model.

5.1.2.1.1. Propagation of the Cooling Front

Initial steady-state simulations before doublet operation show maximum Darcy velocities of 20 cm per year corresponding to permeabilities of 10^{-12} m^2 at maximum (Mottaghy et al., 2011b). This corresponds to a 1D-Péclet number Pe of < 0.1 . For the vertical direction it

²Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek

Table 5.3.: Properties of the 3D reservoir model.

Parameter	Value
Mesh size	$64 \times 41 \times 65$
Resolution	25 m–100 m \times 25 m–100 m \times 17 m
Number of nodes	170 560
Extension	5.5 km \times 3.5 km
Top	1500 m
Bottom	2605 m
Temperature at top	55.8 °C
Basal specific heat flow	67.8 mW m ⁻²
Porosity	15 %
Permeability	10^{-13} m ² – 10^{-11} m ²
Circulation rate	42 L s ⁻¹
Temperature of injected water	40 °C
Stochastic parameter	permeability

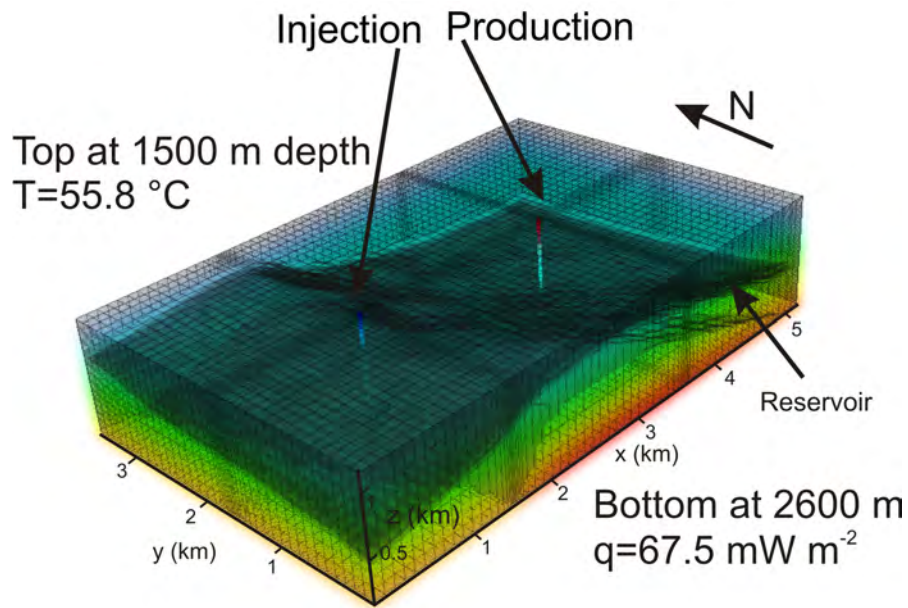


Figure 5.11.: Sketch of the reservoir model showing the grid and the layers as well as the position of the production and injection well. The grid is refined at these locations. The y-axis is rotated by 8° counterclockwise from north. Adapted from Mottaghy et al. (2011b).

is defined for as:

$$Pe = \frac{q_{adv}}{q_{cond}} = \frac{(\rho c)_f v_z L}{\lambda_e}, \quad (5.2)$$

where q_{adv} and q_{cond} are the amount of heat transported by advection and conduction, respectively, v_z is the Darcy velocity in vertical direction (pointing to the surface), and L defines a characteristic length, which is the reservoir thickness here. λ_e is the effective bulk thermal conductivity of the saturated rock. Therefore it can be concluded that the steady-state situation is dominated by conductive heat transport. This in turn implies also purely conductive heat transport as well in the regional model.

Based on the initial condition provided by the steady-state simulation, the transient model simulates the operating system of water production and injection. The simulation extends over 50 years with varying time step length from a few hours at the beginning to one day at later times. This procedure accounts for the higher gradients of thermal and hydraulic parameters during injection and production at the start of the simulation. These gradients decrease later gradually.

The corresponding temperature field shows a significant cooling within a radius of less than 1 km around the injection well (Figure 5.12) for all considered permeabilities (Table 5.3). Longer test runs imply that a first thermal breakthrough will not occur before 90 years after the start the circulation. Furthermore, the fluid from the injection borehole is almost in equilibrium with the surrounding at the producer after this time.

5.1.2.2. Quantification of Uncertainties in Temperature Variation with Time

In order to quantify variation of temperature of the produced water with time as well as its uncertainty, I apply again Sequential Gaussian Simulation (Deutsch & Journel, 1998) to account for the spatial variability of the permeability within the Delft sandstone by sampling from a given permeability probability distribution (Simmelink & Vandeweyer, 2008). For this purpose, the decimal logarithm of permeability is varied stochastically within each grid node of the Delft sandstone within the 3D reservoir model. This provides an ensemble of different realizations of the spatial permeability distribution. Then, heat transport and fluid flow are simulated for each ensemble realization individually. Thus, for each simulation time step and grid node I obtain a temperature distribution which allows quantifying temperature uncertainty.

No observations of state variables are available, such as tracer, pressure or temperature data. Therefore, further constraining of the ensemble is not possible. At this stage, the study is limited to the temperature prediction based on available information on the local rock properties until production data from the drilled boreholes will become available.

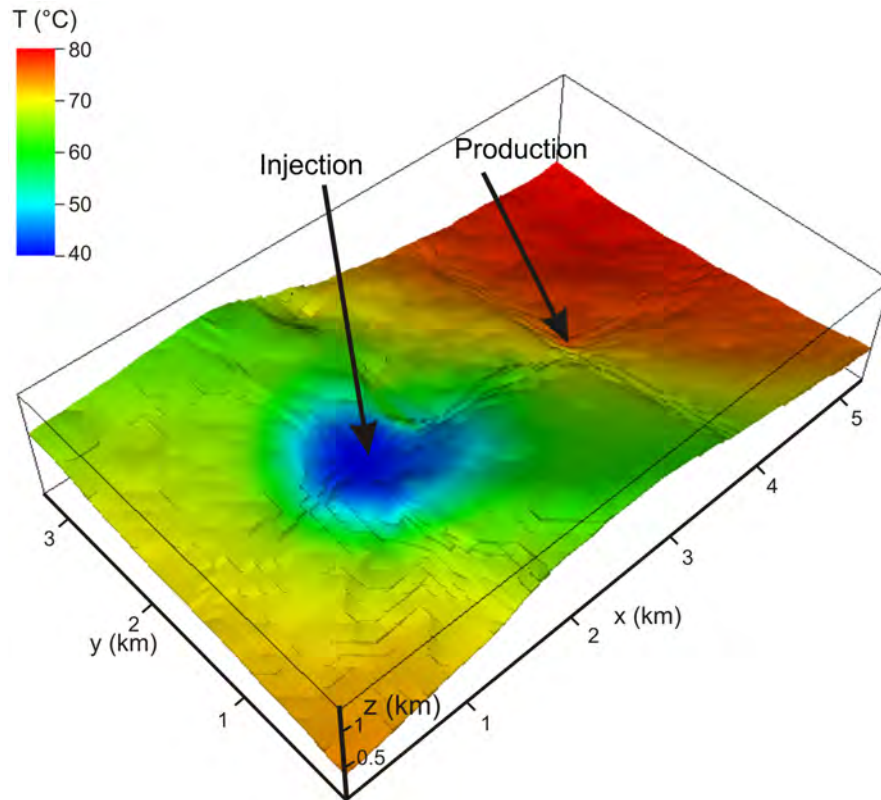


Figure 5.12.: Temperature field inside the Delft Sandstone after 50 years of operation, for a reservoir permeability of $5 \times 10^{-13} \text{ m}^2$. Adapted from Mottaghy et al. (2011b).

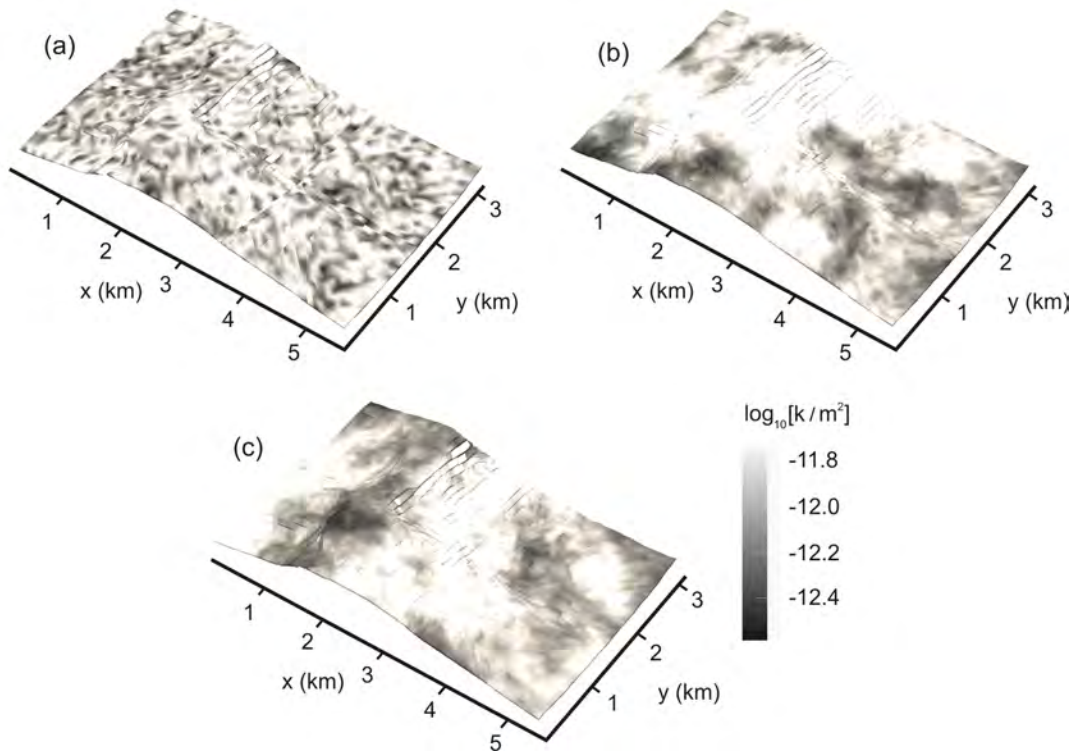


Figure 5.13.: Typical examples for the log permeability fields $\log_{10}[k]$ generated with SGSim for different correlation lengths: 100 m (a); 1000 m (b); 2000 (m) (c).

5.1.2.2.0.1. Target Histogram and Correlation Length

The permeability distribution of the Delft sandstone is assumed to be of triangular shape (Figure 5.10(a)) with values between $0.25 \times 10^{-12} \text{ m}^2$ and $2 \times 10^{-12} \text{ m}^2$ as reported by Simmelink & Vandeweyer (2008). The corresponding decimal logarithmic histogram used for the SGSim simulation (the target histogram) is illustrated in Figure 5.10(b).

Unfortunately, no prior information about geological spatial correlation is available. Thus, the correlation length of the permeability field is unknown. To overcome this lack of information, I implement different horizontal correlation lengths and investigate their effect on the temperature field during the operation of the geothermal doublet. Three correlation lengths are simulated applying SGSim: (a) the dimension of grid cells of our model (100 m), (b) an intermediate length (1000 m) and (c) the distance between injection and production wells (2000 m). The vertical direction is negligible due to the small thickness of the Delft sandstone (55 m on average). Typical examples for the log permeability fields generated with SGSim are illustrated in Figure 5.13.

One ensemble of 500 realizations is generated for each of the three different correlations lengths. Each realization is then simulated over 50 years.

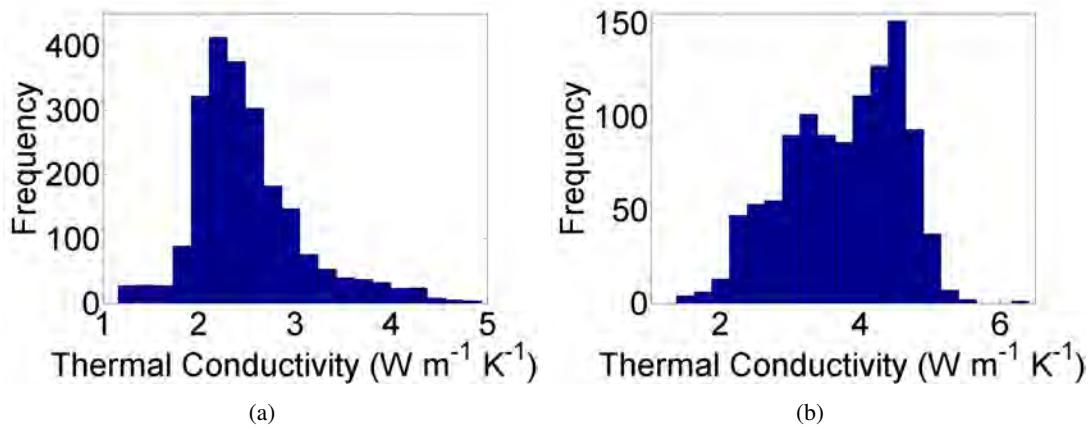


Figure 5.14.: Target distributions of the thermal conductivity in the Lower Cretaceous Supergroup (a) and the Jurassic Supergroup (b) obtained from the logging data. The average thickness of these layers is approximately 1100 m and 550 m, respectively.

5.1.2.2.0.2. Variability of Thermal Conductivity

Regarding the reservoir model, numerical experiments show that the fluid flow between the two wells affects the temperature field more than conductive heat transport. Therefore, an average value of thermal conductivity is used initially for all layers of the model. However, in a second step of this stochastic approach I consider also an additional variation of thermal conductivity based on data observed for each geological layer (see Section 5.1.1). Figure 5.14 shows the thermal conductivity distributions for the Lower Cretaceous Supergroup and the Jurassic Supergroup which comprise the Delft sandstone.

5.1.2.2.1. Results and Discussion

Figure 5.15 illustrates the predicted transient temperature variation over 50 years for all 500 realizations at the production well (red) and at a position at the center of the model within the Delft Sandstone at a depth of about 2000 m (blue) for a correlation length of 1000 m. As expected, the variation of the temperature at the production well is quite low because the cold water front does not reach this position until the end of the simulation. The temperature variation in the center of the model, however, appears to be much stronger. Therefore, the effects of the differing permeability fields can be studied better at the model center. This is obviously due to the smaller distance between injection and production well. Therefore, the closer the distance between production and injection wells, the stronger spatial variability of permeability affects temperature variations, and, hence, the more important become stochastic modeling studies for

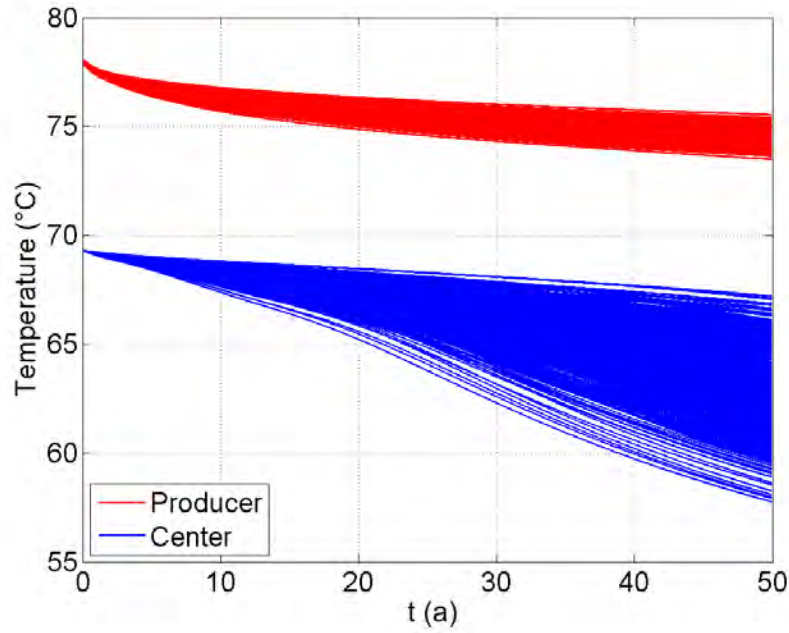


Figure 5.15.: Prediction of transient temperature variation for 50 years of all 500 realizations at the production well (red) and at a position at the center of the model within the Delft Sandstone at a depth of ≈ 2000 m (blue) for a correlation length of permeability of 1000 m.

predicting the long-term behavior of such geothermal installations.

Summarizing the results, Figure 5.16 and Figure 5.17 illustrate the temperature distribution after 50 a of operation at the position of the production well and the center of the model, respectively. Table 5.4 shows the corresponding mean and standard deviation for each ensemble. In contrast to the standard deviations, the mean values of all ensembles do not vary significantly. Whereas there is little difference for correlation lengths of 1000 m and 2000 m, a value of 100 m yields a significantly smaller standard deviation. This is due an averaging effect when using high-frequently spatial permeability variations which come with a small correlation length. Large correlation lengths, in contrast, allow very high or very low permeability in the entire model. Interestingly, the standard deviation for case of a correlation length of 1000 m is even larger than 2000 m. Using an ensemble size of 500 realizations, statistical reasons for this behavior due to small sample size can be excluded because the standard deviation does not change between simulation runs based on different random seeds. More likely, a lack of variation resulting from large correlation length becomes dominant here. Therefore, the effect averaging of realizations for small correlation lengths on the one hand and the lack of variation for large correlation lengths on the other hand results in a maximum standard deviation for the ensemble

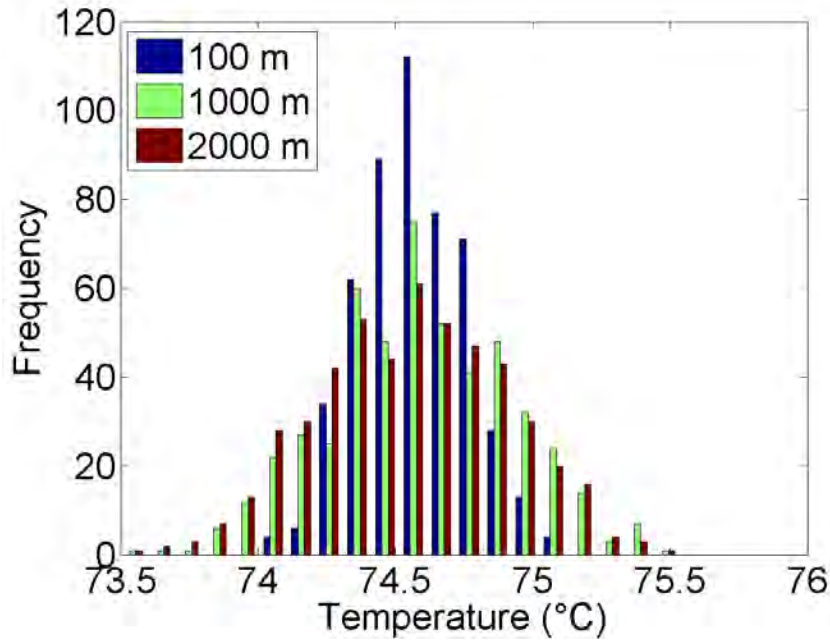


Figure 5.16.: Temperature distribution after 50 years of operation at the production well for different values of the horizontal correlation length of permeability.

simulated for a medium correlation length. In summary, assuming a correlation length of 1000 m results in the largest temperature variations. Therefore, in a conservative approach, this value must be considered for predicting the uncertainty of temperature.

For geothermal projects facing similar geology but closer distance between injection and production borehole, the temperature variation at this center position may be of special interest.

Taking into account the additional influence of thermal conductivity variation yields statistics which are listed in Table 5.5. The effects of a variation of (i) thermal conductivity variation only, (ii) permeability variation only for correlation length of 1000 m and (iii) a combined variation are compared. As illustrated by Table 5.5, thermal conductivity variation appears to be insignificant. Whereas the spatial thermal conductivity distribution plays an important role regarding the initial temperature field of a geothermal reservoir (Section 5.1.1), the uncertainty in the temperature evolution at the production site is dominated by the permeability distribution.

With the information about mean and standard deviation, the uncertainty in temperature prediction based on rock properties is quantified. Due to the unknown spatial characteristics, a correlation length of 1000 m has to be assumed conservatively because it yields the largest standard deviation. Even larger correlation lengths do not affect the temperature histogram significantly. This provides a maximum uncertainty at the production well characterized by a mean of 74.5 °C

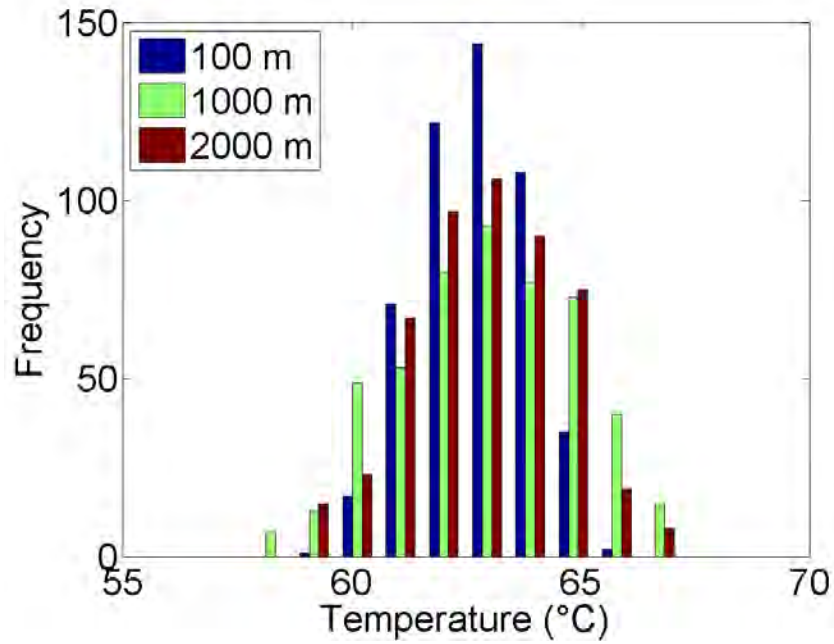


Figure 5.17.: Temperature distribution after 50 years of operation at the center position for different values of the horizontal correlation length of permeability.

Table 5.4.: Temperature prediction given as mean μ and standard deviation σ at the production well and the center of the model and for different correlation lengths a of permeability.

a (m)	Producer		Center	
	μ (°C)	σ (K)	μ (°C)	σ (K)
100	74.6	0.19	62.7	1.16
1000	74.6	0.35	62.9	1.92
2000	74.6	0.34	62.9	1.60

Table 5.5.: Temperature prediction given as mean μ and standard deviation σ for a correlation length a of permeability of 1000 m when also the thermal conductivity variation is considered. SGSim is applied on (i) thermal conductivity only, (ii) permeability only and (iii) both.

Case	Producer		Center	
	μ ($^{\circ}\text{C}$)	σ (K)	μ ($^{\circ}\text{C}$)	σ (K)
(i)	74.6	0.07	62.8	0.36
(ii)	74.6	0.35	62.9	1.92
(iii)	74.5	0.36	62.9	1.94

and a standard deviation of 0.36 K.

5.2. Quantifying Uncertainties for a Free Reservoir Layout in the Northeastern German Sedimentary Basin

Based on the lessons learnt from the last section, I perform now similar simulations in another reservoir. In this study I can choose the well distance. In addition, more detailed seismic information is available and used more comprehensively (Section 5.2.2). Uncertainties involved from the basal specific heat flow are assessed using deterministic inversion (Section 5.2.3). Uncertainties in flow rate arising from uncertain hydraulic parameters are predicted in addition to the statistical moments of temperature (Section 5.2.4). Finally, an alternative doublet system concept is studied based on a single geothermal well (Section 5.2.5).

This case study addresses a depleted oil and gas reservoir in the Northeastern German basin. The basin is potentially interesting for geothermal development and was studied regarding geothermal energy use by Hahne et al. (2011). Potential reservoirs were studied e. g. in Hamburg (Mottaghy et al., 2011a) and Groß Schönebeck (Ollinger et al., 2010; Cherubini et al., 2011). In both cases, numerical reservoir modeling was used to infer information on reservoir properties or transient processes: Mottaghy et al. (2011a) studied the influence of salt domes on the temperature field and proposed target coordinates for a geothermal well. Ollinger et al. (2010) inverted temperature profiles for estimating the regional thermal conductivity. Cherubini et al. (2011) modeled the influence of faults in this region on the temperature field. Rühaak et al. (2010) modeled a similar problem for the Molasse Basin, southern Germany.

Szalaiová (2012) modeled a hypothetical doublet system within a fault zone in crystalline rocks at the KTB-site, a continental 9 km deep borehole in southeastern Germany. Tischner et al. (2010) discussed a single-well concept for deep geothermal systems as an alternative to double-

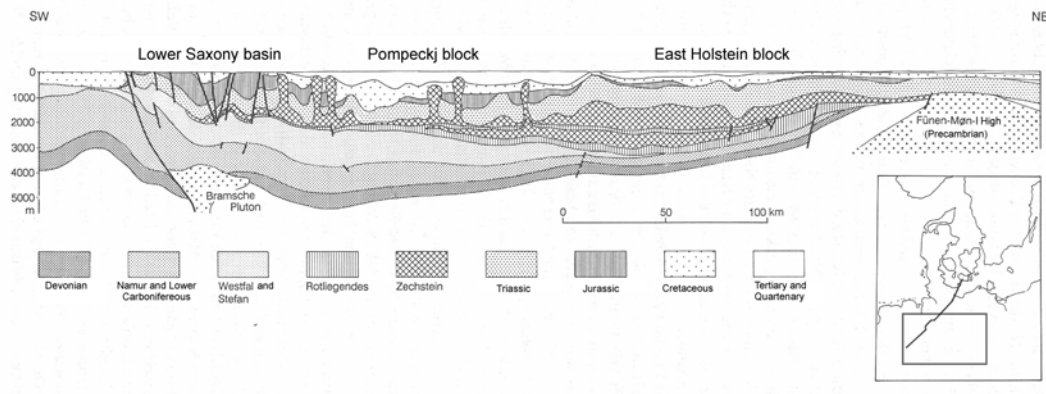


Figure 5.18.: Geologic section along the profile shown in the small map of the Northeastern German basin. The box inside the map indicates the target region. Modified after Walter (2007).

well systems. They tested two concepts: (i) a huff-puff scheme where water is first injected in and then produced from the same target layer, and (ii) circulation between two horizons connected by a stimulated hydraulic fracture. In the present study, I test both, a doublet layout in a single sandstone layer, as well as a single-well concept with circulation in two layers connected by a natural fault.

The target region is located in the center of the Northeastern German sedimentary basin and in the south of the Pompecki block (Figure 5.18). The basin contains several salt structures. Although there are no salt bodies in the subsurface of the location itself, there are three salt domes within a distance of about 10 km, which penetrate into the Lower Cretaceous and Tertiary sediments (Baldschuhn et al., 2001). These salt structures may affect the temperature field in the target region.

In the direct subsurface of the reservoir location, there are several big tectonic faults within the layers deposited from the Middle Triassic down to the Lower Jurassic.

Norden et al. (2008) report an average geothermal temperature gradient of about 35 K km^{-1} for this region. The thermal regime was also modeled by Vosteen et al. (2004) by 2D inversion. They found specific heat flows of 50 mW m^{-2} – 65 mW m^{-2} at a depth of 6 km.

The target layer for the reservoir in this study is the Upper Triassic Rhaetian sandstone at a depth of about 2000 m. It was chosen because an empirical relationship between porosity and permeability is available for the porous sandstone in this horizon (Pape et al., 2005).

5.2.1. Influence of Salt Domes on Temperature?

First, I study whether the temperature field in the target region of the reservoir is affected by salt domes. The available temperature data comprise 39 corrected bottom-hole temperatures (BHT) from boreholes in the region (see Figure 5.19). Data are provided by RWE Dea AG. To account for disturbance of the measured temperature due to the drilling process, the BHT values are corrected, e. g. using the Horner plot method (Horner, 1951). I divide the BHT into two groups: one group at a depth of ~ 1500 m and one at a depth of ~ 2000 m. The only deep BHT from 5000 m is not taken into account here, because lateral temperature variations cannot be derived from a single observation. I compute means and standard deviations for each group and compare the BHT values to these statistical moments. As shown in Figure 5.19, only in the area to the upper-left of the figure, temperatures appear systematically lower. However, this has no impact on the modeling because the area of lateral temperature variation occurs at least 2000 m away from the assumed position of the geothermal well inside the reservoir model. Assuming a Darcy velocity of 10^{-7} m s^{-1} within the low temperature area, these low temperatures will not reach the production well within hundreds of years after start of production in the reservoir. Compared to the results of the simulation (Section 5.2.4), even this Darcy velocity is largely overestimated.

