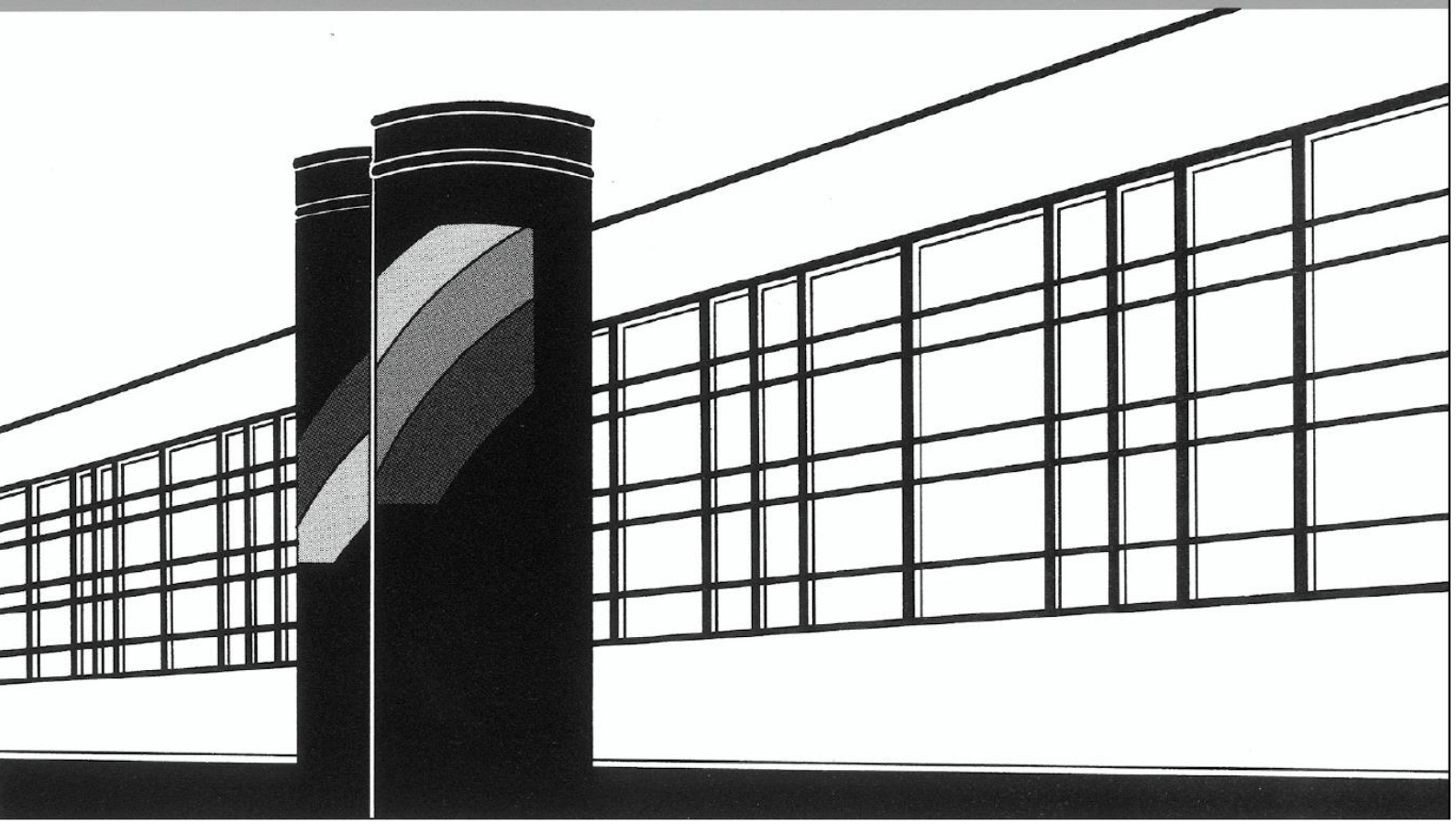
Universität Stuttgart

Institut für Wasser- und Umweltsystemmodellierung

Mitteilungen



Heft 261 Felix Bode

Early-Warning Monitoring Systems for Improved Drinking Water Resource Protection

Early-Warning Monitoring Systems for Improved Drinking Water Resource Protection

Von der Fakultät Bau- und Umweltingenieurwissenschaften der Universität Stuttgart zur Erlangung der Würde eines Doktor-Ingenieurs (Dr.-Ing.) genehmigte Abhandlung

Vorgelegt von
Felix Bode
aus Diepholz

Hauptberichter:	Prof. Dr.-Ing. Wolfgang Nowak
Mitberichter:	Prof. Dr. rer. nat. Dr.-Ing. András Bárdossy
Mitberichter:	Prof. Dr. Patrick M. Reed

Tag der mündlichen Prüfung: 17. September 2018

Institut für Wasser- und Umweltsystemmodellierung
der Universität Stuttgart
2018

Heft 261

Early-Warning Monitoring Systems for Improved Drinking Water Resource Protection

von
Dr.-Ing.
Felix Bode

Eigenverlag des Instituts für Wasser- und Umweltsystemmodellierung
der Universität Stuttgart

D93 Early-Warning Monitoring Systems for Improved Drinking Water Resource Protection

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://www.d-nb.de abrufbar

Bode, Felix:
Early-Warning Monitoring Systems for Improved Drinking Water Resource Protection, Universität Stuttgart. - Stuttgart: Institut für Wasser- und Umweltsystemmodellierung, 2018

(Mitteilungen Institut für Wasser- und Umweltsystemmodellierung, Universität Stuttgart: H. 261)
Zugl.: Stuttgart, Univ., Diss., 2018
ISBN 978-3-942036-65-8

NE: Institut für Wasser- und Umweltsystemmodellierung <Stuttgart>: Mitteilungen

Herausgegeben 2018 vom Eigenverlag des Instituts für Wasser- und Umweltsystem-modellierung

Druck: DCC Kästl e.K., Ostfildern

Danksagung

Ein herzliches Dankeschön geht an meinen Betreuer und Hauptberichter Wolfgang Nowak, der mich jederzeit gefördert und unterstützt hat. Außerdem möchte ich mich bei András Bárdossy und Patrick Reed für ihren Mitbericht bedanken. Ein besonderer Dank geht hierbei an Patrick Reed, der mich für drei Monate in seiner Arbeitsgruppe an der Cornell University in Ithaca aufgenommen hat.

Ein großer Dank geht auch an die gesamte Gruppe des LS3, die jederzeit für fachliche Diskussionen offen waren und hilfreich bei Problemen zur Seite standen. Ein besonderer Dank geht hier an Micha, Anneli, Julian, Sergey und meinen langjährigen Bürokollegen Sebastian, die auch für private Themen immer ein offenes Ohr hatten.

Ebenfalls möchte ich mich beim Zweckverband Landeswasserversorgung und dem DVGW e.V. für die konstruktive und nette Zusammenarbeit bedanken.

Außerhalb der stochastisch hydrogeologischen Welt möchte ich meinen Eltern danken, dass sie es mir ermöglicht haben zu Studieren. Ebenso möchte ich mich bei meinen Freunden bedanken, die jederzeit für einen gesunden Ausgleich gesorgt haben und für mich da waren. Als letztes geht ein ganz besonderer Dank an Alina, die mich immer voll unterstützt hat und mir besonders in stressigen Zeiten Kraft gegeben hat.

Contents

Notation

The following tables show the significant symbols, operators, functions, spaces and abbreviations used in this work. Local notations are explained in the text and are not listed here.

Symbol	Definition	Dimension
Greek Letters:		
α_ℓ	Longitudinal dispersivity	$[L]$
α_t	Transverse dispersivity	$[L]$
β	Rotation angle of ambient flow direction	$[L/L]$
ϵ	Discretization length of the objective space	
$\boldsymbol{\theta}$	Parameter set that is modeled as random variables	
μ	Mean value	
σ^2	Variance	
τ	Travel time	$[T]$
Ω_GD	Area of guaranteed detection	
Ω_PU	Permitted uncertainty area	
Ω_U	Uncertainty area	
Latin Letters:		
A_GD	Size of the area of guaranteed detection	$\left[L^2\right]$
A_PU	Size of the area of permitted uncertainty	$\left[L^2\right]$
B	Aquifer thickness	$[L]$
c	Aqueous contaminant concentration	$\left[M/L^3\right]$
c^crit	Critical concentration at the production well	$\left[M/L^3\right]$
c^det	Chemical detection limit	$\left[M/L^3\right]$
C_thres	Constant concentration value	$\left[M/L^3\right]$
$\mathbf{d}_\mathrm{opt}$	Optimal set of decision variables	
$\mathbf{D}$	Hydromechanic dispersiontensor	$\left[L^2/T\right]$
D_e	Effective diffusion coefficient	$\left[L^2/T\right]$
D_L	Longitudinal dispersion coefficient	$\left[L^2/T\right]$
D_m	Molecular diffusion coefficient	$\left[L^2/T\right]$
D_s	Uncertainty diameter	$[L]$
D_T	Transverse dispersion coefficient	$\left[L^2/T\right]$

$\mathbf{K}$	Hydraulic conductivity matrix	$[L/T]$
L_{id}	List of linear indices containing information about the EDO set	
M	Possible monitoring-well location	
m_0^{IG}	Scaling factor of the inverse Gaussian distribution	$[1/T]$
$\dot{m}_{\mathrm{in}}$	Mass discharge	$[M/T]$
m_{p}	Mass of a single particle	$[M]$
M_0^{IGC}	Scaling factor of the cdf of the inverse Gaussian distribution	$[-]$
P^{det}	Probability of detection	$[-]$
R	Possible contamination source	
t	Time	$[T]$
$\hat{t}^{\mathrm{max}}$	User-defined maximum desirable early-warning time	$[T]$
$\hat{t}_i^{\mathrm{max}}$	Individual maximum achievable early-warning time	$[T]$
$\hat{t}^{\mathrm{min}}$	User-defined minimum desirable early-warning time	$[T]$
$\hat{u}^{\mathrm{min}}$	User-defined minimum early-warning time utility	$[-]$
v	Velocity Scalar	$[L/T]$
$\mathbf{v}$	Velocity vector	$[L/T]$
V	Volume	$[L^3]$
w_{max}	Maximum width of the detectable portion of a plume	$[L]$
x_{max}	x-position of the most distanced intersection point of the detectable portions of two rotated plumes	$[L]$
x_{opt}	x-position where the detectable portion of a plume has its maximum width	$[L]$
x_{tip}	End point of the detectable portion of a plume	$[L]$
y_{thres}	y-position of a concentration isoline	$[L]$

Operators

$\nabla(\cdot)$	Nabla operator
$()\cdot()$	Scalar product
Δ	Differential operator
∂	Partial derivative
$(\cdot)^T$	Transposed vector/matrix

Functions

$c(t)$	Breakthrough curve of an instantaneous mass release
$C(t)$	Breakthrough curve of a continuous mass release
$\tilde{C}(\tilde{x},\tilde{y})$	Concentration distribution in the standard space
$\mathrm{erf}(\cdot)$	Error function

$\operatorname{erfi}(\cdot)$	Imaginary error function
f_{cost}	Objective function: costs
f_{det}	Objective function: detection probability
f_{warn}	Objective function: early-warning time
$\tilde{f}$	Aggregated objective function over multiple realizations/scenarios
$T(x, y)$	Transformation
$P(x, y)$	Probability map
$U_i(\cdot)$	Individual non-linear utility function
$\mathrm{W}_j(\cdot)$	Lambert W function

Spaces and Sets

AS	Approximation set
$\mathbb{E}$	Linear set used as a surrogate for the EDO set
$\mathbf{D}$	Space of permitted solutions
PF	Pareto front
RS	Reference set

Abbreviations

ACO	Ant Colony Optimization
ADE	Advection-Dispersion Equation
cdf	cumulative distribution function
CV	Control Volume
DP	Detection Probability
DE	Differential Evolution
DVGW	German Technical and Scientific Association for Gas and Water
EA	Evolutionary Algorithm
EDO	Elementary Decision Option
EWT	Early-Warning Time
FDM	Finite Difference Method
FEM	Finite Element Method
IG	Inverse Gaussian Distribution
LIR	Linear-Indexing Representation
LP	Linear Programming
LW	Zweckverband Landeswasserversorgung
MC	Monte Carlo
MOEA	Multi-Objective Evolutionary Algorithm
NFE	Number of Function Evaluations
NLP	Nonlinear Programming

PCX	Parent-Centric Crossover
pdf	probability density function
PM	Polynomial Mutation
PSO	Particle Swarm Optimization
PT	Particle Tracking
PTRW	Particle Tracking Random Walk
RPTRW	Reverse Particle Tracking Random Walk
SBX	Simulated Binary Crossover
SPX	Simplex Crossover
UM	Uniform Mutation
UNDX	Unimodal Normal Distribution Crossover
U_Protect	Synthetic test case of a groundwater model
WHO	World Health Organization
Z_Based	Synthetic test case of a groundwater model

Abstract

Motivation and Goal

Safe drinking water is not only a necessary resource for human life but also a human right according to a declaration of the United Nations in 2010 [163]. Hence, to protect one of the largest sources of freshwater, the groundwater, many national guidelines and regulations exist [e.g., 46, 164]. Even in the European Union, where most of the people have access to high-quality freshwater, the politics is still discussing improving the groundwater-protection guidelines [53]. This is in accordance with the World Health Organization (WHO), who claims that also in industrial countries human health is at severe risk because of contaminated water [176]. To control this risk, the WHO suggests water safety plans [30], with monitoring as one of the key aspects. This thesis contributes to an optimization framework for reliable early-warning monitoring systems in drinking-water well catchments for the sake of risk control.

Unfortunately, well catchments of water suppliers include many possible threats (for example gas stations) that require monitoring. Ideally, monitoring can fulfill three objectives in this context: it warns of contamination events (1) reliably, (2) as early as possible, and (3) at low costs. These three objectives, however, are impossible to achieve simultaneously in reality. Instead, many trade-off networks exist that are also called Pareto optimal. These networks contain a compromise between the conflicting objectives and cannot be improved in one of the objectives without sacrificing at least one other objective. The conflict between the objectives makes the optimal-design problem of groundwater monitoring networks challenging, and the complexity even increases when also considering uncertainty in system dynamics and hydrogeological context of the groundwater aquifer.

The overarching goal of this thesis is to contribute to a safe water supply. To achieve this goal, I provide a framework that optimizes monitoring networks in drinking-water well catchments with respect to the three competing objectives introduced above. I approach this goal from different directions leading to the four main contributions introduced below.

Contributions and Conclusions

Formulation of the Optimization Problem First, I formulate the multi-objective optimization problem and develop the corresponding objective functions that mathematically express the objectives mentioned above to enable their quantification and

evaluation. To benchmark my contributions, I develop two test cases: the first abstracts typically used models from water-supply companies, and the second captures key complexities for real-world source water protection in groundwater-based water supplies.

The key conclusion from this part is that multi-objective optimization is an appropriate approach to tackle the monitoring-network optimization problem. It provides detailed insights into the trade-offs between the considered objectives and provides significant added value for decision makers.

Enhancing Performance of the Optimization Algorithm The optimization problem posed above can easily become a massive computational and algorithmic challenge. Therefore, the second contribution tackles performance problems of large, complex, discrete multi-objective optimization problems. The problem is that the often used binary representation of search spaces for discrete optimization problems is limited to a relatively small number of decision-relevant variables. Optimization problems with a large number of decision variables, however, suffer from performance problems in search speed and quality.

Therefore, I develop:

1. a search-space representation that increases the reliability of the optimization, as well as the search speed (efficiency) and search quality (effectiveness).
2. a search-space reduction method that provides efficient and effective search for Pareto-optimal solutions by dramatically reducing the size of the search space.

I found that especially a proper search-space representation during the optimization is key for an enhanced optimization. These findings carry over to a much wider class of optimization problems than the monitoring problem featured in my thesis.

Simplification of Methods for Practical Application In the field of hydro(geo)logy, there is a substantial gap between latest academic developments and practice. Thus, aiming of technology transfer, I develop strategies to apply academic concepts in practice such that they are frugal enough to run on standard personal computers, and that they can be operated with realistically available data. The developed concepts include, among others, performance improvements regarding the simulation of possible groundwater contamination and uncertainty representation, the risk assessment of the possible threats in well catchments, and the modeling of unknown contamination threats.

The main conclusion over these contributions is that it is possible to replace data-hungry complex statistical methods with frugal strategies that are simple enough for

practical application but are still scientifically rigorous. In fact, this effort has enabled a collaboration project with a group of water supply companies, where the developed methods were applied successfully.

Developing Analytical Solutions for Uncertainty Quantification The above three contributions all assume that a numerical simulation model for the well catchment is available, which is not always true. Therefore, I investigate the physical mechanisms of groundwater contaminant transport and their uncertainties that control the optimal placement of monitoring wells. I develop analytical solutions that help (1) understand optimization results given parameter uncertainty, and (2) analyze the effects of the controlling factors and their uncertainty on the optimal monitoring-well placement.

The solutions are provided in closed form, and hence can easily be used for uncertainty analyses and decision support regarding monitoring networks even when no numerical model is available.

Practical Application Finally, for the sake of demonstration, I apply most of the developed methods from the four contributions introduced above on a real case to show that a transfer from academic methods to practical application is possible. Here, I also introduce some communication strategies that help simplify the collaboration between academia and practice.

Kurzfassung

Motivation und Ziel

Sauberes Trinkwasser ist nicht nur eine notwendige Lebensgrundlage für den Menschen. Nach einer UN-Vereinbarung von 2010 ist es auch ein menschliches Recht Zugang zu sauberen Trinkwasser zu haben [163]. Um eine der größten Frischwasserspeicher (das Grundwasser) vor Kontaminationen zu schützen, existieren viele nationale Richtlinie und Gesetze [z. B., 46, 164]. Selbst in der Europäischen Union wird über die Verbesserung des Trinkwasserschutzes diskutiert [53], obwohl fast alle Bewohner Zugang zu sauberem Trinkwasser haben. Das passt zu der Einschätzung der World Health Organization (WHO), die besagt, dass sogar in Industrieländern ein hohes Risiko für die Gesundheit der Menschen durch verschmutztes Trinkwasser besteht [176]. Um das Risiko zu kontrollieren, empfiehlt die WHO Trinkwassersicherheitskonzepte [30], in denen die Überwachung des Risikos zur Kontrolle eine wichtige Rolle spielt. In dieser Arbeit entwickle ich Strategien und Methoden zur Optimierung eines zuverlässigen Frühwarnsystems in Trinkwassereinzugsgebieten zur effektive Risikokontrolle.

Leider gibt es innerhalb der Trinkwassereinzugsgebiete von Wasserversorgungsunternehmen viele mögliche Gefährdungsquellen für das Grundwasser (z. B. Tankstellen), die mit Grundwassermessstellen bzw. Messnetzten überwacht werden sollten. Solche Messnetze erfüllen idealerweise drei Hauptziele: sie warnen den Wasserversorger vor einer Grundwasserverschmutzung (1) zuverlässig, (2) so früh wie möglich, und (3) zu geringen Kosten. Diese drei Ziele sind in der Realität allerdings nicht gleichzeitig erfüllbar. Stattdessen gibt es viele Netzwerkdesigns, die alle unterschiedliche Kompromisse zwischen den drei Zielen abbilden. Diese Kompromisslösungen werden Pareto-optimal genannt und können in keinem Ziel verbessert werden, ohne dass sich gleichzeitig mindestens ein anderes Ziel verschlechtert. Der Konflikt zwischen den Zielen macht dieses Optimierungsproblem zu einer Herausforderung. Die Komplexität steigt noch einmal, wenn das zugrundeliegende System Unsicherheiten in Dynamik und im hydrogeologischem Kontext enthält.

Das große Ziel dieser Arbeit ist es einen Beitrag zum Trinkwasserschutz zu leisten. Um dieses Ziel zu erreichen, entwickle ich Strategien und Methoden, mit denen Messstellennetze in Trinkwassereinzugsgebieten auf der Basis der oben vorgestellten drei Zielkriterien optimiert werden können. Dabei nähere ich mich dem Ziel dieser Arbeit von verschiedenen Richtungen, die zu vier Hauptentwicklungen führen. Diese stelle ich im Folgenden kurz vor.

Beiträge und Schlussfolgerungen

Formulierung des Optimierungsproblems Zuerst formuliere ich das multikriterielle Optimierungsproblem und stelle die zugehörigen Zielwertfunktionen auf. Die Zielwertfunktionen drücken die drei oben vorgestellten Zielkriterien mathematisch aus, um sie quantifizieren und bewerten zu können. Zum Testen und Überprüfen der in dieser Arbeit entwickelten Strategien und Methoden entwickle ich zwei Testfälle. Der Erste ist an typische Grundwassermodelle von Wasserversorgungsunternehmen angelehnt und der Zweite bildet Schlüsselkomplexitäten von Transportsimulation in Grundwasser ab.

Es zeigt sich, dass die multikriterielle Optimierung ein geeigneter Ansatz für das vorliegende Optimierungsproblem ist. Sie bietet einen guten Überblick über die Kompromisse zwischen den betrachteten Zielen und bietet Entscheidungsträgern somit eine gute Grundlage für Entscheidungsprozesse.

Performancesteigerung des Optimierungsalgorithmus Das oben vorgestellte Optimierungsproblem wird schnell zu einer rechenintensiven und algorithmischen Herausforderung werden. Deshalb beschäftige ich mich im zweiten wissenschaftlichen Beitrag mit der Performance der Optimierung bezüglich großer, komplexer, diskreter multikriterieller Optimierungsprobleme. Das Problem ist, dass die häufig gewählte binäre Repräsentation von diskreten Suchräumen auf eine relativ kleine Anzahl von entscheidungsrelevanten Variablen begrenzt ist. Optimierungsprobleme mit einer großen Anzahl von Entscheidungsvariablen haben daher eine schlechte Leistungsfähigkeit hinsichtlich Suchgeschwindigkeit und Qualität.

Deshalb entwickle ich:

1. eine Suchraumdarstellung, die effizientes und zuverlässiges Optimieren ermöglicht.
2. eine Methode zur Reduktion von Suchräumen, die die Suche nach Paretooptimalen Lösungen effizienter und effektiver macht, indem sie den Suchraum signifikant verkleinert.

Daraus lässt sich schließen, dass während der Optimierung eine gut ausgewählte Suchraumdarstellung der Schlüssel für eine leistungsmäßig starke Optimierung ist. Das verbessert signifikant die Zuverlässigkeit der Optimierung, steigert die Suchgeschwindigkeit (erhöhte Effizienz) und die Suchqualität (gesteigerte Effektivität). Diese Ergebnisse lassen sich auf eine große Klasse von Optimierungsproblemen übertragen und gelten nicht nur für das Messnetze-Optimierungsproblem aus dieser Arbeit.

Vereinfachung von Methoden für den praktischen Gebrauch Im Bereich der Hydro(geo)logie gibt es eine erhebliche Diskrepanz zwischen den aktuellen akademischen Entwicklungen und Methoden aus der Praxis. Abzielend auf den Wissenstransfer, entwickle ich in meinem dritten Beitrag Strategien und Methoden, die rechenintensive und datenhungrige akademische Konzepte vereinfachen, so dass sie in der Praxis anwendbar sind. Die vereinfachten Methoden müssen dabei so wenig Rechenleistung benötigen, dass sie auf einem marktüblichen Computer laufen können, ohne wissenschaftlich inkorrekt zu sein. Die behandelten Konzepte umfassen unter anderem die Simulation von möglichen Grundwasserverschmutzungen, Unsicherheitsbewertung, Risikobewertung von möglichen Verschmutzungsquellen in Brunneinzugsgebieten, und die Modellierung von unbekannten Verschmutzungsquellen.

Die wichtigste Schlussfolgerung ist, dass es möglich ist, komplexe statistische Methoden, die viele Daten benötigen, durch einfache Strategien zu ersetzten, die einfach genug für die praktische Anwendung sind. Gleichzeitig genügen sie trotzdem wissenschaftlich sauberen Ansprüchen. Die Ergebnisse konnte ich in einem gemeinsamen Projekt mit einer Gruppe von Wasserversorgungsunternehmen erfolgreich anwenden.

Entwicklung von analytischen Lösungen für Unsicherheitsquantifizierung Die drei oben genannten Beiträge gehen alle davon aus, dass ein numerisches Simulationsmodel des Brunneneinzugsgebietes zur Verfügung steht. Diese Annahme ist nicht immer wahr. Deshalb untersuche ich im vierten Beitrag die physikalischen Mechanismen des Grundwasserschadstofftransports und deren Unsicherheiten, die die optimale Platzierung einer Grundwassermessstelle kontrollieren. Hier entwickle ich analytische Lösungen, die helfen (1) Optimierungsergebnisse bezüglich Platzierung von Messstellen unter Parameterunsicherheit zu verstehen, und (2) die Effekte der kontrollierenden Faktoren und deren Unsicherheiten auf die optimale Platzierung von Grundwassermessstellen zu analysieren.

Die entwickelten Lösungen liegen in geschlossener Form vor und können dadurch einfach für Unsicherheitsanalysen verwendet werden und den Entscheidungsprozess bezüglich optimaler Messstellenpositionierung unterstützen, auch wenn kein numerisches Modell vorliegt.

Praktische Anwendung Zum Abschluss wende ich die entwickelten Methoden aus den oben vier vorgestellten Beiträgen zu Demonstrationszwecken an einem Problem aus der Realität an. Hiermit zeige ich, dass der Transfer von Wissenschaft zur Praxis möglich ist. Außerdem stelle ich einige mögliche Kommunikationsstrategien vor, die eine Zusammenarbeit in Projekten zwischen Wissenschaft und Praxis vereinfachen können.

Chapter 1

Introduction

1.1 Motivation

Clean drinking water is an essential resource and the foundation of life. While many regions worldwide suffer from water shortage, other regions (e.g., Denmark, or Germany) still seem to have an almost unlimited amount of water, such that water quality becomes the major consideration for securing the required freshwater resources. Since ninety-seven percent of the world's usable freshwater is stored in groundwater [141], aquifers are in a special focus of resource protection. In Germany and many other countries, defined water protection zones are common protection measures for drinking-water well catchments and their wellheads. Such zones restrict land-use activities and are specified in national guidelines and regulations [e.g., 46, 164, 173].

However, these restrictions typically cannot remove all possible threats (called *possible contamination sources* in the following) from the well catchments for several reasons:

- Political reasons: Land-use restrictions are trade-offs between conflicting activities that relate to economy, ecology, society, and public/private land-ownership rights. The actual political priorities determine, which activities will be restricted and/or regulated.
- Historical reasons: In centuries-old cities that fall in relatively newly declared protection zones, it is impossible to strictly fulfill all regulations of groundwater protection, although especially urban areas include many possible contamination sources.
- Unknown sources: Possible contamination sources that are unknown in location and existence cannot be removed.
- Uncertainty: The actual outline of the well catchment is affected by changing hydrological conditions. Even if the protection zones are planned in consideration of different hydraulic scenarios, unforeseen geological, hydrological, or hydrogeological conditions might change the actual outline of the well catchment in an unconsidered way.

Thus, very often there is an entire inventory of possible contamination sources in well catchments that puts the groundwater and drinking-water production wells at risk. Hence the risk needs to be assessed and controlled.

To control the risk for groundwater, the World Health Organization (WHO) proposed three-step Water Safety Plans [30]:

1. knowing all relevant possible contamination sources,
2. identifying measures to control these possible sources, and
3. ensuring that they are in fact controlled.

Transposing these steps to well catchments leads to

1. an identification of all possible contamination sources and to assess their risk within well catchments,
2. the installation of monitoring networks to track the groundwater quality prior to extraction, and
3. the actual operation of monitoring networks.

Therefore, the *German Technical and Scientific Association for Gas and Water* (DVGW) recommend in their technical standards the installation of monitoring networks for well catchments [45], and most well catchments are already equipped with monitoring networks. However, the technical standards provide only vague guidance about how these networks should be designed. Specifically, the regulations suggest that numerical groundwater models should be used to plan monitoring networks, but they do not give information about how these models can or should be used to optimize them, not to mention that the regulations do not define 'optimality'. Furthermore, most of the existing networks grew historically. Their individual monitoring wells follow diverse purposes, e.g., monitoring groundwater levels or contamination plumes of existing sources. Hence, they are often inadequate or sub-optimal for rigorously controlling the risk that emanates from the inventory of possible contamination sources. To summarize, water suppliers do not have a tool that helps to optimally (re-)design monitoring networks within their well catchments.

Consequentially, the overall goal of my thesis is to develop and to provide a framework for water suppliers that helps fulfill the three steps of the WHO Water Safety Plans mentioned above. That is, to support the identification and the risk assessment of possible contamination sources within well catchments, to (re-)design optimal monitoring networks, and to reduce the overall costs of such networks for an economic operation. The work towards these goals was supported by the DVGW through a joint project over two years called *Risk-Based Groundwater Monitoring for Well Head Protection Areas* that additionally included water suppliers and stakeholders. The idea of the project was to provide clear guidance in monitoring-network design for water suppliers, considering the demands of the regulations and the needs of the water suppliers. This project strongly influenced the direction of this thesis, because all developed concepts and methods had to be scientifically rigorous, but also applicable in practice on

standard desktop computers. That is, all methods are required to be operable with the expertise and data available in water supply companies, and they have to be computationally efficient. In the following section, I will introduce the problem in detail, and formulate clearly defined research questions and the corresponding approaches.

1.2 Research Questions and Approaches

The research questions and approaches that follow from the motivation and the goal of my work can be categorized into the three general groups *Optimization*, *Transport Simulation*, and *Uncertainty*. In total, seven research questions result from the problem, enumerated below within the three categories.

Optimization

Within the DVGW project introduced above, we defined three main objectives to be fulfilled by an ideal monitoring network in well catchments: (1) a maximum detection probability of contamination from possible contamination sources for a reliable risk control, (2) a maximum early-warning time (early detection of spilled contaminants) to increase the reaction time of water suppliers to install countermeasures, and (3) minimum installation and operation costs of the network.

It is apparent that under fiscal restrictions (third objective) the monitoring-network goals are competing and cannot be fulfilled simultaneously. Figure 1.1 exemplary shows two scenarios of possible groundwater contamination and different possible monitoring networks (labeling can be found in the figure). The color gradient from green (early warning) to red (late detection) qualitatively describes the decrease of early-warning time over distance from possible contamination source to pumping well. For both scenarios, perfect knowledge of the possible contamination plumes from the possible contamination is assumed. Scenario (A) considers only one single possible contamination source. The black squared monitoring well provides maximum detection probability (built within the detectable portion of the plume), a maximum early-warning time (it is close to the contamination source), and low costs (a single monitoring well is sufficient). A different solution with no costs would use the pumping well as monitoring well. This solution would provide maximum detection probability, but minimum early-warning time. Hence, the monitoring network could not be used as early-warning system and would be meaningless for water providers.

Differently to scenario (A), scenario (B) considers three possible contamination sources and all must be detected for a maximum detection probability. Two possible solutions that achieve maximum detection probability are (1) an expensive monitoring network

including the white and blue monitoring wells, or (2) the less expensive monitoring network with a single green monitoring well that monitors all three sources simultaneously in one point. Drawback of the single monitoring well is the lower early-warning time due to the late detection of all three sources. Less expensive trade-off networks that do not seek for maximum detection probability can increase early-warning time for selected possible contamination sources by shifting their focus to selected sources and ignoring the others. Thus, the selected sources can be detected as early as possible (e.g., white monitoring well, or both blue monitoring wells).

In conclusion, the three objectives are conflicting and it is impossible to find the perfect monitoring network that completely satisfies all three objectives simultaneously, i.e., a maximum detection probability, a maximum early-warning time, but no costs. Rather, many trade-off solutions can be found with a different balance between the three objectives. For finding these trade-off solutions, formal optimization can be applied to the problem. Due to the conflicting objectives, the general optimization approach is to formulate the problem as a multi-objective optimization problem and to translate the objectives in quantifiable objective functions. For large and complex problems, however, the optimization might need long wall-clock times to find (eventually) satisfying solutions. Consequently, the first two research questions are:

1. How can I formulate the optimization problem for an effective optimization?

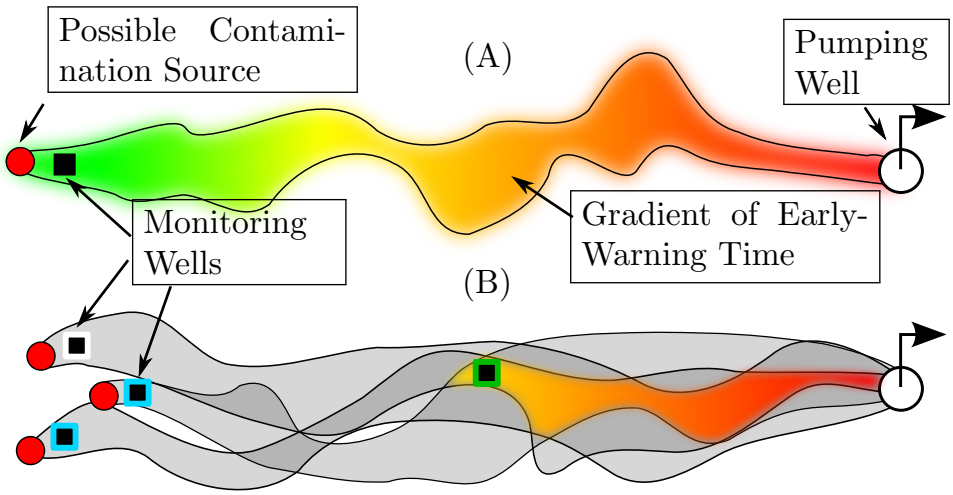


Figure 1.1: Two different scenarios of possible groundwater contamination. Scenario (A): A single possible contamination source can be monitored by a single monitoring well that should be close to the possible contamination source for early warning. Scenario (B): Multiple possible contamination sources need monitoring. Three monitoring networks (green countered square, white and blue contoured squares, white contoured square) satisfy the three objectives to different degrees that cannot be fulfilled simultaneously.

2. How can I speed up the optimization (improve efficiency), improve optimization quality (increase effectiveness), and minimize quality variability of optimization results (increase reliability)?

Transport Simulation

Irrespective of the optimization results, the optimization is based on data from numerical transport simulations. Developing the corresponding transport models is difficult and time-consuming. First, transport relevant model parameters and boundary conditions are subject to uncertainty (see next category below). Second, the possible contamination sources form the source terms of transport simulation. Hence, water suppliers need to collect a lot of information from their well catchments as model input data. Here, the most challenging part is to identify all relevant possible contamination sources and to prioritize them according to their risk for the production wells. This prioritization is important for the monitoring-well design that should especially monitor the most dangerous possible contamination sources. While the identification of possible sources is mostly laborious work, their risk quantification is often impossible. Too many data are unknown, e.g., the probability of failure, or the type and amount of harmful substances that are stored at the location of question. Therefore, the next two research questions are:

3. How can I support the data collection of the water suppliers and help to distinguish between relevant and irrelevant possible contamination sources?
4. How can water suppliers prioritize the relevant possible contamination sources according to their risk without the data required for a quantitative risk assessment?

Uncertainty

In addition to the challenges introduced above, finding optimal positions of monitoring wells is also challenging because various parameters influence the reliability and optimality of a suggested monitoring location and are often subject to uncertainty:

- There may be uncertainty in hydro(geo)logy, i.e., in global and local velocity fields, which can vary in angle and absolute value.
- There may be uncertainty in the exact position of possible contamination sources, or even worse, there may be possible contamination sources that are unknown in existence, hence in location.
- There may be uncertainty in transport-relevant parameters that describe, e.g., dispersion and decay.

- There may be uncertainty in the spilled mass of contamination when a possible contamination source fails.

The resulting research questions are:

5. How can I consider in a computationally feasible way the uncertainty in hydro(geo)logy, transport-relevant parameters, and location of possible contamination sources to get optimization results that are robust against uncertainty?
6. How can I consider unknown possible contamination sources?
7. How can I verify optimization results under uncertainty?

Summarized, in this thesis, I investigate and present a framework that enables water suppliers to optimize their monitoring networks according to the three objectives detection probability, early-warning time, and costs on standard desktop computers. The key approaches cover

- the multi-objective optimization problem formulation,
- the data collection and transport simulation,
- and uncertainty representation for a robust optimization.

1.3 State of the Art

The general field of groundwater monitoring is huge and research is going in diverse directions. Two main categories are the monitoring of groundwater levels and the monitoring of groundwater quality. In this thesis, the second category is relevant and can further be divided into two classes:

1. Monitoring of existing groundwater contamination, and
2. monitoring of potential groundwater threats to safeguard groundwater quality.

For the first class of monitoring networks, the objectives can still be manifold. Typical tasks for existing plumes are to find the shape of the contamination plume [e.g., 9, 108] for better taking countermeasures, to estimate the contaminant discharge [e.g., 99, 145, 161] for a better risk-assessment of the contaminant plume, or to identify contaminant sources [e.g., 66, 119, 139] for permanent and effective counter measures. The second class of monitoring networks focuses on protecting source water for water-supply companies against possible future contamination events. Differently to the first class of monitoring networks, here the networks are typically large-scale and long-term groundwater monitoring networks (see review by Loaiciga et al. (1992) [106] and the comprehensive review in the work of Kollat et al. (2011) [96]).

As already said in Section 1.1, for reducing the risk evolving from these possible future contamination events, the most common measure is to restrict the land use within well protection zones [e.g., 46, 164]. However, with growing urban density and contentious competition over highly valued land uses, it is becoming increasingly difficult to characterize and manage source water well protection zones, especially in rapidly expanding cities. Additionally, there is uncertainty in the actual outline of the well catchment. Thus, there is always an inventory of possible contamination sources in the catchment that remains to be assessed and controlled. Accordingly, there is ample literature on capture zone delineation and its uncertainty [e.g., 80, 120, 153, 167], on aquifer vulnerability [e.g., 4, 44, 184], on (probabilistic) well vulnerability [e.g., 51, 52, 60] and on risk analysis [e.g., 28, 159]. While such works help to evaluate the possible contamination sources to which the production well is exposed, they are not yet helpful in controlling them.

These contamination detection problems are also closely related to early work focused on monitoring of landfills [e.g., 111, 112, 117, 118, 154]. Although early work such as Massmann and Freeze (1987a,1987b) [111, 112] performed a single-objective risk-cost-benefit analysis to optimize and evaluate the quality of monitoring networks, there is a growing trend in the monitoring literature towards formulations that consider multiple objectives and multi-objective optimization. Commonly employed objectives include maximizing the detection probability, minimizing costs, and minimizing the contaminated area, or the volume of contaminated groundwater, respectively, [e.g., 118, 154, 179]. The last of these objectives is indirectly related to the objective of early detection of a contamination that would provide benefits in remediation costs and resource protection. In my approach, however, an early detection gains benefits in the reaction time of water suppliers to install countermeasures for well-head protection of the production wells. The monitoring network is used as early-warning system. Monitoring networks as early warning systems are often associated with disaster management for natural hazards such as tsunami or earthquake warning systems, e.g., Allen and Kanamori (2003) [3]. Alternatively, there is research on detection sensor networks in the signal processing literature [e.g., 115], but without existing applications to well catchments. Therefore, the major differences between my approach and the named studies are:

1. They do not consider the protection of groundwater wells against a whole inventory of possible contamination sources. Instead, they focus on monitoring of a single landfill.

2. My approach considers a significant number of candidate monitoring locations and therefore poses a severely challenging multi-objective combinatorial problem.

1.4 Contributions and Structure of the Work

This thesis is structured into eight chapters that can be divided into three parts:

Introduction and Background After the introduction in the current Chapter 1, Chapter 2 introduces the fundamentals and background of the following chapters. These chapters do not contain any novelties or contributions.

Contributions Chapters 3, 4, 5, and 6 form the main part of this thesis and contain my own contributions:

- In Chapter 3, I formulate the optimization problem and develop the objective functions. These objective functions relate to the efficiency, effectiveness, and robustness of the monitoring network. Here, I also introduce two benchmark test cases that I developed and that are used in the following chapters.
- In Chapter 4, I tackle performance problems of multi-objective optimization algorithms for large and complex, discrete multi-objective optimization problems. The focus is on search-space representation and search-space reduction to achieve efficiency, effectiveness, and reliability of the optimization process.
- In Chapter 5, I develop feasible methods regarding transport simulation, risk assessment, and robust multi-objective optimization for an improved search for optimal early-warning monitoring networks in practice. These methods illustrate the achievable transfer to practice.
- Finally, in Chapter 6, I introduce analytical solutions that help understand optimization results given parameter uncertainty.

Application and Summary In Chapter 7, I apply the optimization framework on a real case from a local water-supply company. I also discuss strategies for a successful transfer from academia to practice, lessons we have learned during the project I introduced in Chapter 1. The last chapter (Chapter 8) contains a summary and the main conclusions of this thesis. Finally, I briefly introduce possible future investigations evolving from the discussed methods in the previous chapters.

Relation to Published Works

This dissertation is based on several publications including Nowak et al. (2015) [126], Bode et al. (2016a) [19], Emmert et al. (2016) [48], Bode et al. (2017) [14], Bode

et al. (2018b) [15], Bode et al. (2018c) [21] (submitted), and Bode et al. (2018a) [13] (in preparation). Chapters 1, 2, and 8 may contain similar and/or identical formulations from these works, but do not contain any scientific novelties and contributions. Therefore, I omit a clear identification in these chapters. The remaining contributing chapters include general but unique references of the sources in the very beginning of each chapter.

I presented most of my results on a number of international conferences, including

- AGU 2013-2016: Bode et al. (2013, 2014c, 2015, 2016b) [10, 16, 22, 20]
- EGU 2015: Bode et al. (2015) [17]
- NGWA 2014: Bode et al. (2014a) [11]
- CMWR 2014: Bode et al. (2014b) [12]
- NUPUS 2013: Bode and Nowak (2013) [18]

Finally, this thesis played an important rule on the methodical level within in the joint research project *Risk-Based Groundwater Monitoring for Well Head Protection Areas* together with the DVGW and the related final report [67]. Following the project, the results should also be used to revise the technical guideline W108 [45].

Chapter 2

Fundamentals

In this chapter, I give a brief introduction to all fundamentals relevant for understanding this thesis. In Section 2.1 I introduce the governing equations of groundwater flow and transport. The following two sections introduce possible approaches to solve the governing equations, ie.e, analytical solutions in Sections 2.2 and numerical methods and established software in Section 2.3. Sections 2.4 and 2.5 are related to locations and levels of uncertainty within the scientific workflow of modeling and to uncertainty quantification. Section 2.6 gives a brief introduction to optimization and an overview of different optimization techniques with the emphasis on multi-objective optimization. Finally, Section 2.7 introduces metrics that help evaluate the performance of optimization algorithms.

2.1 Governing Equations of Flow and Transport in Porous Media

Steady-State Flow in Confined Aquifers

Steady-state groundwater flow in confined aquifers can be described by

$$-\nabla \cdot (\mathbf{K} \nabla h) = q_\mathrm{s} \quad \text{in } \Omega, \tag{2.1}$$

with hydraulic conductivity $\mathbf{K}$ $[\mathrm{L}/\mathrm{T}]$, hydraulic head h $[\mathrm{L}]$, and source and sink term q_s $\left[\mathrm{L}^2/\mathrm{T}\right]$ in the domain Ω [8]. Equation 2.1 is subjected to the general boundary conditions:

$$-(\mathbf{K} \nabla h) \cdot \mathbf{n} = \hat{q} \quad \text{on } \Gamma_1 \text{ and} \tag{2.2}$$

$$h = \hat{h} \quad \text{on } \Gamma \setminus \Gamma_1, \tag{2.3}$$

using the prescribed fluxes $\hat{q}$ and heads $\hat{h}$ on the Neumann boundary Γ_1 and on the Dirichlet boundary $\Gamma \setminus \Gamma_1$. The normal vector $\mathbf{n}$ points outwards on the domain. In confined aquifers, vertical flow is often negligible in relation to the horizontal flow. Then, the conductivity $\mathbf{K}$ can be substituted by the depth-integrated transmissivity $\mathbf{T} = \mathbf{K} \cdot B$, with B $[\mathrm{L}]$ as the aquifer thickness. Equation 2.1 adopts to

$$-\nabla \cdot (\mathbf{T} \nabla h) = q_\mathrm{s}\,, \tag{2.4}$$

and is also known as the two-dimensional, depth-averaged groundwater-flow equation.

Transport of Conservative Tracer

The advective-dispersive transport of conservative tracers is commonly described by the advection-dispersion equation (ADE):

$$\frac{\partial c}{\partial t} + \nabla \cdot (\mathbf{v}c - \mathbf{D}\nabla c) = 0 \quad \text{in } \Omega, \tag{2.5}$$

with concentration c $[\mathrm{M/L^3}]$, time t $[\mathrm{T}]$, velocity $\mathbf{v} = \frac{\mathbf{q}}{n_\mathrm{e}}$ $[\mathrm{L/T}]$ and Darcy velocity $\mathbf{q}$ $[\mathrm{L/T}]$, effective porosity n_e $[-]$ and hydromechanic (or macroscopic) dispersion tensor $\mathbf{D}$ $\left[\mathrm{L^2/T}\right]$ [143]:

$$\mathbf{D} = (\alpha_\mathrm{t}\|\mathbf{v}\| + D_\mathrm{m})\,\mathbf{I} + (\alpha_\ell - \alpha_\mathrm{t})\,\frac{\mathbf{v}\mathbf{v}^T}{\|\mathbf{v}\|}. \tag{2.6}$$

Here, α_ℓ and α_t (both $[\mathrm{L}]$) are the longitudinal and transverse dispersivities, D_m $\left[\mathrm{L^2/T}\right]$ is the molecular diffusion coefficient, and $\mathbf{I}$ is the identity matrix. The boundary conditions for Equation 2.5 are given by:

$$-\mathbf{n}\cdot\mathbf{v}c + \mathbf{n}\cdot(\mathbf{D}\nabla c) = \hat{J} \quad \text{on } \Gamma_2 \text{ and} \tag{2.7}$$

$$c = \hat{c} \quad \text{on } \Gamma \setminus \Gamma_2, \tag{2.8}$$

with $\hat{J}$ as a prescribed normal flux density and $\hat{c}$ as prescribed concentrations. Under steady-state conditions, the first term in Equation 2.5 becomes 0 and the advective-dispersive transport is described by:

$$\nabla \cdot (\mathbf{v}c - \mathbf{D}\nabla c) = 0\,. \tag{2.9}$$

2.2 Relevant Analytical Solutions

Spatial Concentration Distribution

For parallel flow with a uniform dispersion coefficient $\mathbf{D}$ and uniform velocity $\mathbf{v}$, Equation 2.5 can be written as

$$\frac{\partial c}{\partial t} + \mathbf{v}\cdot\nabla c - \nabla\cdot(\mathbf{D}\nabla c) = 0\,. \tag{2.10}$$

For a continuous point-like injection of a conservative tracer, a two-dimensional analytical steady-state solution for Equation 2.10 is given by

$$c\left(x,y\right)=\frac{\dot{m}_{\mathrm{in}}}{B}\frac{1}{\sqrt{4\pi vxD_{\mathrm{T}}}}\cdot\exp\left(-\frac{y^{2}v}{4D_{\mathrm{T}}x}\right). \tag{2.11}$$

Here, the longitudinal dispersion D_{L} $\left[\mathrm{L}^2/\mathrm{T}\right]$ was neglected, because the point source is a continuous injection. Then, the longitudinal dispersion plays only a minor role for the concentration distribution in space [e.g., 73]. For the sake of completeness, a derivation of Equation 2.11 can be found in Appendix A.1. Equation 2.11 provides an expression for the aqueous contaminant concentration c $\left[\mathrm{M}/\mathrm{L}^3\right]$ at any point (x,y) at steady state due to a continuous injection with mass discharge $\dot{m}_{\mathrm{in}}$ $\left[\mathrm{M}/\mathrm{T}\right]$, aquifer thickness B, groundwater-flow velocity v, and transverse dispersion coefficient D_{T} $\left[\mathrm{L}^2/\mathrm{T}\right]$.