For the reservoir region, there is no relevant spot in the region where the temperature is significantly higher, or lower, than the mean temperature of the group. Therefore, I approximate the temperature field as homogeneous in the lateral directions and assume no significant influence of salt domes or other tectonic systems.

5.2.2. The Geometric Model

The geometry of the reservoir is derived from a 3D seismic data set (Figure 5.20(a)) and from a stratigraphic model (Figure 5.20(b)) constructed using Petrel (Schlumberger, 2012) based on about 100 boreholes down to a depth of 2500 m plus one borehole (Z-1) with a depth of 6000 m (Gnjezda, 2009).

The 3D reflection seismic data set, provided by RWE Dea AG, images sediments in the North-eastern German basin in a cube of a horizontal area of about $5 \text{ km} \times 5 \text{ km}$ and a depth of 6 km. The seismic data set consists of 128 inline profiles \times 202 crossline profiles. In Figure 5.20(a), the seismic amplitudes are sign-color-coded and plotted versus recording time. For inferring depth, the seismic velocity model is calibrated based on sonic logs from wells and mean velocities from geophone measurements. With the help of borehole information and logging data from one deep borehole I divide the depth column into 18 major units and groups of layers (Table 5.6). The interfaces between the layers are afterwards identified in the 3D seismic data. To construct an equidistant grid with 50 m spacing for the thermal simulation model and 5 m spacing for the

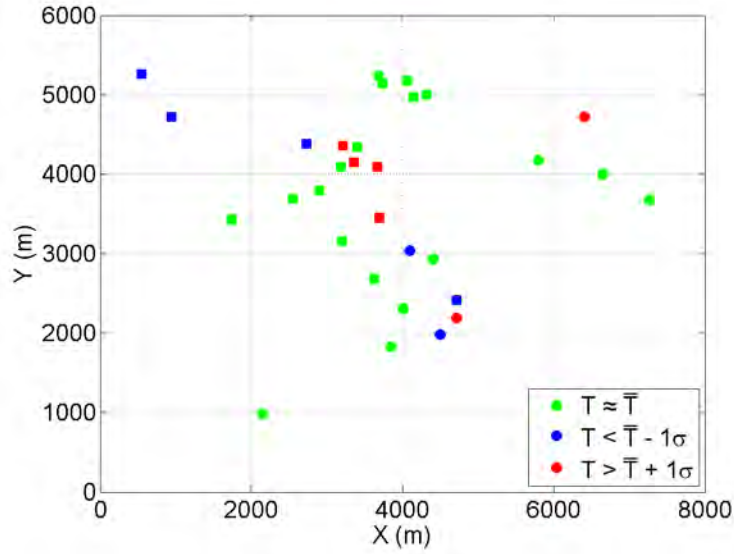


Figure 5.19.: Horizontal projection of the BHT variation in the target region. Colors indicate whether the BHT values scatter within one standard deviation σ around the mean temperature \bar{T} for two different depth ranges: above (dots) and below (squares) a depth of 1700 m.

reservoir models, these horizons are picked on every fifth seismic trace and the picks are interpolated from continuous horizons and discretized onto the equidistant grid. This yields gridded models with information on the unit numbers and, thus, rock properties of the different layers (Figure 5.21). This way, each horizon can be addressed separately for assigning properties. The units of the model and the corresponding rock properties are shown in Table 5.6.

Since the thickness of some of the layers is below seismic resolution, data from borehole Z-1 is essential to calibrate the depths of the picked horizons. This concerns also the target horizon (Table 5.6). Finally, the locations of the layers inside the geometrical model are verified by comparison with information from the stratigraphic model at all borehole positions (Figure 5.20(b)).

The Rhaetian sandstone at a depth range of 1928 m – 1951 m is selected as target horizon for five reasons: (i) it is characterized by an increased porosity of about 10 %; (ii) at least sparse permeability data are available from hydrocarbon exploration, (iii) the steady-state temperature prediction satisfies the minimum requirement of 75 °C, (iv) an empirical relationship between porosity and permeability is available for this horizon (Pape et al., 2005); (v) the target horizon is enclosed by impermeable layers of clay-rich rocks.

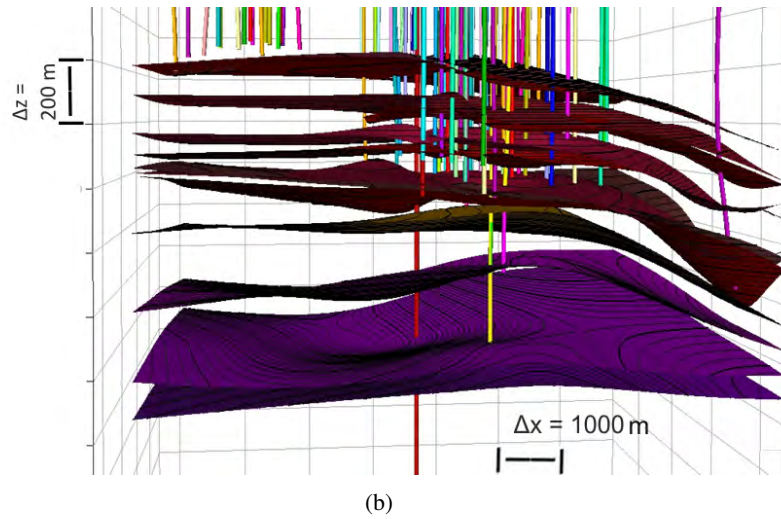
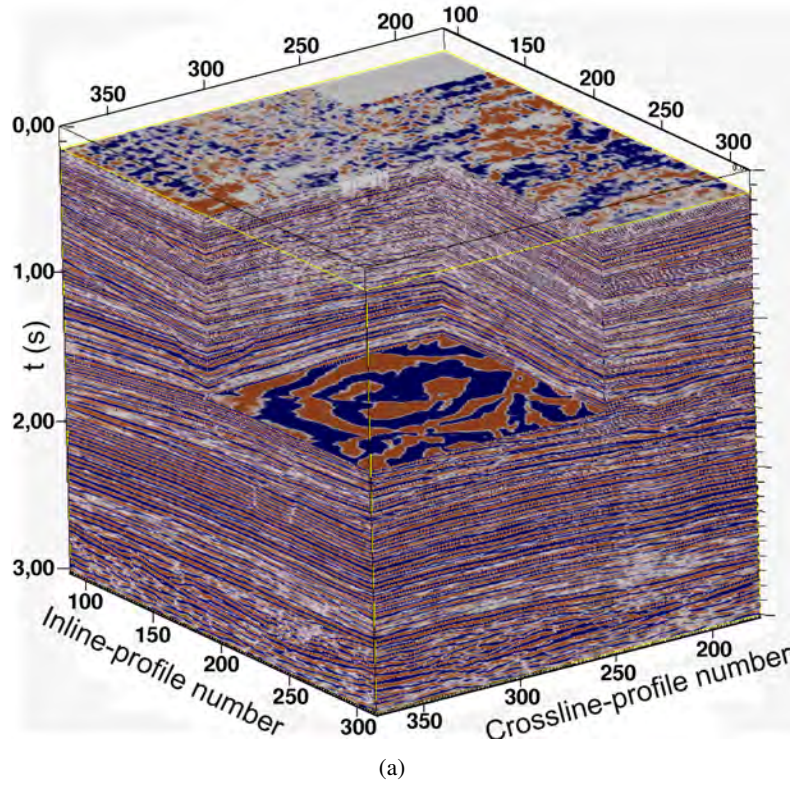


Figure 5.20.: Data sets used to derive a geometric reservoir model: (a) reflection amplitude polarities (red and blue) in a 3D seismic data cube (128 inline profiles \times 202 crossline profiles \times time t), from Vogt et al. (2013), and (b) stratigraphic model (Gnjezda, 2009) produced by Petrel (Schlumberger, 2012) from ~ 100 boreholes.

Table 5.6.: Units of the geometrical model, their corresponding lithology and average rock properties for the models: porosity ϕ , rock matrix thermal conductivity λ_r , and heat generation rate H . Note that I use the actual porosity and thermal conductivity histograms for the stochastic approach, not only mean values. Volumetric heat capacity of the rock for all units is $(\rho c)_r = 2.3 \text{ MJ m}^{-3} \text{ K}^{-1}$ according to Beck (1988).

Unit	Lithology	ϕ (-)	λ_r ($\text{W m}^{-1} \text{ K}^{-1}$)	H ($\mu\text{W m}^{-3}$)
1	sand/clay	0.23	3.95	0.52
2	limestone	0.15	2.64	0.35
3	marl	0.18	2.76	0.59
4	limestone	0.09	2.76	0.31
5	clay-marl	0.12	2.15	0.96
6	sand/clay	0.11	3.16	0.88
6a	target horizon 1 Angulate sandstone	0.11	4.04	0.60
6b	target horizon 2 middle Rhaetian sandstone	0.09	4.09	0.60
7	clay	0.01	2.17	0.90
8	salt	0.01	4.40	.035
9	clay/anhydrite	0.001	2.91	0.88
10	limestone	0.001	2.44	0.60
11	salt	0.001	4.57	0.31
12	limestone	0.001	2.35	0.68
13	clay/anhydrite	0.001	3.06	0.88
14	salt	0.001	5.21	0.32
15	sand/clay	0.001	2.75	1.00
16	salt	0.001	4.88	0.26
17	clay	0.001	2.33	2.09
18	Rotliegend volcanics	0.001	2.02	2.69

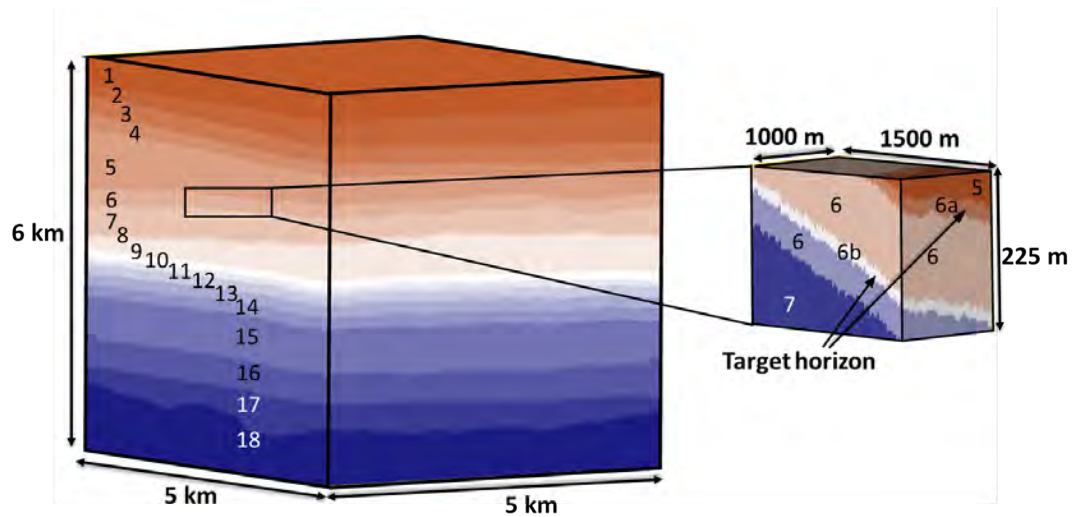


Figure 5.21.: Geologic layers identified from seismic and stratigraphic data used for the thermal model (left) and close-up used for the reservoir model (right). Colors indicate property units of the models. Unit numbers correspond to those in Table 5.6. In the thermal model, layers 6a and 6b are not resolved. For the doublet layout, layer 6b is the target horizon, for the single-well layout, the target layers are 6a and 6b connected by a fault zone. Note that the close-up is shown with vertical exaggeration. From Vogt et al. (2013).

5.2.3. The Thermal Model

Based on the geometrical model and following the approach applied for the reservoir in The Hague, I simulate subsurface temperature in a regional scale model first and then use the results as initial conditions for simulating the transient operating detailed reservoir models.

For obtaining reliable temperature information in the target horizon, I create a coarse 3D model (5 km × 5 km × 6 km) with about 1.4 million grid cells. It comprises all the 18 sedimentary layers down to the crystalline basin (Figure 5.21, Table 5.7). Large-scale models for geothermal reservoirs are essential to model the vertical and lateral variations in the temperature field.

Numerical tests showed that heat advection does not affect the temperature field significantly, because numerous little permeable salt or shale layers inhibit vertical fluid flow on large scale. This agrees with the conclusions of Vosteen et al. (2004). Therefore, I simulate purely conductive, steady-state heat transport.

Table 5.7.: Properties of the 3D thermal model.

Parameter	Value
Mesh size	$100 \times 113 \times 121$
Resolution	$50 \text{ m} \times 50 \text{ m} \times 50 \text{ m}$
Matrix thermal conductivity	$1 \text{ W m}^{-2} \text{ s}^{-1} - 4.5 \text{ W m}^{-2} \text{ s}^{-1}$
Porosity	0.1 %–50 %
Temperature at top	11 °C
Basal specific heat flow	$(77.7 \pm 1.2) \text{ mW m}^{-2}$
Stochastic parameters	porosity, matrix thermal conductivity, basal specific heat flow

5.2.3.1. Rock Properties for the Thermal Model

Rock properties for the thermal model (porosity, thermal conductivity, heat generation rate) are assigned to each layer using a proprietary statistical data set. Corresponding mean values are given in Table 5.6. However, note that I use the actual histograms for the stochastic approach. Underlying rock samples and logging data from the drillhole Z-1 is provided by RWE Dea AG and interpreted by Vogt et al. (2013) according to the approach discussed for the reservoir in The Hague (Figure 5.2).

5.2.3.2. Determining Specific Heat Flow by Inversion

For inverting specific heat flow, I use the arithmetic mean values of the rock property histograms for each of the 18 stratigraphic units in the model (Table 5.6). This yields single constant values for each layer. Now, the temperature field can be computed using the 3D model. However, a critical parameter for calibrating the temperature model is the basal specific heat flow. Therefore, I perform an inversion for specific heat flow using corrected BHT data from 39 available observations following the deterministic Bayesian scheme of Rath et al. (2006) in SHEMAT-Suite. The corrected BHT data were measured from an interval of 1500 m to 2500 m depth, with exception of the 5000 m deep borehole. I take into account given errors (of usually $\pm 5 \text{ K}$), if available. Otherwise, I use typical errors of $\pm 9 \text{ K}$ (Deming, 1989) for the corrected BHT. Norden et al. (2008) provide a value of 75 mW m^{-2} for the regional specific heat flow at 6000 m depth. I use this as *a priori* value with a large assumed error of $\pm 10 \text{ mW m}^{-2}$. I assume an annual mean temperature of 11 °C for the surface and an isolating boundary condition at the lateral model sides.

This inversion yields a basal specific heat flow value of $77.7 \text{ mW m}^{-2} \pm 1.2 \text{ mW m}^{-2}$. This agrees well with the result of Norden et al. (2008), but is significantly larger than the results of Vosteen et al. (2004) based on a 2D model, who proposed $50 \text{ mW m}^{-2} - 65 \text{ mW m}^{-2}$ at a depth

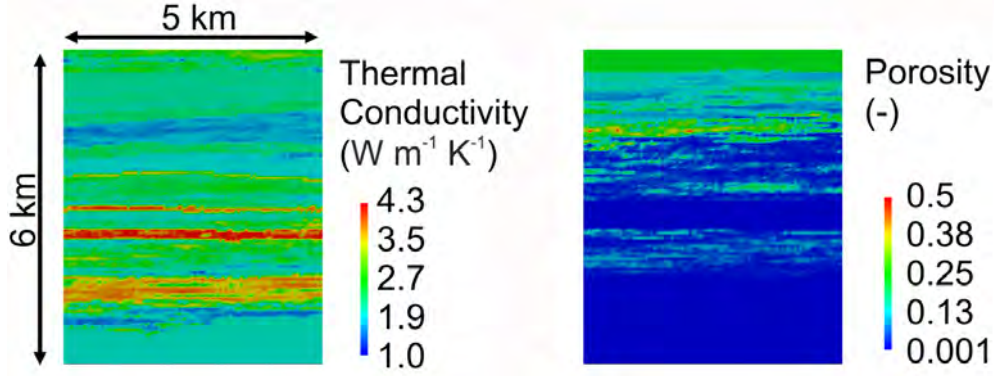


Figure 5.22.: 2D slices of a realization as examples for a stochastic distribution of thermal conductivity and porosity.

of 6 km. However, the 2D section modeled by Vosteen et al. (2004) is about 100 km apart from the reservoir location.

5.2.3.3. Quantifying Uncertainties of the Temperature Field

To quantify temperature uncertainties of the undisturbed reservoir, I analyse 700 realizations of the subsurface. Porosity and thermal conductivity are varied according to their observed distribution measured in borehole Z-1 (see examples in Figure 5.22). Additionally, I vary the basal specific heat flow, using a Gaussian distribution with mean and standard deviation according to the Bayesian inversion.

In order to model the spatial correlation, I assume a spherical variogram for creating stochastic realizations which is a standard type of correlation function in geostatistical modeling (Deutsch & Journel, 1998). For obtaining vertical correlation lengths, I fit a variogram to the inverted thermal conductivity and porosity data in borehole Z-1. This yields vertical correlation lengths in the range of 25 m to 100 m for each geological layer. For the horizontal directions, I obtain correlation lengths between 200 m and 2000 m by fitting variograms to seismic amplitudes provided by the approach described in Section 5.2.2. I infer correlation length for porosity variations from seismic attributes, because seismic reflectivity is a function of seismic impedance variations of the rock, seismic impedance being the product of density and seismic velocity. Both, density and elastic wave velocity are commonly supposed to depend on porosity (e. g. Schön, 1996). I assume the same correlation lengths for thermal conductivity.

Figure 5.23 illustrates the result of the inversion and the uncertainty given by the ensemble of profiles. The remarkable reduction of uncertainty by $\sim 67\%$ by this approach becomes obvious when comparing the results with the original temperature uncertainty from the corrected BHT

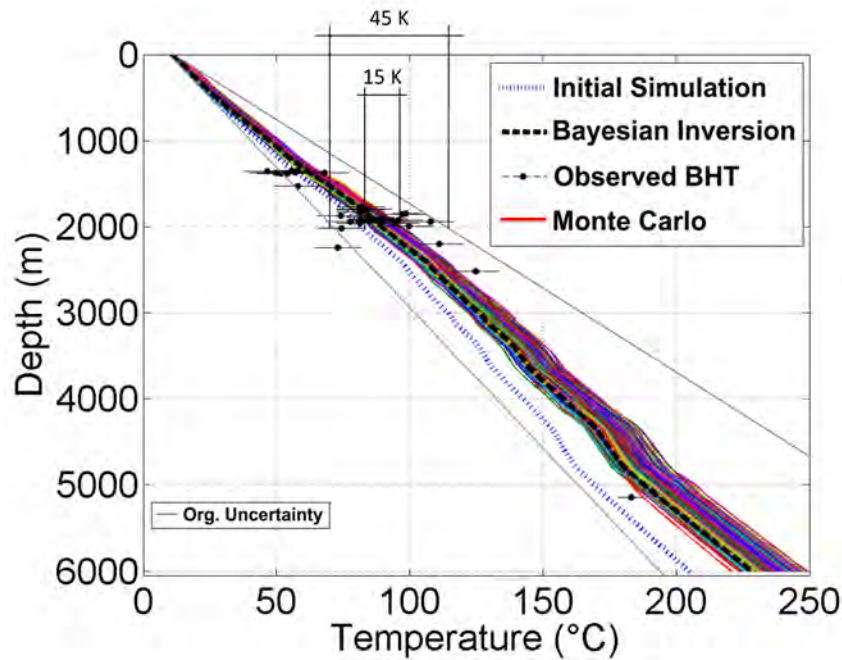


Figure 5.23.: Temperature prediction at the borehole Z-1 including the profile resulting from initial estimates (blue dotted line), Bayesian inversion (black dashed line), and uncertainty quantification (colored lines) based on 700 Monte Carlo realizations. Also shown are corrected BHT data used for the inversion (black dots) and the original temperature uncertainty estimated from the variation of the BHT data (gray lines). Note that only the deepest BHT is located in the borehole, for which the shown profile is simulated. The overall reduction of temperature uncertainty is indicated for a depth of 2000 m.

data. At the center of the target horizon (Rhaetian sandstone) at 2000 m depth, the study predicts a temperature of $87.1^{\circ}\text{C} \pm 1.8\text{ K}$.

5.2.4. Doublet Reservoir Model

Based on the results of the regional-scale thermal model, I analyze the performance of a potential geothermal doublet in the study region (see close-up in Figure 5.21) in terms of transient temperature and pressure variations and their uncertainties. I lay out a hypothetical doublet according to the requirements with regard to pressure and temperature as in the real world heat-use project in The Hague, The Netherlands, with a projected flow rate of 42 L s^{-1} and a temperature of 75°C . In contrast to the project in The Hague, I choose a borehole distance of 500 m (instead of 2000 m) in order to minimize computing time and to maximize the sensitivity of temperature at the production well on the permeability distributions. Additionally, the smaller hydraulic conductivity compared with The Hague yields a larger reservoir pressure gradient. This behavior is

Table 5.8.: Properties of the 3D reservoir model.

Parameter	Value
Mesh size	$50 \times 75 \times 45$
Resolution	$20 \text{ m} \times 20 \text{ m} \times 5 \text{ m}$
Temperature at top	84.5°C
Temperature at bottom	93.0°C
Porosity	$1\% - 20\%$
Permeability	$10^{-18} \text{ m}^2 - 10^{-13} \text{ m}^2$
Circulation rate	42 L s^{-1}
Temperature of injected water	40°C
Stochastic parameters	porosity \rightarrow permeability

mitigated by the smaller well distance.

I simulate the doublet in a detailed reservoir model ($1000 \text{ m} \times 1500 \text{ m} \times 225 \text{ m}$, Table 5.8) as indicated in the close-up of Figure 5.21 with grid cells of $20 \text{ m} \times 20 \text{ m} \times 5 \text{ m}$. The region comprises no faults. The inclined target horizon is vertically discretized with at least three grid cells. Additionally, the target horizon does reach neither the top nor the base of the reservoir model. I assign constant temperatures taken from the larger steady-state thermal model to the top and base of the model. Since the target horizon extends laterally beyond the model boundaries, I fix the initial hydraulic head constant at the model boundaries. This allows fluid in- and outflow across the boundaries. However, numerical tests show that the model is laterally large enough to suffer only little from the influence of the hydraulic boundary conditions. I implement a circulation rate of 42 L s^{-1} and a re-injection temperature of 40°C . 900 numerical time steps for 20 years of simulation time prove to be sufficient for all realizations, because an increased number of time steps does not yield different temperature predictions.

5.2.4.1. Hydraulic Rock Properties

A stochastic approach is applied again for generating 400 realizations of the spatial porosity distribution and, hence, address its heterogeneity and uncertainty. To this end, SGSim is conditioned to a porosity distribution inverted from logging data taken from the Z-1 well. Permeability is the primary rock property controlling fluid flow. Unfortunately, no permeability data are available from any borehole within target region. Therefore, I use an empirical relationship between porosity ϕ and permeability k , available for the Rhaetian sandstone (Pape et al., 2005):

$$k = 0.309(100\phi)^{4.85} \times 10^{-18} (\text{m}^2). \quad (5.3)$$

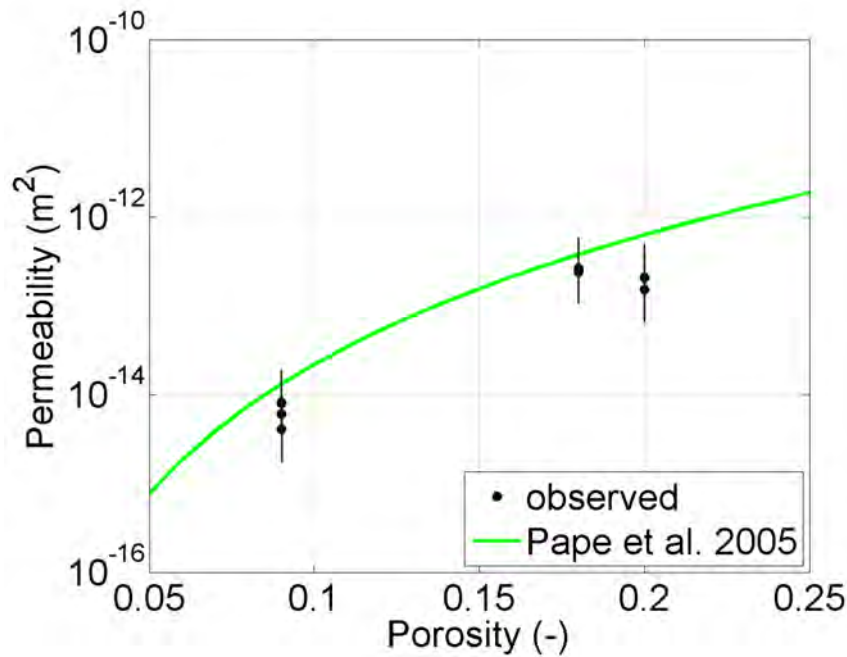


Figure 5.24.: Observed permeability in the Rhaetian sandstone (personal communication by RWE Dea AG) and empirical porosity-permeability relationship of Pape et al. (2005).

Sparse permeability data from adjacent regions and from the same lithology (personal communication by RWE Dea AG) confirm the relationship within the given errors (Figure 5.24). In the following, I use the original parameters of Pape et al. (2005) in equation (5.3) because only three available data points with large error bars do support a reliable data fit. Figure 5.25 illustrates the observed porosity and calculated permeability histograms used with SGSim to create stochastic permeability variations for the modeling.

5.2.4.2. Results and Discussion for the Doublet Layout

Figure 5.26 shows reservoir temperature for one of the 400 Monte Carlo realizations after 20 a of heat production together with the corresponding porosity and permeability field. In Monte Carlo modeling, one single realization contains little information. Therefore, the transient variation of temperature of all 400 realizations is shown in Figure 5.27. By far most of the curves show an early thermal breakthrough, i.e. an arrival of cold water at the production borehole after only a few years. This indicates a high probability for temperatures to fall below the required 75 °C during the operation. A larger borehole distance would result in a later thermal breakthrough. However, a larger distance corresponds to a larger pressure drop at the production borehole.

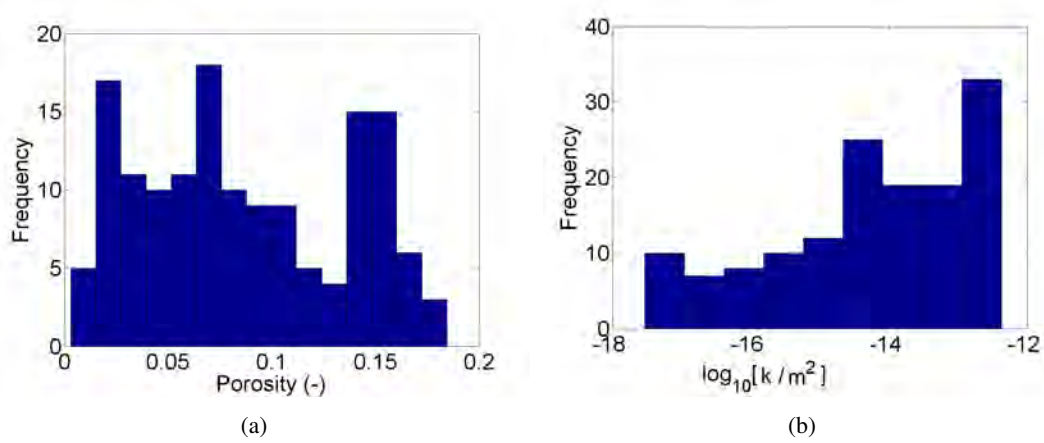


Figure 5.25.: (a) Observed histogram of porosity in the Rhaetian sandstone and (b) corresponding permeability k calculated following Pape et al. (2005).