Breakthrough Curve

For a simple groundwater model with uniform and parallel flow and uniform dispersion coefficient, the shape of a unit-mass breakthrough curve of a point-like and instantaneously injected aqueous contaminant mass can be calculated analytically and is given by the *inverse Gaussian distribution* (IG) [e.g., 57] (cf. Figure 2.1). For these groundwater models the assumption of Fickian transport is valid [e.g., 7], i.e., dispersion can be modeled as Brownian motion. The *probability density function* (pdf) of the inverse Gaussian distribution is given by

$$f\left(t;\mu,\lambda\right)=\left(\frac{\lambda}{2\pi t^{3}}\right)^{\frac{1}{2}}\cdot\exp\left(\frac{-\lambda\left(t-\mu\right)^{2}}{2\mu^{2}t}\right), \tag{2.12}$$

for time $t>0$. It is fully described by the mean $\mu>0$ and the shape parameter $\lambda>0$. The shape parameter λ is a function of the mean and the variance σ^2 and can be calculated as $\lambda=\frac{\mu^3}{\sigma^2}$.

Accordingly, the analytical shape of the breakthrough curve of a continuous point source is given by the *cumulative distribution function* (cdf) of Equation 2.12 (cf. Figure 2.1):

$$\begin{aligned}F\left(t;\mu,\lambda\right)=&\frac{1}{2}\left(1+\operatorname{erf}\left(\frac{\sqrt{\frac{\lambda}{t}}\left(\frac{t}{\mu}-1\right)}{\sqrt{2}}\right)\right)+\exp\left(\frac{2\lambda}{\mu}\right)\\&\frac{1}{2}\left(1+\operatorname{erf}\left(\frac{-\sqrt{\frac{\lambda}{t}}\left(\frac{t}{\mu}+1\right)}{\sqrt{2}}\right)\right)\end{aligned} \tag{2.13}$$

In applications, both functions have to be scaled to calculate the time-dependent concentration distribution $c\,(t)$. The corresponding scaling factor depends on the properties of the contaminant and the release conditions. A detailed explanation on the derivation of this scaling factor can be found in Section 5.1.

2.3 Numerical Solutions of Flow and Transport

Analytical solutions are only valid for strongly simplified problems. Realistic problems, however, are highly complex, hence groundwater flow and transport (Equations 2.1 and 2.5) cannot be solved analytically anymore. Instead, numerical simulation techniques are used.

Numerical methods can roughly be distinguished between *Eulerian* and *Lagrangian*. Eulerian methods apply mass balance on fixed control-volumes to approximate the unknown quantity (e.g., flow velocity of water). Lagrangian methods are based on moving parcels, approximating the unknown quantity by tracking discrete parcels over space and time. Two well-known representatives of the Eulerian method are the finite difference method (FDM) [e.g., 149] and the finite element method (FEM) [e.g., 180]. Representatives of Lagrangian methods are the particle tracking method (PT) [e.g., 130] and the particle-tracking-random-walk method (PTRW) [e.g., 56, 84, 102, 140, 172]. In groundwater modeling, Eulerian methods are usually used to approximate groundwater flow and complex multiphase multicomponent transport processes. Lagrangian

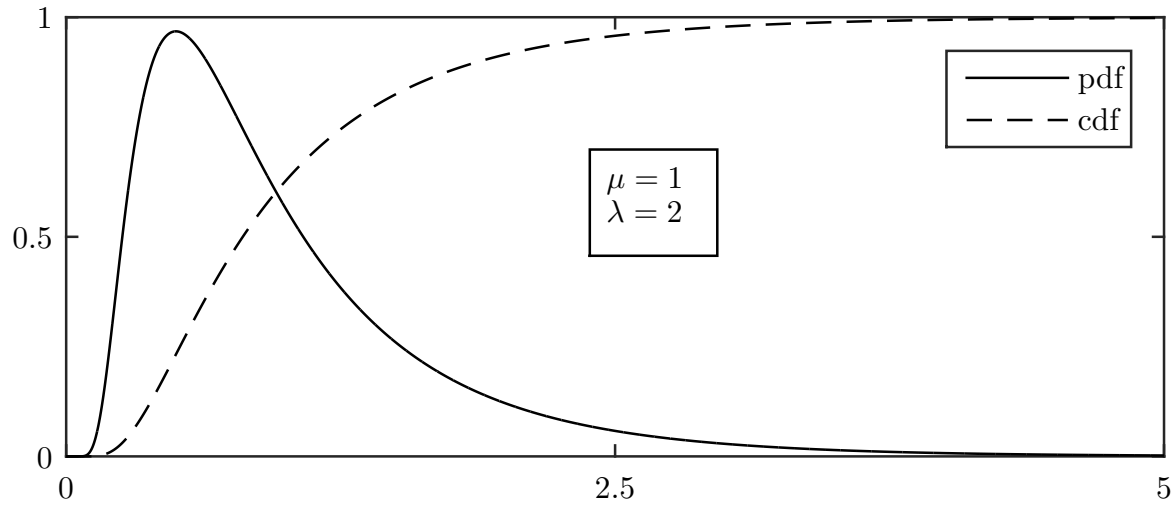


Figure 2.1: Inverse Gaussian distribution: Probability density function (pdf) as analytical solution for the shape of the breakthrough curve of an instantaneous point-like contaminant injection, and the cumulative distribution function (cdf) as solution for the continuous injection.

methods are used for simple transport processes.

In practice, two established software for developing and simulating numerical groundwater models are ModFlow [74] and FeFlow [39, 160]. Both software are widely spread among groundwater-related engineering companies and water suppliers. While ModFlow is based on a finite difference method to solve the groundwater flow equation, FeFlow calculates the velocity field using a finite element method. Most of the velocity fields I use in this thesis are created by an in-house FEM code as described in Nowak et al. (2008) [127].

Solving the flow equation with Eulerian methods (like FEM or FDM) is neither challenging nor computationally expensive. For transport simulation, however, these methods are very expensive and prone to oscillation and numerical dispersion. In contrast, the Lagrangian-based PTRW method is simple to implement, fast, robust, and free of numerical dispersion and oscillation [e.g., 140]. In the following section, I describe this method in detail.

Particle-Tracking-Random-Walk

Unlike particle tracking, PTRW is not a deterministic method as it includes a model to simulate the random diffusion/dispersion behavior of virtual solute particles in groundwater flow. This model assumes Fickian transport laws and, strictly spoken, is only valid on fully resolved velocity fields (i.e., at which the parameterization of dispersion through enhanced diffusion is valid).

The approximation of the steady-state ADE (Equation 2.9) is given by an ensemble of particles driven by advection and dispersion:

$$\mathbf{X}_\mathrm{p}\left(t+\Delta t\right) = \mathbf{X}_\mathrm{p}\left(t\right) + \Delta\mathbf{X}_\mathrm{p}\left(t+\Delta t\right) , \tag{2.14}$$

with the particle position $\mathbf{X}_\mathrm{p}$ [L], the particle displacement $\Delta\mathbf{X}_\mathrm{p}$ [L], time t [T] and time discretization Δt [T]. The particle shift $\Delta\mathbf{X}_\mathrm{p}$ is obtained by

$$\Delta\mathbf{X}_\mathrm{p}\left(t+\Delta t\right) = \mathbf{u}^{*}\Delta t + \mathbf{B}\left(\mathbf{X}_\mathrm{p}, t\right)\sqrt{\Delta t}\cdot\boldsymbol{\xi}\left(t\right) , \tag{2.15}$$

with the deterministic drift $\mathbf{u}^{*} = \left(\mathbf{u}\left(\mathbf{X}_\mathrm{p}, t\right) + \nabla\cdot\mathbf{D}\left(\mathbf{X}_\mathrm{p}, t\right)\right)$ that is determined by the velocity $\mathbf{u}$ and the gradient of the dispersion tensor $\mathbf{D}$, and the stochastic term $\mathbf{B}\left(\mathbf{X}_\mathrm{p}, t\right)\sqrt{\Delta t}\cdot\boldsymbol{\xi}\left(t\right)$, with the displacement matrix $\mathbf{B}$ that scales a vector of standard normally distributed random variables $\boldsymbol{\xi}$. The displacement matrix $\mathbf{B}$ has to fulfill the condition $\mathbf{B}\cdot\mathbf{B}^{\mathrm{T}} = 2\mathbf{D}$. Following Salamon et al. (2006) [140], for three-dimensional problems $\mathbf{B}$ can be calculated as

$$\mathbf{B} = \begin{pmatrix} \frac{u_\mathrm{x}}{|\mathbf{u}|}w_\ell & -\frac{u_\mathrm{x}u_\mathrm{z}}{|\mathbf{u}|\sqrt{u_\mathrm{x}^2+u_\mathrm{y}^2}}w_\mathrm{t} & -\frac{u_\mathrm{y}}{\sqrt{u_\mathrm{x}^2+u_\mathrm{y}^2}}w_\mathrm{t} \\ \frac{u_\mathrm{y}}{|\mathbf{u}|}w_\ell & -\frac{u_\mathrm{y}u_\mathrm{z}}{|\mathbf{u}|\sqrt{u_\mathrm{x}^2+u_\mathrm{y}^2}}w_\mathrm{t} & \frac{u_\mathrm{x}}{\sqrt{u_\mathrm{x}^2+u_\mathrm{y}^2}}w_\mathrm{t} \\ \frac{u_\mathrm{z}}{|\mathbf{u}|}w_\ell & \frac{\sqrt{u_\mathrm{x}^2+u_\mathrm{y}^2}}{|\mathbf{u}|}w_\mathrm{t} & 0 \end{pmatrix} , \tag{2.16}$$

with $w_\ell = \sqrt{2\left(\alpha_\ell \left|\mathbf{u}\right| + D_\mathrm{m}\right)}$ and $w_\mathrm{t} = \sqrt{2\left(\alpha_\mathrm{t} \left|\mathbf{u}\right| + D_\mathrm{m}\right)}$.

Finally, at each location, the temporal evolution of particle densities can be used to estimate breakthrough curves. If the number of particles is large enough, it can be done by direct density estimation. Otherwise, if the number of particles is too small for a direct estimation, one can use a parametric density estimation [e.g., 49], e.g., by using Equations 2.12 and 2.13 introduced in Section 2.2.

Reverse Particle-Tracking-Random-Walk

Reverse particle-tracking-random-walk (RPTRW) is a concept that enables, for instance, to localize the catchment of drinking-water wells with a single simulation. The basic idea behind this concept is to reverse the entire velocity field and inject a unit mass at the pumping wells [e.g., 60, 100, 121]. Then, transport is solved reversely by applying the PTRW method. The boundary conditions have to be adjusted properly, based on the boundary conditions of the forward transport simulation. For both methods (PTRW and RPTRW) I used an already existing code also used in Koch and Nowak (2014, 2015, 2016) [90, 91, 92] and Bode et al. (2016a) [19].

2.4 Sources of Uncertainty Considered in this Work

Decision-supporting models are usually affected by many sources of uncertainty. Walker et al. (2003) [169] generalized uncertainties according to their location in the modeling process and defined the following five types of uncertainty:

1. *Context uncertainty* refers to the problem and the purpose of the model. If problem and purpose are not defined clearly, the model outcome can be highly uncertain. For instance, a model for predicting water levels of a lake can neither be used to forecast water levels of a river nor to predict the damage caused by a flood.

2. *Model uncertainty* describes the uncertainty related (1) to the choice of the conceptual model, i.e., the mathematical equations, and (2) to their computational implementation.

3. *Parameter uncertainty* is associated with the data used for model calibration, as well as the calibration method itself. While some parameters are known with certainty (e.g., universal constants like π), unknown parameters are determined by calibration, hence are affected by measurement errors of the calibration data and the setup used for calibration (method, level of accuracy, grid resolution, etc.).

4. *Input uncertainty* refers to all data that describe the reference system. Input is affected by uncertainty (1) due to *external driving forces* that might underlie natural variabilities (e.g., the evaporation of water is related to solar radiation), and (2) due to a lack of knowledge about system relevant data (e.g., pumping rate of a groundwater extraction well).

5. *Model outcome uncertainty* describes the *prediction error* that is caused by the types of uncertainty introduced above.

In this work, I will only consider parameter and input uncertainty (in the following, I will address both as *parameter uncertainty*) for three reasons:

1. This work is strongly influenced by requirements and limitations of practical applications because results should be used by water suppliers. The models typically used by water suppliers have already been developed with substantial financial investments, and concepts and their mathematical implementation are mostly limited to the software that was decided to be used many years ago. Investigating model uncertainty for concepts that are not predefined in the models of the water suppliers (e.g., dispersion models) would be too time-consuming for the practical use.

2. In this work, model outcome uncertainty is dominated by parameter uncertainty. For instance, uncertainty in ambient flow direction causes large uncertainty in the location of a contaminant plume. Then, for water suppliers, it is more important to identify the possible contaminated area than possible tailing in contaminant breakthrough curves due to different dispersion models.

Therefore, context and model uncertainty is not within the scope of this work. By only considering parameter uncertainty, I assume that all used mathematical equations are sufficiently close to the real processes that govern the system so that models predict the real physical states of the systems reasonably well. The parameters I considered as uncertain are

- the permeability field (heterogeneity concepts: layering, zonations, and geostatistics),
- boundary conditions of flow and transport,
- sink and source terms, and
- transport relevant parameters.

A detailed description of how I considered uncertainty can be found in the relevant sections (Sections 3.5 and 6.3).

When discussing uncertainty, I also follow the terminology of Walker et al. (2003) [169] regarding the *uncertainty level*. They distinguish between three types of uncertainty

levels: *statistical uncertainty*, *scenario uncertainty*, and *recognized ignorance*. Statistical uncertainties can be described by probability distributions and hence are quantifiable. Scenario uncertainty does not have a known probability distribution, but its effects can be investigated by making assumptions about the uncertain parameters to create possible scenarios: scenario uncertainty can describe possibilities but without assigned probabilities. Recognized ignorance contains uncertainty that is known to exist but ignored because neither probability distributions are known nor the behavior such that useful assumptions can be made to express it as scenario uncertainty. These three uncertainty levels are bracketed by the two extrema of *determinism*, where everything is known, and *total ignorance*, where it is unknown what is unknown. With the term *uncertainty*, I refer to scenario uncertainty unless specified otherwise.

2.5 Uncertainty Quantification with Monte-Carlo Methods

In the field of uncertainty quantification, *Monte-Carlo (MC) methods* are often used to analyze the statistical distribution of a deterministic model outcome. While the idea of MC methods is dating back to the Babylonian era [72], Metropolis and Ulam (1949) [116] labeled it and established a systematic use. The basic idea of MC methods is to run a model multiple times with random but equally likely model input parameters and so approximate the distribution of the model outcome. Each set of input parameters is called *MC experiment* and the model outcomes are called *realizations*. Then, the distribution of realizations can be analyzed statistically. Many different MC methods exist, which mostly differ in the sampling rule of the model input parameters [e.g., 137]. The three main steps of MC methods are:

1. **Input parameter set**: Multiple sets of input parameters have to be generated, based on the distribution of the input parameters, and based on a specified sampling rule.
2. **Model run**: Each parameter set triggers a (deterministic) model run.
3. **Statistical analysis**: The set of realizations is evaluated via, e.g., histograms or probability density estimators.

MC methods are computationally expensive and they require a large number of realizations for a good approximation of the model outcome distribution. However, they are intuitive, easy to implement, and independent of the model.

2.6 Optimization

The following introduction to optimization problems is based on the work of Zitzler et al. (2004) [181] and on the book of Talbi (2009) [157]. In general, an optimization problem is the search for the best solution according to a certain criterion from all feasible solutions within the *search space* Ω. Mathematically, the criterion can be expressed as an *objective function* $\boldsymbol{f}$ that maps possible solutions from the search space Ω to the *objective space* $\mathbf{Y}$:

$$\boldsymbol{f} : \Omega \rightarrow \mathbf{Y} \,. \tag{2.17}$$

The outcome of $\boldsymbol{f}$ is an *objective vector* $\mathbf{y} = (y_1, y_2, \ldots, y_m) \in \boldsymbol{Y}$. The challenge is to find a solution $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \Omega$ such that $\boldsymbol{f}(\mathbf{x})$ is either minimal or maximal. In the following, all problems are defined as minimization problems. That is, the global optimum $\mathbf{x}^* \in \Omega$ satisfies the following constraint:

$$\forall \mathbf{x} \in \Omega \,, \boldsymbol{f}(\mathbf{x}^*) \leq \boldsymbol{f}(\mathbf{x}) \,. \tag{2.18}$$

In the case that $\mathbf{Y}$ is a subset of the real numbers ($\mathbf{Y} \subseteq \mathbb{R}$), the problem is a standard single-objective optimization problem with a scalar-valued objective function f. Although there might be more than one optimal solution, Equation 2.18 can be fulfilled because all optimal solution vectors $\mathbf{x}$ are mapped to the same scalar objective value y.

If the objective function $\boldsymbol{f}$ consists of multiple sub-functions f and $\mathbf{Y} \subseteq \mathbb{R}^m$ with $m > 1$, the problem is called a *multi-objective optimization problem*. Then, optimal solutions $\mathbf{x}$ might be mapped to different objective vectors $\mathbf{y}$ such that a global optimum is not necessarily existent. In this case, the comparison of two solutions $\mathbf{x}^1$ and $\mathbf{x}^2$ is more complex and follows the principle of *Pareto dominance* [128]: the assigned objective vector $\mathbf{y}^1$ *dominates* $\mathbf{y}^2$ if no component of $\mathbf{y}^1$ is greater (less good) than a component of $\mathbf{y}^2$ and at least one component of $\mathbf{y}^1$ is smaller (better). Mathematically, this dominance is expressed as $\mathbf{y}^1 \prec \mathbf{y}^2$. In this case, solution $\mathbf{x}^1$ is said to dominate solution $\mathbf{x}^2$. Solutions that are not dominated by any other solution $\mathbf{x} \in \Omega$ are called *non-dominated solutions* or *Pareto-optimal solutions*. The *Pareto set* contains all Pareto-optimal solutions and is a subset of the search space Ω. The *Pareto front* is the image of the Pareto-optimal solutions in the objective space and is a subset of the objective space $\mathbf{Y}$.

There are two main classes of optimization models to formulate and solve the optimization problems: *linear programming* (LP) and *nonlinear programming* (NLP). Further, these models can be divided into *continuous problems* and *discrete problems*, as well as *low-dimensional problems* and *high-dimensional problems*.

Linear programming is a model where the objective function, as well as the constraints of the optimization problem, are both linear. LP is relatively easy to solve and there are efficient and exact optimization algorithms (e.g., the simplex algorithm [29]). The beauty of LP is that the search space (determined by the constraints) and the objective

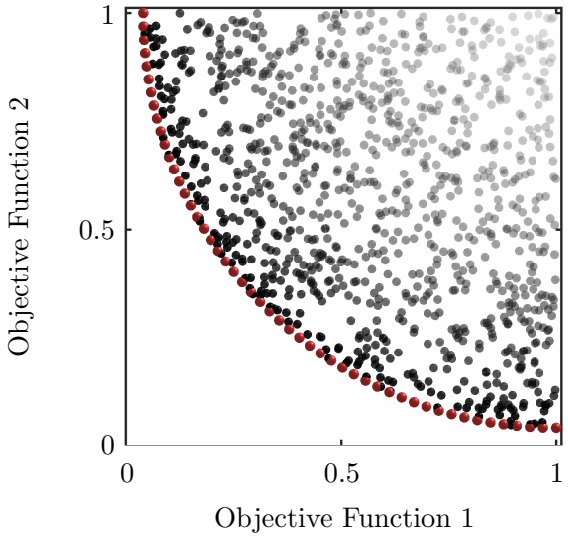


Figure 2.2: Two-dimensional Pareto-front for two objective functions. The Pareto front is a set of non-dominated solutions (red spheres) and dominates all solutions that are not Pareto optimal (dominated solutions in black/gray). The color of the dominated solutions changes from black to transparent gray regarding their distance towards the Pareto front.

function are convex. Hence, either the global optimum can be found on the contact point between objective function and the search space Ω, or the objective function is parallel to the edge of the search space Ω and each found optimum is a global optimum.

Problems with a nonlinear objective function and/or nonlinear constraints have to be solved with nonlinear programming. The search space Ω of NLP problems is a subset of $\mathbb{R}^n$ where n is the number of modifiable influencing parameters of the problem. For low-dimensional problems, n is small, while for high dimensional problems n is a large number. An example of a low-dimensional problem is the optimization of a flight trajectory of a cannonball: for covering the longest possible distance, one can only modifying the shooting angle $(n = 1)$. An example of high-dimensional problems is the optimization of a monitoring network with a large number of possible locations for monitoring wells.

Continuous optimization problems differ from discrete problems in the composition of the search space Ω. Continuous problems are related to continuous search spaces, i.e., the number of solutions is infinite. Discrete problems are often combinatorial problems and are based on a finite set of solutions, although the number of solutions might run towards infinity for large problems.

The choice of the best optimization algorithm to solve the problem depends on the

optimization problem itself (following the *No Free Lunch theorem* by Wolpert (1995, 1997) [177, 178]). For instance, efficient and exact algorithms exist for a continuous, quadratic, convex problem [e.g., in 125]. However, for NLP, it is often unknown whether a problem is convex or not. Thus, many local optima might exist and the search for a global optimum is computational very expensive. Then, so-called *meta-heuristics* come into play that do not aim to find the global optimum, but a sufficiently good approximation. In the following section, I will give a brief overview of common optimization methods with a focus on meta-heuristics.

2.6.1 Optimization Techniques and Optimization Algorithms

The easiest optimization problems are posed by differentiable functions with a closed-form expression of the derivative. To solve such problems, the first derivative locates the extrema and the second derivative defines whether these extrema are local maxima or minima To find the global optimum, local optima just have to be compared.

Functions that are differentiable but do not have a closed-form derivative cannot be optimized analytically. When it is known a priori that the problem has a global optimum but no local optima (e.g., in convex problems), *gradient-based optimization methods* are a good choice to apply. For gradient-based methods, it is not important to know the function to be optimized in an explicit way. However, the function value and the gradient at each point has to be calculable. Then, these methods start from an arbitrary point and always go in the direction of the negative slope.

These methods can also be applied to functions with local optima (e.g., non-convex problems) but either the algorithms need modifications for being able to leave a local optimum, or the search for the global optimum is done multiple times with random starting points. As a result, the solution time for non-convex problems can be very time consuming, especially for large problems. One of the most famous algorithms is the *Gradient Descent* [e.g., in 107, here labeled as Gradient Ascent], which only uses the information of the first derivative. Modifying the Gradient Descent method using also the second derivative results in the well-known *Newton Method* [e.g., 38].

Large-size non-convex problems can be tackled by meta-heuristics. Although there is no guarantee that they will find the global optimum, it has already been shown that they are very efficient and effective. That is, they approximate the global optimum in an acceptable time. Many classifications of meta-heuristics exist, e.g., deterministic and stochastic search, or population-based and single-solution based search.

The *Simulated Annealing algorithm* [24, 85] can be classified as a stochastic and single-solution based search. Stochastic, because the algorithm uses partially randomized decisions in its rules towards the optimal solution, and single-solution based, because

at the same time the algorithm considers only one possible solution. Simulating Annealing is based on the so-called *Hill-Climbing Algorithm*, which tests solutions around the current solution and replaces the current solution when a better one is found in the neighborhood. Simulated Annealing differs from the Hill-Climbing algorithm at the replacement step of the current solution: Simulated Annealing always replaces the current solution if the new solution is better. If the new solution is worse, the current solution will be replaced with a certain probability that depends on the difference between the current and the new solution, and a parameter t. Initially, t has a high value, but over iterations, t decreases (t can be seen as temperature what also explains the name of this algorithm: the temperature cools down (anneals) over time). A high value in t means that the probability of accepting worse solutions is high and the algorithm randomly explores the search space. Contrary, a decreasing t means that the acceptance rate decreases and the algorithm exploits the current region of the search space.

Differently to Simulated Annealing, the algorithm *Tabu Search* by Glover (1986) [62], also contains *memory* of the search process. That is, Tabu Search includes a list of solution candidates the algorithm already explored and refuses to get back to these solutions. Consequently, Tabu Search only works for discrete problems, otherwise, the memorized Tabu-list would be infinite, and therefore useless. Candidate solutions within this list are released after a certain optimization time or a certain number of iteration steps. This list enables the algorithm to leave local optima by forcing it to explore neighboring regions that might be worse but in the direction of a better solution. Extensions of this algorithm include different memory lists, e.g., to avoid the loss of already found good solutions, and to provide a diverse search that supports a sufficient degree of exploration.

Population-based meta-heuristics explore the search space simultaneously with many solution candidates. These solution candidates *communicate* among each other and use swarm intelligence to find the optimal solution. Examples of popular algorithms are the *Ant Colony Optimization* (ACO) [43], the *Particle Swarm Optimization* (PSO) [83], and *Evolutionary algorithms* (EAs).

ACO is based on the efficiency of an ant colony to find the shortest way to a food source. An obstacle that is between the ants and the source must be bypassed, either clockwise, or counter-clockwise. In the beginning, the probability for both sides is the same. If one of the two ways is shorter, over time, the shorter way enables the ants to be faster in collecting the food and the shorter trail is used more often. After a while, the shortcut is full of pheromones, a substance ants are using for communicating and the other ants will also use the shortcut. Thus, to apply ACO, optimization problems have to be formulated as pathways.

PSO is based on fish schools or bird flocks that search for good food places. Each particle (fish or bird) represents a solution candidate of the optimization problem and all

particles explore the search space. The individual change in their locations within the search space follows rules that consider (1) the position of the best solution each individual particle has discovered yet, (2) the position of the best solution the neighboring particles have discovered yet, and (3) the position of the best solution all particles have discovered yet. Together, the particle swarm is moving towards the optimal solution.

EAs are optimization methods that are based on natural evolution. Research on EAs within the last few decades has shown that they can handle large and complex search spaces and are easily and successfully adaptable to multi-objective problems. Thus, EAs designed for solving multi-objective optimization problems are called multi-objective evolutionary algorithms (MOEAs). In this thesis, I use an MOEA to solve the optimization problem. Many studies have shown that MOEAs can handle challenging optimization problems [see reviews of 109, 124]. In the following section, the principle of MOEAs is introduced more into detail. This introduction is based on the tutorial of Zitzler et al. (2004) [181].

2.6.2 Multi-Objective Evolutionary Algorithm - MOEA

As already mentioned above, MOEAs are evolutionary algorithms designed for solving multi-objective optimization problems. They are population-based meta-heuristics and explore the search space with many candidate solutions. The main principles of MOEAs are *selection* and *variation*. Selection describes the competition for reproduction and resources, whereas variation describes the creation of new *individuals* by *recombination* and *mutation*. In MOEAs, an individual is a solution candidate of the optimization problem and a set of individuals are called *population*. The selection process is separated into two different selections: the *mating selection* and the *environmental selection*. While mating selection can roughly be described as a selection process due to assertiveness, the environmental selection is often described as *survival of the fittest*.

Mating selection is separated into two phases. First, the *fitness* of all individuals of the population is determined by the objective functions (i.e., their quality). Second, the *mating pool* is created by a sampling rule that selects the individuals regarding their fitness. An often used sampling rule is the so-called *tournament selection*, where a certain number of individuals is randomly drawn from the population and the individuals are compared by their fitness. The winner of the tournament becomes a member of the mating pool. By doing many tournaments, one obtains a mating pool with many fit solutions.

The mating pool (also called *parent population*) is treated by variation operators that usually are members of the two overall classes recombination and mutation. While recombination creates a predefined number of new individuals (called *children* and *children population*) by combining parts of a certain number of old individuals of the

parent population (called *parents*), mutation modifies individuals by changing small parts of them.

The environmental selection determines the composition of the new population by selecting individuals from the old, parent and children population. That is, MOEAs have a generational *memory* that stores the currently best approximation of the Pareto set and contains information that evolves over generations.

The procedure of MOEAs is illustrated in Figure 2.3: First, an initial population is created, either randomly or predefined with expert knowledge. Second, the basic steps explained above are iterated in a loop (mating selection, variation, environmental selection). Each loop iteration is called a *generation*. The goal of an MOEA is to find the best approximation of the Pareto set (i.e., to minimize the distance of the approximation set to the Pareto set) and to cover the complete Pareto set (i.e., to maximize the diversity of the approximation set). That is, the goal of an MOEA is multi-objective itself: *proximity* and *diversity*. Proximity is mainly related to the mating selection. A good choice of parent population pushes the search of optimal solutions in the right direction. Diversity is related to selection in general. For instance, a selection operator that only draws the same solution from a population leads to no diversity. Both goals, diversity and proximity, are also influenced by the so-called *elitism*, a term that describes how to prevent the loss of solutions from the currently best approximation set. In the following, I introduce strategies to improve the optimization quality (i.e., diversity and proximity) and speed.

Fitness Assignment The fitness assignment of MOEAs can be classified into three basic categories: the aggregation-based, the criterion-based and the Pareto-based strategies. While the aggregation-based strategies aggregate the objective functions to a single fitness function (e.g., via a weighted mean with randomly drawn weights such that each individual has a different direction [71, 79]), criterion-based methods just

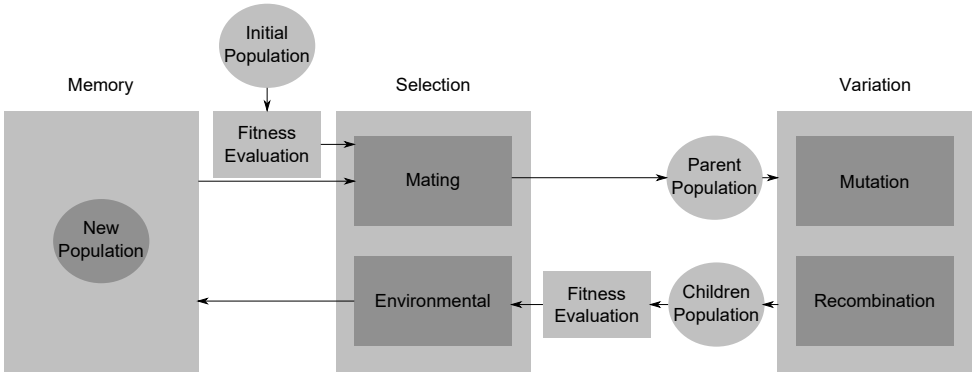


Figure 2.3: Sketch of an MOEA. An optimization run starts with the initial population, followed by the illustrated iterative loop. The iteration is called generation.

use one single objective-function value to assign a fitness value to each individual [e.g., 101, 142]. Goldberg (1989) [63] introduced first Pareto-based strategies, where the fitness of each individual is related to Pareto dominance. Three main approaches evolved from this basis: the *dominance rank* [58] that is related to the number of individuals which dominate the assessed individual, the *dominance count* that is related to the number of individuals that are dominated by the assessed individual, and the *dominance depth* where the current population is subdivided into several fronts (the best is the current approximation of the Pareto front) and the rank is the actual front number [e.g., 33, 152].

Elitism Elitism is an approach that prevents the loss of good solutions due to random effects during the optimization. An often used concept to avoid this loss is using an *archive*. It serves as an external storage and contains all non-dominated solutions and, in some cases, also promising but dominated solutions. Especially for continuous problems, storage problems might occur. Thus, dominance criteria (cf. paragraph above) and density criteria (cf. paragraph below) are applied to reduce the number of solutions kept in the archive.

Diversity Preservation In multi-objective optimization problems, maximizing diversity is important, as the optimal result is not a single solution but a multi-dimensional Pareto front. The basis for most approaches to increase diversity is to acknowledge density information of the population during the selection process. That is, the probability for an individual to be selected decreases when this individual has many neighboring solutions. The density is mostly evaluated in the objective space. Methods for density estimation include *kernel methods* (e.g., used in NSGA [152]) and *nearest neighbor methods* (e.g., clustering in SPEA2 [182]). Basic methods for density estimation can be found in Silverman (1986) [147]. They have in common that each solution must be compared with all other solutions. Two different approaches are the *crowding distance* (used in the NSGA-II [33, 36]) and the *ϵ-dominance archive* [introduced in 103] (a variation is called *ϵ-box dominance archive* [e.g., 35, 94]). The crowding distance estimates the density of solutions surrounding a particular solution only on the same front (dominance depth). The ϵ-box dominance archive discretizes the objective space into hyperboxes (ϵ-boxes with the side-length ϵ). In each ϵ-box, only the non-dominated solution remains (the solution closest to the lower-left corner in a minimization problem). Afterwards, dominance is related to ϵ-boxes and their lower-left corner.

Crossover and Mutation Operators As already mentioned above, the variation step in MOEAs is based on crossover- and mutation-operators. Often used operators are the *Simulated Binary Crossover* [32], *Unimodal Normal Distribution Crossover* [86], *Simplex Crossover* [162], *Polynomial Mutation* [32] and *Uniform Mutation*. These operators

explore the search space of the optimization problem differently. Thus, an adaptive probability-based use of them is recommended [e.g., 168, 69]. The probability of being applied to the individuals is related to the number of produced solutions by each operator within the archive. That is, good performing operators are more likely to be applied again.

Steady-State ϵ-MOEA Deb et al. (2003) [35] introduced a *steady-state ϵ-MOEA* that follows the idea of an ϵ-box dominance archive and differs from classical generational MOEAs in the selection process for the mating. Instead of having just one population that evolves over the generations, a steady-state MOEA consist of two populations, the EA population, and the archive population. The mating happens between two solutions from each population and their sizes remain constant. The archive size is predefined by the number of ϵ-boxes. Details about the selection process and further improvements can be found in Deb et al. (2003) [35], Kollat and Reed (2007) [94], and Hadka and Reed (2013) [69].

2.7 Metrics in Multi-Objective Optimization

Performance metrics evaluate solution strategies of optimization problems. For single-objective optimization problems, the value of the objective function can be used directly as a measure of quality: the smaller the objective value (in a minimization problem), the better the solution. However, unlike single-objective optimization, the quality of an approximation set from a multi-objective optimization cannot be evaluated easily. Although the concept of Pareto optimality defines whether a solution is non-dominated and belongs to the approximation set or not, it becomes more complex to compare the quality of two different approximation sets. For instance, approximation set A could have a good proximity but is not diverse, while set B might be worse in proximity but good in diversity. Thus, the principle of Pareto optimality cannot be used as metric to test whether approximation set A is preferred over set B. An overview of existing methods to measure the performance of multi-objective optimization strategies can be found in Deb (2001) [31], Knowles and Corne (2002) [89], and Zitzler et al. (2003) [183]. Zitzler et al. (2003) [183] showed that optimization problems need at least as many performance metrics as the number of objective functions the problem has, in order to compare two approximation sets unambiguously. However, the main objectives of MOEAs are proximity and diversity of the approximation set [e.g., 23, 37]. Therefore, I introduce three performance metrics that focus especially on diversity and proximity: *generational distance*, *hypervolume*, and *additive ϵ-indicator*. Since the real Pareto front (PF) is usually not known, the approximation sets (AS) are compared with a reference set (RS), which is the best set of solutions known at

this time. The following explanations and notations are based on the paper of Reed et al. (2013) [134].

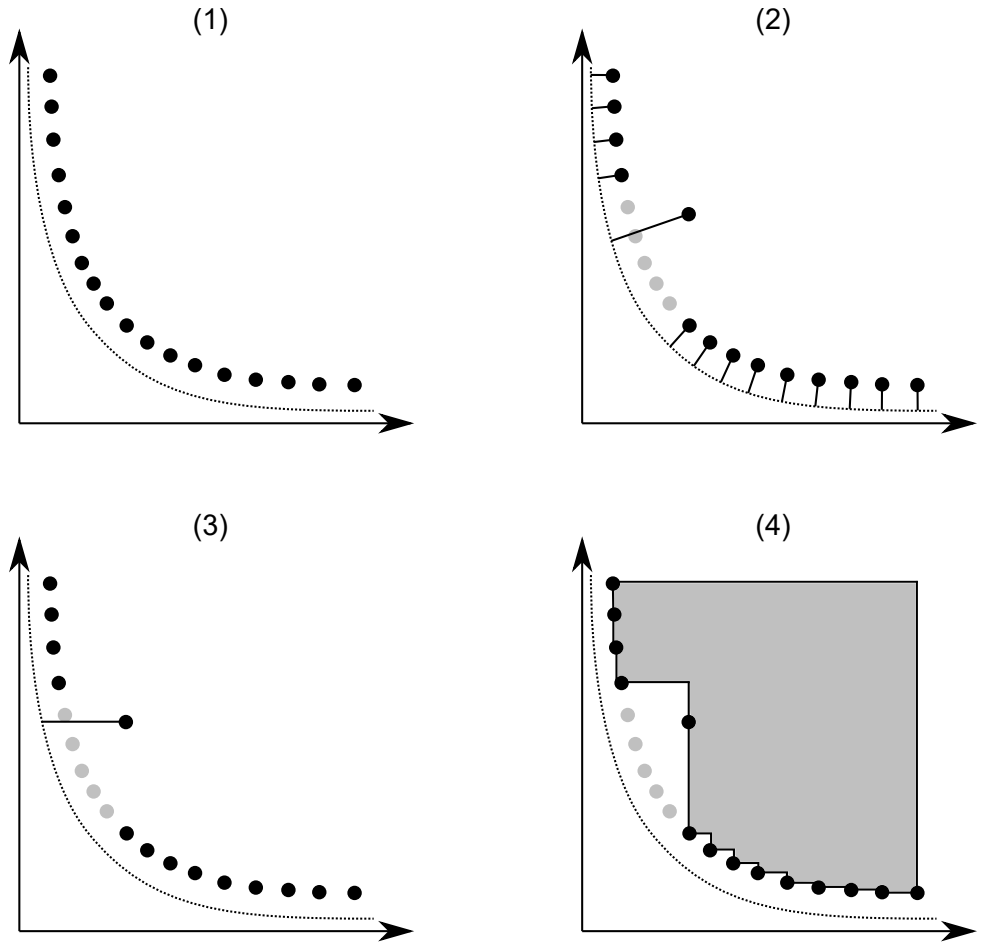


Figure 2.4: Metrics to measure the quality of an approximation set. Multi-objective problem with two objective functions (vertical and horizontal axes). **Dashed line**: Reference set. **Black dots**: Approximation set. **(1)**: Reference set and a good approximation set. **(2)**: Generational distance of an approximation set with a gap. The missing solutions are marked in gray. The Euclidean distance from the approximation-set solutions to the reference set is illustrated as black lines. The gap has no influence on this indicator. Its value only increases slightly due to the quality of the worse solution. Thus, it is an indicator only for proximity. **(3)**: The additive ϵ-indicator is the distance of the worst solution to the nearest solution of the reference set (illustrated as the black line). A small value means that the approximation set is good in all of its solutions and hence it is consistent. **(4)**: The hypervolume (shaded area) measures the volume that is covered by the approximation set. A bad diversity/proximity decrease the hypervolume and hence it is a measure of both proximity and diversity. This figure is based on the illustration of Hadka and Reed (2012) [68].

Generational Distance The generational distance [e.g., 165, 166] (sub-figure (2) in Figure 2.4) measures the average Euclidean distance of the approximation-set solutions to the nearest solutions of the reference set. This proximity is represented by the indicator I_{G}, calculated as follows:

$$I_{\mathrm{G}}\left(AS, RS\right) = \frac{\sqrt{\sum_{\mathbf{x} \in AS} \mathbf{d}_{\mathbf{x}}^2}}{N_{\mathrm{S}}^{AS}} \tag{2.19}$$

with

$$\mathbf{d}_{\mathbf{x}} = \min_{\mathbf{y} \in RS} \sqrt{\sum_{i=1}^{N_{\mathrm{OF}}} \left(f_i\left(\mathbf{x}\right) - f_i\left(\mathbf{y}\right)\right)^2}.$$

Here, N_{S}^{AS} is the number of solutions within the approximation set, N_{OF} is the number of objective functions, $\mathbf{d}_{\mathbf{x}}$ is the minimum Euclidean distance between the solutions of the approximation set $(\mathbf{x})$ and the reference set $(\mathbf{y})$, and f_i, $i \in N_{\mathrm{OF}}$ is the set of objective functions.

Additive ϵ-Indicator The additive ϵ-indicator [183] (sub-figure (3) in Figure 2.4) is neither a good measure for proximity nor for diversity, but for consistency [68]. It measures the smallest distance an approximation set has to move towards the reference set such that all solutions of the reference set would be dominated by the approximation set. This distance is given by the worst approximation-set solution, i.e., the distance that the solution is required to "improve" in order to dominate its nearest neighbor solution of the reference set:

$$I_{\epsilon}\left(AS, RS\right) = \max_{\mathbf{y} \in RS} \min_{\mathbf{x} \in AS} \max_{1 \leq i \leq N_{\mathrm{OF}}} \left(f_i\left(\mathbf{x}\right) - f_i\left(\mathbf{y}\right)\right). \tag{2.20}$$

However, additive ϵ-indicator is very sensitive to gaps within the approximation set because gaps dramatically increase the worst distance, even when the rest of the approximation set might have a good proximity.

Hypervolume The hypervolume [183] (sub-figure (4) in Figure 2.4) is a measure for proximity and diversity. It measures the volume of the dominated objective space by the approximation set. As the hypervolume is the result of a volume integral $\int_V$, it is computationally expensive to calculate. Hypervolume indicator I_{H} is the hypervolume of the approximation set normalized by the hypervolume of the reference set:

$$I_{\mathrm{H}}\left(AS, RS\right) = \frac{\int_V \alpha_{\mathrm{AS}}\left(\mathbf{x}\right) dx}{\int_V \alpha_{\mathrm{RS}}\left(\mathbf{y}\right) dy} \tag{2.21}$$

with

$$\alpha_{\mathrm{AS}}\left(\mathbf{x}\right) = \begin{cases} 1 & \text{if } \exists\, \mathbf{z} \in AS \text{ such that } \mathbf{z} \preceq \mathbf{x} \\ 0 & \text{otherwise} \end{cases}$$

and $\alpha_{\mathrm{RS}}(\mathbf{y})$ likewise. That is, $\alpha(\cdot)$ contains the dominated objective space. The hypervolume indicator gives a good impression of the quality of an approximation set because it combines proximity and diversity such that good performance in I_{H} means that the approximation set is diverse and converged.

Chapter 3

The Optimization Problem: Reliable Early-Warning Monitoring

Most of the content of this chapter has been published in the Journal TRANSPORT IN POROUS MEDIA under the title *Optimization for Early-Warning Monitoring Networks in Well Catchments Should Be Multi-objective, Risk-Prioritized and Robust Against Uncertainty* [19]. I am reusing parts of the text and figures from this publication by the kind permission of the publisher Springer.

As stated in Chapter 1, groundwater abstraction wells are commonly protected by zones of restricted land use. Such wellhead protection areas typically do not cover the entire well catchment, and numerous possible contamination sources that are within the catchment remain uncovered. Each of them could release contaminants at any time that could hit the well sooner or later and put the quality of supplied water at risk.

In this context, it seems fortunate that most well catchments are equipped with monitoring networks (at least in countries like Germany and Denmark). Such networks, however, often grew historically while following diverse purposes that changed with time, e.g., monitoring the groundwater level, or identifying the shape of existing contamination plumes. Thus, they are often inadequate (or at least suboptimal) as reliable and cost-efficient risk control mechanism according to modern standards. A prioritization according to the severity of the perceived risk is sometimes done implicitly (through adding monitoring wells to monitor the possible contamination sources that are currently perceived as the most severe ones), but seldom in a coordinated fashion. In such situations, a re-optimization of the existing monitoring networks is advisable. Contrary, there are well catchments in other countries without any monitoring networks for controlling possible contamination sources. The absence of monitoring networks often coincides with the absence of corresponding budgets or regulations. However, this is in sharp contrast with the recommendation of the World Health Organization to use a risk control structure [30]. To initialize cost-minimal risk control in such cases, again, optimization of monitoring networks is advisable. Thus, in any case, one could achieve or increase effectiveness and cost efficiency of monitoring-based risk control through formal, risk-prioritized optimization.

When optimizing monitoring networks for such purposes, it is apparent that the goals

of monitoring are manifold and often competing [e.g., 136]. In the current context, they should include at least the following three objectives:

1. to reliably detect contaminant spills for a reliable risk-control level,
2. to detect new contaminant spills as early as possible after they have occurred, because countermeasures such as installing additional water treatment steps require time for implementation, and
3. to minimize costs for installation and operation of the monitoring network.

In the following, I will call these three objectives *detection probability*, *early-warning time*, and *costs*. It is intuitive that these objectives are partly competing (a cost-efficient monitoring network cannot offer a high detection probability and a long early-warning time). Hence, concepts of multi-objective optimization are appropriate [e.g., 110].

To reliably monitor possible contamination sources under uncertain conditions, scenario analyses or Monte-Carlo simulation should be used to cover predictive uncertainty. Probabilistic coverage of uncertainty in predicting the yet non-existent contaminant transport from the possible contamination sources is substantial. Appropriate treatment of uncertainty and a corresponding formulation of the multi-objective optimization problem will make the optimization results robust against considered uncertainty.

In this chapter, I introduce the optimization problem. Here, the primary focus is on developing the objective functions, i.e., the mathematical translation of the objectives. With this intention, I begin with defining the requirements the transport simulation has to meet (Section 3.1). On this basis, I can formulate problem-specific objective functions (Section 3.2). In Section 3.3, I classify the optimization problem according to well-known problem classes and I show in Section 3.4 how to extend the optimization problem for uncertain conditions. In Section 3.5, I introduce two test cases that I will use as benchmark through the entire work. Finally, in Section 3.6, I will show and exemplarily discuss selected results for monitoring network optimization.

3.1 Model Setup

The optimization approach assumes that a sufficiently well-calibrated simulator for flow and transport in the well catchment is available. The extension to account for parameter uncertainty will be discussed in Section 3.4. In practical applications, such a simulator will typically solve the groundwater flow equation at steady state, and the advection-dispersion equation for contaminant transport (cf. Section 2.1). We also assume that there is a list of known possible contamination sources that typically includes, e.g., housing areas with oil tanks, sewer systems and wastewater treatment plants, industrial sites, and so forth.

First, we discretize the catchment area (model domain) with a fine spatial mesh of possible monitoring well locations M_j, $j = 1, \ldots, n_{\mathrm{M}}$. For each possible location, the drilling costs should be known, or at least estimable. All possible contamination sources R_i, $i = 1, \ldots, n_{\mathrm{R}}$ within the model domain are relevant in the following. We assume that conservative estimates of contaminant mass m_i in case of spill events are available for each possible contamination source R_i, and that the corresponding chemical detection limit c_i^{det} is known.

Second, starting from the possible contamination sources, we solve instantaneous-release forward transport problems to simulate the contaminant plumes that would emerge in case a possible contamination source actually released contaminants into the subsurface. Following the discussion of risk estimation in well catchments in Enzenhoefer et al. (2015) [50], we distinguish the different spatial extent of possible contamination sources (e.g., point sources, line sources and areal sources versus point sources with uncertain position along a line or within an area) through corresponding source geometries. From all these transport simulations, we store the following data:

$$\delta_{ij} = \begin{cases} 1 & \text{a contamination released at possible contamination source } R_i \\ & \text{exceeds the detection limit } c_i^{\mathrm{det}} \text{ at monitoring candidate} \\ & \text{location } M_j \\ \\ 0 & \text{no detectable concentration at } M_j. \end{cases} \tag{3.1}$$

If $\delta_{ij} = 1$, we can also obtain the travel time τ_{ij}^{det} between contaminant release at possible contamination source R_i and first exceedance of the detection limit c_i^{det} at monitoring-well candidate M_j. We also extract a travel time $\tau_{i,\mathrm{well}}$ that marks the first arrival of a critical concentration c_i^{crit} at the production well. The difference yields the respective early-warning time t_{ij}:

$$t_{ij} = \begin{cases} \max\left(\tau_{i,\mathrm{well}} - \tau_{ij}^{\mathrm{det}}, 0\right) & \text{early-warning time } \left(\text{if } \delta_{ij} = 1\right) \\ 0 & \text{no early warning } \left(\delta_{ij} = 0\right). \end{cases} \tag{3.2}$$

Further, we extract the duration $\Delta t_{ij}^{\mathrm{vis}}$ for which the contaminant plume from possible contamination source R_i is visible at a monitoring-well candidate M_j, i.e., the duration for which the simulated plume from R_i exceeds the detection limit c_i^{det} at the position of M_j.