Finally, I evaluate the number of ensemble realizations meeting the requirements for pressure and temperature for estimating the success probability of a hypothetical geothermal project.

By using a Monte Carlo ensemble, I obtain a best estimate (ensemble mean) of the temperature and its uncertainty (ensemble standard deviation) after 20 years of operation. However, the energy output depends not only on temperature, but also on the flow rate $q_w = vA$, where A is the well surface. The specific discharge v depends on reservoir pressure drop with respect to equation (3.1). Given a certain flow rate, insufficient permeability k may result in pressure drops which are not achievable by common pumps. Therefore, pressure is also of major importance for evaluating the risk of the geothermal project. This pressure can be evaluated stochastically from the Monte Carlo ensemble of realizations. To this end, I calculate the pressure inside the well p_w for each realization from the corresponding simulated block pressure p_b of the grid cell in the numerical model which represents the production well. I follow the semi-analytical approach of Peaceman (1983) for the case of fluid production:

$$p_w = p_b - \frac{q_w \mu_f}{2\pi k \Delta z} \ln \left(\frac{r}{r_w} \right). \quad (5.4)$$

Here, the flow rate q_w is $0.042 \text{ m}^3 \text{ s}^{-1}$, Δz the vertical cell size is 5 m, r_w the well radius is 0.08 m, and r is the equivalent radius of the cell, defined for cubic cells of size Δx and isotropic permeability by:

$$r = 0.14\sqrt{2}\Delta x, \quad (5.5)$$

with $\Delta x = 20 \text{ m}$ in this case.

Figure 5.28 illustrates the variation of production well pressure with time in comparison with

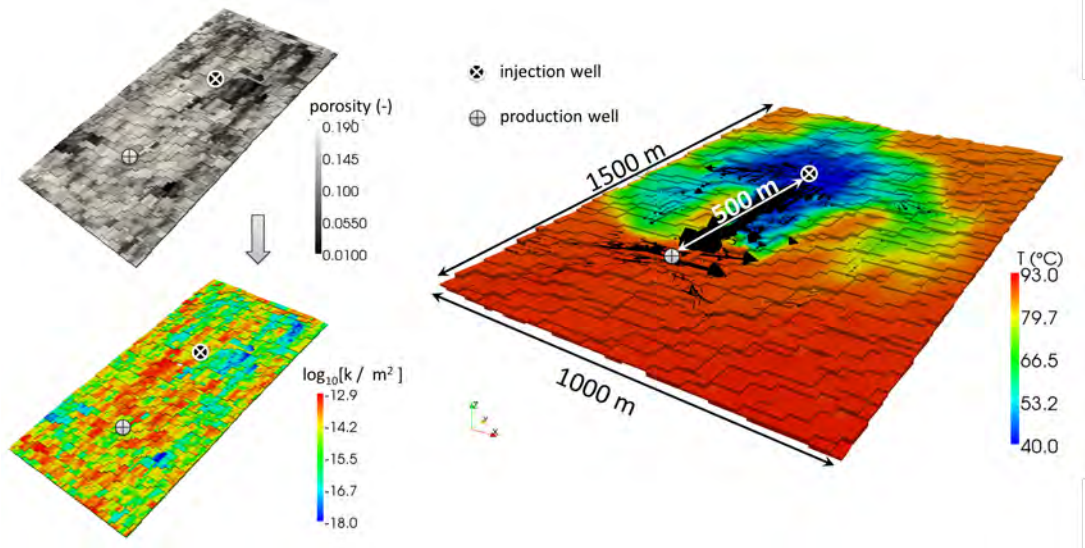


Figure 5.26.: Reservoir porosity, corresponding permeability after Pape et al. (2005), and corresponding temperature for one Monte Carlo realization after 20 a of heat production. Velocity arrows indicate the flow direction. Heterogeneity in temperature result from heterogeneities in permeability, not from boundary effects.

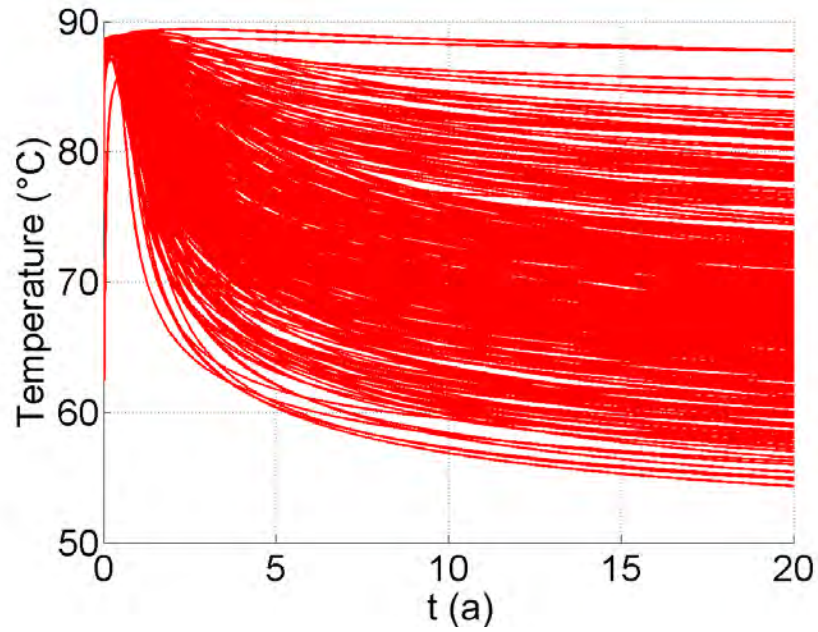


Figure 5.27.: Variation of temperature with time at the production well for all 400 realization of the Monte Carlo ensemble.

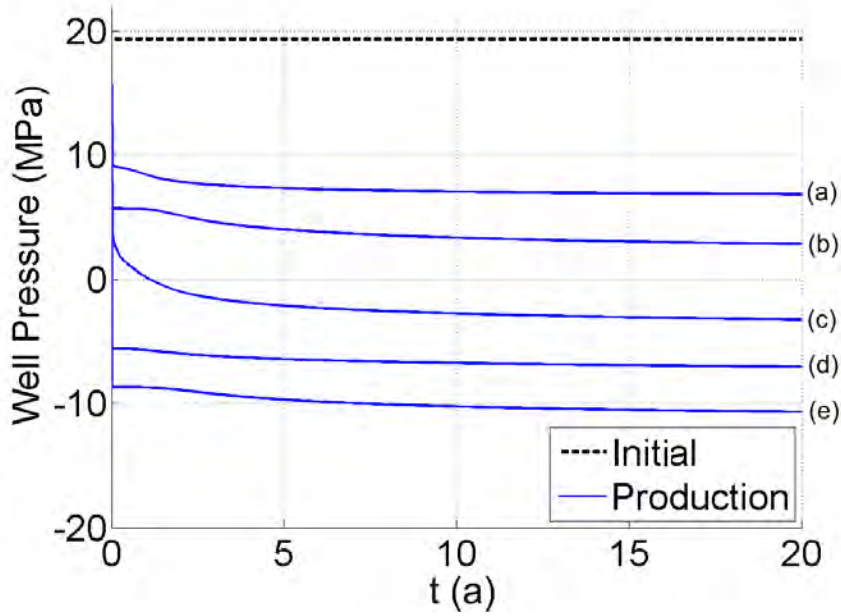


Figure 5.28.: Variation of production well pressure with time at the production well for five sample realizations of the Monte Carlo ensemble (a-e) and initial reservoir pressure before start of operation.

the initial reservoir pressure. As in the entire ensemble pressure varies from -80 MPa to 20 MPa, I show only five sample realizations to demonstrate the following findings: Curves (a) and (b) in Figure 5.28 shows that the well pressure drops instantaneously by ~ 10 MPa when production starts and decreases slowly further with time. This slow decrease is a result of the predefined constant flow rate as the low permeability of the reservoir allows no sustainable recharge. Thus, the well pressure becomes smaller. In the case of curve (c), the pressure would become negative after two years due to this effect according to equation (5.4). Of course, in a real-world reservoir, negative absolute pressure can not be realized. In this case, the given circulation rate of $0.042 \text{ m}^3 \text{ s}^{-1}$ cannot be sustained. For curves (d) and (e), the realized permeability does not even allow positive pressures at the very beginning of production. For this modeling, I use negative pressures as indicator for failing of a particular ensemble member when realizing the given pumping rate.

The average productivity index \overline{PI} of the N ensembles members is defined by the flow rate q_w divided by the pressure drop $p_i - p_{w,j}$:

$$\overline{PI} = \frac{1}{N} \sum_{j=1}^N \frac{q_w}{p_i - p_{w,j}}, \quad (5.6)$$

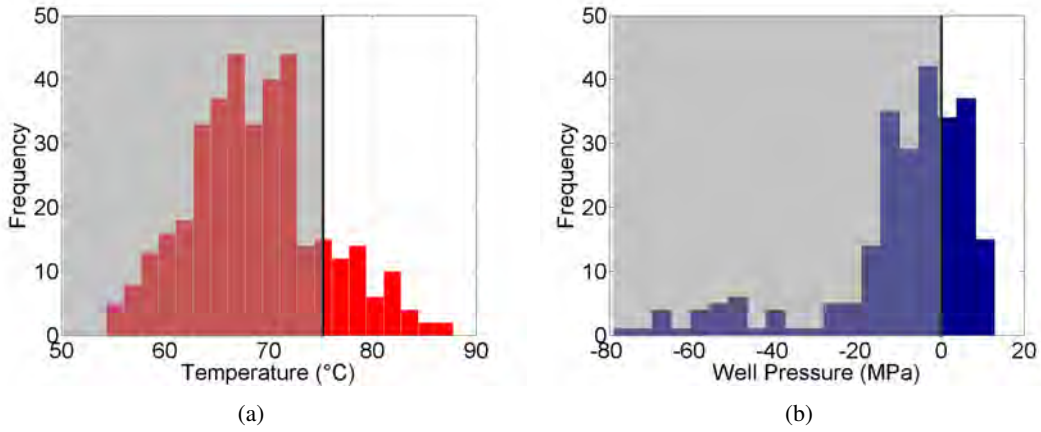


Figure 5.29.: (a) Temperature and (b) pressure distribution according to the stochastic ensemble after 20 a of operation. Also shown are the projected requirements for a geothermal installation (vertical lines). For the temperature requirement (a) the threshold is supposed to be similar to the heat-use project in The Hague, Netherlands (Vogt et al., 2010). For well pressure (b), negative values indicate that the flow rate requirement of the heat-use project cannot be reached. Use of feasible pumps is assumed.

where p_i denotes the initial reservoir pressure and j the realization index. For this case, I find an average $\overline{PI} = 2.2 \text{ L s}^{-1} \text{ MPa}^{-1}$, which is about two thirds of the productivity index of the geothermal installation at Soultz-sous-Forêts, France, (Tester, 2006). However, the installation at Soultz operates at about half of the production rate assumed for this hypothetical installation. When comparing the productivity indices of the porous reservoir and the Soultz fractured granite reservoir, it has to be considered that a linear relationship between pressure and flow rate are assumed (Darcy flow in equation (3.1). In particular in fractured hard rock, turbulent flow may occur which disturbs this linearity.

Figure 5.29 shows the probability distributions for temperature and well pressure. Based on these histograms, I can quantify the probability for a successful geothermal installation designed for 20 years of operation. I identify all realizations featuring a temperature above 75 °C and non-negative pressures at the required flow rate of 42 L s^{-1} as successful and promising for operating the reservoir.

Now, I can determine how many realizations meet the requirements for temperature (16.3 %), for pressure (22.8 %) and both combined (1.6 %). As mentioned above, a good connectivity corresponds to a small pressure drop but also to an early thermal breakthrough. This yields a very low probability of success (1.6 %) for the hypothetical project. Therefore, I recommend not to use the Rhaetian sandstone in this area as a reservoir for a geothermal doublet within the defined requirements.

5.2.5. Single-Well Reservoir Model

During seismic interpretation, different fractures and faults were detected within the regional model. Due to the predicted insufficient performance for the target sandstone, now I also study an alternative design using a fluid circulation along one of these detected fault zone connecting the two sandstone aquifers middle Rhaetian sandstone and Angulate sandstone (indicated in the close-up in Figure 5.21). The latter is located about 150 m above the Rhaetian sandstone. This scenario has the additional advantage of requiring only one well for operation. Since drilling costs are a major part of the project cost, saving one borehole may make a project more attractive for investors.

For modeling this single-well layout, detailed information is needed about the structure of faults, in particular the targeted one (Figure 5.30). Complex fracture networks offer potential pathways for fluids or may act as barriers. Detailed knowledge as well as an assessment of its reliability is essential. To detect these fractures Vogt et al. (2013) apply an automatic fracture detection algorithm which can handle the large amount of data. The algorithm is based on the fact, that sediments mostly feature plain coherent reflectors in which fractures manifest themselves as lateral disturbances. Therefore they can be detected using a coherence-based algorithm described by Gersztenkorn & Marfurt (1999) including a dip calculation based on the gradient structure tensor (Bakker, 2002). The result of this analysis is a data cube that contains parameters describing the seismic signal coherence along the reflecting horizons. Fault displacements in these horizons are identified by lateral minima of the coherence field. The computation is performed in the following way: In a small sub-volume scanning of the data cube, the coherence is calculated following a dip correction. The result is a 3D coherence cube, gridded in the same way as the geometrical model with binary information on each sub-volume, whether it is part of a fracture structure or not. The fracture inventory of the study location is classified into two domains: an upper zone with a great amount of small cracks and a lower zone, near the target horizon, which is dominated by larger fractures.

As an additional verification (Vogt et al., 2013) checked the size-frequency distribution of the automatically detected faults that turn out to be compatible with fractal distributions typically found for geologically mapped fault systems (Turcotte, 1997).

I assume thermally isolated fluid flow within the geothermal well for upstream and downstream (a best-case assumption). This allows the injection of cold water in the upper (Angulate) sandstone, circulation through the fault zone and the production of warm water from the lower (Rhaetian) sandstone using just one well.

Since I have no data for estimating the hydraulic characteristics of the fault zone, I cannot quantify, as previously, the uncertainty of reservoir performance in a Monte Carlo scheme. In contrast, I use a simplified approach based on different possible connectivity cases. To this end,

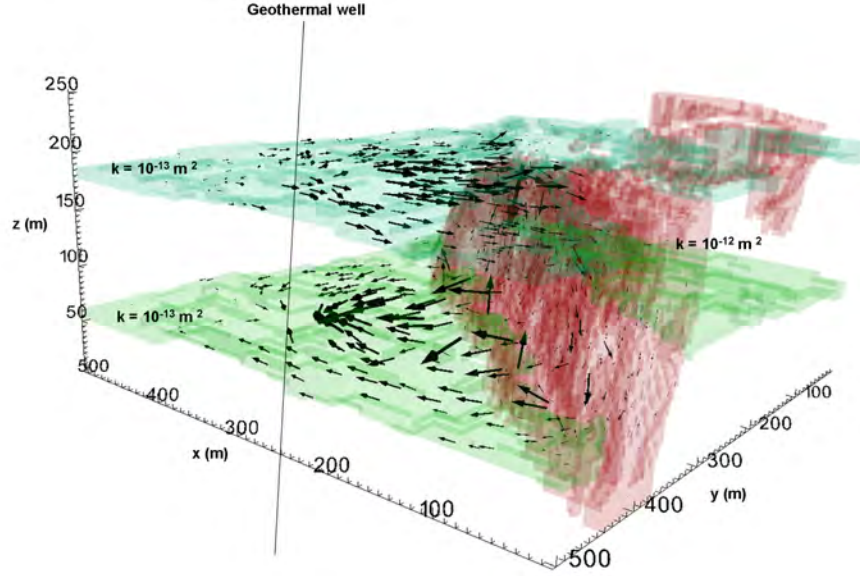


Figure 5.30.: Single-well-layout: geothermal well at a distance of 250 m from the fault zone, Angulate sandstone (dark green) and middle Rhaetian sandstone (light green), fault zone (red), and resulting flow field (velocity arrows only qualitative). Permeabilities k are assumed as constant within each layer and the fault zone.

I study the sensitivity of pressure and temperature variation with time on different average fault zone permeabilities and porosities. I simulate two different scenarios: (i) ~ 250 m and (ii) ~ 40 m distance between well and fault zone. Geometrically the fault zone is modeled by a minimum of three cells of dimension $5 \text{ m} \times 5 \text{ m} \times 5 \text{ m}$. As the real lateral extension of the fault zone is unknown, I assume these dimensions to avoid disturbance of the modeled fault zone by an insufficiently fine discretization. Model parameters are summarized in Table 5.9. Figure 5.30 illustrates the well position near the fault zone in scenario (i) and shows the resulting flow field in the two target layers as well as within the fault zone.

Again, the simulation extends over 20 years, discretized into 15 000 numerical time steps of equal length of about 12 hours. For all models, I assume a circulation rate of 42 L s^{-1} and a homogeneous permeability of 10^{-13} m^2 for the target layers. Initial hydraulic heads and temperatures are taken from the steady-state temperature model (see Section 5.2.3). I study different permeabilities (10^{-14} m^2 — 10^{-11} m^2) and porosities (5 % — 20 %) for the fault zone.

Potential mechanical or chemical effects (such as permeability or porosity changes) due to the re-injection of reservoir brine into a different horizon are ignored.

Table 5.9.: Properties of the 3D reservoir model.

Parameter	Value
Mesh size	$101 \times 101 \times 51$
Resolution	$5 \text{ m} \times 5 \text{ m} \times 5 \text{ m}$
Temperature at top	76.5°C
Temperature at bottom	86.5°C
Porosity	5 %–20 %
Permeability	10^{-14} m^2 – 10^{-11} m^2
Circulation rate	42 L s^{-1}
Temperature of injected water	40°C
Stochastic parameters	none

5.2.5.1. Results and Discussion for the Single-Well Layout

The results of these simulations for scenario (i) are shown in Figure 5.31 for (a) and (b) temperature and well pressure, respectively, assuming various permeabilities and porosities for the fault zone. Obviously, also in this single-well scenario, an early thermal breakthrough occurs after about 2.5 years of operation time. All hydraulic scenarios show a similar temperature variation with time. Surprisingly, a high permeability of 10^{-11} m^2 results in higher temperatures after 20 a of operation. This might be explained in the following way: The steeper drop at the beginning due to a good hydraulic connectivity is offset because hot water from deeper parts of the fault is produced at later operation time resulting in somewhat warmer production fluid at later times. For $k = 10^{-14} \text{ m}^2$, the simulation is numerically unstable due to the low resulting pressure. A production scenario would fail, as for $k = 10^{-13} \text{ m}^2$ and $k = 10^{-12} \text{ m}^2$ because well pressure drops below zero. This indicates that the circulation rate cannot be sustained after about 7 years and 15 years, respectively, as discussed above. Only the best-case scenario with a permeability of 10^{-11} m^2 and a porosity of 20 % maintains a sufficient well pressure for continuous heat production. However, for a real installation, the well needs to be placed further away from the fault or the distance between the chosen targeted sandstone layers should be larger to avoid an early thermal breakthrough.

Figures 5.32 compares the results for scenario (ii) where the well is closer to the fault zone in comparison with the best result for scenario (i) with a permeability of 10^{-11} m^2 . In scenario (ii), the length of the direct flow paths is shorter by about 445 m. In terms of pressure, this scenario is more suitable. A similar pressure is obtained even though permeability is one order of magnitude lower. However, due to the short distance the thermal breakthrough occurs much earlier. The temperature declines to 60°C about 10 years earlier. Thus, a geothermal installation is not

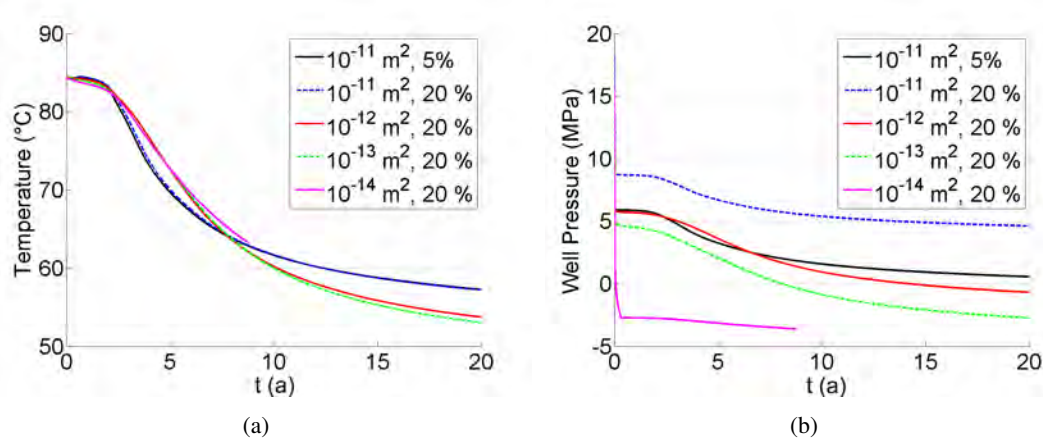


Figure 5.31.: Transient variation of (a) temperature and (b) well pressure for different hydraulic properties (permeability and porosity of the fault zone) for scenario (i) at the production level of the single-well layout.

feasible under the given requirements, but could well be operated if lower temperatures are satisfying for direct heat use. To make use of a fault for a geothermal reservoir in geologic conditions similar to the studied ones, I therefore recommend a distance of at least 250 m between well and fault.

5.3. Discussion

For a hypothetical installation in both, single-well or doublet scenarios, the small thickness of the target sandstones combined with insufficient permeabilities does not support a successful heat production. Even a larger distance between the wells of a doublet for avoiding the early thermal breakthrough is unlikely to help, because of the expected larger pressure drop resulting from a larger distance.

To overcome this obstacle, an alternative approach is conceivable. It may comprise, besides hydraulic or chemical stimulation, a single-well concept using fluid-flow along a fault. Such fault can be identified by seismic interpretation. According to my findings, a fault permeability of 10^{-11} m^2 and a porosity of 20 % or more are most suitable for operation in terms of variation of pressure and temperature with time. In contrast to the doublet (Section 5.2.4), suitable well pressures due to high permeabilities do not prevent high temperatures, because hot water may be produced from deeper regions of the fault. However, the distance between the target layers in a real installation should be significantly larger than the one in this study as long as the pressure drop, i.e. the permeability distribution, allows this. I recommend a distance of at least 250 m between well and fault. In addition, an isolated piping in the well may be required when using a

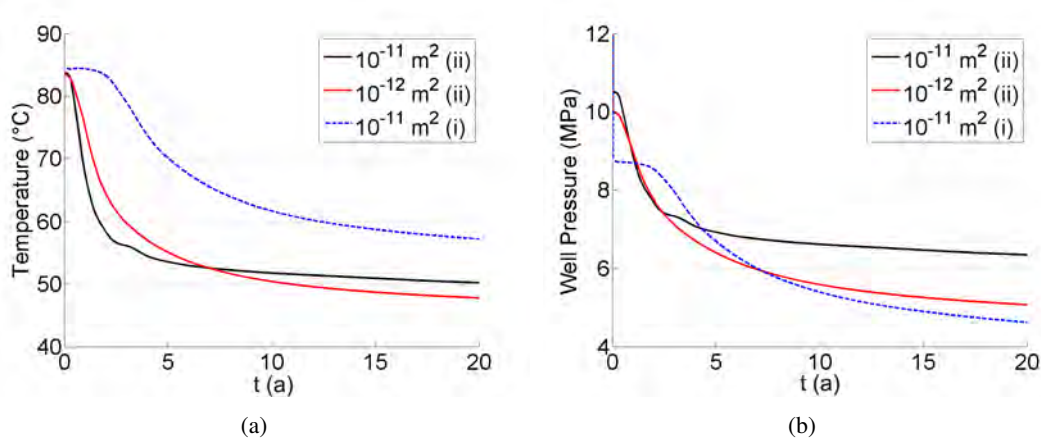


Figure 5.32.: Transient variation of (a) temperature and (b) well pressure for different hydraulic properties (permeability and porosity of the fault zone) for scenarios (i) and (ii) at the production level of the single-well layout. Porosity is 20 % in all cases.

one-well design to avoid cooling of the produced water within the well.

When interpreting results, it must be mentioned that I do not account for chemical reactions within the reservoir. A temperature drop due to cold water injection may also change the reservoir's chemical equilibrium, resulting in chemical precipitation or dissolution. The latter may clog the pore space, resulting in an increase of permeability. This effect may alter the flow regime (Pape et al., 2005) significantly. However, this will require additional numerical simulation of chemical reactions similar to the implementation by Clauser (2003).

An additional approximation is the neglect of any stress field, which could yield information on whether faults are more likely to be open or sealed for fluid flow. Further changes of pressure due to forced circulation may affect the permeability, in particular fault permeability. Accounting for this will require the simulation of geomechanics, e. g. following the approach of Watanabe et al. (2010).

Production temperatures are estimated at the bottom of the well. Due to cooling on the way up to the surface, I very roughly estimate a temperature drop of about 1.0 K for this case based on computing heat losses in a analytical pipe flow scenario (as in Section 5.1.1.2.1).

After quantifying uncertainty in sedimentary rocks, now I investigate granitic rocks. This comprises also identifying possible fluid pathways and comparing different techniques for parameter estimation.

Chapter 6.

Appraisal of Methods for a Stochastic Parameter Estimation in a Fractured Crystalline Reservoir

Usually, direct information on subsurface properties of geothermal reservoirs is derived from only a small number of boreholes. This applies in particular for deep reservoirs in the crystalline basement, where no data are available from exploration for oil or gas. To overcome this limitation, several inversion techniques exist for estimating the properties in the entire reservoir (e. g. Carrera et al., 2005; Hendricks Franssen et al., 2009).

I study the potential of the two approaches, massive Monte Carlo and the Ensemble Kalman Filter, to estimate the hydraulic properties and their uncertainties and heterogeneities of an Enhanced Geothermal System (EGS) at Soultz-sous-Forêts (Gérard et al., 2006), France, in the Lower Rhine Graben. Here, at approximately 5000 m depth an engineered reservoir was created between 2000 and 2007. However, estimating permeability is not easy while quantifying geothermal reservoir properties and their uncertainties is crucial for EGS reservoirs.

The EGS reservoir at Soultz-sous-Forêts has been discussed in several studies. Gérard et al. (2006) e. g. provided an overview of the project. Studies on brine-rock-interaction were performed by Baechler & Kohl (2005) and André et al. (2006). Baujard & Bruel (2006) highlighted the importance of fluid density for the development of EGS reservoirs using finite-volume simulations in a discrete fracture network. Delépine et al. (2003) derived an estimate of the reservoir's permeability based on the interpretation of microseismic events. The effects of the hydraulic fracturing on the permeability field at Soultz were studied by Kohl & Mégel (2007), who could explain the associated major seismic processes in the reservoir by modeling the rock's hydro-mechanical response of the rock matrix to massive hydraulic injections. Tenzer et al. (2010) identified geomechanical facies for the EGS sites at Soultz (France) and Bad Urach (Germany). Kolditz & Clauser (1998) discussed the characteristics of analytical reservoir models with simplified geometries and stochastic reservoir models with complex geometries. They presented a

deterministic numerical model of the HDR site at Rosemanowes (UK).

At Soultz, a tracer circulation test was performed in 2005 for studying the hydraulic connectivity between injection borehole GPK3 and the two production boreholes GPK2 and GPK4 (Figure 6.1). Based on these data, this study will provide information on the permeability distribution in the EGS reservoir after hydraulic stimulation. To this end, I investigate on the one hand possible additional pathways and hence, address the non-uniqueness of the problem (Section 6.4). On the other hand, I perform an inverse study by fitting a reservoir model to the observed tracer concentrations (Section 6.5). Finally, I compare results from different approaches (Section 6.6).

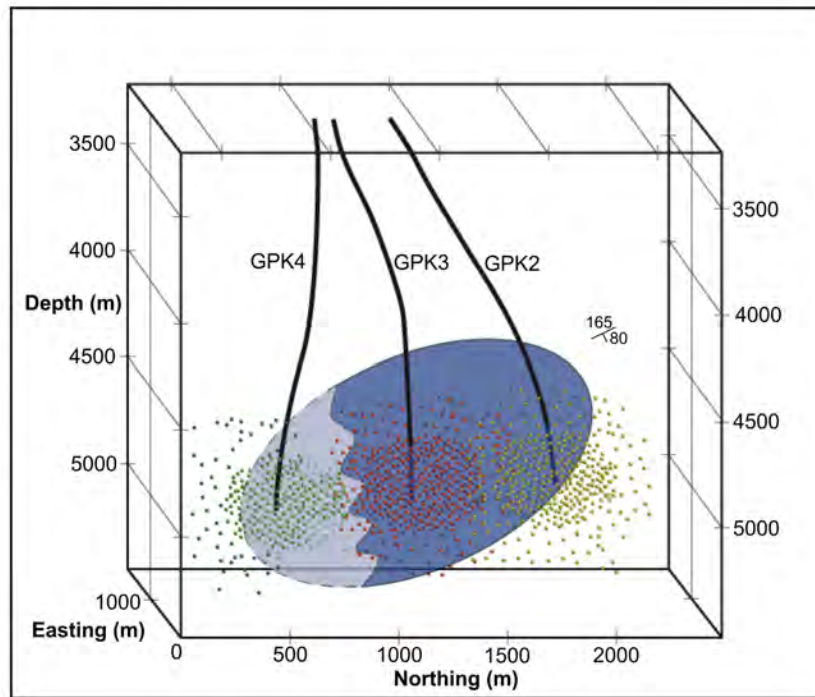


Figure 6.1.: Sketch of the borehole configuration at Soultz-sous-Forêts in 2005, adopted from Gessner et al. (2009). The fault is shown in dark blue which is intersected by the boreholes GPK2 and GPK3. The green, red and yellow dots indicate zones affected by the hydraulic stimulation of the boreholes. The part of the fault zone shown in light blue illustrates the unknown continuation towards GPK4.

6.1. The Geological Framework at Soultz-sous-Forêts

Soultz-sous-Forêts is located in the Rhine Graben which is part of the European Cenozoic Rift System. It developed during the Tertiary along pre-existing Permian troughs concomitant to the

European Alpine collision that formed the Alps. It has a pronounced Graben structure characterized by normal faults which can be followed in seismic sections down to 4000 m. Faulting is listric, displaces the Cretaceous sediments, and continues into the granitic basement where it shallows out (Cloetingh et al., 2006). On top, the Quaternary fluvial sediments of the Rhine cover the faulted sediments.

At Soultz, three major boreholes GPK2, GPK3, and GPK4 were drilled between 1987 and 2004. They intersect the sedimentary cover and reach to granitic basement. The boreholes GPK2 and GPK3 intersect a pre-existing fault at a depth of approximately 4700 m, while GPK4 misses the fault (Figure 6.1). This fault dips approximately 80° striking to 255° N (Gessner et al., 2009).

The granite can be divided into two classes. The first one is fresh granite, hardly affected by alteration. Therefore, it is dense, almost impermeable, and does not support fluid circulation. The second one is associated with highly fractured zones surrounded by the fresh granite. It shows signs of strong hydrothermal alterations. This may produce pore space implying that these regions may support fluid flow (Ledéseret et al., 2010). Thus, fresh and altered granite affect the hydraulic properties of the reservoir in different ways. However, the distribution of these rock types and their individual influence on the flow paths between the wells cannot be identified easily based only on drilling information and borehole experiments.

6.2. The Circulation Experiment of 2005

A long-term tracer circulation test was performed between the boreholes GPK2, GPK3, and GPK4, for characterizing the effects of hydraulic stimulation on the reservoir at a depth of 5000 m. From July to December 2005 a fluid volume of about $209\,000\text{ m}^3$ was injected into GPK3 and $165\,000\text{ m}^3$ and $40\,000\text{ m}^3$ were produced from GPK2 and GPK4, respectively (Sanjuan et al., 2006), resulting in a nearly equilibrated mass balance.

A mass of 150 kg of 85 % pure fluorescein was used as a tracer. The fluorescein was dissolved in 0.95 m^3 of fresh water yielding a concentration of 0.389 mol m^{-3} . This fluid was injected into GPK3 over a period of 24 hours. Simultaneously, geochemical fluid monitoring was started at GPK2 and GPK4. Figure 6.2 shows the measured fluorescein concentration in GPK2 and GPK4 versus time. Fluorescein was first detected in GPK2 4 days after the injection into GPK3. In GPK4, in contrast, fluorescein was first detected only 24 days after the injection. The average pumping rates were 11.9 L s^{-1} in GPK2, 15 L s^{-1} in GPK3, and 3.1 L s^{-1} in GPK4, already indicating a reduced water supply to GPK4. In combination with the tracer arrival times, this indicates the heterogeneous character of the hydraulic pathway, providing a much better connectivity between GPK3 and GPK2 than between GPK3 and GPK4.

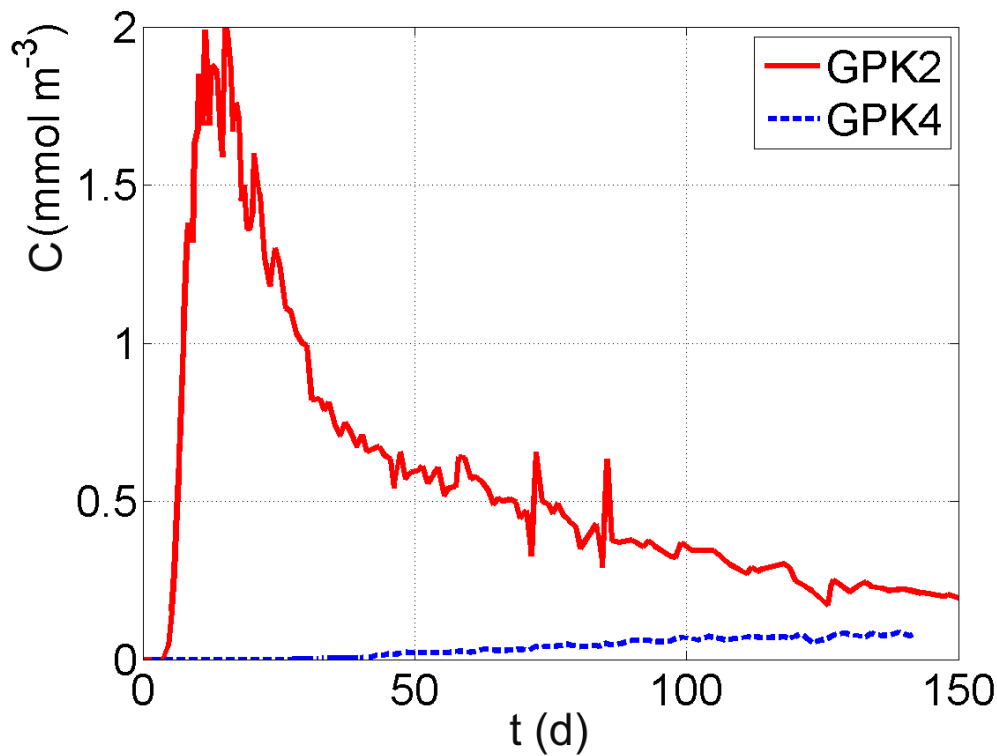


Figure 6.2.: Tracer concentration in GPK2 and GPK4 in response to injection in GPK3.

During the circulation test concentration was measured every 30 min. The maximum tracer concentration (breakthrough time) measured in the produced fluid was 1.9 mmol m^{-3} in GPK2, recorded 9 to 16 days after the injection (Figure 6.2). After 5 months, it dropped to approximately 0.2 mmol m^{-3} . No clear maximum tracer concentration could be seen in GPK4 even after 5 months of production. The final measured value was close to 0.09 mmol m^{-3} . The relative uncertainty of the measured tracer concentration was estimated to fall between 10 % – 15 %. During the entire experiment, only 23.5 % of the tracer was recovered (Sanjuan et al., 2006), indicating fluid exchange with the surrounding and an open system.

Besides the tracer concentration several other parameters were recorded during the circulation test, such as complete logs of temperature, pressure, and concentration recorded in GPK2, GPK3, and GPK4.

A first interpretation of the tracer experiment of 2005 was suggested by Sanjuan et al. (2006) based on an analytical model, which fitted the first 20 days of the measurement assuming three possible circulation loops for the tracer within the reservoir. The full 150 days of the curve were modeled by Blumenthal (2007), who simulated the tracer transport in a 2D numerical model. He

showed the dependency of the peak concentration with the variation of porosity and permeability. Nevertheless, the 2D model did not fit the tracer curve satisfactorily due to the assumed flow in a horizontal plane.

A numerical 3D model by Kosack et al. (2010) used a gradient-based deterministic Bayesian inversion technique (Rath et al., 2006). Hydraulic parameters were estimated in only two homogeneous flow zones. The model provided a nearly perfect fit of the tracer curve between GPK2 and GPK3 over the entire time of the experiment, capturing both, the short- and the long-term behaviors well. However, since model simplifications are inevitable in a reservoir of this dimension, no study so far provided a clue whether small-scale variations of the tracer curve are due to secondary flow paths or caused by measurement errors. The zones in the approach of Kosack et al. (2010) reflect only the main fault and the surrounding rock. Additional stimulated zones were considered in an approach comprising four zones. Again, this provided a nearly perfect fit of the tracer curve. This demonstrates that different geometrical models can provide comparably good fits and only products of parameters, e. g. porosity and permeability, can be obtained by the inversion. This is due to the high correlation associating low permeability with high porosity and vice versa in the zones.

In contrast to Kosack et al. (2010), Gentier et al. (2010a) and Gentier et al. (2010b) followed an approach motivated by identifying geological and tectonic structures, fitting permeabilities to an identified fracture network and a stochastic distribution of fractures. The fit was obtained by adjusting the parameters of a forward model, not by numerical inversion. For the first time in a numerical model of the Soultz reservoir, they fitted the tracer curves at GPK2 and GPK4 simultaneously. However, despite their detailed representation of geological evidence in the numerical model, the simultaneous fit obtained was by far inferior to the ones provided for the tracer concentration at GPK2 alone by the models with rigorously simplified geometries (Kosack et al., 2010). In addition, the elaborate model design of Gentier et al. (2010a) may be too complex to be used routinely for history-matching of geothermal reservoirs.

A simultaneous fit of the GPK2 and GPK4 tracer curves was done successfully by Held (2011) based on a discrete network of identified faults and objects identified by microseismic observations. Based on the results and financial modeling, production scenarios were optimized.

Recently, Radilla et al. (2012) fitted the tracer data from both wells simultaneously and successfully based on an equivalent stratified medium approach, too. Like Vogt et al. (2012) and this work, they found a narrow and very permeable connection as well as a second wide connection (corresponding to the surrounding rock in the approaches of Kosack et al. (2010) and Vogt et al. (2012)) between GPK3 and GPK2, and a wide but little permeable connection between GPK3 and GPK4.

6.2.1. Motivation for Applying Massive Monte Carlo and EnKF Approaches

In this study I apply two alternatives to the very simplified geometry of Kosack et al. (2010) on the one hand and detailed complex fracture networks of Gentier et al. (2010a) on the other hand. As one alternative, following Evensen (2003), I update numerically an ensemble of heterogeneous stochastic reservoir models in an EnKF approach. As an advantage, this approach also quantifies heterogeneously distributed uncertainty. As another alternative, I investigate possible additional pathways in the stimulated zones that can explain the observed tracer concentration during the circulation test of 2005 in a Monte Carlo approach. These simulations address the fact that the main fault in the reservoir could be totally or partially closed. This means that assuming this fault as the main pathway may oversimplify the problem and, consequently, other pathways for the fluid and tracer transport in the Soultz fracture network may be ignored.

My approach of studying the non-uniqueness of the inverse problem can set a light on permeable objects which are not visible as faults in the wells or in the microseismic cloud, too. These object may represent high permeable zones which are present before stimulation will not be affected by shearing. Therefore, not seismic events will occur in this zones.

A similar stochastic approach for uncertainty modeling in crystalline rock was described by Watanabe et al. (2010) for the EGS at Bad Urach, Germany. They also created an ensemble of Monte Carlo reservoir realizations. In contrast to this study, they used assumed permeability distributions before hydraulic fracturing, and accounted for the subsequent permeability enhancement by multiplying by a factor (decreasing with the distance from a borehole) after ensemble generation. However, when comparing Bad Urach with Soultz, it has to be taken into account that the reservoir in Bad Urach is located in a dense gneiss formation, not in granite. Gneiss has an inherent anisotropy. In contrast to Gentier et al. (2010a), Watanabe et al. (2010) used an equivalent porous medium to simulate fracture networks of EGS. This approach was also followed by McDermott et al. (2006) and Szalaiovà (2012). McDermott et al. (2006) characterized a fracture zone in crystalline rock in a forward simulation of a hydraulic-geomechanical 3D model at the KTB-site (a continental super-deep borehole) in southeastern Germany. Szalaiovà (2012) performed simulations based on SHEMAT-Suite (Rath et al., 2006) (as applied here) and, thus, characterized groundwater flow and heat advection in sedimentary and hard rocks at the KTB site. In the presented study, I also assume an equivalent porous medium for the simulations. However, turbulent effects can not be captured by this approach.

6.3. Modeling Approach

In contrast to sedimentary regimes, permeability in crystalline rocks is controlled by fractures. The basic idea behind the approach of the present study is to assign a heterogeneous permeability

field to the grid cells of a 3D model for representing zones of different fracture density. I use the algorithm of Sequential Gaussian Simulation (SGSim, see Chapter 3.4) for generating a number of random realizations of the decimal logarithm of the permeability field according to the distribution shown in Figure 6.3. The histogram shows a bimodal Gaussian distribution. Here, the high peak on the left represents the combined effect of background permeability on the scale of grains and low fracture density of the solid rock. The smaller peak on the right represents the fractured cells of the reservoir. Using this approach allows to capture fractures of very different scales: from major fracture zones cross-cutting the granite to intra-crystalline micro-fractures mechanically weakening the solid rock as reported by Dezayes et al. (2010). Further, I fixed high permeability values of 10^{-12} m^2 at the production and injection cells to enforce high permeability at the centers of hydraulic stimulation.

For this case, the correlation length a defines the range of similar fracture density and hence permeability. Magnitude and orientation of the maximum correlation length in different model directions should reflect typical fracture length and orientation, respectively. For evaluating a reliable value for a , I consider the induced micro-seismicity around GPK3 during the hydraulic stimulation performed in July 2000. According to Delépine et al. (2003), the total number of microseismic events drops to $\frac{1}{e}$ of the original value at a distance of 350 m. This distance is supposed to be a minimum correlation length in all directions. To honor an anisotropy along the main fault zone, I allow a maximum correlation length of 800 m in the horizontal directions of the model.

An initial study is performed based on a previous study by Kosack et al. (2010) to calibrate the model parameters by roughly fitting the tracer curve of GPK2 manually. Thus, I obtain the prior assumption for permeability distribution as well as value ranges. I assign a porosity of 5×10^{-4} . This low value is required for obtaining a fit of the tracer breakthrough time at GPK2 — as shown by numerous numerical tests that we performed.

Tracer concentration is set to zero at all boundaries of the model. Thus, tracer is allowed to cross all model boundaries and flow into the infinite reservoir around the model. This reflects the true conditions at Soultz where a connection of the stimulated zone to an extended groundwater system has been identified (Sanjuan et al., 2006). This explains the low tracer recovery of only 23.5 %. The hydraulic head is set to the initial values at all boundaries before production to allow also mass transport through the boundaries. The normal component of the specific heat flow is set to zero at the lateral sides, assuming they are thermally isolated. At the top, a Dirichlet boundary condition is specified for constant temperature and at the base a Neumann boundary condition for constant specific heat flow (Table 6.1).

The reservoir is considered at a depth range between 4500 m and 5500 m. In principal, the model consists of two zones: a stimulated zone with a heterogeneous permeability field in the

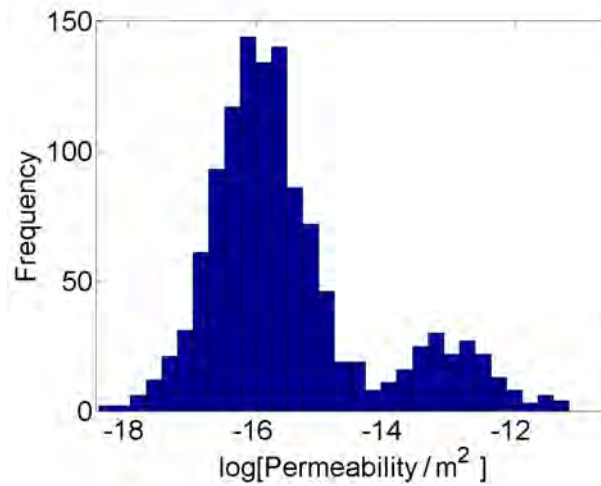


Figure 6.3.: Assumed log permeability distribution in the stimulated zone.

shape of an ellipsoid and a surrounding zone with a homogeneous permeability of 10^{-18} m^2 . The parameters and boundary conditions of the forward model are listed in Table 6.1.

For the fluorescein tracer, I consider the actual injection and pumping rates, temperature, and pressure data as discussed in Section 6.2 and use these data as time-dependent boundary conditions at all three well positions during this simulation. Moreover, the simulation considers the re-injection of the produced fluid into GPK3 with its chemical tracer concentration and heat content. Injected tracer concentration is implemented as as time-dependent boundary conditions at the grid cell corresponding to the injection well. I take into account the mutual dependencies of fluid properties on the one hand, and head and temperature on the other hand. This results in a higher computational cost of each forward simulation. Numerical tests without accounting for these dependencies yield insufficient results.

6.4. Massive Monte Carlo Study of the Fractured Reservoir

The model consists of $21 \times 36 \times 21$ grid cells on a regular grid with a cell size of 50 m in each direction. It comprises only GPK2 and GPK3. Experiments with finer computational grids yield similar results with respect to the calculated state variables, such as tracer concentration, but require a significantly larger computing time.

To account for the relatively weak decrease of tracer concentration with time at GPK2 (Figure 6.2), I suppose different tracer pathways. Therefore, I assume permeability as the parameter of major influence. Detailed information is given in Section 6.4.2.3 about the influence of a varying porosity on the results.

Table 6.1.: Hydraulic and thermal properties and boundary conditions of the 3D model.

Parameter	Value
Porosity	5×10^{-4}
Permeability	10^{-18} m^2 – 10^{-12} m^2
Rock compressibility	10^{-8} Pa^{-1}
Dispersion length	$5 \text{ m}^2 \text{ s}^{-1}$
Thermal conductivity	$2.5 \text{ W m}^{-1} \text{ K}^{-1}$
Heat capacity	$2.06 \text{ MJ kg}^{-1} \text{ K}^{-1}$
Radiogenetic heat production	$10^{-10} \text{ W m}^{-3}$
Hydraulic head at top	4 500 m
Temperature at top	185 °C
Specific heat flow at bottom	80 mW m^{-2}
Stochastic parameter	permeability

As mentioned in Section 6.2 and seen in Figure 6.2 the measured concentration in GPK4 is low compared with GPK2. This, as well as the pumping rates, suggests a very low connectivity between GPK3 and GPK4. Because of this fact, I consider first the circulation between GPK2 and GPK3 only.

6.4.1. Identification of Fitting Realizations

When applying Monte Carlo techniques, usually there is no interest in a single realization because it contains no probability information. Even mean and standard deviation of the entire ensemble may be misleading because important information may be discarded when interpreting only these statistical values. Therefore, the probability distribution containing all realizations needs to be investigated when using massive Monte Carlo methods (e. g., for a parameter of interest, such as temperature at the position of a borehole).

In this study, however, I investigate single realizations and groups of realizations with similar features in order to reveal general characteristics of the Soultz reservoir. Different possible permeability distributions are identified which provide optimal fits of the tracer data. Thus, different possible pathways are identified. Very different realizations can fit the observed tracer concentration with few common features. Therefore, presenting the average of all successful realizations as result is not regarded as meaningful.

I use three different misfit functions in order to select successful realizations: (i) root mean square error, (ii) Kolmogorow-Smirnow test, and normalized Kolmogorow-Smirnow test. However, judging fit quality just from objective values without any other information may be misleading. Therefore, visual control of the simulated produced tracer concentration and the 3D permeability field is also part of this analysis. I use the equally weighted *root mean square error* for the complete curve.

6.4.1.1. Root Mean Square Error

The *root mean square error* (RMSE) E is a standard measure for the misfit of simulated d_i^{sim} data for each realization i and observed data d^{obs} for n concentration data values:

$$E_i = \sqrt{\frac{1}{n} \sum_{j=1}^n (d_{i,j}^{sim} - d_j^{obs})^2}. \quad (6.1)$$

Small values of E_i correspond to a high quality of fit.

6.4.1.2. Kolmogorow-Smirnow Test

The Kolmogorow-Smirnow test is not as powerful as the *root mean square error* to quantify the absolute misfit. However, it is sensitive to the shape of the measured curve. That is, it will account for additional extrema or differences in the sign of the slope of the curve. Similar to the *root mean square error*, a misfit value \mathbb{K}_i can be defined for each realization i . The k th simulated cumulative tracer data point

$$\mathbb{F}_i(k) = \sum_{j=1}^k d_{i,j}^{sim} \quad (6.2)$$

is compared with its corresponding observed data point

$$\mathbb{S}(k) = \sum_{j=1}^k d_j^{obs} \quad (6.3)$$

by building upper and lower differences:

$$\mathbb{T}_i^u(k) = |\mathbb{F}_i(k) - \mathbb{S}(k)| \quad (6.4)$$

and

$$\mathbb{T}_i^l(k) = |\mathbb{F}_i(k) - \mathbb{S}(k-1)|. \quad (6.5)$$

The misfit value is thus obtained by the maximum of these values:

$$\mathbb{K}_i = \max(\mathbb{T}_i^u, \mathbb{T}_i^l). \quad (6.6)$$

6.4.1.3. Normalized Kolmogorow-Smirnow test

To identify simulated curves with fitting shapes, but wrong absolute values, I normalized the Kolmogorow-Smirnow test by dividing all observed and simulated values by their corresponding maximal value $\max(d_{obs})$ and $\max(d_{sim}^i)$, respectively. Thus, the misfit value \mathbb{N}_i is defined.

6.4.2. Modeling Results

Initially, 10^4 ensemble members with different spatial permeability distributions are created using the SGSim algorithm and propagated in time for simulating the circulation test. The simulation runs on 88 parallel threads for approximately 10 days on 88 Intel® Xeon® E5420 2.5 GHz processors in a ScaleMP architecture¹. These 10^4 realizations are evaluated with respect to their ability to fit of the observed tracer curve.

Then the *root mean square error* of the tracer concentration is calculated for every single realizations i . Visual inspection showed that realizations with an $E_i > 0.35 \text{ mmol m}^{-3}$ did not fit the tracer curve successfully. Thus, all realizations with an $E_i > 0.35 \text{ mmol m}^{-3}$ are discarded from the ensemble, leaving 120 realizations. Additionally, the ensemble is controlled visually. Realizations not fitting the shape of the tracer curve are discarded, i.e. realizations with multiple distinct extrema or a distinct re-rising tail. Even though visual control involves some degree of subjectivity, the human brain is an excellent tool for shape identification. Furthermore, the group of discarded realizations is investigated using the Kolmogorow-Smirnow test. Based on the result, again a visual control is performed to identify realizations with a successful fit. These realizations are re-integrated into the ensemble.

Just 49 realizations remain in the ensemble after this selection process. This corresponds to 0.5 % of the ensemble, showing that realizations fitting the observations are generated with a very low probability. This may result from a required occurrence of discrete pathways in crystalline rock, in contrast to more continuous hydraulic conductivity in porous rocks.

The tracer curves of the 21 best-fitting ensemble members are shown in Figure 6.4. Alternatively, the normalized Kolmogorow-Smirnow test identifies curves that yield a good fit when adjusting the tracer amplitude. Thus, arrival times and curve shape are in good agreement with the measurement, but not the total tracer mass. Figure 6.5(b) shows the curve which is adjusted

¹Here, the processors of a multiprocessor high-performance computer are connected using a virtual shared main memory. This allows for an OpenMP parallelization strategy.

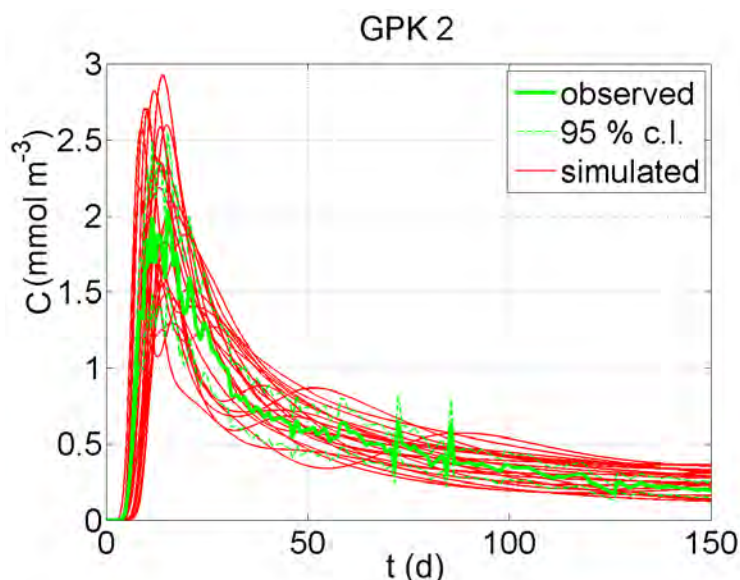


Figure 6.4.: Simulated and observed tracer concentration curve C with time t at GPK2 during the circulation test of 2005. The best fitting 21 realizations and also the 2σ error band (95 % confidence level) of the observation are shown.

by normalization for one realization and the original one (Figure 6.5(a)). A very good fit can be obtained simply by scaling. The difference between the curves can be attributed to a larger amount of tracer reaching GPK2 in the model. An obvious explanation is that the connecting zone in the model is merely too wide. The permeability range inside the connecting zone is in the same correct range as defined by the other successful models (Figure 6.4). However, while the width of the flow path has a significant influence on the arriving tracer mass, it has to be noted that adjusting the area of the flow channel changes the entire mass transport system. Therefore, inferring permeability from scaled concentration data may yield spurious fits. Because of this limitation the realizations identified using the normalized Kolmogorov-Smirnov test are not included in the following study but may give further evidence for the classification.

Moreover, it has to be noted that heterogeneities smaller than 50 m are not captured by the model. Nevertheless, I account for the influence of the small-scale flow paths within a grid block by applying an analogous porous medium approach. That is, effective permeabilities and porosities are provided for the grid blocks, not evaluating single fractures or fissures. I focus here on large-scale connections as candidates for flow paths.

The permeability and velocity fields of the original 49 best-fitting ensemble members (Figure 6.4) are investigated in detail using 3D virtual reality techniques, which allow an appropriate exploration of a 3D model (see Appendix B). Thus, the characteristics of the model can be revealed.

According to this investigation, I classify the realizations into three groups:

- Group 1: This group features a single, direct, narrow, and more or less straight path between GPK3 and GPK2, with a width of no more than 50 m, with a high permeability of at least 10^{-12} m^2 surrounding GPK3. This region may be responsible for the tailing of the curve. The surrounding of GPK2 has little influence on the tracer curve here. The high permeability found in the surrounding of GPK3 results in a spreading of the tracer into the periphery. In general, the best-fitting realizations belong to this group;
- Group 2: This group shows a wide direct path with more than 50 m width. A high permeability around GPK2 is typical in this case. Here, the flow mainly occurs in the connection between the boreholes. This group is characterized by the largest available subsurface heat exchanger surface;
- Group 3: This group is characterized by multiple straight or curved pathways. Mostly one path is dominant with additional secondary paths with lower velocities. Because of multiple pathways, additional extrema within the main peak or a re-rising tail are not unusual for this group. However, this does not necessarily mean that this group fails the fit, because similar shapes can also be found in the observed data at least in the main peak but with minor amplitude.