3.2 Optimization Problem Formulation

I formulate the multi-objective optimization problem at hand as follows:

$$\mathbf{d}_{\text{opt}} = \underset{\mathbf{d}\in\mathbf{D}}{\arg\min} \left[f_{\text{det}}, f_{\text{warn}}, f_{\text{cost}}\right]. \tag{3.3}$$

Here, $\mathbf{d}_{\text{opt}}$ is the optimal set of decision variables $\mathbf{d}$, which characterize the planned monitoring system (e.g., number and positions of monitoring wells, filtering depth and screened window, frequency of sampling). $\mathbf{D}$ is the space of permitted designs (e.g., restricted to accessible positions within the catchment, maximum admissible installation and operation costs), and f_{det}, f_{warn}, and f_{cost} are the objective functions that assign goal attainment levels to each design $\mathbf{d} \in \mathbf{D}$. These three objective functions are explained in the following.

I formulate each objective as a minimization problem, i.e., f_{det} is the probability of not detecting any of the emitted plumes, f_{warn} evaluates the time lost between contaminant spill and detection, and f_{cost} specifies the costs of installation and operation. Furthermore, I normalize all objective functions to the interval $[0, 1]$, such that the theoretical optimal value in all individual objectives is zero. While f_{det} falls between zero and one by definition, I express f_{warn} through a utility function that I chose to be normalized to that interval, and I normalize the costs function f_{cost} through division by a maximum admissible cost value.

3.2.1 Detection Probability

The values δ_{ij} from Equation 3.1 express, which contaminant plumes from the possible contamination sources R_i can in principle be detected by the monitoring wells M_j. However, the monitoring wells are only sampled in time intervals of $\Delta t_j^{\text{sample}}$ while the plumes exceed the detectable concentration c_i^{det} only for a duration $\Delta t_{ij}^{\text{vis}}$. This fact causes a probability that a plume may pass unnoticed through a monitoring well position due to unfortunate timing. To account for this effect, I adjust from δ_{ij} to a probability of detection P_{ij}^{det} as follows:

$$P_{ij}^{\text{det}}(\mathbf{d}) = \min\left(\frac{\Delta t_{ij}^{\text{vis}}}{\Delta t_j^{\text{sample}}(\mathbf{d})}, 1\right) \cdot \delta_{ij}. \tag{3.4}$$

The aggregation of the obtained probabilities P_{ij}^{det} over all possible contamination sources $R_i, i = 1 \ldots n_{\text{R}}$ and over all monitoring wells suggested from the candidates $M_j, j = 1 \ldots n_{\text{M}}$ by any given design $\mathbf{d}$ is multi-objective itself and will be discussed in Section 3.2.4.

3.2.2 Early-Warning Time

Early-warning time in context of supply-well protection is the time period between the first detection of a groundwater contaminant and its arrival time at the production wells. In other words, it is the remaining time for water suppliers to install countermeasures against the imminent contamination. Hence, a sufficiently long early-warning time for each possible contamination sources is desirable. Nonetheless, for the optimization, a simple linear relation between individual early-warning time values and the overall early-warning performance of a monitoring network is not suitable: the achievable early-warning time values for different possible contamination sources may vary between none and several decades, and an increase in early-warning time of remote sources is less valuable for water suppliers as an increase for close sources.

Therefore, I use a non-linear utility function $U_i(t)$ (see Figure 3.1) for each possible contamination source R_i (illustrated by the different solid black lines). The utility $U_i(t)$ is normalized by the individual maximum achievable early-warning time $\hat{t}_i^{\max}$ that is defined through the travel time $\tau_{i,\text{well}}$ from possible contamination source R_i to the production wells. A monitoring network that can achieve $\hat{t}_i^{\max}$ as early-warning time obtains a utility of one for source R_i. This specification serves to ensure that a monitoring well that achieves the maximum possible early-warning time for source R_i is rated with the maximum possible utility value.

Short early-warning times for supply-well contaminations are critical for a reliable freshwater production and each increment of these early-warning times can be very valuable. Short early-warning times can either be related to possible contamination sources close to the production wells or they are caused by a bad performance of the monitoring well. To distinguish these two scenarios, I introduce a user-defined minimum desirable early-warning time $\hat{t}^{\min}$ and a corresponding early-warning time utility $\hat{u}^{\min}$ that determines the benefit of $\hat{t}^{\min}$. $\hat{t}^{\min}$ defines whether an early-warning time is critical (orange and red area in Figure 3.1) and $\hat{u}^{\min}$ defines whether the performance of a monitoring well is bad (red area in Figure 3.1). That is, the larger $\hat{u}^{\min}$, the steeper $U_i\left(t \leq \hat{t}^{\min}\right)$, and the larger the increase in utility for extended early-warning times towards the maximum achievable early-warning time for possible contamination source R_i. If the travel time $\tau_{i,\text{well}}$ is larger than a user-defined maximum desirable early-warning time $\hat{t}^{\max}$, $\hat{t}_i^{\max}$ is set to $\hat{t}^{\max}$, such that longer early-warning times have no additional benefit for the water supplier (green area in Figure 3.1). In other words, $\hat{t}^{\max}$ is the maximum still useful early-warning time and monitoring wells within the green area provide maximum utility. The performances of monitoring wells within the yellow area in Figure 3.1 are between sufficient and very good. These monitoring wells can reach large early-warning times and high utility values. Usually, from $\hat{t}^{\min}$ to $\hat{t}^{\max}$, the utility should grow slower compared to the utility of the orange/red area. Early-warning times within this corridor still feature an increasing early-warning time utility, but these extension are less important and can be seen as an extra benefit.

Summarized, for remote possible contamination sources, monitoring wells should fall into the green or the yellow area. Close possible contamination sources with short travel times towards the production wells should be monitored as early as possible (orange area). In either case, the red area should be avoided by monitoring wells.

Technically, I assign $U(t = 0) = 0$, $U(t = \hat{t}^{\min}) = \hat{u}^{\min}$ and $U(t \geq \hat{t}^{\max}) = 1$. This yields as utility function:

$$U_i(t) = \begin{cases} 1 & \forall t : t \geq \hat{t}^{\max} \\ \frac{\left(1-\hat{u}^{\min}\right)\cdot\left(t-\hat{t}_i^{\max}\right)}{\hat{t}_i^{\max}-\hat{t}^{\min}} + 1 & \forall t : \hat{t}^{\min} \leq t \leq \hat{t}_i^{\max} \wedge \hat{t}_i^{\max} < \frac{\hat{t}^{\min}}{\hat{u}^{\min}} \\ \frac{t}{\hat{t}_i^{\max}} & \forall t : t \leq \hat{t}^{\min} \vee \hat{t}_i^{\max} \geq \frac{\hat{t}^{\min}}{\hat{u}^{\min}}\,. \end{cases} \tag{3.5}$$

Changes in all three user-defined values ($\hat{u}^{\min}$, $\hat{t}^{\min}$, and $\hat{t}^{\max}$) might influence the final results more or less distinctively. For instance, $\hat{u}^{\min}$ defines the steepness of the utility-curve between 0 and $\hat{t}^{\min}$. Therefore, it can be seen as a catchment-specific weighting factor for early or late detection, hence it needs to be chosen carefully.

3.2.3 Costs

The objective function for costs, f_{cost}, is a summation of predefined and predictable costs, like depth-specific drilling costs that are spatially variable, and operation costs

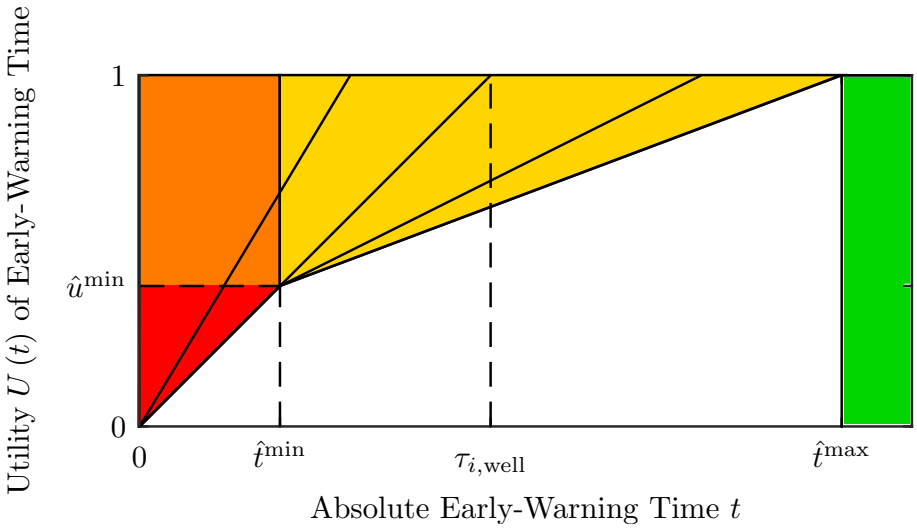


Figure 3.1: Case-specific utility function $U_i(t)$ for early warning defined in Equation 3.5. Each solid line represents the utility as a function of early-warning time of possible contamination sources that differ in $\tau_{i,\text{well}}$. The colors represent four different characteristic areas of this function.

that are composed of costs per sampling round within a given budget period:

$$f_{\text{cost}} = \sum_{l}^{N_C} C_l \,, \tag{3.6}$$

with N_C different expenses C_l [€]. Consequential costs are not included. Examples for consequential costs are costs due to a shut down of production wells as consequence of exceeding a threshold concentration, or high remediation costs due to a late detection of a contamination and a large contaminated area. Since some consequential costs can be extremely high, they are a strong motivation for installing reliable early-warning monitoring networks. However, a consequential-cost function would complicate the optimization problem dramatically and economic models would need to be added to forecast potential financial damages. These are two reasons why consequential costs are not considered in this work, but the extension would be technically straightforward: they could be incorporated as separate objective functions without influencing any other implementations discussed here.

3.2.4 Evaluation of a Monitoring Network - Aggregation Rules

For any given design **d**, there is a list $L_{\mathbf{d}}$ of monitoring wells to be installed. From that list, several monitoring wells $M_j, j \in L_{\mathbf{d}}$ could have non-zero detection probabilities P_{ij}^{det} for any of the given possible contamination sources R_i. However, these monitoring wells could differ in early-warning time and detection probability, and we need to define a rule how to identify a unique value for detection probability and early-warning utility for each possible contamination source R_i. We need such a rule to avoid that the monitoring wells have to be sampled for the whole list of possible contaminants (alternatively, this could also be solved by optimization).

There is no general rule that the monitoring well with the best early-warning time for a possible contamination source R_i will automatically have the best value in detection probability for this R_i. As a rule to identify a unique monitoring well per possible contamination source when rating the network performance, I use a method of best compromise. That means, I select the one monitoring well position M_j for the possible contamination source R_i out of the list $L_{\mathbf{d}}$ that maximizes the product of detection probability and early-warning time t_{ij}:

$$\ell_i = \underset{j \in L_{\mathbf{d}}}{\arg\max} \left[P_{ij}^{\text{det}} t_{ij} \right] . \tag{3.7}$$

Finally, I average the values of P_{ij}^{det} with $j = \ell_i$ over all possible contamination sources R_i to obtain the overall detection probability $P_{\mathbf{d}}^{\text{det}}$ of the monitoring network and de-

fine:

$$f_{\text{det}}(\mathbf{d}) = 1 - \underbrace{\frac{1}{n_{\text{R}}} \sum_{i=1}^{n_{\text{R}}} P_{i\ell_i}^{\text{det}}}_{P_{\mathbf{d}}^{\text{det}}}. \tag{3.8}$$

As argued in Section 3.2.2, the achievable early-warning time values for different possible contamination sources may vary between none and several decades, such that a simple linear relation between individual early-warning time values and the overall early-warning performance of the monitoring network is not suitable. Instead, I work with a non-linear utility function $U_i\left(t_{ij}\right)$. Thus, in analogy to f_{det}, I define:

$$f_{\text{warn}}(\mathbf{d}) = 1 - \underbrace{\frac{1}{n_{\text{R}}} \sum_{i=1}^{n_{\text{R}}} U_i\left(t_{i\ell_i}\right)}_{U_{\mathbf{d}}^{\text{warn}}}, \tag{3.9}$$

where $U_{\mathbf{d}}^{\text{warn}}$ is the overall early-warning utility of the monitoring network defined by $\mathbf{d}$.

3.3 Classification of the Optimization Problem

After formulating the optimization problem, it can be classified using the terms defined in Section 2.6. The optimization problem is based on a finite search space that contains all combinations of potential monitoring well locations (cf. Section 3.1), hence it is a *discrete combinatorial problem*. The problem is *high dimensional*: each possible monitoring well location has the potential to influence the outcome of the optimization (at least changes in cost function). It is a *multi-objective optimization problem* with three competing objectives with the objective functions defined above. The objective functions are highly *non-linear* and the problem is very likely *non-convex*. The problem formulation is yet deterministic. However, it can easily be extended to a stochastic optimization problem (considering uncertainty), as I will present in the following section.

3.4 Considering Uncertainty for Reliable Monitoring Networks

The objective functions f_{det} and f_{warn} use the simulation-based data specified in Section 3.1. Hence, they rely on model-based predictions about flow and transport in the catchment, which are subject to uncertainty. The uncertainty arises from many sources, but I only focus on parameter uncertainty (as I discussed in Section 2.4),

which includes the description of flow boundary conditions. Uncertainty in transport boundary-conditions is already considered by the possible contamination sources.

When not accounting for uncertainty, the optimized monitoring network would most likely achieve a lower performance (when applied in practice) than predicted by the optimization. Vice versa, accounting for uncertainty at least to some extent will make the optimized monitoring networks robust against the considered parameter uncertainty.

Thus, I propose to perform Monte-Carlo simulations if the pdfs of the considered uncertain parameters are known, scenario analyses otherwise. Each scenario or Monte-Carlo realization $k, k = 1, \ldots, n_\mathrm{k}$ yields its own values δ_{ijk}, $\tau_{ijk}^{\mathrm{det}}$, $\tau_{ik,\mathrm{well}}$ and $\Delta t_{ijk}^{\mathrm{vis}}$ for the quantities defined in Section 3.1. Then, we replace the affected objective functions f_{det} and f_{warn} in Equation 3.3 by adequate statistics over the objective function values $f_{\mathrm{det}}^{(k)}$ and $f_{\mathrm{warn}}^{(k)}$ obtained per realization:

$$\mathbf{d}_{\mathrm{opt}} = \underset{\mathbf{d}\in\mathbf{D}}{\arg\min} \left[\tilde{f}_{\mathrm{det}}, \tilde{f}_{\mathrm{warn}}, f_{\mathrm{cost}}\right], \tag{3.10}$$

where $\tilde{f}$ could denote, e.g., the largest value or a high percentile (since the objective functions are formulated for a minimization problem) or the arithmetic mean. The expected value is a risk-neutral approach to optimization under uncertainty [e.g., 54], while working with extremes or percentiles is a risk-averse and even more robust approach.

3.5 Benchmark

To demonstrate and discuss the proposed methods and approaches, I set up two synthetic application scenarios as test cases. The first test case is inspired by typical models used for well catchment management. It consists of a domain divided into a few zones with different hydraulic conductivity values. I will refer to it as *Z_Based* from now on, as zonation-based model. The second test case is abstracted from a real-world case and captures key complexity of urban source water protection. Hence, I refer to it as *U_Protect* in the following, as urban water protection.

Both test cases represent a single geological layer of an aquifer as a quasi-three-dimensional domain (15,000m $\times$ 7,000m $\times$ 10m for Z_Based and 15,000m $\times$ 7,500m $\times$ 10m for U_Protect). The regional flow direction is approximately from east to west, defined through Dirichlet conditions at all domain boundaries. The prescribed boundary values follow geometrically from a uniform gradient that is specified through its absolute value and its orientation relative to the east-west axis. For simplification, all possible contamination sources are assumed to be point sources with a pulse release of

their contaminant. Each source is represented by 1000 particles released as point-like and instantaneous injection (particles are injected within an area of 5×5 m^2). In the following, I describe both test cases in detail.

3.5.1 Z_Based

As mentioned above, the model domain of Z_Based is divided into several different hydraulic conductivity zones with the intention to reflect typical groundwater models of water suppliers (e.g., developed with ModFlow). All relevant model parameters of Z_Based are provided in Table 3.1, and a system sketch is shown in Figure 3.2. Near the eastern boundary of the model domain, there is a gallery of 15 pumping wells with 50 meters spacing between neighboring wells. The figure also shows the location of the considered possible contamination sources as red circles.

For covering the robustness aspect mentioned in Section 3.4, Z_Based works with four different hydraulic scenarios (cf., Table 3.2) that differ in the strength of the regional head gradient, in its orientation angle, and in the overall pumping rate of the well gallery. These scenarios represent aspects of hydrological uncertainty. They do not include the other sources of uncertainty listed in Section 2.4 to keep the application scenario straightforward. The different scenarios yield travel times from the most distant possible contamination source to the well of up to 80 years. Again, for the sake of simplicity, I assume conservative tracer transport. This assumption may seem crude, but it is justifiable as a worst-case scenario for contaminant impact on the well, i.e., unretarded transport and no degradation.

I discretize the domain with rectangular, equispaced cells sized 10m $\times$ 10m $\times$ 10m and simulate groundwater flow with the standard-Galerkin finite element code already used in Nowak et al. (2008) [127]. For all subsequent transport simulations, I employ the Particle-Tracking Random Walk (PTRW) code used earlier by Enzenhoefer et al. (2014) [51] and by Koch and Nowak (2014) [90]. I choose PTRW due to its ease of implementation and its absence of numerical dispersion [e.g., 84, 102, 140].

Table 3.1: Transport-relevant parameters of Z_Based. The hydraulic conductivity value provided here is the arithmetic mean of all zone-wise values.

Average conductivity K [m/s]	Porosity n [$-$]	Molecular diffusion coefficient D_m [m^2/s]	Longitudinal dispersivity α_ℓ [m]	Transverse dispersivity α_t [m]
3.15×10^{-4}	0.35	1×10^{-9}	3	0.3

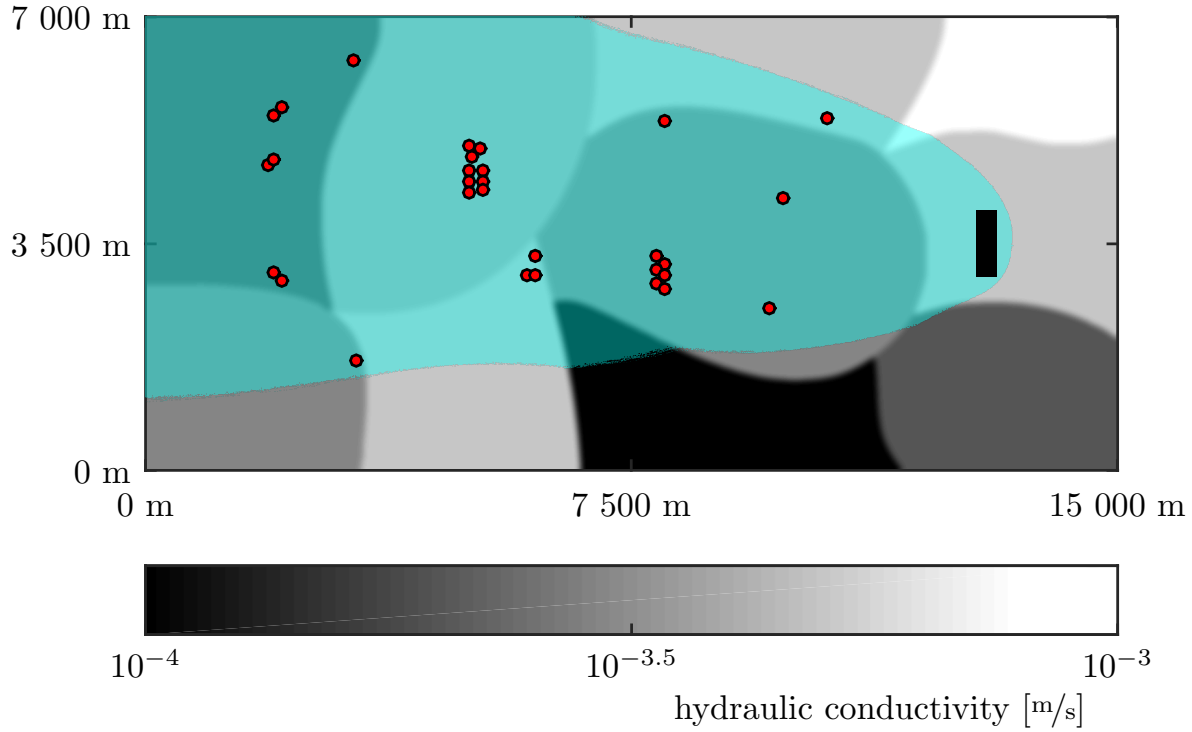


Figure 3.2: Model domain of Z_Based. Grayscale: zonation and values of hydraulic conductivity; Black rectangle: position of the well gallery; Blue overlay: union of all scenario-wise well catchments (see Table 3.2); Red circles: possible contamination sources.

Table 3.2: Definition of four different hydraulic scenarios

Hydraulic scenario	Gradient strength [−]	Gradient angle [°]	Pumping rate [l/s]
1	0.01	0	15
2	0.01	15	15
3	0.015	0	15
4	0.01	0	7.5

3.5.2 U_Protect

U_Protect is a test case that captures the key complexity of a real-world case. It includes:

- a random conductivity field instead of zonations,
- urban areas with a high density of possible contamination sources, and economic restrictions in installation of monitoring wells and countermeasures, and
- inaccessible areas, where the feasibility of installing monitoring wells is not given.

The conductivity field is defined on a fine numerical mesh (10m $\times$ 10m $\times$ 10m) by $Y = \log(\mathbf{K})$. I assume local isotropy, that is, $\mathbf{K}$ simplifies to the scalar K. The field follows the underlying discretized, two-dimensional multi-Gaussian random space function:

$$p(\mathbf{Y}) \propto \exp\left(-\frac{1}{2}(\mathbf{Y}-\mu_Y)^T \mathbf{C}^{-1}(\mathbf{Y}-\mu_Y)\right), \tag{3.11}$$

with $\mathbf{Y}$ as vector of Y-values at all locations within the domain. The random space function is defined by the mean μ_Y, the variance σ_Y^2 and a stationary covariance function $C(l)$ that is used to complete the covariance matrix $\mathbf{C}$ in Equation 3.11. As covariance model we use the Matérn covariance function [113]:

$$C(l) = \frac{\sigma_Y}{2^{\kappa_m-1}\Gamma(\kappa_m)}\left(2\sqrt{\kappa_m}l\right)^{\kappa_m} B_{\kappa_m}\left(2\sqrt{\kappa_m}l\right), \tag{3.12}$$

with l as an effective separation distance between x_1 and x_2:

$$l = \sqrt{\left(\frac{\Delta x_1}{\lambda_1}\right)^2 + \left(\frac{\Delta x_2}{\lambda_2}\right)^2}.$$

Here, $\Gamma(\cdot)$ is the Gamma function, $B_{\kappa_m}(\cdot)$ the modified Bessel function of the third kind [e.g., 1], κ_m is a shape parameter that controls the shape of the covariance function and λ_i with $i = 1, 2$ is the correlation length. All relevant model parameters of U_Protect are provided in Table 3.3, and a system sketch is shown in Figure 3.3.

On the eastern side at $x = 13000$ m, eleven pumping wells are installed, all along a north-south line at $y \in [3500, 4000]$ m, and each with a pumping rate of 30 l/s. The distance between neighboring pumping wells is 50 m. Within the test case there are three different types of areas:

1. The *open field* with no limitation regarding the installation of monitoring wells (illustrated in Figure 3.3 as gray/white structure),
2. *urban areas* (illustrated in Figure 3.3 as maps) where it is not possible to install monitoring wells, either because it is too expensive or because there are strong conflicts (e.g., mobility, shopping, ...), and

Table 3.3: Transport-relevant parameters of U_Protect. The hydraulic conductivity value provided here is the arithmetic mean of all values.