However, this classification is not unique. Hybrid realizations are assigned to two classes. I consider secondary extrema within the simulated tracer curve as consistent with the observed data, as long as they fall within a 95 % confidence range defined by the measurement error. One typical example for each group is illustrated in Figure 6.6.

6.4.2.1. Characteristics of possible flow paths

Identifying basic common characteristics of all groups is possible by exploring the permeability field and fluid velocities. General findings for flow paths and thus heat exploitation in the Soultz-sous-Forêts reservoir can be summarized as follows:

- On the one hand, boreholes must be connected through a main path with permeabilities of at least 10^{-14} m^2 fitting the tracer curve. If permeability, on the other hand, exceeds about 10^{-13} m^2 , the tracer arrives too early, with too large an amplitude (based on the given porosity assumption) or without a distinct tailing;
- If the connection zone between the two boreholes is not wide enough, fluid paths need to exist in the surrounding of GPK2 with a permeability of at least $3 \times 10^{-12} \text{ m}^2$. This

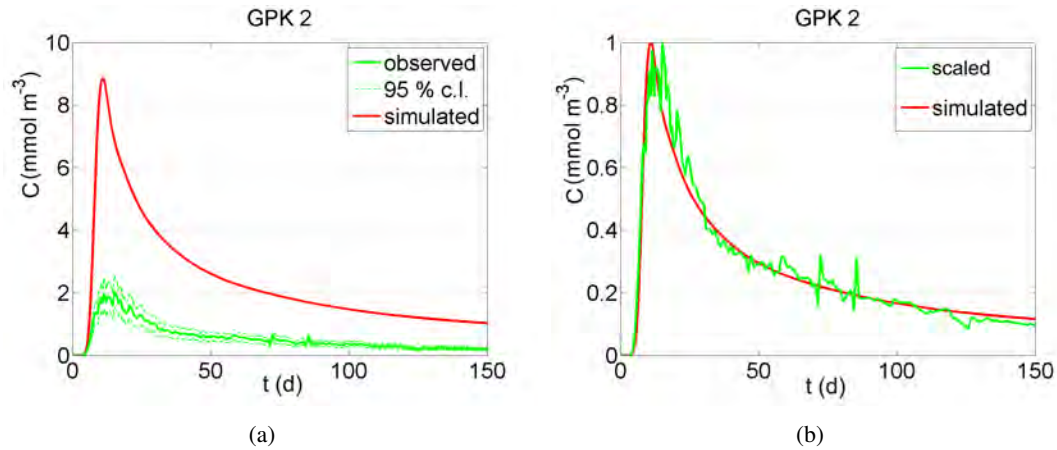


Figure 6.5.: Tracer concentration curve C with time t of realization number 2949 in comparison with the observed data and its 95 % confidence level: (a) original and (b) normalized.

shows that connate water may also be produced, which was also suspected in other studies (Sanjuan et al., 2006);

- In all modeled cases, parts of the tracer are collected and trapped in dead ends of connectivity pathways. This tracer is lost for the circulation test. However, the main tracer loss — only 23.5 % recovered in total (Sanjuan et al., 2006) — is supposed to result from flow into the large drainage system beyond the model boundaries;
- The periphery of the model beyond 500 m distance of the main flow path has little influence on the tracer curve and can hence not be characterized. The same is true for highly permeable zones which act as dead ends for the flow. Fluid stored in this dead ends does not take part in the closed circulation system. Dead end zones can also be expected for the natural reservoir with the unfortunate consequence, that the stimulated zone does not entirely act as a heat exchanger for the fluid. This fact limits the lifetime of the reservoir, even though increasing fluid viscosity evolving from reservoir cooling may result in extended circulation paths and, in turn, longer lifetime;
- The tracer is transported on pathways not wider than 300 m. This is additional evidence for the fact that the main flow does not encompass the entire stimulated region. However, because of model resolution, this is valid only for the large-scale circulation paths;
- Even though a large drainage system is assumed to exist at Soultz (Dezayes et al., 2010), I conclude that extended flow paths may account only for the tail. The main flow responsible for the peak of the tracer curve takes a short direct path between the wells. This is

evidence for an at least partially open fracture surrounded by secondary pathways;

- The modeling approach suffers from an increasing fit uncertainty in the periphery and from model resolution, which does not cover small-scale secondary flow paths. Thus, even though the major characteristics and geometrical limitations are revealed, some small-size and far-field reservoir characteristics are certainly not resolved.

As stated by Sanjuan et al. (2006), for the Soultz reservoir one path is assumed to explain the tracer curve peak and another one to explain the tail. Figure 6.6(c) illustrates the different pathways responsible for the peak and tail of the tracer curve. The narrow direct connection between GPK3 and GPK2 responsible for the peak is characterized by clearly visible velocity arrows, whereas the wide, curved connection responsible for the tail is characterized by a swarm of small velocity arrows. In further investigations, I therefore use the Kolmogorow-Smirnow test to identify realizations that fit only the shape of the peak or the tail. This provides hints about the most likely permeability distribution, as well as quantities of path permeability ranges and velocities involved.

The 10 best realizations for each case are identified to estimate Darcy velocities. All realizations with a satisfying fit of the peak yield a characteristic mean Darcy velocity on the main flow path in a range of $10^{-6} \text{ m s}^{-1} - 3 \times 10^{-6} \text{ m s}^{-1}$. Not surprisingly, the velocities tend to be higher if the pathway is curved. This main pathway has a width of 50 m or less. The reliability of this result becomes obvious when comparing the good agreement of the simulated traveled distance resulting from the tracer particle velocities, calculated by dividing Darcy velocity by porosity, with the observed breakthrough time. However, resulting particle velocities are based on a simplified porosity and the corresponding dependency between tracer arrival and porosity is discussed in Section 6.4.2.3.

Additionally, the pathways responsible for the shape of the tail are investigated. The Darcy velocity next to the path creating the slowly decreasing tail is in the range of $3 \times 10^{-7} \text{ m s}^{-1} - 10^{-6} \text{ m s}^{-1}$ in a region with a 150 m – 200 m in diameter around the main path. In regions with a diameter of 150 m – 300 m, the Darcy velocity is $5 \times 10^{-8} \text{ m s}^{-1} - 3 \times 10^{-7} \text{ m s}^{-1}$. Even though velocities of 10^{-6} m s^{-1} along narrow permeable paths are possible in the periphery of single realizations, in general, a trend of decreasing velocity is observed in the outer regions and therefore a corresponding areal spread of the flow. Here, the tail arrival times are also in good agreement with the particle velocities when considering correspondingly longer tortuous pathways. However, the models trend to overestimate the concentration toward the tail slightly. This may result from porosity effects, which are discussed in the following section (or from sorption effects and natural decay of fluorescein not considered in the model).

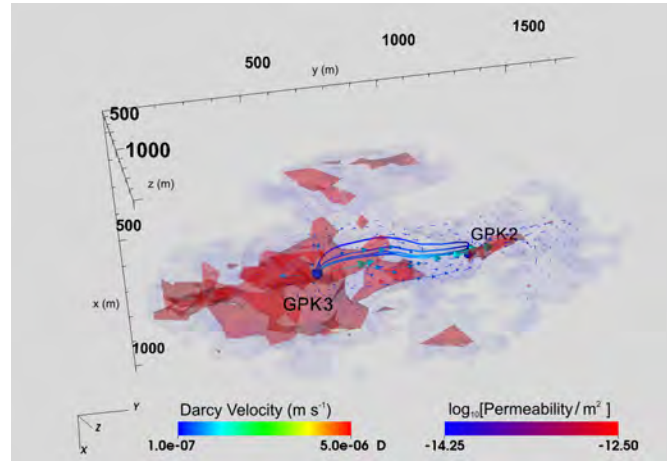
In each of the realizations a separate peak and tail analysis shows that secondary flow paths occur in close vicinity to the main path. Moreover, distinct secondary flow pathways often yield additional- or second-order extrema in the tracer concentration curve additionally to the main peak. Undulations in the observed tracer curve may be caused by these paths. However, the undulations are not clearly visible because of noise caused by minor pressure changes and measurement errors. This makes tortuous secondary paths possible involving small amounts of fluid. But secondary flow paths close to the main path seem more likely. The grid resolution of 50 m edge length of the grid cells allows a temporal resolution of about 7 hours. Correspondingly, the tracer observations are smoothed to one observation per day. A high observation frequency cannot be resolved in the model. Therefore, small-size secondary flow paths in the vicinity of the production hole also cannot be resolved by this model. Additionally, it should be taken into account that a long tracer travel time results in increasing uncertainty of the fit. This applies here in particular for the tail of the curve. Consequently, the far field at Soultz is fitted with less accuracy.

6.4.2.2. Comparing Results with Discrete Fracture Network Approach

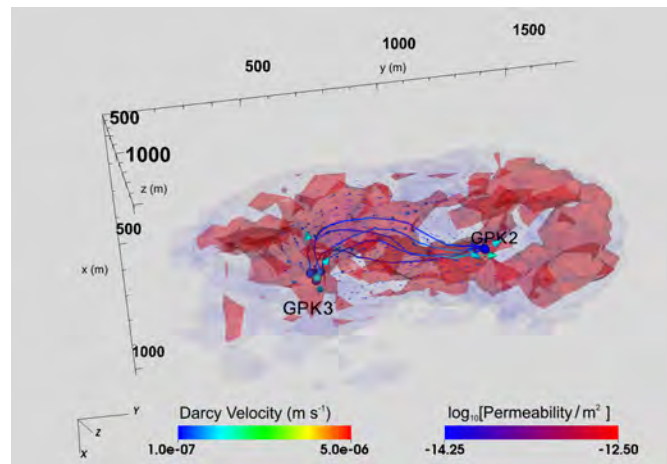
In order to validate the results, I compare the 49 successful realizations with the discrete fracture network used by Gentier et al. (2010a) to model the tracer experiment of 2005. Unfortunately, a quantitative comparison is not possible because I do not have access to the actual 3D data of the fracture model. However, I compare a 2D map of trajectories of 100 simulated particles in the fracture network with 2D slices of the permeability fields (Figure 6.7). As a direct correlation between trajectory density and permeability seems reasonable, this comparison makes certain realizations more likely, as the one I chose for the figure. By visual inspection, I identify six out of 49 realizations that fit the trajectories map. That is, they are characterized by an increased permeability in the southwest and northeast of GPK3. In addition, they show a direct flow path which is slightly tortuous towards to the right-hand side of the flow direction. From these six realizations, four are from Group 1, and one from each of the Groups 2 and 3. This indicates an open fracture between the boreholes.

6.4.2.3. The Influence of Porosity

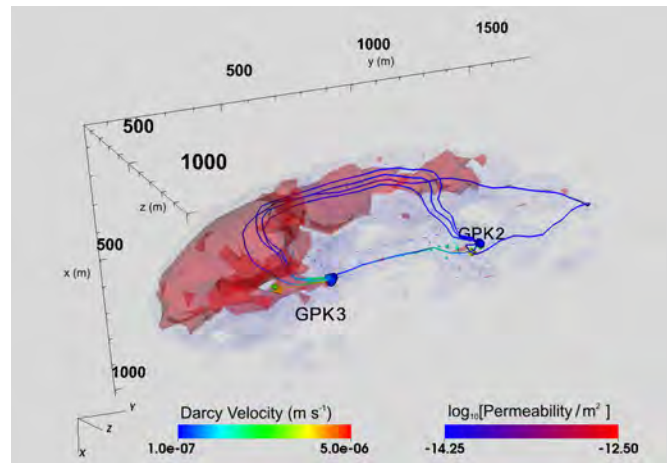
I focused on permeability as the most significant parameter in this study. Taking porosity into account as an additional free parameter would result in too many free parameters that cannot be fitted by the Monte Carlo approach. However, the influence of porosity on the particle velocity, i.e. the ratio of Darcy velocity and porosity, and therefore on the tracer arrival time, is not negligible. Blumenthal (2007) and Gessner et al. (2009) studied the effect of different porosities



(a) Group 1



(b) Group 2



(c) Group 3

Figure 6.6.: Examples for each group of fitting realizations: (a) direct, narrow path; (b) wide path; (c) multiple paths.

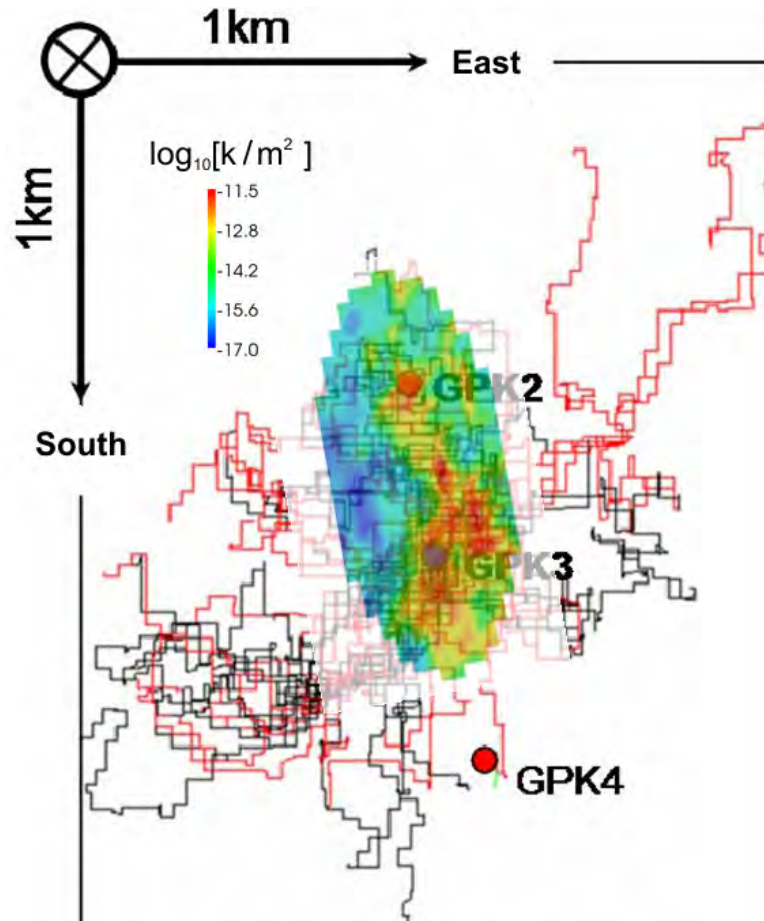


Figure 6.7.: Comparison of the permeability field k of one good-fitting Monte Carlo realization from the ensemble (colored horizontal 2D cross-section), with the trajectories of 100 simulated particles in the fracture network of Gentier et al. (2010a) (lines). High permeability regions of the approach match with regions of dense trajectories. The transparent white area around the permeability field indicates the model boundaries of the present study.

on the simulated tracer curve at Soultz. They stated that the peak of the tracer curve arrives earlier when the porosity in the connecting rock decreases, and the peak concentration increases when the porosity of the surrounding stimulated zone decreases. I find a similar behavior in my numerical models. As discussed by Kosack et al. (2010), only the ratio of permeability k and porosity ϕ can be resolved, not the absolute values.

Ledéseret et al. (2010) reported a porosity of 0.4 % – 4 % for the Granite at Soultz based on laboratory measurements under atmospheric pressure. However, the natural in-situ values at 5000 m depth can be assumed to be much smaller because of overburden pressure. Using the theory of elastic pore space stiffness (Mavko et al., 1998) and an aspect ratio of the ellipsoidal pore space of 0.1, a surface porosity of 2 % in granite can be assumed to close between 5 and 10 km depth. In order to fit the tracer arrival times, I find that porosity in the crystalline rock at 5000 m depth needs to be very small, in the range of $10^{-4} - 10^{-3}$. This agrees well with the inverse modeling results of Kosack et al. (2010). Hence, I assume a constant value $\phi = 5 \times 10^{-4}$ for the stochastic permeability study.

However, around the borehole, numerous measures have been applied for increasing the hydraulic connectivity opening small cracks. Therefore, in this section, I investigate the effect of varying porosity. Tracer concentration curves are plotted in Figure 6.8 for one selected, well-fitting Monte Carlo permeability realization and different porosities. Obviously, variation in porosity has a non-negligible influence. Permeability may vary nevertheless by some orders of magnitude and, hence, is still the dominant parameter in the modeling.

Because of this impact of porosity on the simulated tracer, it has to be stated that all the results for permeability are valid only for the given porosity value of 5×10^{-4} . They may differ moderately by less than half an order of magnitude, taking into account porosities in the range of $10^{-4} - 10^{-3}$.

I assume here that because of the fracture-dominated permeability, no large voids in the rock are created and, hence, porosity is not significantly affected by the stimulation. However, in order to address a relation between permeability and porosity, I implemented a relationship between porosity and permeability according to Naderi Beni (2011):

$$\phi = \phi_0 \left(\frac{k}{k_0} \right)^{1/3}. \quad (6.7)$$

Here, ϕ denotes the porosity and k the corresponding permeability. Reference values are $\phi_0 = 5 \times 10^{-4}$ and $k_0 = 10^{-13} \text{ m}^2$ and are obtained from estimated values inside the direct flow path according to the results. Other reference permeabilities, such as 10^{-15} m^2 or 10^{-17} m^2 , yield a large misfit. Note that is relationship of permeability and porosity is deviated for porous rocks and, hence, can provide only rough estimates for granite rock. An alternative formulation from

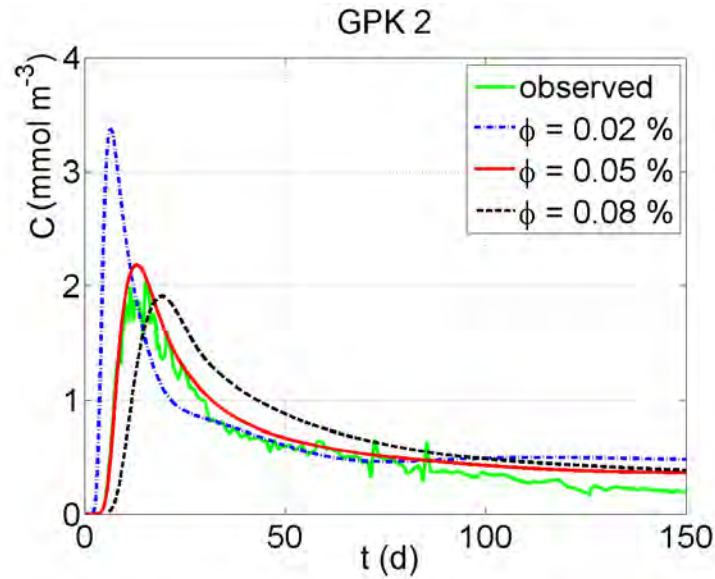


Figure 6.8.: The influence of porosity variations on the results. Tracer concentration C with time t at different porosities ϕ .

Pape et al. (1999) for shallow undisturbed crystalline rock yielded no realistic values for porosity in the deep stimulated granite within the Soultz reservoir.

One arbitrary realization of each group is simulated using Eq. 6.7. The result is shown in Figure 6.9.

Linking porosity to permeability has only a minor effect for all three groups. For the cases of wide and multiple flow paths, the effect on the behavior of the tail of the tracer curve is stronger than on the arrival part, which illustrates the more pronounced effect of the far field in these cases.

6.4.3. Discussion

The massive Monte Carlo study reveals common ranges of permeabilities, Darcy velocities, and corresponding possible geometries. I show quantitatively different possible fluid pathways and therefore provide limits for parameters and flow paths. Nevertheless, the method cannot fit the GPK4 and GPK2 tracer curve simultaneously. Moreover, it does not yield the fit quality of previous studies with simplified geometry.

The results suggest that the simplified geometrical approach of Kosack et al. (2010) with a direct connection and an adjacent stimulated zone is in no contradiction to the most likely pathways found here. However, other geometries cannot be excluded involving numerous different pathways, such as Gentier et al. (2010b).

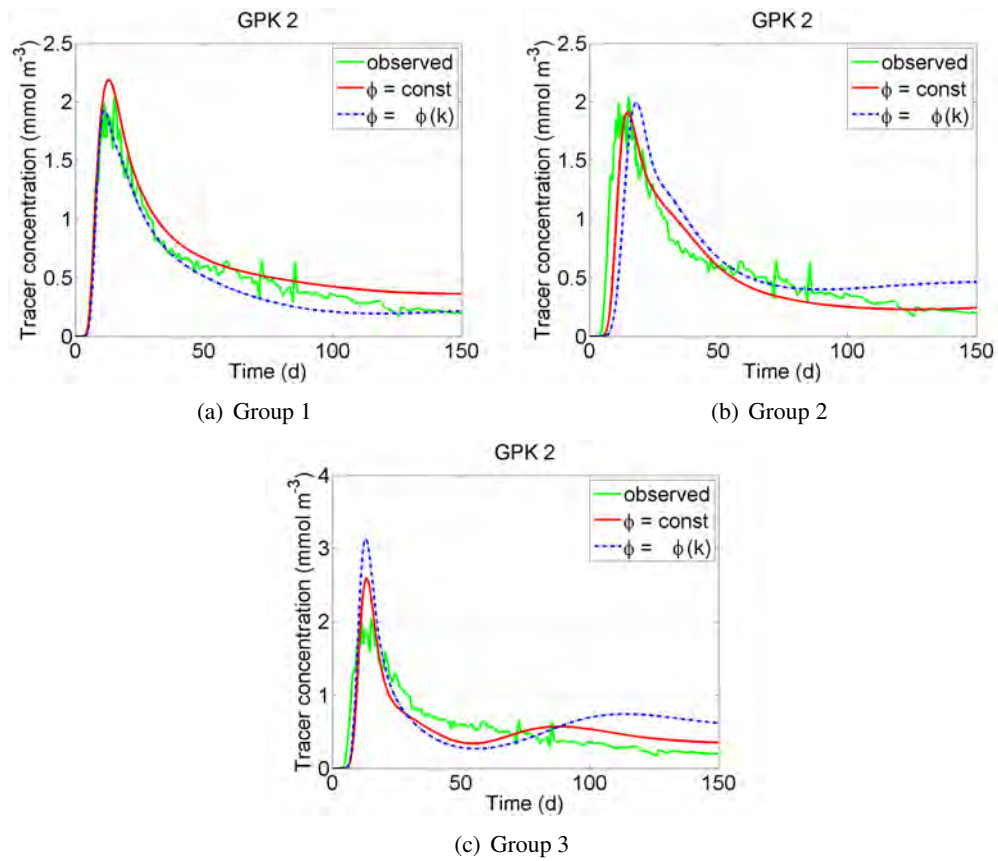


Figure 6.9.: Examples for tracer concentration curves in the original (red lines) and resulting from porosity-permeability relation in Eq. 6.7 (blue dashed lines). One realization is given of each group of tracer flow path classification (Figure 6.6).

Not unexpected, the massive Monte Carlo approach does not provide a fit of similar quality as Kosack et al. (2010). This would eventually not be true if the ensemble size is increased by some orders of magnitude. However, the corresponding demand in computing power cannot justify this approach. Nevertheless, the major advantage of this approach as it is applied here lies in investigating the general characteristics of an EGS reservoir. This helps significantly to evaluate the reservoir's rock property distributions and to reveal previously unsuspected flow paths.

From a geological point of view, the main fracture zone at Soultz-sous-Forêts is a good candidate for the primary path. However, secondary paths are also required for explaining the tail of the tracer curve if a high porosity in the deep crystalline bedrock is ruled out. This seems to be likely because of the general behavior of porosity in the crystalline crust (Clauser et al., 1999). In summary, I propose a partially disturbed main fault zone of heterogeneous permeability distribution surrounded by directly adjacent secondary pathways as the most likely scenario for the Soultz reservoir.

Stochastic modeling using all three wells remains unsuccessful because of the large parameter space to be sampled, combined with the required fit of a second tracer curve. The combination of a good hydraulic connectivity to GPK2 and bad connectivity to GPK4 in the way it is observed at Soultz is seldomly realized by the SGSim algorithm. Therefore, no information about the hydraulic connections involving GPK4 can be identified using this method.

Only 0.5 % of the realizations remained in the ensemble after the selection process. This indicates exploring the parameter space using the massive Monte Carlo method is computationally very demanding for this numerical problem. Permeability fields which fit the data are seldom realized in the ensemble for crystalline rock with discrete fluid pathways. Therefore, I apply the Ensemble Kalman Filter method in the following. This allows for considering also data measured in GPK4 and for an effective inversion in terms of computational effort.

6.5. Ensemble Kalman Filter Study of the Fractured Reservoir

As the Ensemble Kalman Filter requires fewer realizations, the model comprises now all three boreholes GPK2, GPK3, and GPK4. Therefore, it is larger than before, comprising $21 \times 48 \times 21$ grid cells on a regular grid with a cell size of 50 m in each direction. This allows fitting the tracer concentration curves of GPK2 and GPK4 simultaneously.

Chapter 4.3 describes the feasibility of an inversion of the basic characteristics of the flow paths inside a Soultz-like synthetic reservoir using the EnKF based on data from the three boreholes. I use a significantly larger ensemble size than for the previous test case comprising 880 realizations for avoiding underestimating the variability which may contaminate the results.

For avoiding unphysical or unlikely values, I limit the permeability parameter range to 10^{-12} m^2 .

10^{-18} m^2 . The simulation time is discretized into 3000 numerical time steps of 3 h duration each. For a total simulation time of 150 days, an EnKF update is performed once a day.

6.5.1. Results and Discussion

Again, the initial homogeneous mean permeability is $\sim 10^{-14} \text{ m}^2$ expect for the injection and production cells. The initial standard deviation is about two orders of magnitude. Figure 6.10 illustrates the estimated reservoir log permeability as ensemble mean and standard deviation after one EnKF run with 150 update steps (so-called global iterations), and after re-initializing the EnKF a second and third time. The first global iteration already reflects the major characteristics of the reservoir: a narrow, well-connected zone between GPK3 and GPK4 and an ill-connected one between GPK3 and GPK4, and their corresponding permeabilities. The uncertainty is below one order of magnitude only along the direct flow path. As there is no more borehole information available, this is as expected. However, the fit is unsatisfactory, which becomes obvious when comparing the simulated tracer curves obtained from the single ensemble realizations with the measured concentration (Figure 6.11(a,b)).

After three global EnKF iterations, the fit quality is much improved (Figure 6.11(c,d)). Further, this reveals more characteristics of the ensemble: The narrow high-permeable path between GPK3 and GPK4 ($\sim 10^{-13} \text{ m}^2 - 10^{-12} \text{ m}^2$) is embedded into a low-permeability zone near GPK2 ($\sim 10^{-17} \text{ m}^2$) and a medium permeability zone near GPK3 ($\sim 10^{-15} \text{ m}^2 - 10^{-16} \text{ m}^2$), preventing the flow to spread into the periphery. In contrast, the medium permeability zone extends all the way to the periphery of GPK4. This allows a number of secondary flow paths in this region, where tracer is stored until it reaches the production wells, contributing to the concentration at GPK4 or to the tail part of the GPK2 concentration curve.

However, we cannot determine whether the presence of a number of artifacts beyond the flow paths result from a realistic signal or indicate filter inbreeding, an unwanted loss of ensemble variability (see Chapter 3.5), when using three global iterations. This must be taken into account when judging the results.

In contrast to GPK2, the slopes of the observed and simulated tracer concentration at GPK4 rarely match. This is due to the different magnitudes of the two observed concentrations. Data assimilation is dominated by the much stronger signature of the GPK2 tracer curve.

6.5.2. Updating a Non-Gaussian Distribution

I find that during the first global iteration of the EnKF assimilation, the bimodal shape of the permeability distribution transforms into a Gaussian distribution (Figure 6.12(a)). Thus, the bimodal characteristic of a fractured reservoir is lost to some extent. This was also observed by

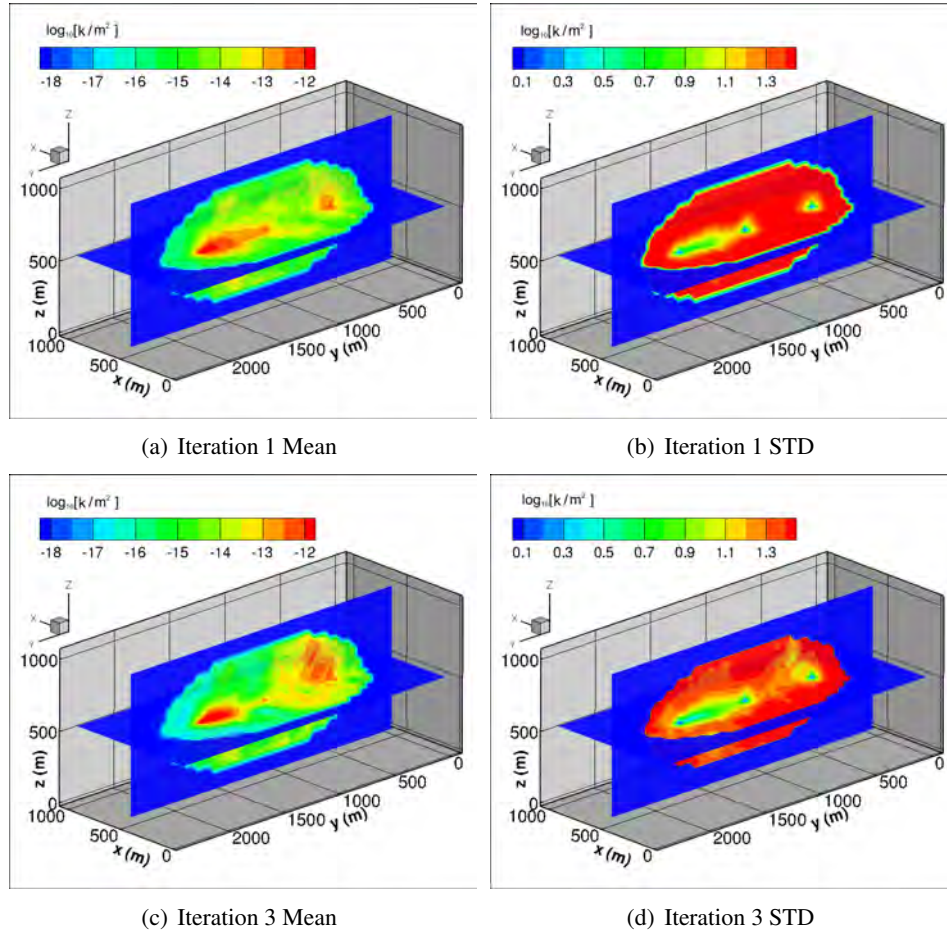


Figure 6.10.: Estimation results for the reservoir permeability k : ensemble mean after the first (a) and third (c) global EnKF iteration and the corresponding ensemble standard deviations (b,d).

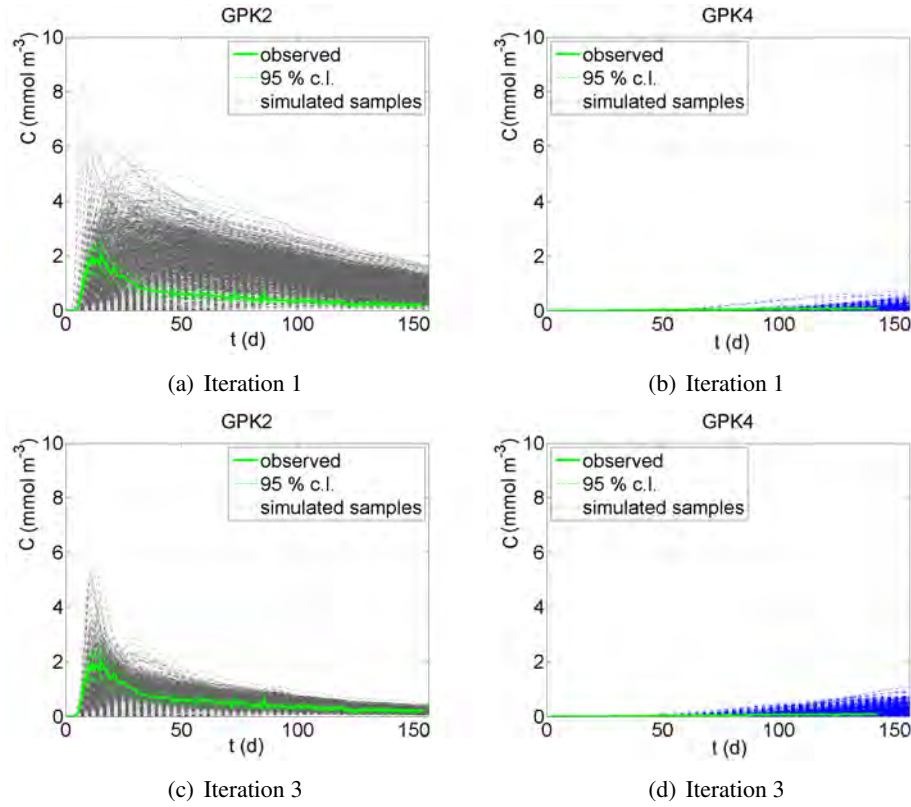


Figure 6.11.: Tracer concentration C curves with time t observed at the production wells GPK2 and GPK4 for one (a,b) and three (c,d) global EnKF iterations, respectively. The curves are computed in a reinitialized run using the final permeability field without any further ensemble update run performed after the EnKF data assimilation run. The figure shows a comparison of all 880 simulated ensemble members and the observation. Also shown is the 95 % confidence level of the observation.

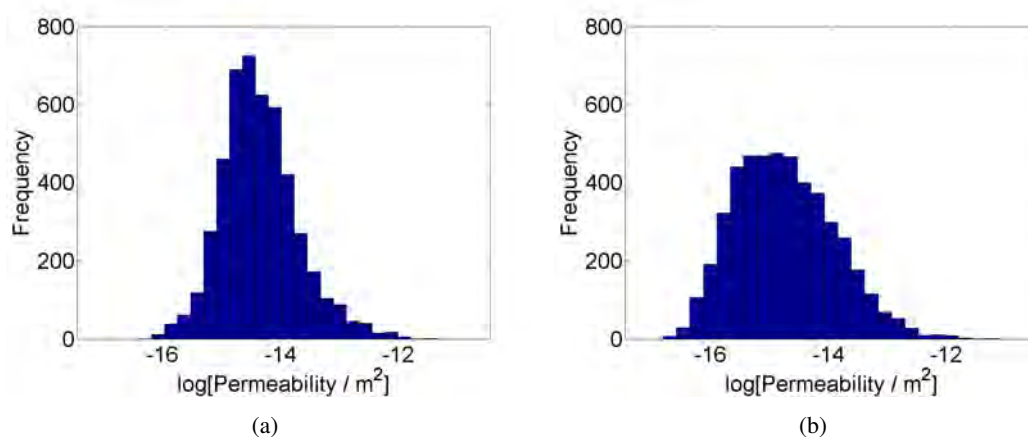


Figure 6.12.: Ensemble mean log permeability distribution in the stimulated zone after one (a) and three (b) global iterations of EnKF assimilation.

Zhou et al. (2011). The Gaussian shape of the estimated distribution changes only little during additional global iterations (Figure 6.12(b)). Just the mean is changed (by 0.5 % of the log permeability value) to slightly smaller values, whereas the maximum values remain constant. The latter is required to realize the main fracture zone between GPK2 and GPK4. Therefore, mainly the spatial distribution within the reservoir is modified during additional global iterations.

The geological setting of the high permeability fracture zone in a less permeable surrounding within the Soultz reservoir demands a bimodal probability distribution for permeability. The filter formalism of the EnKF transforms this distribution into a Gaussian representation during assimilation. I try to use also an initial Gaussian distribution, but since there is no proper information for any Gaussian permeability histogram available, this always yields worse results. Therefore, I keep the bimodal initial distribution and allow the filter formalism to adjust the proper Gaussian representation.

The lost of the bimodal shape during the assimilation still allows fitting permeability and identifying fluid pathways. However, this behavior may result in smoothing of possible distinct boundaries between zones of different permeability.

For future EnKF studies, a transformation of the parameters and states with non-Gaussian error distribution into Gaussian space during the EnKF update step (called Normal-Score Ensemble Kalman Filter) is expected to yield improved results for the bimodal log permeability distribution studied here. This technique follows Zhou et al. (2011), who pioneered this approach and applied it successfully to fluvial sediments, also characterized by a bimodal log hydraulic conductivity distribution. Therefore, this method is of particular interest for reservoirs characterized by channel sands or fractures.

6.5.3. Long-Term Performance Prediction

A long-term prediction of temperature is of special interest in geothermal reservoir modeling, as the arrival of the injected cold water front (thermal breakthrough) controls the lifetime of the installation. In this case, warm water is produced at GPK2 and GPK4, whereas cold water is injected into GPK2 to maintain reservoir pressure. Therefore, a cold water front is propagating from the injection borehole to the production boreholes. On its way, the injected water is heated. I use the estimation result (the ensemble of updated realizations) from the previous data assimilation for predicting the transient variation of temperature within the reservoir as well as its uncertainty.

As the diffusive process is very different in case of tracer and heat transport (by pore space and pore-matrix-mixture, respectively), it is questionable to predict the long-term heat transport by inversions of tracer data. Different fluid pathways or permeability distributions may cause similar tracer curves, but very different temperature predictions. Only an uniquely identified permeability distribution would allow a correct prediction of both, tracer and heat transport. However, by applying a Monte Carlo approach, the non-uniqueness of estimated permeability field is automatically taken into account as the ensemble yields all kinds of possible permeability realizations. Thus, the temperature prediction under uncertainty is valid even though the transport phenomena are different.

Head and temperature are simulated for 50 years, using the circulation rate in the year 2010 of 30 L s^{-1} (Gentier et al., 2010b). This volume flow rate is distributed into production rates 17.5 L s^{-1} at GPK2 and 12.5 L s^{-1} at GPK4, respectively. For this prediction, an ensemble of 440 realizations is computed from the original ensemble. Numerical experiments show that a large ensemble size does not yield significantly different results. Note that as fluid density is a function of temperature, an implemented volume flow rate may violate the mass balance and, hence, result in an increasing reservoir pressure during the simulation. This is due to the temperature difference between injected and produced water. As the temperature and pressure equations are coupled in the simulation, this may accordingly affect the performance prediction. However, in this case hydraulic head varies by 0.02 % over time, indicating a reliable simulation.

Figure 6.13 illustrates the variation of temperature with time at the two production boreholes for 50 years of operation. Some realizations show an increasing temperature in the first years, indicating that hot water from deeper parts of the reservoir is produced. A thermal breakthrough at GPK2 is not predicted by any ensemble member during the simulated operation time. This is a positive indication for the installation. In contrast, at GPK4 an early arrival of the cold water front seems possible. This behavior is illustrated also in the histograms of production temperatures after 50 years (Figure 6.14). An average temperature of $(205 \pm 1.8) ^\circ\text{C}$ is predicted for GPK2. The temperature distribution is Gaussian. At GPK4, however, the temperature distributes

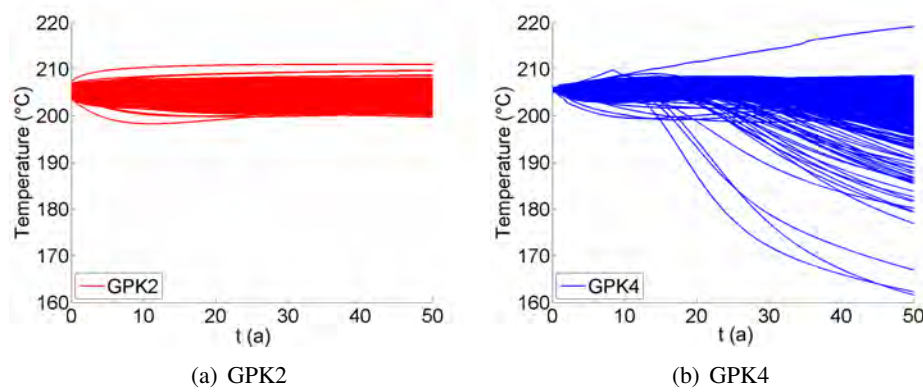


Figure 6.13.: Performance prediction for 50 years of operation for the production boreholes GPK2 (a) and GPK4 (b). Single extreme values for GPK4 are not shown.

is non-Gaussian with a slightly smaller average temperature of $(202 \pm 6.3) ^\circ\text{C}$. The significantly different standard deviations of both predictions result from the higher quality in the case of the GPK2 tracer fit compared to the GPK4 fit as explained before.

It must be mentioned that the actual measured temperature of the produced water at the surface is currently only $155 ^\circ\text{C}$ (Gentier et al., 2010b). In contrast, the prediction is performed for a production from a depth of 5000 m neglecting any cooling of the water on its way to the surface. Further, the significantly cooler measured production temperature may indicate that cold water is produced from shallower parts of the reservoir along flow paths not realized in the simulations. This must be taken into account when comparing predicted and measured temperatures at the begin of production. Simply re-normalizing the results based on the measured temperature should be possible if pipe-cooling dominates. However, if water is produced from shallower depths, this is not reliable.

Note, that the predictions are only valid for the permeability field as it were in the year 2005, when the tracer circulation test was performed. Any further interference e. g. through further hydraulic or chemical stimulation would change the parameter fields and hence, the temperature prediction. However, the method shown here for predicting the reservoir performance can be applied as long as the ensemble of realizations is updated with new data after a change in reservoir conditions.

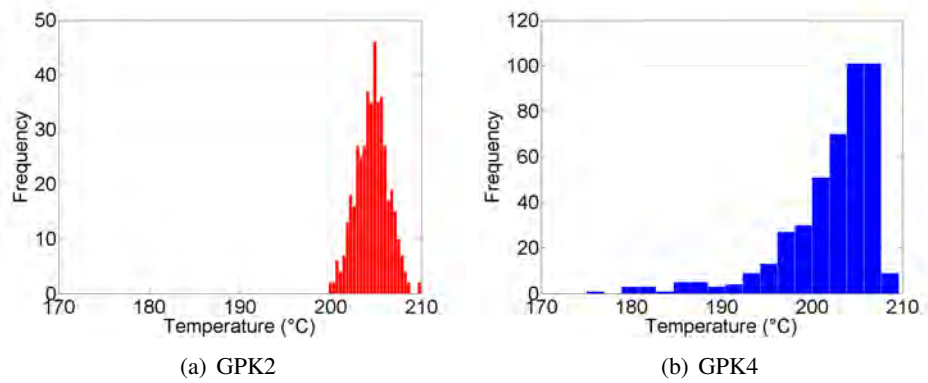


Figure 6.14.: Predicted temperature distribution after 50 years of operation for the production boreholes GPK2 (a) and GPK4 (b).

6.6. Comparing Bayesian Inversion, Massive Monte Carlo, and Ensemble Kalman Filter

I find that all three inverse simulation methods, deterministic Bayesian inversion, massive Monte Carlo and Ensemble Kalman Filter, provide useful estimates of the permeability field at Soultz.

Deterministic inversion as described in Kosack et al. (2010) yields a nearly perfect fit of the tracer curve and allows quantifying parameter uncertainties and their dependencies very well. Additionally, it converges fast. However, this approach is limited to a small number of parameters, requires specifying a reservoir architecture, and therefore is limited to well defined homogeneous reservoir geometries. This need for constrained prior assumptions introduces an important additional source of uncertainty. Further, deterministic predictors are very vulnerable to local solutions, missing the correct global minimum of the objective function when dealing with uncertain prior information.

In contrast, a massive Monte Carlo method with constraining post-processing suffers much less from this problem because it explores the entire parameter space. This applies also for the EnKF used in this study. As an advantage, both massive Monte Carlo and Ensemble Kalman Filter can estimate a large number of parameters. Therefore, an estimation is possible not only for an average parameter value in a geological unit but for the recovery of the entire heterogeneous parameter field. Nevertheless, the quality of fit is worse for the stochastic methods, and a rigorous quantification of uncertainty is possible only for deterministic and EnKF inversion, not for the massive Monte Carlo approach. Many realizations (10^4) are necessary for sampling the parameter space and, hence, for applying the massive Monte Carlo method successfully. This usually requires large computing times.

Both, deterministic Bayesian and massive Monte Carlo approaches cannot fit the reservoir properties well when adding also the observed tracer concentration in GPK4 to the inversion. This is due to the heterogeneity of the permeability field in this case. On the one hand, the corresponding complex reservoir geometry is not easy to model when simplifying the flow paths to apply the deterministic approach of Kosack et al. (2010). On the other hand, the massive Monte Carlo approach requires sampling from too large a parameter space for this application, which cannot be realized in reasonable computing time.

The Ensemble Kalman Filter proves a powerful third alternative to both approaches, combining the advantages of fast convergence and rigorous quantification of uncertainty in a Bayesian framework with the handling of many parameters, heterogeneity, and robustness against local solution. It is most practical for applications with unknown reservoir geometry, heterogeneous parameter distribution and when uncertainty needs to be quantified. However, it does not provide the same quality of fit as deterministic inversion with the used ensemble size, i.e. 880 realizations. Nevertheless, when comparing the entire ensemble and its uncertainty, the tracer concentration curve average is captured well by the EnKF curves. Further, non-uniqueness of possible inverse solutions of this particular problem is not as well addressed by the EnKF as by the massive Monte Carlo approach. The latter illustrates a large number of different possible pathways at Soultz, which all fit the tracer concentration of GPK2.

As an important additional advantage, the EnKF approach allows estimating the entire Soultz reservoir consisting of the three boreholes GPK2, GPK3, and GPK4 — in contrast to the other approaches presented in this section.

A summary of the advantages and disadvantages of all three inversion approaches is given in Table 6.2

Table 6.2.: Comparison of three different approaches for parameter estimation for the application at Soultz.

Gradient-Based Inversion	Massive Monte Carlo	Ensemble Kalman Filter
+ nearly perfect fit	+ requires little prior information on geological structure	+ requires little prior information on geological structure
+ simple fracture geometry consistent with data	+ accounts for heterogeneity	+ accounts for heterogeneity
+ yields information on resolution and sensitivity	+ illustrates non-uniqueness of the inverse problem	+ yields inherently information on uncertainty
		+ Additional estimation of state variables
- only applicable to few structural units	- little information on resolution	- sensitive to filter parameters (e. g. damping, system error)
- structural units need to be defined beforehand	- high computational effort	
- vulnerable to local solutions	- joint inversion of different data complicated	

Chapter 7.

Summary and Conclusions

In this thesis, I study different stochastic tools for estimating reservoir parameters as well as states, at the same time quantifying their uncertainties. Thus, I identify physical processes within geothermal reservoirs and subsurface characteristics. In addition, based on the studies of this thesis, I provide a workflow for stochastic geothermal reservoir modeling:

- Use seismic data to construct a geometrical model.
- Identify rock property histograms for each geological unit measured in wells or in the laboratory.
- Perform initial simulations or use temperature data to investigate the ratio of advective and conductive heat transport.
- Use gradient-based deterministic inversion to estimate the basal specific heat flow and its uncertainty.
- Perform stochastic (ensemble-based) simulations on a regional scale based on rock property histograms and on the uncertain basal specific heat flow. This quantifies temperature uncertainties at the target location.
- Use the results obtained as boundary conditions for a detailed reservoir model.
- Perform transient simulation of the single-well / doublet / triplet operation.
- Calibrate the reservoir model using production data such as tracer, pressure, or temperature. This can be done e.g. by using constraining post-processing or the Ensemble Kalman Filter method.
- Perform stochastic long-term simulations in order to quantify the variation of reservoir state variables such as pressure and temperature with time.

- Continuously update the calibrated reservoir model when new production data becomes available.

By following this workflow, the operator of a geothermal installation can base his or her investment decisions on a quantitative estimation of the risk of project failure. At the same time, the presented modeling results help avoiding the oversizing of geothermal installations which may result in lower project cost.

In a first step, synthetic test cases prove the usefulness of the stochastic methods of massive Monte Carlo and Ensemble Kalman Filter inversion for estimating heterogeneous rock property and reservoir state distributions based on only few observation wells (3 – 20).

I perform Monte Carlo modeling based on geostatistical algorithms for capturing heterogeneities and quantifying uncertainties, in this case for well pressure and production temperature of geothermal doublet installations in sedimentary reservoir rocks at The Hague, The Netherlands (i), and in the Northeastern German basin (ii). The Monte Carlo approach of Sequential Gaussian Simulation combined with constraining post-processing helps to quantify uncertainty much more accurately than by specifying minimum-maximum estimates (value ranges) of the involved rock properties or extrapolated bottom-hole temperature (BHT) data. In addition, I show the importance of accounting for heterogeneity of rock parameters, which results in significant variations of production temperature [case (i): Chapter 5.1.2; and case (ii): Chapter 5.2] and well pressure [case (ii)] with time.

Constraining techniques are particularly effective in filtering realizations after stochastic simulation of hydraulic properties in systems where advective heat transport is significant. Unfortunately, they are not based on Bayesian principles. Therefore, calibrating the realizations to BHT data is done by Bayesian deterministic inversion to estimate the basal specific heat flow for the reservoir inside the Northeastern German basin, case (ii). For these problems with low dimensionality (1D), Bayesian deterministic inversion is useful in particular because of its effective convergence towards the optimal estimate. Large numbers of parameters ($> 1\,000$ for estimating heterogeneous parameter fields) can only be addressed by Monte Carlo approaches. As another alternative, the Ensemble Kalman Filter is applied for calibrating transient models of the Soultz reservoir. As an advantage, the Ensemble Kalman Filter follows a Bayesian scheme rigorously. Thus, the Ensemble Kalman Filter provides the correct probability density function inherently. However, this method is limited to transient problems, requiring monitoring data as input.

Stochastic modeling based on hydraulic rock properties in detailed reservoir models allows predicting the transient variation of the energy output of a doublet including uncertainty. This comprises temperature only for reservoir (i), and both, temperature and pressure for reservoir (ii).

Chapter 7. Summary and Conclusions

For case (i), the reservoir models in The Hague show encouraging results since no significant thermal breakthrough can be expected. The temperature of the extracted water is predicted to be around $(74.6 \pm 0.35)^\circ\text{C}$ after the first 50 years of production. First measurement yielded a temperature of 76°C before start of operation.

For case (ii), the reservoir in the Northeastern German basin, comprehensive seismic data allows combining seismic processing and statistical rock property for stochastic geothermal reservoir modeling. Information on permeability for the reservoir in The Hague is sparse, therefore I vary reservoir permeability within reasonable bounds. In contrast, for the reservoir in Germany an empirical relation from Pape et al. (2005) allows to calculate permeability from porosity. Here, I predict a very low probability of success (1.6 %) after 20 years of operation. Thus, the probability of a potential project failure can be estimated without the loss of investment costs for drilling a geothermal well.

Finally, I study the hydraulic properties and flow paths of the fracture system in hard-rock granite at the Enhanced Geothermal System test site at Soultz-sous-Forêts. Here, I show an equivalent porous medium approach for the interpretation of the Soultz tracer experiment. This is an alternative to the discrete fracture network approach presented previously for fitting GPK2 and GPK4 simultaneously (Gentier et al., 2010a).

Here, the fractured rock is assumed as equivalent porous medium. This is justified by damage zone of small cracks and fissures surrounded by a single fracture, and by the fact that at Soultz a fracture network was stimulated, not a single penny-shaped fracture. However, in single fractures, turbulent flow may occur. Resulting effects on the flow field may be not captured by the linear relationship of Darcy flow. In addition, the effect of fractal fracture geometry is not considered here. It has been studied by e. g. Kühnle (2011) on a discrete fracture model. Assuming a fractal framework, no correlation lengths would exist. In my case, Darcy flow in an equivalent porous medium is controlled by effective permeability resulting from fracture densities. Therefore, the use of correlation lengths is still appropriate to account for similarities in fracture density. Thus, only effective values for permeability can be computed in this study neglecting the influences of rough fractal fracture surfaces.

A massive Monte Carlo approach is performed to address the non-uniqueness of the inverse problem. It yields possible pathway geometries classified into three characteristic groups (narrow path, wide path, multiple paths) and corresponding permeabilities as well as filtration velocities. However, this approach only fits the GPK2 tracer concentration observation.

In contrast, an Ensemble Kalman Filter approach provides a simultaneous fit of the GPK2 and GPK4 concentrations. A comparison between gradient-based Bayesian inversion performed by Kosack et al. (2010), massive Monte Carlo, and Ensemble Kalman Filter shows that the Ensemble Kalman Filter outperforms the massive Monte Carlo approach in terms of fit quality

on the one hand, and the gradient-based inversion with respect to heterogeneity of the hydraulic parameter field on the other hand. In addition, the Ensemble Kalman Filter allows for much smaller ensemble sizes than the massive Monte Carlo approach and therefore a significant faster computation time.

When taking into account the results of both combined, massive Monte Carlo and Ensemble Kalman Filter approaches, fluid transport through a single direct fault between GPK3 and GPK2 seems likely. However, other preferential pathways are still possible and simplifications done by Kosack et al. (2010) must be kept in mind when comparing results with their nearly perfect fit. In addition, these simplifications do not allow for fitting the GPK4 concentration curve jointly with the curve of GPK2. This is possible only with the Ensemble Kalman Filter approach or the discrete fracture network approach of Gentier et al. (2010a).

Using the Ensemble Kalman Filter estimation results, a long-term performance prediction for the reservoir can be provided (for reservoir conditions of 2005), including an uncertainty estimate.

By comparing three approaches for parameter estimation (gradient-based Bayesian inversion, massive Monte Carlo, and Ensemble Kalman Filter, see Table 6.2), I conclude that:

- Gradient-based Bayesian inversion is optimal for problems with well known geometry and constrained parameter space. In this case, it yields very good fits.
- Massive Monte Carlo is optimal to explore very heterogeneous reservoirs and the non-uniqueness of the fits. However, the required data analysis may be too complex for continuously calibrating a reservoir.
- The Ensemble Kalman Filter combines the advantages of both approaches and is preferable for all cases where geometry is not well-defined or the spatial parameter distribution is very heterogeneous.

According to my results, the operations of the geothermal systems at Soultz and The Hague are sustainable (see also Axelsson (2010)) in terms of transient temperature drop which is limited to few Kelvin during operation time for these cases. In contrast, the doublet system in the Northeastern German basin is characterized by possible large temperature drops of about 20 K.

When applying constraining post-processing or the Ensemble Kalman Filter approach for model calibration, usually data is sparsely available. But how much boreholes are needed for a successful calibration? The synthetic studies show that three wells are sufficient for capturing the very basic characteristics of a reservoir's permeability field (Chapter 4.3), but no unique solution of the inverse problem can be found (Chapter 6.4). For 2D problems, five wells seem to be sufficient for capturing the characteristics of a the reservoir permeability. This is the

same number for constraining post-processing (Chapter 4.1) and Ensemble Kalman Filtering (Chapter 4.2). Of course, this number depends on the heterogeneity and the dimensions of the reservoir and will be larger for complex 3D permeability structures. Nevertheless, according to the results of the present study, I very roughly recommend five equally distributed wells as requirement for a sufficient parameter estimation. In addition, tracer data yield better fits than temperature data.

In summary, the presented stochastic modeling sequences provide an ensemble of calibrated reservoir models which allow for adjusting the exploitation strategy with regard to profit or sustainability or a compromise of both.

7.1. Outlook

The operation of the doublet in The Hague started in March 2012. This will allow for comparing some results of this study with in-situ measurements. This in turn will enable calibration of the reservoir model in terms of incorporating production data and, hence, update the temperature prediction of the numerical simulations, e. g. by using the EnKF.

For the reservoir in the Northeastern German basin, heterogeneous porosity can be identified from seismic attributes (e. g. Schön, 1996). Results from this approach may be compared with the results of the stochastic approach applied in this dissertation. A similar approach was already followed by Szalaiová (2012), who identified hydraulic parameters in basement rock using seismic measurements.

Following the studies on economic modeling of geothermal installations by Knaut (2012), the impact of uncertain produced temperature on financial modeling will be further investigated.

For future EnKF studies of reservoirs characterized by bimodal parameter distributions, such as reservoirs containing channel sands or fractures, the Normal-Score Ensemble Kalman Filter (Zhou et al., 2011) is expected to provide improved results.