Average conductivity K [m/s]	Porosity n [−]	Molecular diffusion coefficient D_m [m²/s]	Longitudinal dispersivity α_ℓ [m]	Transverse dispersivity α_t [m]
0.08	0.3	1×10^{-9}	1	0.1

3. *inaccessible areas* (illustrated in Figure 3.3 as a forest) where it is not possible to install monitoring wells, because it is too expensive in time, effort and costs.

For representing uncertainty in hydro(geo)logy, I vary (a) the ambient flow direction of the groundwater, and (b) the conductivity field. The variations in ambient flow direction are controlled by the Dirichlet boundary condition: the general flow direction is corrected by an angle β drawn from a normal distribution $\mathcal{N}\left(\mu_\beta, \sigma_\beta^2\right)$ with mean $\mu_\beta = 0$ and variance $\sigma_\beta^2 = 10^2$.

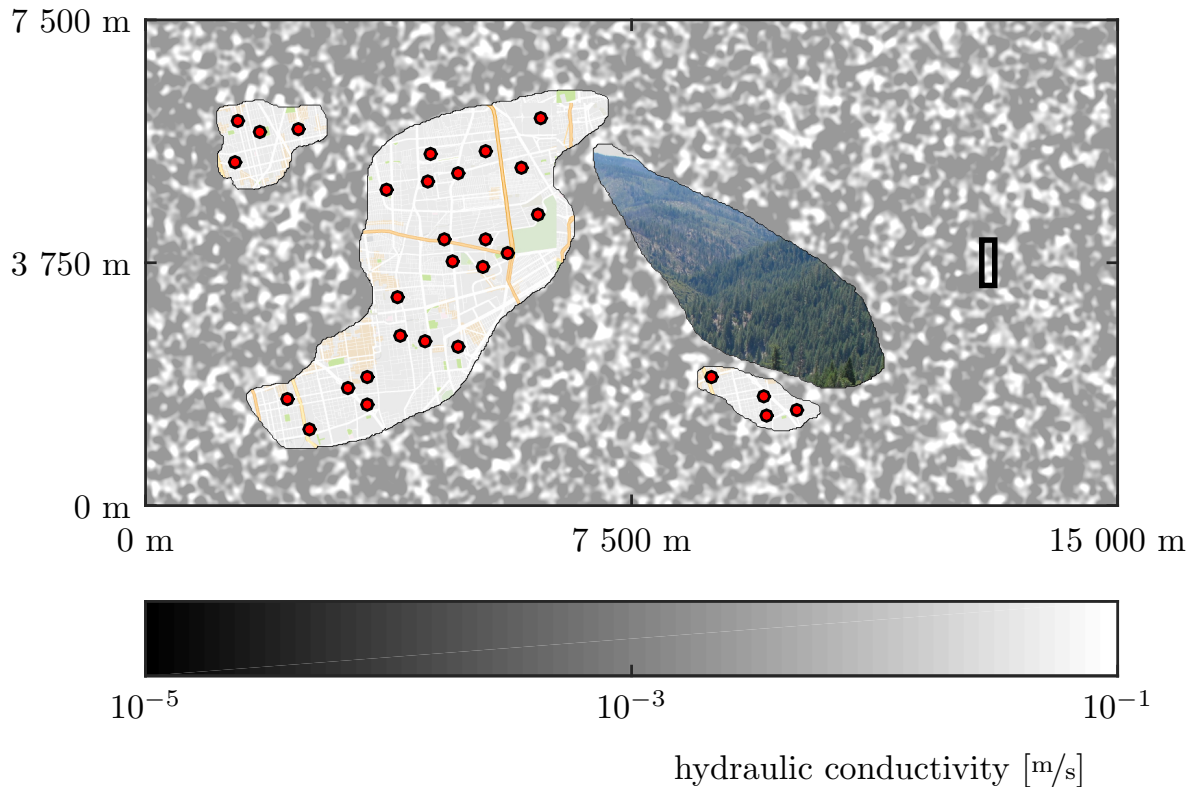


Figure 3.3: Model domain of U_Protect. Grayscale: values of hydraulic conductivity; Black rectangle: position of the well gallery; Maps: urban areas; Forest: difficult accessible area; Red circles: possible contamination sources within urban areas.

The variations in the conductivity field are controlled by the covariance function $C(l)$ (Equation 3.12). Following the Bayesian geostatistical approach [e.g., 87], also the covariance function is uncertain. Therefore, I model the structural parameters $\boldsymbol{\theta} = \left[\sigma_Y^2, \lambda_i, \kappa_m\right]$ as random variables and the covariance function $C(l)$ itself becomes a random function governed by the random parameter $\boldsymbol{\theta}$: $C(l, \boldsymbol{\theta})$.

3.6 Exemplary Results

In this section, I exemplary discuss an optimization result of the optimization problem defined above. Therefore, I use optimization outcomes from the following Chapter 4, where all relevant optimization settings can be found in Table 4.4. I discuss the results using Figure 3.4 and Figure 3.5.

Figure 3.4 shows an approximation of the Pareto front of the optimization problem (red spheres) and the projections of the three-dimensional approximation set onto the planes of the coordinate system. Figure 3.5 shows two different monitoring networks for U_Protect. One is marked with black crosses and the other one is marked with green pluses. In the following discussion, I refer to the networks as CROSSES and PLUSES. The two selected monitoring networks CROSSES and PLUSES are highlighted as green and black spheres, respectively, in Figure 3.4, according to their colors in Figure 3.5.

Figure 3.4 clearly reveals that the three considered objectives are competing. One can achieve a full coverage of all possible contamination sources at maximum early warning utility, but only with a budget that allows installing 9 monitoring wells (black sphere of the approximation set). At decreasing costs, one can either maintain maximum detection probability (back/right edge of the approximation set) while losing early-warning time, or one can try to lose less early-warning functionality but restrict that functionality to a smaller number of possible contamination sources (back/left edge of the approximation set). In general, for this problem there is a list of four geometric aspects in the Pareto front (and approximation sets) that will hold for any catchment:

1. The idealistic goal of benefits without costs (blue sphere in Figure 3.4) can never be reached.

2. Maximum detection probability and yet minimum costs are only possible by directly monitoring the mixed water of the pumping-well gallery (bottom front/right corner of the objective space). Since I did not consider this case as relevant in our application, the optimization was not forced to search for Pareto optima at this part of the approximation set.

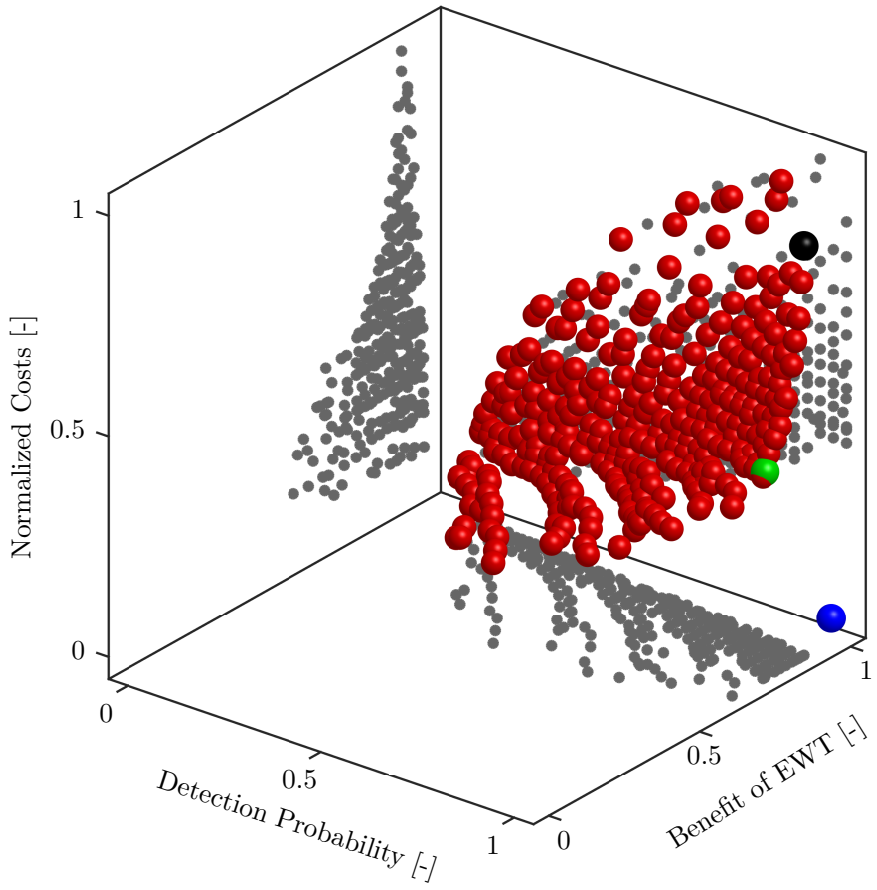


Figure 3.4: Approximation of the Pareto front for the multi-objective optimization problem on U_Protect. Blue sphere: theoretical optimum; Red spheres: non-dominated solutions (approximation set of the Pareto front); green and black sphere: performance of the selected monitoring networks shown in Figure 3.5 (green: PLUSES; black: CROSSES). Gray dots projections of the approximation set onto the planes of the coordinate system

3. The solution with almost no costs and a very poor detection of the possible contamination sources (bottom/left edge of the objective space) is to place one individual monitoring well at the cheapest individual drilling position and depends only on the complexity of f_{cost}. This selection defines the bottom back/bottom left tip of the Pareto front.

4. The top/right back tip of the Pareto front is given by the most expensive solution that places monitoring wells close to the possible contamination sources.

These four facts define the fundamental geometry of the Pareto front. Only the curvature of the Pareto front towards the idealistic unattainable optimum depends on the properties of the investigated catchment and the positions of its possible contamination sources. Contamination sources arranged in a line transverse to the dominant flow direction would lead to a strong competition between the objective functions (and hence to a low curvature of the Pareto front) because non-overlapping plumes prohibit the monitoring of several possible contamination sources with only a few monitoring wells. Possible contamination sources arranged in a longitudinal line would lead to one overlapping plume, with almost no competition between costs and detection. Clustered possible contamination sources would fully remove the competing character of all objective functions because these sources could be monitored with good early-warning time by just one or a few monitoring wells very close to the spill location.

The selected monitoring networks shown in Figure 3.5 both have a detection probability of 100%, i.e., they reliably monitor all possible contamination sources. However, they starkly differ in their number of monitoring wells (9 for CROSSES, 5 for PLUSES), and hence in costs (88% vs. 41% of maximum costs). The high performance of PLUSES in detection probability at relatively low costs can be achieved by the strategy to monitor the well catchment at positions where several predicted contaminant plumes coincide. As a result, the network PLUSES is shifted towards the pumping wells, and further away from the possible contamination sources. In contrast, the more expensive network CROSSES forms a barrier wall close to the urban areas, where all possible contamination sources occur. Consequently, the performance in early-warning time of CROSSES outperforms the one of PLUSES.

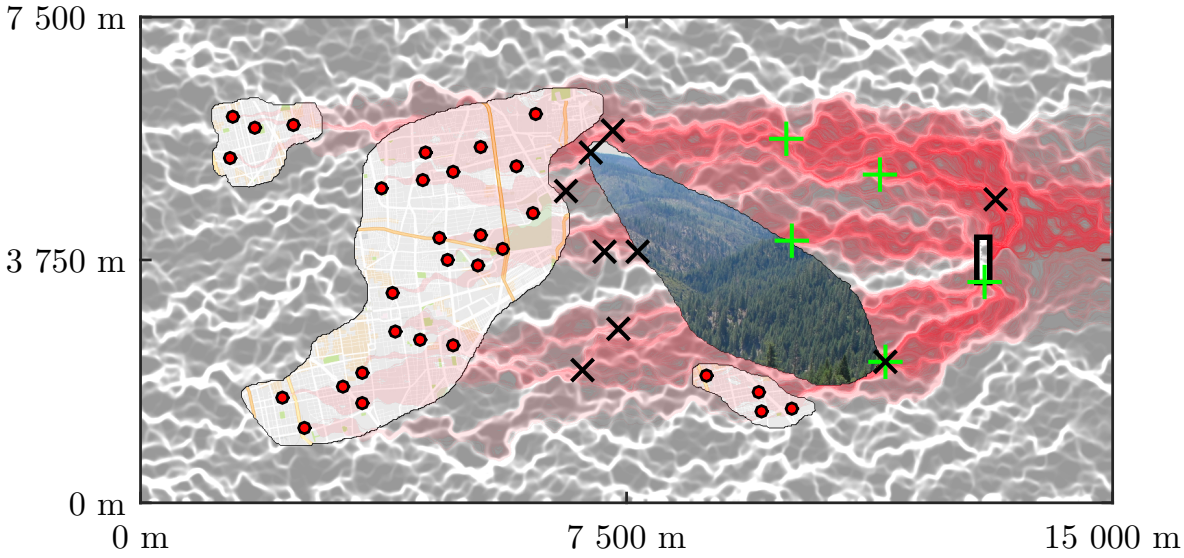


Figure 3.5: Model domain of U_Protect (cf. Figure 3.3). Grayscale: values of velocity field; Magenta overlay: possible contamination plumes (degree of transparency is equivalent to the number of overlapped plumes); Black crosses/green pluses: possible monitoring networks.

Chapter 4

Representation and Reduction of the Search Space

This chapter is based on the submitted manuscript *Search Space Representation and Reduction Methods to Enhance Multi-Objective Water Supply Monitoring Design* to the Journal WATER RESOURCES RESEARCH [21]. I am reusing parts of the text and figures from this manuscript.

This chapter builds on the general problem formulation from the previous chapter. It contributes enhanced approaches for managing large search spaces to make optimization faster (computationally less expensive) without sacrificing the quality of the attained trade-off solutions. These methods work for combinatorial optimization problems with a precalculable search space and approach the problem of speed by reducing or re-organizing the search space. Fortunately, my problem formulation allows for precalculating the transport simulations for each possible contamination source and monitoring well such that the mathematical space of possible monitoring-well positions and a few relevant properties are known before searching for optimal monitoring networks. From this available information, all possible monitoring networks can be created, hence the search space is precalculable. The strategies introduced in this chapter focus on two key aspects of the design problem:

1. a more compact representation of the search space that can be addressed with evolutionary multi-objective search operators, and
2. a formal mathematical screening of candidate decisions to reduce the search space.

While the representation of the search space affects the optimization algorithm in use, the search-space reduction is a pre-processing step with a stand-alone application. Both general ideas are outlined in the following two paragraphs.

Representation of the search space. We assume that the optimization is based on precalculated numerical transport simulations from all possible contamination sources towards the production wells. Then, we have a discrete approximation of our model domain (the well catchment) and a finite set of possible monitoring-well positions. The

search space consists of the power set of all possible monitoring-well positions, hence its size can easily reach intractable dimensions. For instance, the number of possible monitoring networks with up to 15 monitoring wells arranged from 100 possible monitoring-well positions yields a search space larger than 3×10^{17} possible combinations (without repetition). Therefore, usually, the search space (power set) is not defined explicitly, but implicitly by all possible monitoring-well positions. The set of all possible positions can reproduce all combinations the search space contains. For the sake of generality, I use the term *elementary decision options* or *EDOs* to refer to candidate monitoring-well positions, and the *EDO set* to define the entire mathematical space of all possible monitoring-well positions. A common representation of the EDO set during the optimization is the *binary representation* (cf. Section 4.2.1). However, state-of-the-art evolutionary multi-objective algorithms struggle with the binary representation in combination with large EDO sets that have complex interdependent decision variable sensitivities on the objective functions [95, 146]. The number of discrete decision variables directly correlates with the total number of EDOs (N_{EDO}) as well as with the computational scaling of the search tool [94]. To reduce the number of decision variables, I introduce the *linear-indexing representation* in Section 4.2.2 that is based on the *random keys* representation proposed by Bean (1994) [6]. The basic idea of the random keys representation is that the search for optimal solutions takes place on an easier surrogate space that is more appropriate to the translation and variation operators of heuristic optimizers than the original (binary) mathematical space. The linear-indexing representation underlies the random keys representation and serves to decouple the number of decision variables from the exponential growth of the power set of EDOs.

Reduction of the search space. Evolutionary multi-objective optimization represents a balance between robustness to challenging mathematical problem features (nonconvexity, discreteness, stochasticity, nonlinearity, etc.) and the degree to which the algorithms maintain their ability to provide high-quality Pareto-approximation sets. A core control on this balance is the number and type of decisions that compose the search space. A common strategy to more effectively address this balance is to reduce the search-space size. A number of different reduction strategies have been established in the literature. An overview of existing methods can be found in the position paper of Maier et al. (2014) [109]. Many of these methods use adaptively refining grids, either on binary encoding [e.g., 64, 144, 175], or numerical grid-based dynamic emulation [e.g., 5, 148]. Alternatively, Walters and Lohbeck (1993) [170], as well as Walters and Smith (1995) [171], focus on screening solutions that are not valid or feasible. Building on a similar logic, I propose decision space reduction strategies that exclude EDOs (along with large portions of the search space in the corresponding power set) that are limited in their influence on the optimization. Simple exclusions are already possible given expert knowledge about the optimization problem. For example, in ground-

water monitoring, one could remove candidate monitoring wells where it would be impossible to detect contamination. Usually, the expert-driven reduction of the set of EDOs yields a search space that still remains very large. Further reduction can be done by excluding redundant EDOs from the set. The check for redundancy is based on comparing pairs of EDOs. If one EDO dominates the other or both are equally good, the better option remains in the set, and the other one can be excluded. There are different approaches for implementing this comparison. It can be done in a systematic manner (e.g., all EDOs are compared among each other), or it can be done with a heuristic, randomized screening method.

As the overall optimization problem of this thesis is combinatorial, a reduction of single possible monitoring-well positions would reduce the number of their combinations dramatically. For instance, a relatively small problem with an EDO set of 100 possible monitoring-well positions and a search space that includes combinations of 1 to 5 monitoring wells, can be simplified by more than 40% when 10 positions of the EDO set are proven to be irrelevant and removed (from $\sim 79 \times 10^6$ to $\sim 47 \times 10^6$ possible solutions). That is, the optimization algorithm can either search more intensely within the remaining 60% candidate solutions for optimal solutions, or it can keep the intensity level, but reduce optimization time.

The presented representation and reduction methods in this chapter significantly enhance the effectiveness, efficiency, and reliability of the optimization. The proposed framework shifts focus to the most impactful monitoring design decisions while also enhancing decision makers understanding of key performance trade-offs. In combination, the proposed representation and reduction techniques have significant promise for enhancing the size and the scope of combinatorial monitoring problems that can be explored.

The remainder of this chapter provides a more detailed explanation of our proposed representation and reduction contributions. The following section introduces the used optimization algorithm (the Borg MOEA [69]) and its performance metrics. Section 4.2 describes the different ways in which the search space within the optimization was represented. In Section 4.3, I define the reduction strategies explored in this study. Here, I also generalize the methods to show that they can be used for a broad class of optimization problems. I provide a list of conditions that a problem has to meet to reliably speed up and to improve the optimization procedure with our reduction methods. Section 4.4 provides the details of the computational experiment used to benchmark the contributions. Section 4.5 presents the results and discussion of the tested search space representation and reduction methods. Finally, in the last section (Section 4.6), I summarize and conclude this study.

4.1 Methods

4.1.1 Multi-Objective Optimization Algorithm

Reed et al. (2013) [134] showed that the Borg MOEA [introduced in 69] has competitive to superior performance relative to existing state-of-the-art MOEAs on a diverse suite of multi-objective water resource applications including a monitoring benchmark problem. These results motivated my use of the Borg MOEA. The key characteristics of the Borg MOEA are:

1. the use of an ϵ-dominance archive [103] that enhances the convergence and the diversity of the approximation set,
2. the use of ϵ-progress [69] to evaluate the search progress and to avoid stagnation in search,
3. the adaptive re-sizing of the population [75, 93] to provide a population large enough to improve the current approximation set but small enough to still be efficient in calculation time,
4. the adaptive selection of multiple recombination operators [168], based on the current solution-composition of the approximation set and
5. that Borg is a steady-state ϵ-MOEA [35] that avoids computationally expensive operations and can be easily parallelized.

Borg includes six different recombination operators: Simulated Binary Crossover (SBX) [32], Differential Evolution (DE) [155], Parent-Centric Crossover (PCX) [34], Simplex Crossover (SPX) [162], Unimodal Normal Distribution Crossover (UNDX) [86], and Uniform Mutation (UM). Additionally, Polynomial Mutation (PM) [32] mutates the offsprings generated by SBX, DE, PCX, and SPX.

4.1.2 Performance Metrics

I use the normalized hypervolume as the performance metric, where the hypervolume of an approximation set is normalized by the hypervolume of the reference set (cf. Section 2.7). The hypervolume is a good indicator of the proximity and diversity of attained approximation sets. Achieving hypervolumes close to one signals that the approximation set is meeting the same level of diversity and proximity as the reference set. Hypervolume also provides a rigorous means of ranking alternative problem formulations regarding their performance. Overall, we can evaluate MOEA performance in terms of *effectiveness*, *efficiency*, and *reliability*. An effective optimization means that

the hypervolume of an approximation set is close to one. Efficiency describes the ability of the search algorithm to quickly find good solutions, which can be measured by the number of function evaluations (NFE) or the optimization runtime necessary for a specific hypervolume attainment. Reliability refers to minimizing the variation in attained search results across random-seed trials. For calculating the hypervolume, I use the diagnostic framework of Hadka and Reed (2012) [68].

4.2 Representation of Search Spaces

When tackling an optimization problem with an MOEA, algorithmic representation of the search space is fundamentally important as it shapes the scope of search as well as the appropriateness of search operators (i.e., mating, mutation, and selection). Here, I only discuss spatially discretized problems as illustrated in Figure 4.1. Again, the search space is implicitly defined via an EDO set that contains all elementary decision options, while the explicit definition would be the power set of all EDOs (i.e., all possible combinations of the elementary decision options).

Solutions within the search space can be addressed in different ways that directly depend on the representation of the EDO set. In many optimization problems, it is appropriate to represent the search space by the physical meaning of the EDOs directly (in the current problem the physical meaning are the coordinates of the possible monitoring-well positions). However, the EDO-set representation is required to be closed under the genetic operators of the optimization algorithm. That is, when the genetic operators are applied to EDOs, the resulting elements should be EDOs as well, hence a part of the EDO set. Otherwise, the genetic operators could generate invalid solutions.

Figure 4.2 shows the meaningful physical EDO set of U_Protect (see Section 3.5.2) as gray area on the bottom plane. It is obvious that this EDO set is not closed against the genetic operators when using coordinates to represent the search space because the genetic operators could easily enter invalid (not gray) areas. A possible closing strategy is to make the search area convex by overestimating it to include the void space (in Figure 4.2 the interspersed white spaces between gray areas in the bottom rectangle). However, the complexity of the optimization problem would increase unnecessarily and the search would be likely to degrade in effectiveness and efficiency. Thus, I present two different representation approaches in the following two sections.

4.2.1 Binary Representation

The *binary representation* of search spaces is often used in discrete optimization problems [e.g., 135] and many optimization algorithms are preset for a binary optimization

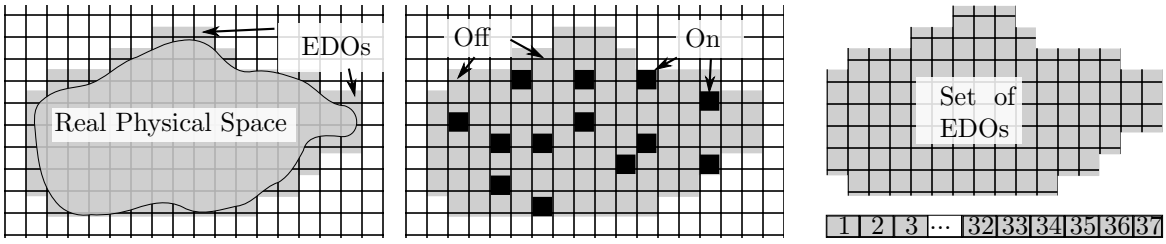


Figure 4.1: Basic steps to represent the search space in the current test case (from left to right). First, the physical space is approximated by numerical simulation. The set of EDOs is given by the mesh discretization and implicitly defines the search space. Second, with binary representation, all EDOs are decision variables and switch between two statuses: on or off. That is, only activated decision variables determine the outcome of the solution. Third, with linear-indexing representation, the set of EDOs (here in total 37 EDOs) is reshaped to a list. The number of decision variables is independent of the number of EDOs and only depends on the desired maximum number of monitoring wells.