In case of a small ensemble size compared to the number of free parameters, as is usual in this kind of problems, limited space for solution search and spurious correlation may cause an excessive decrease of the ensemble variance. This may result in filter divergence (so-called filter inbreeding). Besides damping as applied in this thesis, covariance localization (Oliver & Chen, 2011; Zhang & Oliver, 2011) can help to remedy filter inbreeding in future studies. Localization methods reduce the impact of distant observations on a grid point to be updated by reducing the amplitude of long-range spurious correlations. This increases effectively the ensemble solution space and allows using small ensemble sizes for estimation of high-dimensional problems.

Further, accounting also for the mechanical coupling of the relevant processes and approximating the effect of thermal stress on the rock mass or on fault hydraulic parameters will increase

Chapter 7. Summary and Conclusions

information on the reservoir, in particular for stimulated reservoirs (Watanabe et al., 2010). In addition, this coupling may allow an EnKF-prediction of reservoir permeability based on micro-seismic events following the approach of Tarrahi & Jafarpour (2012).

In addition, the impact of a background current in South-North direction (as proposed by Held (2011)) on the inversion results will be investigated.

Also a simulator for discrete fracture networks could be used as forward propagator within the EnKF frame, allowing to account also for influences from turbulent fluid flow. There, also rough fractal fracture surfaces can be implemented resulting in different particle velocities along a fracture.

Additional data from different sources such as self potential data (Darnet et al., 2006) will be assimilated for more comprehensive EnKF reservoir parameter estimation.

The installation at Soultz and The Hague are supposed to be operated in a sustainable way. However, the influence of uncertain thermal rock properties has to be considered for analysing the renewability of the reservoirs in the future following the work of O'Sullivan et al. (2010), i.e. for estimating the temperature recovery time of the entire reservoir after the end of production.

In the future, the studied tools and modeling sequences may be applied for numerical prediction of the behavior of geothermal reservoirs during all stages of exploration, development, and exploitation.

Bibliography

- Aanonsen, S. I., Naevdal, G., Oliver, D. S., & Reynolds, A. C., 2009. The Ensemble Kalman Filter in reservoir engineering – a review, *Society of Petroleum Engineers Journal*, **14**(3), 393–412.
- André, L., Rabemanana, V., & Vuataz, F. D., 2006. Influence of water-rock interaction of the deep reservoir at Soultz-sous-Forêts, France, *Geothermics*, **35**(5-6), 507–531.
- Axelsson, G., 2010. Sustainable geothermal utilization – case histories; definitions; research issues and modelling, *Geothermics*, **39**, 283–291.
- Baechler, D. & Kohl, T., 2005. Coupled thermal-hydraulic-chemical modelling of enhanced geothermal systems., *Geophysical Journal International*, **161**(2), 533–548.
- Bakker, P., 2002. *Image structure analysis for seismic interpretation*, Doctoral dissertation, Delft University of Technology.
- Baldschuhn, R., Binot, F., Fleig, S., & Kockel, F., 2001. *Geotektonischer Atlas von Nordwestdeutschland und dem deutschen Nordsee-Sektor*, Geologisches Jahrbuch A, 153, Schweizerbarth, Stuttgart.
- Baujard, C. & Bruel, D., 2006. Numerical study of the impact of fluid density on the pressure distribution and stimulated volume in the Soultz HDR reservoir, *Geothermics*, **35**, 607–621.
- Beardmore, G. R. & Cull, J. P., 2001. *Crustal Heat Flow*, University Press, Cambridge.
- Beck, A. E., 1988. Methods for determining thermal conductivity and thermal diffusivity, in R. Hänel, L. Rybach, & L. Stegena (eds.), *Handbook of terrestrial heat-flow density determination*, pp. 87 – 124, Kluwer, Dordrecht.
- Berger, A., Richter, B., Imolauer, K., Hepp, K., Schwien, M., Heckelmann, M., Ueltzen, M., & Nein, T., 2009. Erfolgreiche Projektumsetzung Geothermie Unterhaching, Rödl & Partner GbR, Nürnberg, [http://www.geothermie-unterhaching.de/cms/geothermie/web.nsf/gfx/73F32393CB88135BC1257663003B5456/\\$file/Projektdokumentation_Roedl.pdf](http://www.geothermie-unterhaching.de/cms/geothermie/web.nsf/gfx/73F32393CB88135BC1257663003B5456/$file/Projektdokumentation_Roedl.pdf), retrieved: 03/01/2012.

Bibliography

- Bertani, R., 2012. Geothermal power generation in the world 2005–2010 update report, *Geothermics*, **41**, 1–29.
- Bischof, C. H., Carle, A., Corliss, G. F., Griewank, A., & Hovland, P. D., 1992. ADIFOR: Generating Derivative Codes from Fortran Programs, *Scientific Programming*, **1**(1), 11–29.
- Bischof, C. H., Carle, A., Khademi, P., & Mauer, A., 1996. ADIFOR 2.0: Automatic Differentiation of Fortran 77 Programs, *IEEE Computational Science & Engineering*, **3**(3), 18–32.
- Blumenthal, M., 2007. *Numerische Modellierung hydraulischer und thermischer Prozesse im tiefen Wärmereservoir bei Soultz-sous-Forêts, Frankreich*, unpublished Diploma thesis, Institute for Applied Geophysics and Geothermal Energy, E.ON Energy Research Center, RWTH Aachen University, http://www.eonerc.rwth-aachen.de/global/show_document.asp?id=aaaaaaaaabqwfufu, retrieved: 03/01/2012.
- BMWi, 2012. Die Energiewende in Deutschland, Bundesministerium für Wirtschaft und Technologie (BMWi), Berlin, <http://www.bmwi.de/DE/Mediathek/publikationen,did=475210.html>, retrieved: 08/17/2012.
- Bücker, C. & Rybach, L., 1996. A simple method to determine heat production from gamma-ray logs, *Marine and Petroleum Geology*, **13**(4), 373–375.
- Burgers, G., Leeuwen, P., & Evensen, G., 1998. Analysis scheme in the ensemble Kalman filter, *Monthly Weather Review*, **126**, 1719–24.
- Caers, J., 2005. *Petroleum Geostatistics*, Society of Petroleum Engineers, Richardson, TX.
- Campos, O. M., 2002. *Spatial Distribution Of Calcite Concretions Inferred From Borehole Image Data And Their Use In Reservoir Modeling: An Application To Breitbrunn Field*, Master's thesis, The University of Texas, Austin TX, <http://www.pge.utexas.edu/theses02/campos.pdf>, retrieved: 05/22/2012.
- Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., & Slooten, L. J., 2005. Inverse problems in hydrogeology, *Hydrogeology Journal*, **13**, 206–222.
- Celaya, M. & Wahr, J., 1996. Aliasing and noise in core-surface flow inversion, *Geophysical Journal International*, **126**(2), 447–69.
- Chen, Y. & Zhang, D., 2006. Data assimilation for transient flow in geologic formations via ensemble Kalman filter, *Advances in Water Resources*, **29**, 1107–1122.

Bibliography

- Cherubini, Y., Cacace, M., & Scheck-Wenderoth, M., 2011. Assessment of the impact of faults on the coupled fluid and heat transport in a geothermal site (Groß Schönebeck, NE-German Basin): First results from 3d finite element simulations, *Geophysical Research Abstracts*, **13**, EGU2011–2514, <http://meetingorganizer.copernicus.org/EGU2011/EGU2011-2514.pdf>, retrieved: 03/01/2012.
- Chiles, J. P. & Delfiner, P., 1999. *Geostatistics: Modeling Spatial Uncertainty*, Wiley & Sons, New York.
- Clauser, C. (ed.), 2003. *Numerical Simulation of Reactive Flow in Hot Aquifers. SHEMAT and processing SHEMAT*, Springer, Heidelberg-Berlin.
- Clauser, C., 2006. Geothermal Energy, in K. Heinloth (ed.), Landolt-Börnstein, Group VIII "Advanced Material and Technologies", Vol. 3 "Energy Technologies", Subvol. C "Renewable Energies", pp. 480–595, Springer, Heidelberg-Berlin.
- Clauser, C., 2011. Thermal storage and transport properties of rocks, ii: thermal conductivity and diffusivity, in H. K. Gupta (ed.), *Encyclopedia of Solid Earth Geophysics*, Vol. 2, pp. 1431–1448, Springer, Dordrecht.
- Clauser, C., Popov, Y., & Kukkonen, I. T., 1999. Heat transfer processes in the upper crust - a detailed geothermal and hydrological study in the area of the Kola deep hole, Russia, Final report, intas-93-273-ext, Archiv Nr. 119469, Leibniz Institute for Applied Geosciences, Hannover, Germany.
- Cloetingh, S., Cornu, T., Ziegler, P., & Beekman, F., 2006. Neotectonics and intraplate continental topography of the northern alpine foreland, *Earth-Science Reviews of Geophysics*, **74**(3-4), 127–196.
- Cohn, S. E., 1997. An introduction to estimation theory, *Journal of the Meteorological Society of Japan*, **74**, 63–75.
- Cooley, R. L., 2004. A theory for modeling ground-water flow in heterogeneous media, Professional Paper 1679, U.S. Geological Survey (USGS), Reston VA, <http://pubs.usgs.gov/pp/1679/report.pdf>, retrieved: 05/16/2012.
- Darcy, H., 1856. *Les fontaines publiques de la ville de Dijon*, Dalmont, Paris.
- Darnet, M., Marquis, G., & Sailhac, P., 2006. Hydraulic stimulation of geothermal reservoirs: fluid flow, electric potential and microseismicity relationships, *Geophysical Journal International*, **166**, 438–444.

Bibliography

- Marsily, G. de, 1986. *Quantitative Hydrogeology*, Academic, Orlando.
- Delépine, N., Cuenot, N., Rothert, E., Parotidis, M., Rentsch, S., & Shapiro, S. A., 2003. Characterization of fluid transport properties of the Hot Dry Rock reservoir Soultz-2000 using induced microseismicity, *Journal of Geophysics and Engineering*, **1**, 77–83.
- Deming, D., 1989. Application of bottom-hole temperature corrections in geothermal studies, *Geothermics*, **18**(5/6), 775–786.
- Deutsch, C. V. & Journel, A. G., 1998. *GSLIB. Geostatistical software library and user's guide*, University Press, Oxford.
- Dezayes, C., Genter, A., & Valley, B., 2010. Structure of the low permeable naturally fractured geothermal reservoir at Soultz, *Comptes Rendus Geoscience*, **342**(7–8), 517–530.
- Diersch, H. J.-G. & Kolditz, O., 2002. Variable-density flow and transport in porous media: approaches and challenges, *Advances in Water Resources*, **25**, 899–944.
- Dong, Y., Gu, Y., & Oliver, D. S., 2006. Sequential assimilation of 4D seismic data for reservoir description using the ensemble Kalman filter, *Journal of Petroleum Science and Engineering*, **53**, 83–99.
- Doveton, J. H. & Cable, H. W., 1979. Fast matrix methods for the lithological interpretation of geophysical logs, *Computers & Geology*, **3**, 101–116.
- Elfeki, A. & Dekking, M., 2001. A Markov Chain Model for Subsurface Characterization: Theory and Applications, *Mathematical Geology*, **33**(5), 569–589.
- Evensen, G., 1994. Sequential data assimilation with a nonlinear quasigeostrophic model using Monte Carlo methods to forecast error statistics, *Journal of Geophysical Research*, **99**(C5), 10143–62.
- Evensen, G., 2003. The Ensemble Kalman Filter: theoretical formulation and practical implementation, *Ocean Dynamics*, **53**, 343–367.
- Evensen, G., 2009. The ensemble Kalman filter for combined state and parameter estimation, *IEEE Control Systems Magazine*, **29**(3), 83–104.
- Field, D., 1987. Relations between the statistics of natural images and the response properties of cortical cells, *Journal of Optical Society of America*, **4**(12), 2379—2394.
- Fishman, G., 1996. *Monte Carlo. Concepts, Algorithms, and Applications*, Springer, New York.

Bibliography

- Förster, A., 2001. Analysis of borehole temperature data in the Northeast German Basin: continuous logs versus bottom-hole temperatures, *Petroleum Geoscience*, **7**(3), 241–254.
- Frey, M. & Milles, U., 2007. Geothermische Stromerzeugung in Landau, *BINE Projektinfo*, **14/2007**, ISSN 0937-8367, http://www.geox-gmbh.de/media/Downloadbereich/projekt_1407internet-x.pdf, retrieved: 09/03/2012.
- Gauch, H. G., 2003. *Scientific Method in Practice*, University Press, Cambridge.
- Gavalas, G. R., Shah, P. C., & Seinfeld, J. H., 1976. Reservoir history matching by Bayesian estimation, *Society of Petroleum Engineers Journal*, **16**(6), 337–350.
- Gentier, S., Rachez, X., Peter-Borie, M., Tran Ngoc, T. D., & Souque, C., 2010a. Un modèle d'écoulement et de transport du réservoir géothermique profond de Soultz-sous-Forêts (France), in Proceedings of the Journées Nationales de Géotechnique et de Géologie de l'Ingénieur (JNGG2010), Grenoble, France, July 07–09, http://jngg2010.hmg.inpg.fr/sites/meca_des_roches/GENTIER.pdf, retrieved: 03/01/2012.
- Gentier, S., Rachez, X., Tran Ngoc, T. D., Peter-Borie, M., & Souque, C., 2010b. 3D flow of the medium-term circulation test performed in the deep geothermal site of Soultz-sous-Forêts (France), in R. Horne (ed.), Proceedings of the 2010 World Geothermal Congress, Bali, Indonesia, April 25–30, International Geothermal Association, Bochum, Germany, <http://b-dig.iie.org.mx/BibDig/P10-0464/>, retrieved: 03/01/2012.
- Gérard, A., Genter, A., Kohl, T., Lutz, P., Rose, P., & Rummel, F., 2006. The deep EGS (Enhanced Geothermal System) project at Soultz-sous-Forêts, Alsace, France, *Geothermics*, **35**(5-6), 473–483.
- Gersztenkorn, A. & Marfurt, K. J., 1999. Eigenstructure-based coherence computations as an aid to 3-D structural and stratigraphic mapping, *Geophysics*, **64**(5), 1468–1479.
- Gessner, K., Kühn, M., Rath, V., Kosack, C., Blumenthal, M., & Clauser, C., 2009. Coupled process models as a tool for analysing hydrothermal systems, *Surveys in Geophysics*, **30**(3), 133–162.
- Gnjezda, D. K., 2009. *Generierung eines geologischen Untergrundmodells zur geothermischen Reservoirsimulation für das Produktionsfeld Hohne*, unpublished bachelor thesis, Institute for Applied Geophysics and Geothermal Energy, E.ON Energy Research Center, RWTH Aachen University, http://www.eonerc.rwth-aachen.de/aw/cms/website/zielgruppen/gge/publications_gge/ablage_struktur/2009/~wgg/

Bibliography

- Generierung_eines_geologischen_Untergrundmodells/?lang=en, retrieved: 09/01/2012.
- Griewank, A., 2000. *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia PA.
- Gu, Y. & Oliver, D. S., 2006. The ensemble Kalman filter for continuous updating of reservoir simulation models, *Journal of Energy Resources Technology*, **128**(1), 79–87.
- Hahne, B., Thomas, R., & the gebo Geosystem Team, 2011. Combined geoscientific investigations of geothermal reservoir characteristics in Lower Saxony, germany, *Geophysical Research Abstracts*, **13**, EGU2011–10054, <http://meetingorganizer.copernicus.org/EGU2011/EGU2011-10054.pdf>, retrieved: 03/01/2012.
- Häring, M. O., Schanz, U., Ladner, F., & Dyer, B. C., 2008. Characterisation of the Basel 1 Enhanced Geothermal System, *Geothermics*, **37**(5), 469–495.
- Hartmann, A., Rath, V., & Clauser, C., 2005. Thermal conductivity from core and well log data, *International Journal of Rock Mechanics & Mining Sciences*, **42**, 1042–1055.
- Hartmann, A., Pechinig, R., & Clauser, C., 2007. Petrophysical analysis of regional-scale thermal properties for improved simulations of geothermal installations and basin-scale heat and fluid flow, *International Journal of Earth Sciences*, **42**(7-8), 1042–1055.
- Held, S., 2011. *Numerische Berechnungen auf Basis eines geologischen 3D Modells zur optimalen Bewirtschaftung eines geotherischen Reservoirs*, unpublished Masters thesis, Institut für Angewandte Geowissenschaften, Karlsruher Institut für Technologie.
- Hendricks Franssen, H. J. & Kinzelbach, W., 2008. Real-time groundwater flow modeling with the Ensemble Kalman Filter: Joint estimation of states and parameters and the filter inbreeding problem, *Water Resources Research*, **44**, W09408.
- Hendricks Franssen, H. J., Alcolea, A., Riva, M., Bakr, M., van de Wiel, N., Stauffer, F., & Guadagnini, A., 2009. A comparison of seven methods for the inverse modelling of groundwater flow. application to the characterisation of well catchments, *Advances in Water Resources*, **23**(6), 851–872.
- Hermanrud, C., Cao, S., & Lerche, I., 1990. Estimates of virgin rock temperature derived from BHT measurements: Bias and error, *Geophysics*, **55**(7), 924–931.
- Herrera, G. S., 1998. *Cost effective groundwater quality sampling network design*, Ph.D. thesis, The University of Vermont, Burlington VT.

Bibliography

- Hill, M. C. & Tiedeman, C. R., 2006. *Effective Model Calibration: With Analysis of Data, Sensitivities, Predictions, and Uncertainty*, Wiley, New York.
- Horner, D., 1951. Pressure build-up in wells, *Proceedings of the Third World Petroleum Congress*, **34**(316), 503–521.
- Houtekamer, P. L. & Mitchell, H. L., 1998. Data assimilation using an ensemble Kalman filter technique, *Month Weather Review*, **126**, 796–811.
- Hunt, R. J., Doherty, J., & Tonkin, M. J., 2007. Are models too simple? Arguments for increased parameterization, *Ground Water*, **45**(3), 54–262.
- Iwanowski-Strahser, K., Vogt, C., Arnold, J., Szalaiovà, E., & Rabbel, W., 2011. Seismic assessment of the geothermal potential of North German Basin sediments – from seismic to a geothermal model, *Geophysical Research Abstracts*, **13**, EGU2011–11035, <http://meetingorganizer.copernicus.org/EGU2011/EGU2011-11035.pdf>, retrieved: 03/01/2012.
- Jafarpour, B. & Tarrahi, M., 2011. Assessing the performance of the ensemble Kalman filter for subsurface flow data integration under variogram uncertainty, *Water Resources Research*, **47**, W05537.
- Jain, C., Vogt, C., & Clauser, C., 2012. Optimized layout of engineered geothermal systems and potential in Germany, *Renewable Energy*, submitted.
- Kalman, R. E., 1960. A New Approach to Linear Filtering and Prediction Problems, *Transaction of the ASME – Journal of Basic Engineering*, **82**(D), 35–45.
- Kitanidis, P. K., 1997. *Introduction to Geostatistics, Applications in Hydrogeology*, University Press, Cambridge.
- Knaut, A., 2012. *Effects of temperature uncertainty on the financial valuation of geothermal projects*, unpublished Diploma thesis, Institute for Future Energy Consumer Needs and Behavior, E.ON Energy Research Center, RWTH Aachen University.
- Kohl, T. & Mégel, T., 2007. Predictive modeling of reservoir response to hydraulic stimulations at the European EGS site Soultz-sous-Forêts, *International Journal of Rock Mechanics & Mining Science*, **44**, 1118–1131.
- Kolditz, O. & Clauser, C., 1998. Numerical simulation of flow and heat transfer in fractured crystalline rocks: Application to the Hot Dry Rock site in Rosemanowes (UK), *Geothermics*, **27**, 1–23.

Bibliography

- Kosack, C., Vogt, C., Rath, V., & Marquart, G., 2010. Stochastic estimates of the permeability field of the Soultz-sous-Forêts geothermal reservoir – comparison of Bayesian inversion, MC geostatistics, and EnKF assimilation, *Geophysical Research Abstracts*, **12**, EGU2010–4383, <http://meetingorganizer.copernicus.org/EGU2010/EGU2010-4383.pdf>, retrieved: 03/01/2012.
- Krymskaya, M. V., Hanea, R. G., & Verlaan, M., 2009. An iterative ensemble Kalman filter for reservoir engineering applications, *Computational Geosciences*, **13**(2), 235–244.
- Kühnle, N., 2011. *Hydraulische Modellierung einer Kluft mit selbst-affinen Oberflächen*, unpublished Diploma thesis, Institut für Angewandte Geowissenschaften, Karlsruher Institut für Technologie.
- Lawniczak, W., Hanea, R., Heemink, A., & McLaughlin, D., 2008. Multiscale ensemble filtering for reservoir engineering applications, *Computational Geosciences*, **13**(2), 245–254.
- Ledéser, B., Hebert, R., Genter, A., Bartier, D., Clauer, N., & Grall, C., 2010. Fractures, hydrothermal alterations and permeability in the Soultz Enhanced Geothermal System, *Comptes Rendus Geoscience*, **342**(7–8), 607–615.
- Li, S., Zhang, C., Yin, Y., Yin, T., & Yan, S., 2008. Stochastic modeling of reservoir with multi-source, *Earth Science Frontiers*, **15**(1), 196–201.
- Lund, J. W., Freeston, D. H., & Boyd, T. L., 2011. Direct utilization of geothermal energy 2010 worldwide review, *Geothermics*, **40**(3), 159–180.
- Manzella, A., 2010. Technological challenges in exploration and investigation of EGS and UGR, in R. Horne (ed.), Proceedings of the 2010 World Geothermal Congress, Bali, Indonesia, April 25–30, International Geothermal Association, Bochum, Germany, <http://b-dig.iie.org.mx/BibDig/P10-0464/>, retrieved: 03/01/2012.
- Marquart, G., Wolf, A., Rath, V., & Vogt, C., 2012. Stochastic estimate of permeability in geothermal reservoir simulation using the ensemble Kalman filter method, *Geothermics*, submitted.
- Mavko, G., Mukerji, T., & Dvorkin, J., 1998. *The Rock Physics Handbook. Tools for Seismic Modelling of Porous Media*, University Press, Cambridge.
- McDermott, C. I., Lodemann, M., Ghergut, I., Tenzer, H., Sauter, M., & Kolditz, O., 2006. Investigation of coupled hydraulic-geomechanical processes at the KTB site: pressure-dependent characteristics of a long-term pump test and elastic interpretation using a geomechanical facies model, *Geofluids*, **6**, 67–81.

Bibliography

- Montzka, C., Moradkhani, H., Weihermuller, L., Canty, M., Hendricks Franssen, H. J., & Vereecken, H., 2011. Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter, *Journal of Hydrology*, **399**(3-4), 410–421.
- Moore, C. & Doherty, J., 2006. The cost of uniqueness in groundwater model calibration, *Advances in Water Resources*, **29**(4), 605–623.
- Moradkhani, H., Hsu, K., Gupta, H. V., & Sorooshian, S., 2005a. Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using particle filter, *Water Resources Research*, **41**, W05012.
- Moradkhani, H., Sorooshian, S., Gupta, H. V., & Houser, P. R., 2005b. Dual state–parameter estimation of hydrological models using ensemble Kalman filter, *Advances in Water Resources*, **28**, 135–147.
- Mosegaard, K. & Sambridge, M., 2002. Monte Carlo analysis of inverse problems, *Inverse Problems*, **18**, 29–54.
- Mottaghy, D., 2007. *Heat transfer processes in the upper crust: influence of structure, fluid flow, and paleoclimate*, Doctoral dissertation, RWTH Aachen University.
- Mottaghy, D., Pechnig, R., Taugs, R., Kröger, J., Thomsen, C., Hese, F., & Liebsch-Dörschner, T., 2011a. Geothermal 3-D modeling Hamburg and surrounding areas: Temperature prediction and reservoir simulation, *EGU2011-8684-1*, **13**, EGU2011–8684–1, <http://meetingorganizer.copernicus.org/EGU2011/EGU2011-8684-1.pdf>, retrieved: 03/01/2012.
- Mottaghy, D., Pechnig, R., & Vogt, C., 2011b. The geothermal project Den Haag: 3D numerical models for temperature prediction and reservoir simulation, *Geothermics*, **40**(3), 199–210.
- Naderi Beni, A., 2011. *Multi-phase, Multi-species Reactive Transport Modeling as a Tool for System Analysis in Geological Carbon Dioxide Storage*, Doctoral dissertation, RWTH Aachen University.
- Neuman, S. P. & Wierenga, P. J., 2002. A comprehensive strategy of hydrogeologic modeling and uncertainty analysis for nuclear facilities and sites, technical report NUREG/CR8605, University of Arizona for U. S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, Washington DC, <http://www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr6805/cr6805.pdf>, retrieved: 05/16/2012.

Bibliography

- Norden, B., Förster, A., & Balling, N., 2008. Heat flow and lithospheric thermal regime in the Northeast German Basin, *Tectonophysics*, **460**(1-4), 215–229.
- Nowak, W., 2005. *Geostatistical methods for the identification of flow and transport parameters in the subsurface*, Doctoral dissertation, Universität Stuttgart.
- Nowak, W., 2009. Best unbiased ensemble linearization and quasi-linear Kalman ensemble generator, *Water Resources Research*, **45**, W04431.
- Oliver, D. S. & Chen, Y., 2011. Recent progress on reservoir history matching: a review, *Computational Geosciences*, **15**(1), 185–221.
- Ollinger, D., Baujard, C., Kohl, T., & Moeck, I., 2010. Distribution of thermal conductivities in the Groß Schönebeck (Germany) test site based on 3D inversion of deep borehole data, *Geothermics*, **39**(1), 46–58.
- O’Sullivan, M., Yeh, A., & Mannington, W., 2010. Renewability of geothermal resources, *Geothermics*, **39**, 314–320.
- Papadopoulos, D., Herty, M., Rath, V., & Behr, M., 2011. Identification of uncertainties in the shape of geophysical objects with level sets and the adjoint method, *Computational Geosciences*, **15**(4), 737–753.
- Pape, H., Clauser, C., & Iffland, J., 1999. Permeability prediction based on fractal pore-space geometry, *Geophysics*, **64**(5), 1447–1460.
- Pape, H., Clauser, C., Iffland, J., Krug, R., & Wagner, R., 2005. Anhydrite cementation and compaction in geothermal reservoirs: Interaction of pore-space structure with flow, transport, p-t-conditions, and chemical reactions, *International Journal of Rock Mechanics and Mining Sciences*, **42**, 1056–1069.
- Paradigm, 2011. *SKUA and GOCAD User Guide*, Paradigm Ltd., Houston TX, http://www.kxcad.net/skua-gocad/2011/SKUA_GOCAD/UserGuide/11WellPlanning.pdf, retrieved: 05/15/2012.
- Paschen, H., Oertel, D., & Grünwald, R., 2003. Möglichkeiten geothermischer Stromerzeugung in Deutschland, work report 84, Büro für Technikfolgenabschätzung beim Deutschen Bundestag, <http://www.tab-beim-bundestag.de/de/publikationen/berichte/ab084.html>, retrieved: 03/01/2012.