(e.g., traditional genetic algorithms [174]). This representation follows the criteria of closure through the use of integer operators such as rounding because all EDOs can only be zero or one $(\text{EDO} \in [0, 1])$. Especially for combinatorial problems, the binary representation is intuitive. For instance, consider the optimal design of monitoring networks: a monitoring-well position can be switched on or off and the MOEA is searching for the best on/off combinations (cf. middle sub-figure in Figure 4.1). That is, each solution is a list of on/off decisions and contains the entire set of EDOs, all decision variables, respectively.

Although the binary character of the decision variables appears simple, this representation is limited to problems with a relatively small number of decision variables (< 200) on currently available machines. The presence of fiscal constraints and physics-based interdependencies between the binary decisions makes the monitoring problem an extremely difficult analog to the classic combinatorial Knapsack (i.e., with non-deterministic polynomial time hardness or NP-hard [95, 146]).

4.2.2 Linear-Indexing Representation

As a means of reducing the number of decision variables and to better manage problems with a large number of EDOs, I introduce the *linear-indexing representation* of the search space. For a discrete problem, the set of EDOs can be transformed into a list L_{id} of linear indices that just contain the information about the original location of the EDOs within the EDO set (cf. right sub-figure in Figure 4.1). The approach is to address the list L_{id} via an index i ($i \in \mathbb{E}$ and $\mathbb{E} = \{1, 2, \ldots, N_{\text{EDO}}\}$) and use the linear

set $\mathbb{E}$ as a surrogate for the actual EDO set. With that, the index i becomes a decision variable and the number of considered decision variables is limited by the user to the maximum number of EDOs $(N_{\max})$ per solution (e.g., a maximum number of monitoring wells given by an upper-cost constraint). We complete our set of decision variables with a set of binary decision variables that define whether a drawn EDO (via the index i) is part of the solution candidate or not. With this, we achieve that the number of decision variables is constant, even if a solution addresses less than the number $N_{\max}$ of EDOs. We can guarantee the closure of the EDO set under the genetic operators by the use of integer operators for both types of decision variables. For a combinatorial optimization problem, linear-indexing representation reduces the number of decision variables dramatically compared to the binary representation, if the maximum desired number of decision variables $N_{\max}$ is small compared to the EDO set. For instance, if the upper limit of a planned monitoring network is five wells, only ten decision variables are required: five that address the actual locations of the monitoring wells and another five that switch the five drawn monitoring wells on or off.

For efficient and effective optimization runs, the sorting of the EDOs within the list L_{id} is crucial. This is because, when using evolutionary algorithms, proximity in the search space should imply proximity in the objective space. In Figure 4.2 the EDO set of U_Protect is illustrated. Here, the gray area on the lower plane illustrates the binary representation. The colored area with the gradient from upper back to lower front illustrates the linear-indexing representation. The values on the z-axis represent the position of the EDOs within L_{id}. All EDOs are sorted by their distance to the pumping wells. Thus, neighboring EDOs within the EDO set that might have similar characteristics, are also contiguous within L_{id}. A more advanced sorting strategy that works without educated guess about the properties of the problem could sort the EDOs directly based on similarity characteristics in the objective space, e.g., using clustering via a k-means algorithm [76], or according to travel times (f_{warn}) from/to sources or wells.

4.3 Search Space Reduction Methods

In the following sections, we discuss possible reduction methods for precalculated, finite, and discrete search spaces. We explore reduction methods with a range of properties including ease of implementation, strength of reduction, and preservation of Pareto-optimal solutions. The three reductions are summarized as follows:

1. The *fundamental reduction method* requires knowledge about the optimization problem and about the numerical-model results that provide the pre-calculated data for the optimization. This method removes obviously irrelevant EDOs from the EDO set.

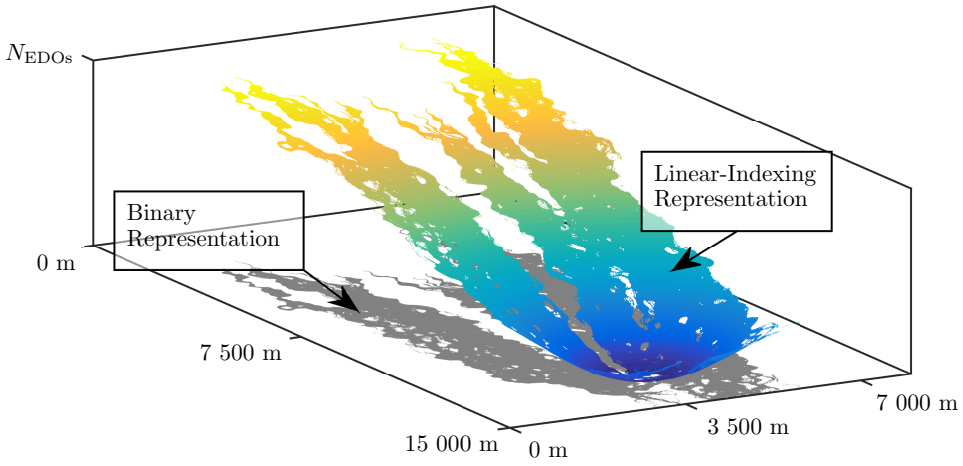


Figure 4.2: Example of the two discussed methods how to represent the search space. The binary representation is illustrated as a gray area where a contamination can be detected. The linear-indexing representation is illustrated as colored surface. The z-values describes the position within the list of all EDOs (i.e., it corresponds to the sorting of the EDOs). Thus, the sorting can influence the efficiency of the optimization and can be done in different ways. In this illustration, the EDOs are sorted by their distance to the pumping wells.

2. The *grid-based reduction method* is often used for optimization problems and is easy to handle. It discretizes the EDO set using a grid and represents all EDOs within the same grid cell by the best single EDO that is determined by user-defined rules.

3. The *comparison-based reduction method* compares the goal attainment of two arbitrary EDOs and removes the dominated one. We present two different implementations, the *optimal reduction*, and the *optimization-based reduction*.

Their key characteristics are summarized in Table 4.1.

4.3.1 Fundamental Reduction

The search space for many optimization problems has a high potential for simple and fast reduction. Reduction can be considered at two fundamental levels.

First, external constraints can be used to a priori screen out EDOs that are not favorable. For example, in the U_Protect test case (cf. Section 3.5), all monitoring-well positions that trigger high installation and/or operation costs, can be excluded, also when there is a dangerous or infeasible installation of a monitoring well (e.g., in cities, or anhydrite

Table 4.1: Key Characteristics of the Presented Reduction Methods

Reduction Method	Easy to implement	Reduction potential	Possible loss of Pareto-optimal solutions
Fundamental	Yes	Varies	No
Grid-Based	Yes	High	Yes
Optimal	No	Varies	No
Optimization-Based	No	High	Yes

layers that could grow when coming into contact with water). These areas are shown as maps or forest in Figure 3.5. When doing this reduction too generously, there is a risk of over-restricting the EDOs and eliminating important decision alternatives from the search space. Strong restrictions within the search space might lead to the loss of solutions that would otherwise be part of the Pareto-optimal set. That being said, extremely large decision spaces with complex constraints could require dramatically more search or even render search operators invalid.

Second, for precalculated problems, the meaningful set of useful EDOs can be determined a priori and the search is restricted to the useful EDOs. In our case, the precalculated dataset clearly identifies all positions where a contamination could be detected, and those where no contaminant plume from any of the possible contamination sources will ever pass by. That is, using the complete model domain as the EDO set is pointless, because many of its EDOs only cause costs but do not contribute to a monitoring-network solution. These useless EDOs can be excluded easily and quickly from the EDO set without the risk of losing relevant Pareto-optimal solutions from the search space. This reduction leads to the fingered geometry of the remaining EDO positions in Figure 4.2, also shown in Figure 3.5 as semi-transparent red areas downgradient of the considered possible contamination sources.

4.3.2 Grid-Based Reduction

The basic idea of the grid-based reduction is to represent the actual EDO set by a reduced set of representative EDOs. The choice of the representative EDOs is based on a regular or irregular coarse-meshed grid that more coarsely discretizes the original EDO set. Then, all EDOs within a discretization element of the coarse grid are represented by just one EDO. Optimization problems that are strongly focused on a spatial EDO set frequently employ this reduction method [e.g., 5, 96, 105]. The main drawback is that potentially well-performing EDOs get lost that could better contribute to global solutions than their representative EDOs. There is no guarantee that the lost EDOs

are not part of any Pareto-optimal solution. However, this method is a very effective reduction method and easy to implement.

Here, I implemented this reduction method on U_Protect and used a grid resolution that was coarse enough to downsize the EDO set such that the binary representation could be applied. I used three decision criteria to find the representative EDOs:

1. all EDOs that could only detect one possible contamination source were directly removed,
2. in each discretization element, among all remaining EDOs, I kept only the one that could detect the largest number of up-gradient possible contamination sources, and
3. in the case that two locations could monitor the same number of possible contamination sources, I kept the EDO that was closer to the center of area of all detectable plumes.

I did not use the early-warning time as a third decision criterion because, within one discretization element, the differences of the EDOs are relatively small compared to the overall travel time of the contaminants to the production wells.

4.3.3 Comparison-Based Reduction Methods

While the grid-based reduction method is a brute-force method that eliminates EDOs without checking whether any EDO might be part of Pareto-optimal solutions, the following two comparison-based reduction methods identify only irrelevant EDOs and exclude them from the EDO set.

The main idea is to check, whether a single EDO x_j can either be replaced by another EDO x_k because they have the same effect or because EDO x_k dominates EDO x_j. For the check whether EDO x_j can be excluded from the EDO set or not, all objective values $f_i(x)$ (with $i = 1, 2, \ldots, N_{\mathrm{OF}}$ and N_{OF} as the number of objective functions) must fulfill the following three conditions:

$$\begin{aligned} &1) \quad \boldsymbol{f}(x_k) \preceq \boldsymbol{f}(x_j) \text{ and} \\ &2) \quad \boldsymbol{f}(x_k) \preceq \boldsymbol{f}(x_j, x_k) \text{ and} \\ &3) \quad \boldsymbol{f}(\mathbf{X}) \preceq \boldsymbol{f}(x_j, \mathbf{X}) \text{ or } \boldsymbol{f}(x_k, \mathbf{X}) \preceq \boldsymbol{f}(x_j, \mathbf{X}) \ . \end{aligned} \tag{4.1}$$

Here, $\boldsymbol{f}$ denotes the set of objective functions f_i and $\mathbf{X}$ is an arbitrary subset of EDOs without $x_{j,k}$ $\left(\mathbf{X} \subseteq \text{EDO set} \backslash \{x_j, x_k\}\right)$. In detail, Equation 4.1 means:

1. If EDO x_k is not dominated by EDO x_j (described by $\preceq$), and
2. x_k is also not dominated by the set of both EDOs $\left[x_j, x_k\right]$, and
3. if any arbitrary EDO subset $\mathbf{X}$ is not dominated by the set $\left[x_j, \mathbf{X}\right]$, or if the set $[x_k, \mathbf{X}]$ is not dominated by the set $\left[x_j, \mathbf{X}\right]$, then x_j can be removed from the EDO set.

In the case when all three constraints are fulfilled, the EDO x_j cannot be a part of any Pareto-optimal solution, because the gain in goal attainment will always be higher with x_k for any solution.

While the first and the second constraints in Equation 4.1 are intuitive, the third constraint is just for completeness. Many objective functions already include the third constraint if the first and the second are fulfilled (e.g., all monotonic objective functions). If the third constraint must be evaluated for each comparison step, the comparison-based methods are not feasible anymore, because the number of all arbitrary subsets $\mathbf{X}$ is almost as large as the explicit search space.

Comparison-based methods only work in a meaningful manner for discretized problems and finite EDO sets. Otherwise, they would suffer infinite runtime. In the following, I will present two different methods that use the EDO-comparison concept: the *optimal reduction method* and the *optimization-based reduction method*. The main difference between the two methods is that the optimal reduction method globally and systematically identifies all irrelevant EDOs, and the optimization-based reduction method only checks the relevance of EDOs relative to already checked EDOs. Unchecked EDOs cannot be a part of the resulting reduced EDO set. In other words, for the first method, it can be guaranteed that no EDO will be excluded that is part of a Pareto-optimal solution set, while for the second method it can only be guaranteed that the remaining EDOs are better than the checked and excluded EDOs. However, there might be better EDOs within the unchecked but excluded EDOs.

Optimal Reduction

As already said above, the optimal reduction method uses the EDO comparison globally. The degree of EDO-set reduction and the runtime of this method depends on the characteristic of the EDO set. Two characteristics are of special relevance: (1) the spatial variability of similar EDOs, and (2) the variability of objective values of all EDOs:

1. EDO sets with many spatially clustered redundant EDOs can significantly be reduced in relatively short time.
2. EDO sets with highly variable EDOs in objective values cannot be reduced strongly, because the EDOs can cooperatively complete their individual weaknesses.

3. EDO sets with many redundant EDOs that are spatially variable would lead to a long runtime because the ratio of comparison operations and eliminations of EDOs would clearly be on the side of comparison operations.

While the first case is ideal from the view of reduction, the second and the third cases are suboptimal. However, at least for the third case, it is possible to shift the ratio between comparison operations and eliminations of EDOs to the elimination side. Following the first characteristic from above, the order of comparison is crucial for computational efficiency. For example, the probability that two potential monitoring wells M_1 and M_2 monitor the same possible contamination sources is high when M_1 and M_2 are neighbors. Then, M_1 might be dominated by M_2 (or vice versa) and together they might not provide any improvements; M_1 can be excluded from the EDO set. If M_1 is not contiguous to M_2 and they monitor different possible contamination sources, together they always improve the overall objective value for detection probability. When acknowledging this behavior, for the third case the organization of EDOs can be manipulated regarding their objective values such that clusters of similar EDOs are neighboring upon re-organization. Alternatively, all EDOs can be sorted according to their distance to the current reference EDO (cf. Figure 4.2) to identify clusters of similar EDOs.

A special case of EDO sets with highly variable EDOs in the objective space (second case) is an EDO set with a small value range. Many EDOs provide almost the same quality, but cannot be removed from the EDO set following the rules defined in Equation 4.1. Then, an effective computational efficiency booster for the optimal-reduction method is to modify the objective space of the fundamental EDO set in a way that EDOs with different, but almost equal, qualities are getting the same values (called *being homogenized* in the following). This approach reduces the number of EDOs within the comparison steps (cf. Equation 4.1). More EDOs can be rejected immediately because only one of multiple EDOs with the same quality (after the homogenization) must be a part of the final reduced EDO set. For the homogenization of the fundamental EDO set, we discretize its objective space in hyperboxes with a side length ϵ similar to the ϵ-box dominance concept (cf. Section 2.6.2). Then, we can either apply the ϵ-box dominance concept and only accept a single non-dominated solution per hyperbox, or we randomly choose a solution per box (equivalent to the round function). The difference between the two methods is illustrated in Figure 4.3. With an increasing discretization length ϵ, the EDO set would be more homogeneous in objective values, and the reduction would be more efficient. However, both approaches (ϵ-box concept and round-function) would nullify the guarantee that the reduced EDO set contains all Pareto-optimal solutions. Still, but depending on the discretization length ϵ, the resulting EDO sets could be a very good approximation of the optimal reduced EDO set. Therefore, I suggest applying these methods only if the EDO set is very indifferent in the objective values of the EDOs, such that a combination of two EDOs has almost no impact on the joint objective value.

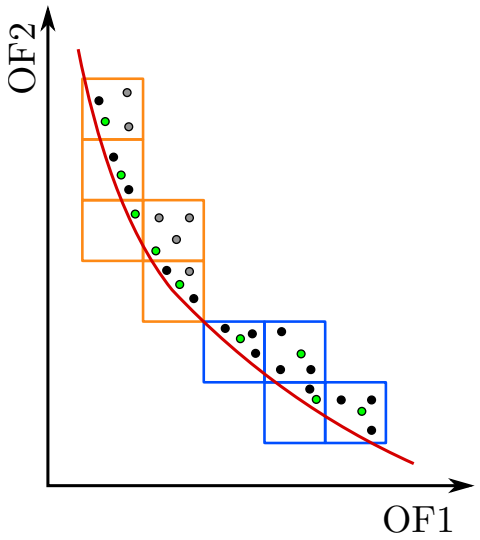


Figure 4.3: Differences between the ϵ-box dominance concept and the round function (colored boxes discretize the objective space) on an example with the two objective functions OF1 and OF2. The red line illustrates the real but unknown Pareto front. The dots are candidate solutions. Each box is represented by one solution that becomes selected over multiple other candidate solutions per box (black dots). The difference between the ϵ-box dominance concept and the round function is that the ϵ candidates are deterministically selected out of all non-dominated box solutions (green and black dots, the gray dots are dominated) and the round function handles all candidate solutions within a box equally and also dominated solutions can be selected.

Optimization-Based Reduction

The *optimization-based reduction method* uses formal optimization algorithms to approximate the EDO set of the optimal reduction method introduced directly above, hence is an alternative but an approximate approach for complex and large EDO sets. The required optimization algorithm for the reduction can be the same as for the entire optimization problem. However, it might be necessary to adopt the objective functions of the overall problem, if the objective functions aggregate over multiple single values. With the aggregation, information will be lost and EDOs will be wrongly excluded from the EDO set. For instance, in Equation 4.2 two EDOs are being compared by the aggregating objective function f_*. Although the single values of both EDOs differ, they have the same aggregated objective value (f_*^1 and f_*^2) and together, they would increase the aggregated objective value:

$$
\begin{aligned}
f_*^1 &= \text{mean}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0.5 \\
f_*^2 &= \text{mean}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0.5 \\
f_*^{1,2} &= \text{mean}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1.
\end{aligned} \tag{4.2}
$$

Therefore, when working with aggregating objective functions, the comparison must take place on the non-aggregated level, such that only EDOs are eliminated if all single values are dominated by another EDO.

The objective functions of our optimization problem aggregate over the single values for all possible contamination sources. However, to evaluate the performance of monitoring wells, it also matters, which possible contamination source they are able to monitor. That means, we have to segregate our objective functions into objective functions for each possible contamination source. This modification leads to an increase in dimensionality of the optimization problem, but the segregated objective functions are easy to evaluate.

Although the approximation of the optimal reduced EDO set can save computing time, there are two drawbacks compared to the optimal reduction of the EDO set: (1) it is an approximation, thus EDOs might be excluded from the EDO set that could be part of a Pareto-optimal solution, and (2) the quality of the EDO set depends on the runtime of this pre-optimization, generating yet another trade-off between efficiency and effectiveness.

4.4 Computational Experiment

The computational experiment exploits the U_Protect test case described in Section 3.5 with and without considering uncertainty. The U_Protect benchmarking problem captures key source water protection monitoring challenges as to contribute generalizable insights when distinguishing our proposed approaches for reduction and representation of the search space.

Without considering uncertainty First, I simulated the contaminant evolution from all 30 possible contamination sources towards the pumping well (using the particle-tracking-random-walk method) and recorded all relevant data that were needed for the objective function (cf. Section 3.2). Second, on the resulting set of EDOs, I applied

the fundamental reduction (Section 4.3.1) to get the *base EDO set*. Third, on this base set, I tested the remaining reduction methods introduced in Section 4.3. Then, I formulated seven different problem formulations varying the different EDO sets and the representation method (linear-indexing representation (LIR) versus binary representation (BR)). A summary of the problem formulations is shown in Table 4.2.

With considering uncertainty First, I performed around 300 realizations of the contaminant evolution of all 30 possible contamination sources varying hydro(geo)logic parameters and ambient flow direction among all realizations. Second, I aggregated the resulting separated EDO sets using the arithmetic mean over the single EDOs and realizations (cf. Section 3.4). Third, I applied the fundamental reduction to get the base EDO set for the uncertainty case. Finally, I tested the optimal reduction method together with the linear-indexing representation and varied the homogeneity of the EDO-set objective space. A summary of the problem formulations is shown in Table 4.3

All reduction methods were performed on a single Intel Core i5-4590 CPU at 3.30 GHz and 32GB RAM. Both experiments were optimized with the parallelized master-slave

Table 4.2: Summary of the Different Problem Formulations Without Considering Uncertainty

#	Repr.	Redu. Meth. 1	Redu. Meth. 2	Redu. Meth. 3	Size of EDO set	Random Seeds
1	LIR	Fundamental	-	-	219,122	500
2	LIR	Fundamental	Optimal	-	1,189	200
3	LIR	Fundamental	Optimal	Grid	150	200
4	LIR	Fundamental	Optimization	-	1,479	200
5	LIR	Fundamental	Grid	-	148	200
6	BR	Fundamental	Optimal	Grid	150	200
7	BR	Fundamental	Grid	-	148	200

Table 4.3: Summary of the Different Problem Formulations With Considering Uncertainty

#	Repr.	Redu. Meth. 1	Redu. Meth. 2	Booster	Size of EDO set	Random Seeds
8	LIR	Fundamental	-	-	436,986	30*
9	LIR	Fundamental	Optimal	-	69,202	200
10	LIR	Fundamental	Optimal	Round	801	200

* Reduced number of runs due to extended wall-clock time of 600 s.

version of the Borg MOEA introduced in Hadka and Reed (2015) [70] on a cluster (Sun Grid Engine version 6.2u3 as batch-queuing system) with 7 compute nodes. Each node includes eight Quad-Core AMD Opteron Processors 2376 at 2.3 GHz and 32GB of RAM. Subtracting the communication cores, in total 49 cores remained for the optimization.

For each problem formulation, we started the optimization multiple times (random-seed analysis) to account for the stochastic nature of the optimization and evaluate the reliability of its search (cf. last column in Tables 4.2 and 4.3). Runtime dynamics were compiled by tracking the evolution of the approximation sets over the number of function evaluations (NFE). That is, the evolution of the approximation set was tracked every 1,000 function evaluations. An exception is the base case with uncertainty consideration. Due to its long wall-clock time, the evolution was tracked every 50,000 function evaluations. The ϵ precisions for the ϵ-box dominance were set to \$1,000 for the cost function f_{cost} and to 0.01 for detection probability f_{det} and early-warning time utility f_{warn}. All default settings for the Borg MOEA (crossover and mutation parameters, ϵ precision, wall-clock time, etc.) are summarized in Table 4.4.

For analyzing the different reduction and representation methods, I used the diagnostic framework of Hadka and Reed (2012) [68] to calculate the evolution of the hypervolume over number of function evaluations. I use the merged approximation set over all optimization runs and problem formulations as reference for the best possible approximation set.

Table 4.4: Default Settings of the Borg MOEA

Parameter	Value	Parameter	Value
PM Rate	1/L*	PCX Zeta	0.1
PM Distribution Index	20	UNDX nr. of Parents	10
SBX Rate	1	UNDX nr. of Offspring	2
SBX Distribution Index	15	UNDX Eta	0.35
DE Crossover Rate	0.1	UNDX Zeta	0.5
DE Step Size	0.5	Initial Population Size	100
UM Rate	1/L	Injection Rate	0.25
SPX nr. of Parents	10	Wall-Clock Time	14 s
SPX nr. of Offspring	2	Nr. of Cores	49
SPX Epsilon	3	f_{cost} Epsilon	1000
PCX nr. of Parents	10	f_{det} Epsilon	0.01
PCX nr. of Offspring	2	f_{warn} Epsilon	0.01
PCX Eta	0.1	NFE Output Interval	1000

* L: Number of decision variables.

4.5 Results and Discussion

I investigate the effects of the different problem formulations (listed in Tables 4.2 and 4.3) in terms of their resulting effects on the efficiency, effectiveness, and reliability of the MOEA solution framework. Initially, I demonstrate the benefits of the linear-indexing representation in Sections 4.5.1 and 4.5.2. Next, Section 4.5.2 also analyzes how search-space reduction can influence optimization results positively and negatively. In Section 4.5.4, I demonstrate and discuss the effects of search-space representation and wall-clock time reduction for selected problem formulations. In Section 4.5.5, I explore the effects of selected search-space reduction methods on resulting trade-off analysis and decision making aspects of the source water monitoring problem. Finally, Section 4.5.6 investigates the effect of considering uncertainty on the optimal reduction method.