Bibliography

- Peaceman, D. W., 1983. Interpretation of well-block pressure in numerical reservoir simulation with non-square grid blocks and anisotropic permeability, *Society of Petroleum Engineers Journal*, **23**(3), 531–543.
- Pryor, R. W., 2011. *Multiphysics Modeling Using COMSOL V.4A First Principles Approach*, Transatlantic Publishers, London.
- Radilla, G., Sausse, J., Sanjuan, B., & Fourar, M., 2012. Interpreting tracer tests in the enhanced geothermal system (egs) of soultz-sous-forêts using the equivalent stratified medium approach, *Geothermics*, **44**, 43–51.
- Rall, L. B., 1981. *Automatic Differentiation: Techniques and Applications*, Springer, New York.
- Rath, V., Wolf, A., & Bucker, M., 2006. Joint three-dimensional inversion of coupled groundwater flow and heat transfer based on automatic differentiation: sensitivity calculation, verification, and synthetic examples, *Geophysical Journal International*, **167**, 453–466.
- Reichle, R. H., McLaughlin, D. B., & Entekhabi, D., 2002. Hydrologic Data Assimilation with the Ensemble Kalman Filter, *Monthly Weather Review*, **130**, 103–130.
- Remy, N., 2005. S-GeMS: The Stanford Geostatistical Modeling Software: A tool for new algorithms development, in O. Leuangthong & C. V. Deutsch (eds.), *Geostatistics Banff 2004*, Springer, Dordrecht.
- Rings, J., Huismann, J. A., & Vereecken, H., 2010. Coupled hydrogeophysical parameter estimation using a sequential Bayesian approach, *Hydrology and Earth Science Discussions*, **14**, 545–556.
- Ristic, B., Arulampalam, S., & Gordon, N., 2004. *Beyond the Kalman Filter: Particle Filters for Tracking Applications*, Artech House Inc, Norwood MA.
- Robert, C. P. & Casella, G., 2004. *Monte Carlo statistical methods*, Springer, New York.
- Rühaak, W., Rath, V., & Clauser, C., 2010. Detecting thermal anomalies within the Molasse Basin, Southern Germany, *Hydrogeology Journal*, **18**(8), 1897–1915.
- Rybach, L., 2010. "The Future of Geothermal Energy" and Its Challenges, in R. Horne (ed.), *Proceedings of the 2010 World Geothermal Congress, Bali, Indonesia, April 25–30*, International Geothermal Association, Bochum, Germany, <http://b-dig.iie.org.mx/BibDig/P10-0464/>, retrieved: 03/01/2012.

Bibliography

- Sakov, P. & Oke, P., 2008. A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters, *Tellus*, **60A**(2), 361–371.
- Sambridge, M. & Mosegaard, K., 2002. Monte Carlo methods in geophysical inverse problems, *Reviews of Geophysics*, **40**(3), 3.1–3.29.
- Sanjuan, B., Pinault, J., Rose, P., Gérard, A., Brach, M., Braibant, G., Crouzet, C., Foucher, J., Gautier, A., & Touzelet, S., 2006. Tracer testing of the geothermal heat exchanger at Soultz-sous-Forêts (France) between 2000 and 2005, *Geothermics*, **35**(5-6), 622–653.
- Schafmeister, M.-T., 1999. *Geostatistik für die hydrogeologische Praxis*, Springer, Heidelberg-Berlin.
- Schlumberger, 1999. *Eclipse 100 Users Course*, Schlumberger GeoQuest, Houston TX, http://www.fanarco.net/books/reservoir/E100_Manual.pdf, retrieved: 05/15/2012.
- Schlumberger, 2012. Petrel 2011, <http://www.slb.com/services/software/geo/petrel.aspx>, retrieved: 06/27/2012.
- Schön, J. H., 1996. *Physical properties of rocks: Fundamentals and principles of geophysics*, Pergamon, Oxford.
- Seiler, A., Evensen, G., Skjervheim, J.-A., Hove, J., & Vabø, J., 2011. Using the enkf for history matching and uncertainty quantification of complex reservoir models, in L. Biegler, G. Biros, O. Ghattas, M. Heinkenschloss, D. Keyes, B. Mallick, L. Tenorio, B. van Bloemen Waanders, & K. Willcox (eds.), *Computational Methods for Large-Scale Inverse Problems and Quantification of Uncertainty*, pp. 247–271, Wiley, Chichester, UK.
- Simmelink, H. J. & Vandeweyer, V., 2008. Geothermie Den Haag Zuid-West, 2e fase geologisch onderzoek, technical report 034.72157, Netherlands Organisation for Applied Natural Science Research (TNO), Utrecht.
- Simmelink, H. J., Vandeweyer, V., & Ramaekers, J., 2007. Geologisch locatie-specifiek onderzoek voor het Business Plan Geothermie Den Haag Zuid-West, technical report 2007-U-R1118/B, Netherlands Organisation for Applied Natural Science Research (TNO), Utrecht.
- Szalaiová, E., 2012. *Seismic Assessment of Geothermal Potential – Concept and Application Case Study of the German Continental Deep Drilling Site (KTB)*, Doctoral dissertation, Christian-Albrechts-Universität zu Kiel.

Bibliography

- Tarantola, A., 2004. *Inverse problem theory. Methods for model parameter estimation*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia PA.
- Tarrah, M. & Jafarpour, B., 2012. Inference of permeability distribution from injection-induced discrete microseismic events, *Water Resources Research*, in press, <http://www.agu.org/journals/pip/wr/2012WR011920-pip.pdf>, retrieved: 09/04/2012.
- Tenzer, H., Park, C. H., Kolditz, O., & McDermott, C. I., 2010. Application of the geomechanical facies approach and comparison of exploration and evaluation methods used at Soultz-sous-Forêts (France) and Spa Urach (Germany) geothermal sites, *Environmental Earth Sciences*, **61**(4), 853–880.
- Tester, J. (ed.), 2006. *The Future of Geothermal Energy*, Massachusetts Institute of Technology, Boston MA, <http://geothermal.inel.gov>, retrieved: 03/01/2011.
- Tippett, M. K., Anderson, J. L., Bishop, C., Hamill, T. M., & Whitaker, J. S., 2003. Ensemble square root filters, *Month Weather Review*, **131**, 1485–1490.
- Tischner, T., Evers, E., Hauswirth, H., Jatho, R., Kosinowski, M., & Sulzbacher, H., 2010. New concepts for extracting geothermal energy from one well: The GeneSys-project, in R. Horne (ed.), Proceedings of the 2010 World Geothermal Congress, Bali, Indonesia, April 25–30, International Geothermal Association, Bochum Germany, <http://b-dig.iie.org.mx/BibDig/P10-0464/>, retrieved: 03/01/2012.
- Turcotte, D. L., 1997. *Fractals and Chaos in Geology and Geophysics*, University Press, Cambridge.
- van Balen, R. T., van Bergen, F., de Leeuw, C., Pagnier, H., Simmelink, H., van Wees, J. D., & J.M.Verweij, 2000. Modelling the hydrocarbon generation and migration in the West Netherlands Basin, the Netherlands, *Netherlands Journal of Geosciences*, **79**(1), 29–44.
- Vogt, C. & Widera, A., 2012. Ensemble Kalman filter assimilation of temperature data to estimate permeability in geothermal reservoirs, *Geophysical Research Abstracts*, **14**, EGU2012–495, <http://meetingorganizer.copernicus.org/EGU2011/EGU2012-495.pdf>, retrieved: 05/03/2012.
- Vogt, C., Mottaghy, D., Wolf, A., Rath, V., Pechnig, R., & Clauser, C., 2010. Reducing temperature uncertainties by stochastic geothermal reservoir modeling, *Geophysical Journal International*, **181**(1), 321–333.

Bibliography

- Vogt, C., Marquart, G., Kosack, C., Wolf, A., & Clauser, C., 2012. Estimating the permeability distribution and its uncertainty at the EGS demonstration reservoir Soultz-sous-Forêts using the ensemble Kalman filter, *Water Resources Research*, **48**, W08517.
- Vogt, C., Iwanowski-Strahser, K., Marquart, G., Arnold, J., Mottaghy, D., Pechinig, R., Gnjezda, D., & Clauser, C., 2013. Modeling contribution to risk assessment of thermal production power for geothermal reservoirs, *Renewable Energy*, **53**, 230–241.
- Vosteen, H.-D., Rath, V., Schmidt-Mumm, A., & Clauser, C., 2004. The thermal regime of the Northeastern-German Basin from 2-D inversion, *Tectonophysics*, **386**(1-2), 81–95.
- Vrugt, J. A., Diks, C. G. H., Gupta, H. V., Bouten, W., & Verstraten, J. M., 2005. Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, *Water Resources Research*, **41**, W01017.
- Walter, R., 2007. *Geologie von Mitteleuropa*, Schweizerbart, Stuttgart, Germany.
- Watanabe, N., Wang, W., McDermott, C. I., Taniguchi, T., & Kolditz, O., 2010. Uncertainty analysis of thermo-hydro-mechanical (THM) coupled processes in heterogeneous porous media, *Computational Mechanics*, **45**(4), 263–280.
- Weiskopf, D. & Erlebacher, G., 2004. Overview of flow visualization, in C. D. Hansen & C. R. Johnson (eds.), *The Visualization Handbook*, pp. 261–278, Academic, Burlington MA.
- Wolf, A., 2011. *Ein Softwarekonzept zur hierarchischen Parallelisierung von stochastischen und deterministischen Inversionsproblemen auf modernen ccNUMA-Plattformen unter Nutzung automatischer Programmtransformation*, Doctoral dissertation, RWTH Aachen University.
- Wolf, P. & Felix, M., 2009. Rahmenkonzept Tiefengeothermie Freistaat Sachsen, Freistaat Sachsen, Landesamt für Umwelt, Landwirtschaft und Geologie, http://www.smul.sachsen.de/umwelt/download/TGT-Rahmenkonzept_Sachsen_100809.pdf, retrieved: 03/01/2012.
- Wolter, M., 2010. *Navigation in Time-Varying Scientific Data*, Doctoral dissertation, RWTH Aachen University.
- Xu, C., Dowd, P. A., & Mohais, R., 2012. Connectivity analysis of the Habanero Enhanced Geothermal System, in *Proceedings of the 37th Workshop on Geothermal Reservoir Engineering*, Stanford, January 30 – February 1, Vol. 37, Stanford University, CA, <http://www.certain.com/system/profile/web/index.cfm?PKWebId=0x335199c38d>, retrieved: 05/22/2012.

Bibliography

- Xu, T., Sonnenthal, E., Spycher, N., & Pruess, K., 2006. TOUGHREACT – a simulation program for non-isothermal multiphase reactive geochemical transport in variably saturated geologic media: Applications to geothermal injectivity and co2 geological sequestration, *Geothermics*, **32**(2), 145–165.
- Yilmaz, Ö., 1991. *Seismic Data Processing*, Society of Exploration Geophysicist, Tulsa OK.
- Zhang, Y. & Oliver, D. S., 2011. Evaluation and error analysis: Kalman gain regularization versus covariance regularization, *Computational Geosciences*, **15**(3), 489–508.
- Zhou, H., Gómez-Hernández, J. J., Hendricks Franssen, H. J., & Li, L., 2011. An approach to handling non-gaussianity of parameters and state variables in ensemble Kalman filtering, *Advances in Water Resources*, **34**(7), 844–864.

Appendix A.

Tool Implementation and Parallelization

A.1. SGSim

During the construction of the geometrical model for the simulation, groups of cells are combined into units corresponding to geological layers. Each unit is defined by a set of up to 13 rock properties, depending on the particular choice of problem to be solved. When the stochastic module is run for a target property, the corresponding property distribution is simulated with the SGSim algorithm as if the unit would comprise the entire model. This is done in each unit for a specific set of rock properties. The individual cell values – specified rock properties – are assigned according to the generated SGSim model distribution. Thus, cells of a unit no longer have identical properties. As a result, each geological unit in the model is simulated with respect to its measured value distribution. For different rock properties, the stochastic simulation may be performed in a linear or logarithmic mode. For permeability simulation in particular, the logarithmic mode is useful. This property initialization process is repeated for each realization. Then, fluid flow and heat transport are simulated which requires usually the larger part of the computation time.

The grids for stochastic and transport simulation need to be similar to avoid geometric discrepancies. The sums of all cell boundary positions in one dimension are not allowed to differ more than one 1 %. However, a grid refinement is possible for the fluid flow and heat transport simulations, e.g. around boreholes. Because SGSim works only with equidistant grids, the property values in each node of the coarser grid are assigned to the corresponding node of the finer resolution grid. To avoid upscaling problems, the grid used for stochastic simulation is always the coarser one.

Because all realizations are computed independently, the algorithm suits itself very well for parallelization. Here, a nested parallelization is implemented which means that each individual realization is simulated in parallel by a group of computing threads and, additionally, the different realizations are computed by different groups of threads. As a consequence, the code can be run on several processor cores which allows fast computation, even for large ensembles. Details

of the parallelization technique are described in Wolf (2011).

The stochastic variation of boundary conditions with SGSim is also implemented in a similar way to the simulation of the rock properties.

A.2. Ensemble Kalman Filter

The Ensemble Kalman Filter method, as implemented in SHEMAT-Suite, mainly depends on the computation of hundreds or thousands of different model realizations. Therefore a multi-level (nested) OpenMP parallelization (Wolf, 2011) is implemented to overcome the computational costs.

To evaluate the quality of the parallelization, simulations are done on 11 computing nodes, each of them are equipped with two Intel[®] Xeon[®] E5420 processors at 2.5 GHz. The nodes are connected via DDR Infiniband and each processors has four processor cores. The innovative part of this system is the vSMP[™] software of the company ScaleMP[™]. The vSMP[™] software creates virtually a single shared memory system, which is needed for the OpenMP parallelization. This node configuration utilizes up to 88 processor cores for the computation of 88 different model simulations at the same time.

An EnKF model test computation needs the computation of 880 realizations in overall three separated EnKF iterations. Each iteration consists of four steps, namely:

1. stochastic initialization or reading of the last iteration results;
2. the actual EnKF iteration (including 880 realization computations);
3. the re-computation of the ensemble without data assimilation (including 880 realization computations); and
4. the computation of one mean realization.

The third step allows for comparing observation and simulation results. Between the three EnKF iterations intermediate results are evaluated. Overall around $T_p = 12.6$ days are needed only for the real computation with the 11 nodes. Furthermore, the computation time for one EnKF iteration with only one single realization is evaluated to estimate the computational speedup. It turns out that the theoretical computation time for a single processor core (no parallelization, but 880 realizations) would take $T_s = 912$ days (~ 3 years) based on this evaluation. Thus, the computational speedup T_s/T_p is around 72. This is a very good computation time reduction for such a kind of vSMP[™] architectures.

Appendix B.

Virtual Reality Visualization

Virtual reality (VR) techniques allow exploring numerical 3D models in an intuitive and comprehensive way (Weiskopf & Erlebacher, 2004; Wolter, 2010). In geothermal studies under uncertainty, the VR tools can be applied for visualizing different Monte Carlo realizations or time slices as well as state or parameter standard deviations at the same time. I used 3D VR projection to explore the reservoir model at The Hague (Chapter 5.1.2) and the pathways inside the Soultz reservoir (Chapter 6.4) on a single projection wall or inside a CAVE using 5 projecting walls at the same time.

By tracking the position of the modeler optically, making use of threshold values and volume rendering, and visualizing the model using stereoscopic 3D projection, the modeler can explore features of model geometries as well as of model state and parameter fields, which are not obvious in 2D desktop visualization. 2D exploring on normal screens is usually limited to slices of the model. As a consequence, interconnections of e.g. preferential flow paths may be overlooked or tracer concentration maxima may be hidden during the exploration. In contrast, in virtual reality environments, the modeler is free to move intuitively inside the data cube. This is of particular importance in heterogeneous models as they are studied in this thesis. Figures B.1 and B.2 illustrate the CAVE virtual reality environment.

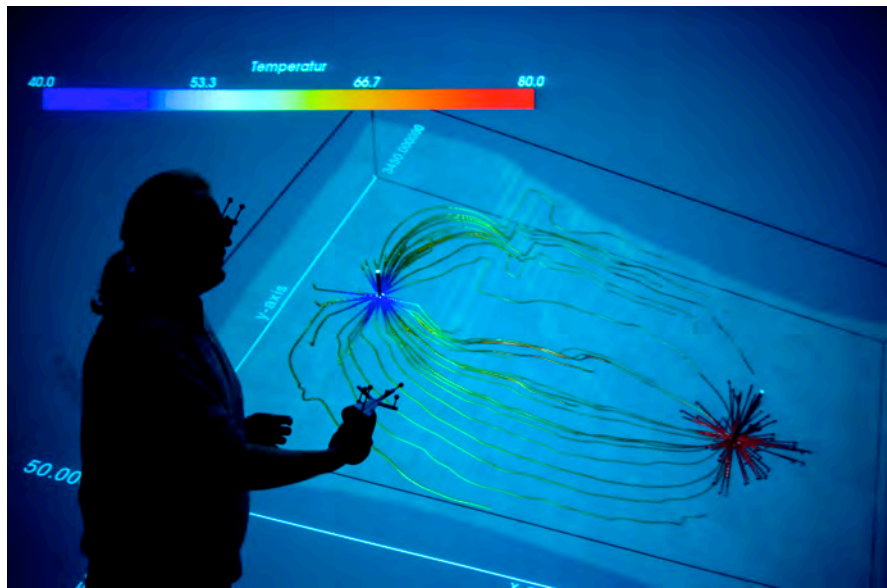


Figure B.1.: Injection and production wells, filtration velocity streamlines, and modeler inside the CAVE 3D virtual reality environment projecting the reservoir model in The Hague, Netherlands (see Chapter 5.1.2).

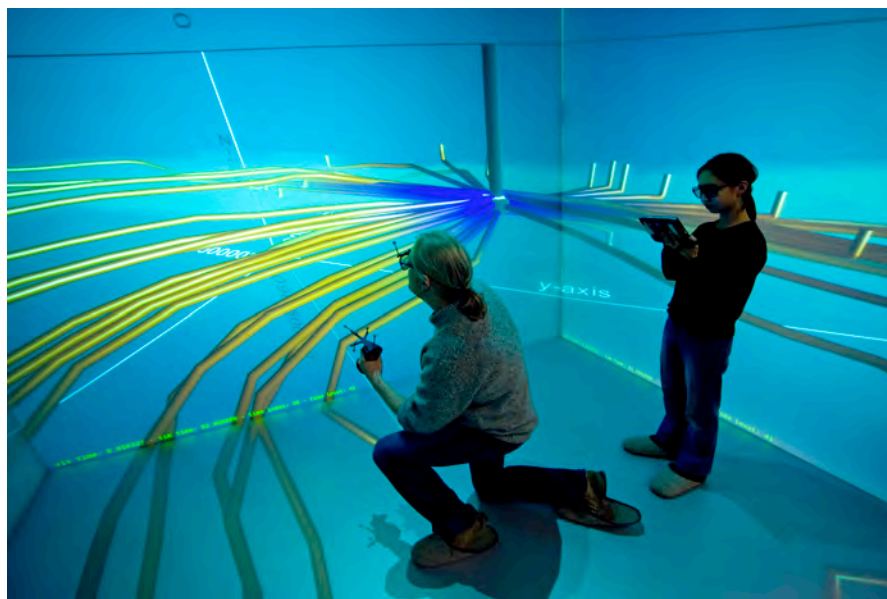


Figure B.2.: Injection well, filtration velocity streamlines, and two modelers inside the CAVE 3D virtual reality environment projecting the reservoir model in The Hague, Netherlands (see Chapter 5.1.2). The picture looks disturbed because the projection is computed for the user's head position.

Acknowledgments

Many people contributed to the successful completion of this dissertation — wherever they coauthored the scientific work, discussed the content, provided helpful advice and support, or just put some fun into the job.

First, I want to thank Prof. Christoph Clauser for providing me the opportunity to work in the fields of numerical modeling and of geothermal energy at his institute, which was actually my major research interest after diploma. His enthusiasm and knowledge in the field of geothermal energy, motivation for publishing papers, broad advice during my research helped substantially to do the research and to finish this thesis.

Also, I thank Prof. Thomas Kohl for agreeing to be co-advisor of this dissertation and Prof. Christoph Hilgers for being in the examination committee.

In particular, I owe my deepest gratitude to Gabriele Marquart for her never-ending patience in answering my questions and discussion various problems with profound experience in numerical modeling of the subsurface — and for all chit-chat about everything under the sun. And for the cookies ...

Darius Mottaghy and Volker Rath never hesitated to share their wide knowledge which helped me a lot avoiding the pitfalls in reservoir simulation and inverse modeling. I am indebted to Andreas Wolf, who never lost patience during hundreds of SHEMAT questions and programmed a lot of ideas into the code I would never manage to realize by myself.

For fruitful scientific and non-scientific discussion, optimism and fatalism at the right moment, cake and transport, ears and hands, Bulettos and helicopters, fun and friendship, and a really nice working atmosphere I thank my fellow doctoral students of the "Doktorandenrunde", Johanna Bruckmann, Henrik Büsing, Annick Fehr, Huang Simin, Alexander Michalski, Ali Naderi, Christ Nordlund, Jan Niederau, Ines Rick, Johanna Smaczny, Jan Volkmann, and my other colleagues from GGE, Lothar Ahrensmeier, Birgit Alewelt, Frank Bosch, Michal Culicki, Jouad Doghmi, Anozie Ebigbo, Jannis Frigge, Natalia Günther, Charitra Jain, Peter Jaxy, Christoph Klein, Norbert Klitzsch, Christian Kosack, Sandra Krause, Ute Kreutz, Andreas Koch, Li Shiyuan, Oliver Mohnke, Rachel Jorand, Justus Schwan, Rainer Schütt, Heinz Triebe, André Widera, and Karen Willbrand. It was a good time at GGE.

I thank the IT team and the EONERC administration for all support.

Further, I thank the *MeProRisk* working group for the nice cooperation in revealing characteristics of geothermal reservoirs, in particular Katja Iwanowski-Strahser and Eva Szalaiová for the seismics and the nice meetings, as well as Irene Tedjo-Palczynski for virtual reality models.

For cooperating in the interesting topic of transferring temperature uncertainties to uncertainties in financial modeling, I thank Christiane Rosen and Andreas Knaut.

I thank Harrie-Jan Hendricks Franssen for a lot of answers to questions regarding the EnKF and various reviewers which optimized my journal papers.

For providing data for my numerical modeling, I am indebted to RWE Dea AG (Hamburg, Germany), IF Technology (Arnheim, the Netherlands) and TNO Built Environment and Geosciences (Delft, the Netherlands) for providing data and granting permission to publish results. Frank Schoof and Bas Pittens (Aardwarmte Den Haag, The Netherlands) kindly provided information on the production temperature measurement of the project in the Hague.

My research was funded by the the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU), grant FKZ 0327563, and by the German Ministry of Education and Science (BMBF), grant 03SF0326A.

For proofreading of this thesis and for discussions about ups and downs in a researcher's life, I thank Henrik Büsing, Lydia Hanßen, Christina Hopfgarten, Christian Kosack, Gabriele Marquart, Chris Nordlund, Jan Volkmann, Judith Vogt, and Marc Wolter.

Thank you, 'Mittwochrunde' aka 'Montagsrunde': Alex, Christoph, Dietmar, Hannah, Judith, Klaus, Lydia, Marc, Nils, Tobias for the right kind of distraction.

I warmly thank my parents, my parents-in-law, and my brother for all kinds of support.

Mostly, I thank my "moon and stars", my wife Judith and my children Yannik, Macha and Fionn — for their love and inspiration during hours without science.

Curriculum Vitae

Persönliche Daten

Name	Christian Vogt
Geburtstag	30. Oktober 1979
Geburtsort	Mechernich
Staatsangehörigkeit	deutsch
Familienstand	verheiratet

Schulbildung

1986 – 1990	Grundschule Kommern
1990 – 1992	Gymnasium am Turmhof, Mechernich
1992 – 1999	St. Angela-Gymnasium, Bad Münstereifel
06/1999	Abitur

Zivildienst

1999 – 2000	Pflegemitarbeiter der Intensivstation, Eifelhöhenklinik Marmagen
-------------	---

Studium

2000 – 2002	Studium des Maschinenbaus (Diplom) an der RWTH Aachen University
2002 – 2007	Studium der Physik (Diplom) an der RWTH Aachen University
09/2004	Vordiplom in Physik
10/2006 – 10/2007	Diplomarbeit am III. Physikalischen Institut B, RWTH Aachen University: ”Aufbau eines Messplatzes zum Test akustischer Sensoren für den Nachweis ultra-hochenergetischer Neutrinowechselwirkungen”
11/2007	Diplom in Physik

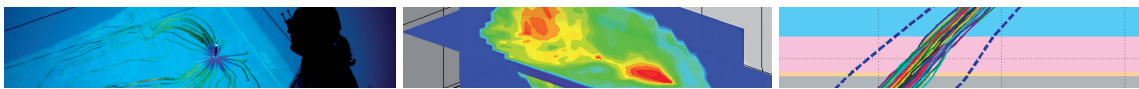
Berufserfahrung

01/2008 – 03/2008	Wissenschaftlicher Mitarbeiter am III. Physikalischen Institut B, RWTH Aachen University
seit 03/2008	Wissenschaftlicher Mitarbeiter zum Zweck der Promotion am Institut für Applied Geophysics and Geothermal Energy, E.ON Energy Research Center, RWTH Aachen University

Publications in Peer-reviewed Journals

- IceCube Collaboration: Ackermann, M., et al., 2008. Search for Ultra High-Energy Neutrinos with AMANDA-II, *Astrophysical Journal*, **675**, 1014-1024.
- The IceCube Collaboration: Abbasi, R., et al., 2008. Solar Energetic Particle Spectrum on 13 December 2006 Determined by IceTop, *Astrophysical Journal Letters*, **689**, L65–L68.
- Vogt, C., Laihem, K., and Wiebusch, C., 2008. Speed of Sound in Bubble-Free Ice, *Journal of the Acoustical Society of America*, **124**(6), 3613-3618.
- The IceCube Collaboration: Abbasi, R., et al., 2009. Search for Point Sources of High Energy Neutrinos with Final Data from AMANDA-II, *Physical Review D*, **79**, 062001.
- The IceCube Collaboration: Abbasi, R., et al., 2009. The IceCube Data Acquisition System: Signal Capture, Digitization, and Timestamping, *Nuclear Instruments and Methods*, **A601**, 294–316.
- The IceCube Collaboration: Abbasi, R., et al., 2010. Measurement of Sound Speed vs. Depth in South Pole ice for Neutrino Astronomy, *Astroparticle Physics*, **33**, 277-286.
- Vogt, C., Mottaghy, D., Wolf, A., Rath, V., Pechinig, R., and Clauser, C., 2010. Reducing temperature uncertainties by stochastic geothermal reservoir modeling, *Geophysical Journal International*, **181**(1), 321–333.
- Mottaghy, D., Pechinig, R., and Vogt, C., 2011. The geothermal project Den Haag: 3D numerical models for temperature prediction and reservoir simulation, *Geothermics*, **40**(3), 199–210.
- Vogt, C., Kosack, C., and Marquart, G., 2012. Stochastic inversion of the tracer experiment of the EGS demonstration reservoir in Soultz-sous-Forêts — revealing pathways and estimating permeability distribution, *Geothermics*, **42**, 1–12.
- Vogt, C., Marquart, G., Kosack, C., Wolf, A., and Clauser, C., 2012. Estimating the permeability distribution and its uncertainty at the EGS demonstration reservoir Soultz-sous-Forêts using the ensemble Kalman filter, *Water Resources Research*, **48**, W08517.
- Vogt, C., Iwanowski-Strahser, K., Marquart, G., Arnold, J., Mottaghy, D., Pechinig, R., Gnjezda, D., and Clauser, C., 2012. Modeling contribution to risk assessment of thermal production power for geothermal reservoirs, *Renewable Energy*, **53**, 230–241.
- Marquart, G., Wolf, A., Rath, V., and Vogt, C., 2012. Stochastic estimate of permeability in geothermal reservoir simulation using the ensemble Kalman filter method, *Geothermics*, submitted.
- Jain, C., Vogt, C., and Clauser, C., 2012. Optimized Layout of Engineered Geothermal Systems and Potential in Germany, *Renewable Energy*, submitted.

Aachen, 05.09.2012



A major obstacle for further development of geothermal energy use is the risk of project failure due to uncertain flow rate and temperature of geothermal installations. In this work, stochastic methods and modeling strategies are developed and applied for predicting the variation of pressure, temperature, and their uncertainty with time within geothermal reservoirs based on observed thermal and hydraulic rock property distributions. This comprises stochastic forward and inverse modeling approaches for simulating heat and tracer transport as well as fluid flow numerically. The approaches reduce the corresponding a priori uncertainties of perturbed parameters and states drastically by 50% – 67% in case of temperature at a depth of 2000 m, depending on the target location. Furthermore, the spatial distribution of permeability is estimated as well as its uncertainty by applying the stochastic assimilation technique of Ensemble Kalman Filtering on production data for sedimentary rocks and fractured hard rocks. This addresses structure and parameter heterogeneity within the reservoir. The presented work flows constitute a method for creating calibrated reservoir models based on data which will allow the operators of a geothermal installation to compute production scenarios optimized with respect to profit or sustainability.