4.5.1 Linear-Indexing Representation is more Reliable and Efficient

At first, let us compare the linear-indexing representation against the binary representation in terms of reliability and efficiency. Figure 4.4 shows attained hypervolume with increasing numbers of function evaluations (NFE) for the linear-indexing and binary search-space representations (Scenarios 5 and 7). The 'blue' hypervolume performance dynamics are the result of the linear-indexing representation and the 'red' is based on the binary representation. The solid and the dashed lines represent the median (over all 200 random-seed trials) for the respective representations. The shaded areas represent variations in attained hypervolume, representing the 10% and the 90% percentiles across the random-seed trials. Figure 4.4 highlights that the linear-indexing representation aids in attaining highly reliable (low variance) MOEA search performance. In contrast, the binary representation clearly increases the difficulty of the search problem yielding a much higher variance in hypervolume dynamics.

In terms of efficiency, two characteristics of the hypervolume evolution in Figure 4.4 are important: (1) the initial hypervolume level, and (2) the gradient of the hypervolume over NFE. A relatively high initial hypervolume indicates the existence of highly diverse valid solutions at the beginning of the optimization, such that the algorithm does not have to invest function evaluations for searching them. A large positive gradient of the hypervolume indicates a fast improvement of the optimization results.

As can be seen in Figure 4.4, the problem formulation based on the linear-indexing representation outperforms the binary problem formulation in both reliability and efficiency. The hypervolume performance differences between the two search-space representations result due to the enormous difference in their resulting number of decision variables. The large number of decision variables within the binary representation (149) yields a vast decision space that requires a large NFE to approximate

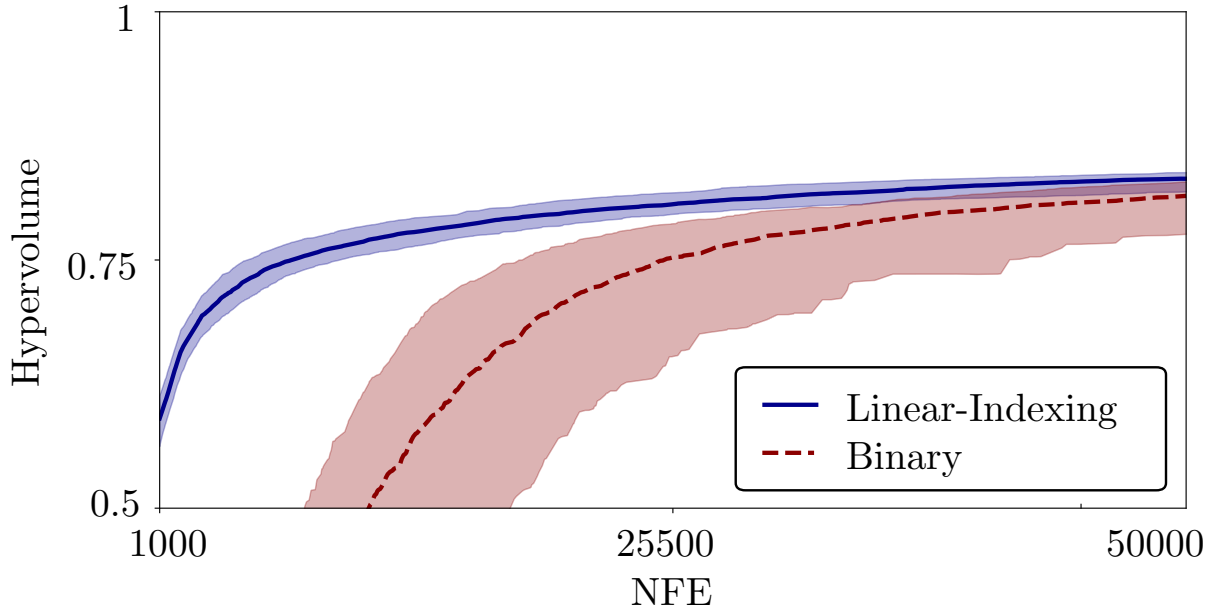


Figure 4.4: Runtime dynamics: Hypervolume over function evaluations (NFE) for two problem formulations that only differ in search-space representation (Blue (Scenario 5): Linear-indexing representation; Red (Scenario 7): Binary representation). The used EDO set includes 148 EDOs and is grid-based reduced with a grid size of 500×500 m^2 (cf. Section 4.3.2). The shaded areas represent the variation in hypervolume of all performed random-seed trials and are bounded by the 10% and 90% percentiles. The solid and the dashed lines represent the median for the respective problem formulation.

Pareto-optimal monitoring trade-offs. In contrast, the linear-indexing formulation has a relatively small set of decision variables (31) and better captures feasible monitoring solutions. That is, both representation methods address the power set of EDOs, but linear-indexing representation ignores all EDO combinations that are larger than a specified maximum number. Thus, the linear-indexing formulation has an immense initial advantage compared to the binary formulation. This advantage persists over the entire illustrated NFE range in Figure 4.4.

4.5.2 Linear-Indexing Representation Can Handle Large EDO Sets

Next, we examine the behavior of the linear-indexing representation regarding large EDO sets. Figure 4.5 shows comparative hypervolume dynamics as a function of NFE for alternative implementations of the linear-indexing representation that vary in their different EDO sets. The different EDO sets originate from the fundamental set and several options for search-space reduction. Figure 4.4 already demonstrates that the linear-indexing representation improves both efficiency and reliability of search compared to the binary representation. The linear-indexing representation is successfully addressing large EDO sets without significant losses in MOEA performance. Direct comparison with the binary representation is not tractable given the binary representation is limited to EDO sets with a maximum of about 1,000 EDOs on the available machines. The largest EDO set within our test-case framework, however, possesses more than 200,000 EDOs (i.e., the full fundamental-reduced set labeled as 'Full' in Figure 4.5 and plotted in orange). The other EDO sets used in Figure 4.5 contain 1,500 EDOs or less (cf. Table 4.2) and will be discussed in detail later on.

The 'Full' problem formulation (Scenario 1) is less computationally efficient; its hypervolume converges only slowly. Additionally, the 'Full' problem formulation initiates with weak monitoring solutions with low initial hypervolume performance when compared to the other illustrated problem formulations. The main reason for these drawbacks is that only a small percentage of the 200,000 EDOs is actually useful, and the remaining EDOs are limited in contributing to Pareto-optimal solutions due to the hardness of the optimization problem: the number of bad and dominated solutions within the initial search population is high and the optimization algorithm needs more NFE to identify a pool of promising EDOs.

Despite these shortcomings, the fact that linear-indexing representation can manage such large EDO sets is definitively useful and attractive for the optimization: First, for some optimization problems, the 'Full' EDO set cannot be reduced, or the reduced EDO set is still too large for alternative representations like the binary one. Second, reduction algorithms can be complex to implement, and in some instances, they can be too computationally burdensome. Nevertheless, in terms of effectiveness, a search-space reduction is often advisable. Therefore, in the following, I compare different

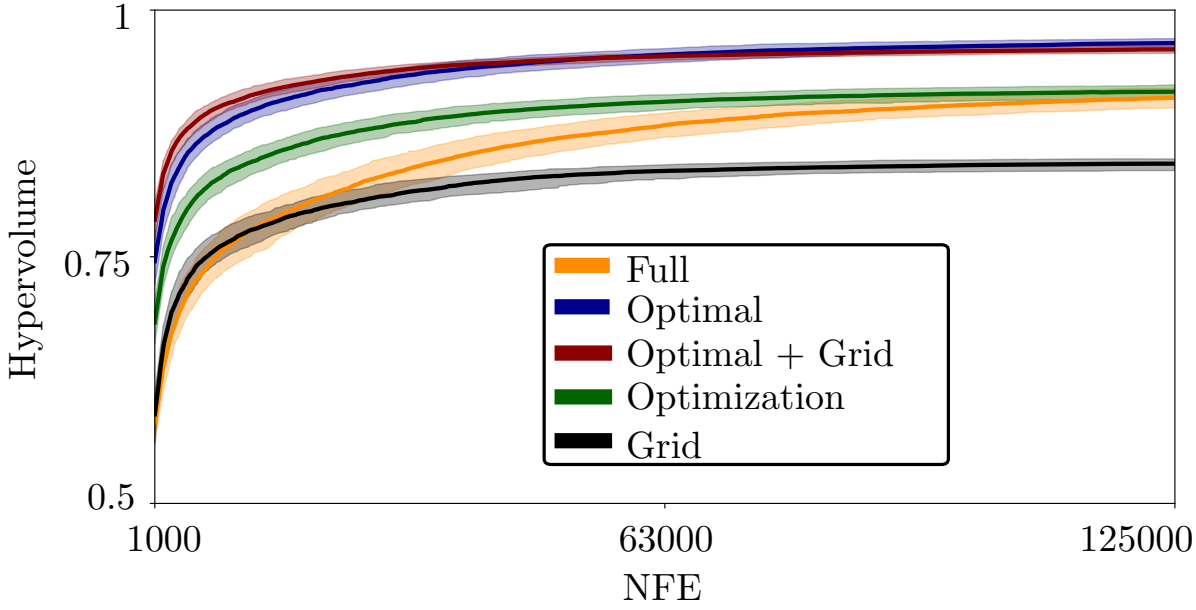


Figure 4.5: Runtime dynamics: Hypervolume over function evaluations (NFE) for five problem formulations that differ in the used EDO set (Scenarios 1-5). All problem formulations are based on the linear-indexing representation. As in Figure 4.4, the shaded areas represent variations in hypervolume and are bounded by the 10% and 90% percentiles. The solid lines represent the median for the respective problem formulation.

reduction methods and the impacts of their reduced EDO sets on search efficiency and effectiveness.

4.5.3 Search-Space Reduction Increases Efficiency

In Figure 4.5, the problem formulations with small EDO sets show a strong increase in hypervolume during the first 5,000 NFE. From this perspective, the reduced EDO sets enable a highly efficient optimization, but for an unbiased evaluation, the invested NFE and/or wall-clock times of the different reduction algorithms must be included to the analysis. Unfortunately, for reduction methods, their wall-clock times and required number of function evaluations are difficult to generalize across problems. Generalization challenges emerge due to (1) the optimization problem, specifically the EDO-set size and EDO properties (e.g., the objective values for single EDOs), and (2) on algorithm choices (e.g., grid discretization in the grid-based method).

For the sake of completeness, the wall-clock times of the reduction algorithms for U_Protect can be found in Table 4.6. For U_Protect all reduction methods were relatively fast, but experiments on different problems have shown that the wall-clock time is subject to large fluctuations, especially for the 'Optimal' reduction method. However, the 'Full' problem formulation needs around 33,000,000 NFE to reach the same hypervolume as the 'Optimal' problem formulation after 500,000 NFE (not illustrated in Figure 4.5). It is very likely that for more complex problems than U_Protect, this discrepancy increases, hence more time can be invested into an effective EDO-set reduction.

However, Figure 4.5 shows that an EDO-set reduction can be worthwhile. Here, the 'Optimal' and the 'Optimal + Grid' formulations are clearly superior and outperform the 'Full' problem formulation in both efficiency and effectiveness. The 'Optimization' problem formulation is, at least, more efficient compared to the 'Full' formulation, but the 'Grid' formulation fails and can neither outperform the 'Full' formulation in effi-

Table 4.5: Hypervolumes for different problem formulations after 1,000, 5,000, 125,000, and 500,000 NFE

Label	1,000	5,000	125,000	500,000
Full	0.582	0.713	0.911	0.934
Optimal	0.746	0.856	0.967	0.974
Optimal + Grid	0.788	0.880	0.960	0.964
Optimization	0.683	0.799	0.917	0.925
Grid	0.590	0.728	0.844	0.850

ciency, nor in effectiveness. In the following, I discuss the reasoning behind the different performances of the presented problem formulations.

The 'Grid' problem formulation (Scenario 5) is based on a grid-based reduced EDO set with a grid size of 500×500 m^2. Following the rules defined in Section 4.3.2, all EDOs within the same grid cell are represented by a single EDO and the others are removed from the EDO set. The final selected EDO subset contains the best compromise EDOs per grid cell, but the selection neglects combinatorial effects with EDOs from the same and other grid cells. Hence, many EDOs are removed that would have otherwise been a part of the Pareto-optimal solutions. Therefore, it is impossible for the 'Grid' formulation to reach high hypervolume values. The limits of its EDO set can be seen in Figure 4.5 (black).

The 'Optimal' and the 'Optimization' reduction strategies (Scenarios 2 and 4) are based on EDO sets that are reduced with the comparison methods (cf. Section 4.3.3). While the 'Optimal' EDO set guarantees to contain all EDOs that are part of Pareto-optimal solutions, the 'Optimization' EDO set only approximates the 'Optimal' EDO set without any guarantee of optimality. The 'Optimal + Grid' EDO set (Scenario 3) is a grid-based reduced subset of the 'Optimal' EDO set. Its representative EDOs are all relevant, but other, also relevant EDOs, are removed by the additional grid-based reduction. From Figure 4.5, it is clear that for the U_Protect test case the 'Optimal' and the 'Optimal + Grid' problem formulations perform significantly better than the 'Optimization' problem formulation. Both reach a similar high final value in hypervolume (cf. Table 4.5), with a slight advantage for the 'Optimal' formulation. This is not surprising considering that the 'Optimal' search space contains all Pareto-optimal solutions, in contrast to the search spaces of the 'Optimal + Grid' formulation. For longer optimization run times, this gap in hypervolume would clearly grow, because the 'Optimal' formulation has (at least) the potential to reach the ideal hypervolume of one. However, the performance of the 'Optimal + Grid' formulation is quite strong, especially compared with the 'Grid' formulation.

The large difference between the 'Grid' and the 'Optimal + Grid' formulation puts the 'Grid' reduction method into a different perspective. Not only the reduction method itself determines the quality of the reduced EDO set, but also the quality of the underlying (to be reduced) EDO set. Using the 'Optimal' EDO set as basis for the 'Grid' reduction enables an EDO selection out of the best EDOs, while the 'Full' EDO set as basis mainly contains bad EDOs. As a conclusion, the 'Grid' reduction method can be useful for a reduction of large but high-quality EDO sets. In contrast, EDO sets with many redundant and/or dominated EDOs should not be reduced by the 'Grid' method.

The last three sections clearly showed the benefits of linear-indexing representation versus binary representation, and the benefits of EDO-set reduction in general. However, the performed hypervolume runtime-dynamic analysis lacks of a differentiated

Table 4.6: Wall-Clock Times of the Reduction Methods used on U_Protect

Optimal	Optimal + Grid	Optimization	Grid
42 s	47 s	30 s	8 s

view of the optimization results, because it averages over the entire Pareto-set approximation. Hence, it cannot identify local strengths and weaknesses in objective space of the approximation sets. But, for instance, outperforming local parts of approximation sets might be of special interest for stakeholders and decision makers, if these parts belong to the subjectively preferred regions of the objective space. The hypervolume analysis also cannot provide any information about the physical space, but the choice for a certain problem formulation is often connected with its impact on the effective physical solutions. For example, in many cases, uncertainty of optimization-relevant input data blurs the optimization results, and decision makers might need an alternative solution for the objectively favorite one. Therefore, in the following two sections, I compare the objective and physical space of different problem formulation approximation sets and show the resulting effects and benefits.

4.5.4 Representation Impacts on Trade-off Analyses

Building on the results of Section 4.5.1, this section investigates the geometrical consequences (the shape of approximative Pareto sets) for estimating and understanding monitoring trade-offs that emerge by specifying the alternative search space representations. The left column of Figure 4.6 shows runtime dynamics (1 s, 3.5 s, 7 s) for the evolving approximated Pareto-fronts for detection probability (DP, x-axis), early-warning time utility (EWT, y-axis) and costs (Costs, z-axis). Figure 4.6 compares the problem formulations for the 'Optimal + Grid' EDO set where the red solutions exploited the linear-indexing representation (Scenario 3), and the green solutions utilized the binary representation (Scenario 6). Blue solutions were found with both representations. For a better visualization, I project the approximation sets on the respective two-dimensional planes. The theoretically ideal point is in the bottom right back corner. Please note that the y-axes only report values between 0.5 and 1 for the early-warning time utility.

After 1 second, the red approximation set of the linear-indexing formulation already covers a broad range of the objective space, while the green approximation set from the binary representation only covers the upper right edge. The poor diversity of the green front is caused by the difficulty of the search space problem as implicit to the binary representation. The probability of finding valid solutions consisting of a maximum of 15 EDOs is low, the probability of finding low-cost monitoring networks (less than 15 EDOs) is even lower. Hence, the binary representation causes exploration of the

objective space to initially be approached from the high-cost upper-right edge in Figure 4.6, representing the high-cost zone of monitoring alternatives. Even the few green binary representation alternatives found early are fully dominated by the broader red linear-indexing front, which very rapidly captures less expensive solutions with equal early-warning time utility and equal detection probability.

Given more time for search (3.5 and 7 seconds), the linear-indexing red front has already translated the space to find a Pareto approximate front, allowing the algorithm shift its focus to capturing a more diverse suite of solutions along the full extent of the objectives' trade-offs. The binary representation's green front also improves with additional search to cover larger portions of the objective space. After 7 seconds, portions of both fronts become equal in the compromise region (blue spheres in the middle of the fronts). The extents of the linear-indexing red front, however, still fully dominate those found with the binary representation. The sparsity of the binary green front in the lower back left of the scatter plot has direct negative impacts for decision support given this region should capture trade-off solutions with low costs, low detection probability, but high early-warning time utility. These solutions are low-cost monitoring networks (only few monitoring wells) that closely monitor selected possible contamination sources.

In the right column of Figure 4.6, both problem formulations are based on linear-indexing representation but exploit different EDO sets. Here, the red front was found using the 'Optimal' reduced EDO set (Scenario 2), and the green front was found using the 'Full' EDO set (Scenario 1). As already seen in the left column, with the linear-indexing representation both approximation fronts explore large parts of the objective space. Over runtime, both converge, but the 'Full' EDO set green front is mostly dominated by the 'Optimal' reduced EDO set formulation. Thus, as already shown in Figure 4.5, the hypervolume of the 'Optimal' red front has a larger extent compared to the hypervolume of 'Full' green front. In contrast to the binary problem formulation from the left column of Figure 4.6, here a shorter wall-clock time would not lead to significant consequences. However, it shows the benefits of the 'Optimal'-reduction methods for the approximation. From Figure 4.5, it can be seen that the 'Optimal' red approximation set reaches the same hypervolume after 8% of the illustrated NFE ($\sim 15,000$ NFE) as the 'Full' green approximation set after 125,000 NFE. That is, to get similar optimization results, decision makers can decide between long wall-clock times and using the 'Full' problem formulation, or reducing the 'Full' EDO set with the 'Optimal'-reduction method and also reduce wall-clock time of the optimization. These differences would be expected to matter more as monitoring problems increase in size and shift towards rapid response operational warning systems for water supplies.

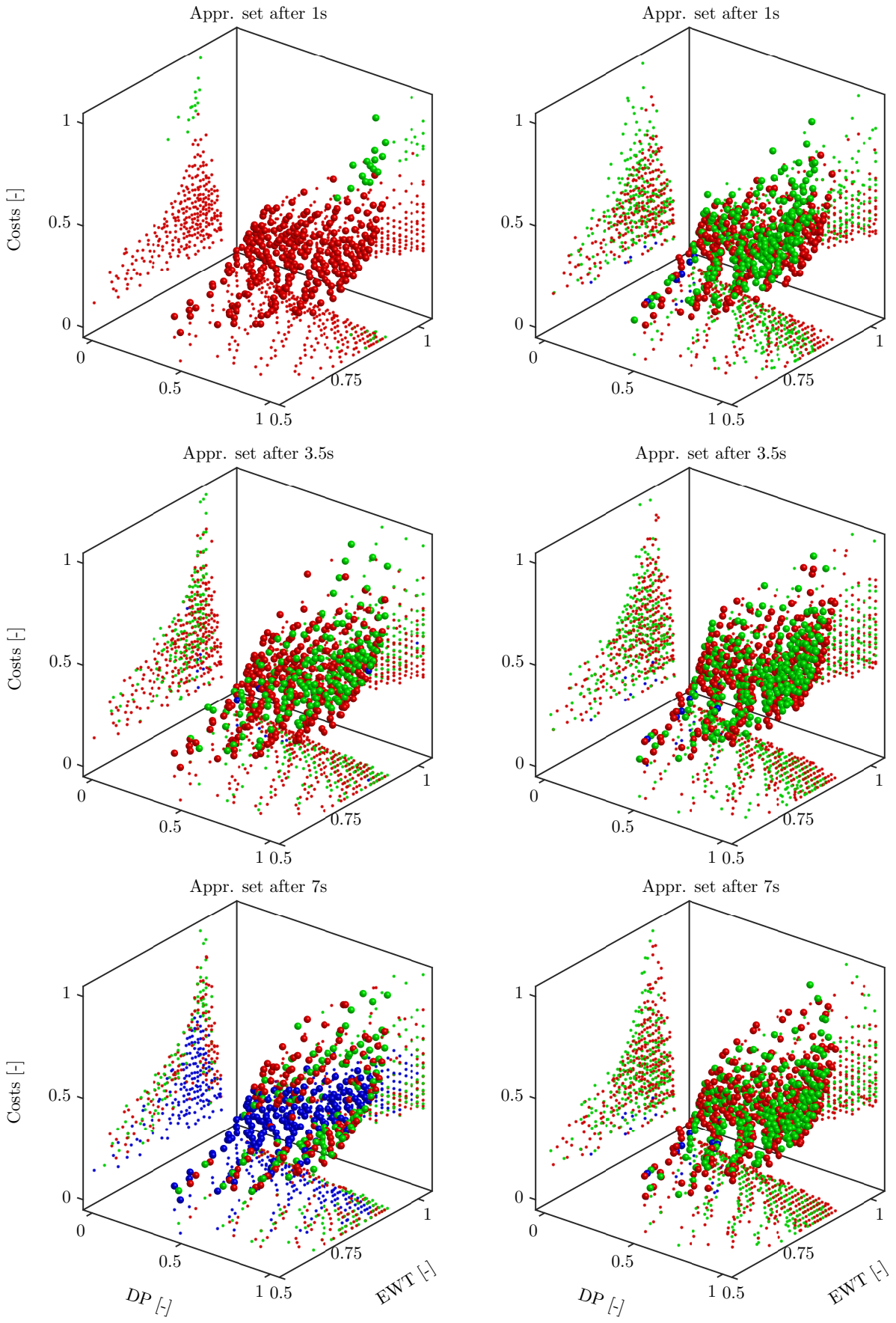


Figure 4.6: Figure description on next page.

Figure 4.6: Runtime dynamics of approximation sets. **Left column:** 'Optimal + Grid' reduced EDO set with linear-indexing representation (red, Scenario 3) and binary representation (green, Scenario 6); Mutual solutions of both problem formulations in blue. **Right column:** 'Full' EDO set (green, Scenario 1) and 'Optimal' EDO set (red, Scenario 2), both with linear-indexing representation; Mutual solutions of both problem formulations in blue.

4.5.5 Monitoring Impacts from Problem Formulation

Up to this point, I have investigated the hypervolume runtime dynamics and the evolution of approximation sets in the objective space. In this section, I demonstrate the endpoint monitoring effects of the different candidate problem formulations. Figure 4.7 shows the resulting monitoring-well positions (blue bars) for the U_Protect test case (cf. Figure 3.5). The height of the blue bars is proportional to the frequency a position occurred in the solutions that compose the approximation sets. The frequency is normalized by the number of solutions per set. Here, I compare three problem formulations that are based on the linear-indexing representation (Scenario 1: 'Linear, Full', Scenario 2: 'Linear, Optimal', Scenario 5: 'Linear, Grid', sub-figures (A), (B), (C) in Figure 4.7), and one problem formulation based on the binary representation (Scenario 7: 'Binary, Grid', sub-figure (D)).

Figure 4.7 clearly illustrates similarities and differences of the approximation sets for the four considered problem formulations. The 'Linear, Full' and 'Linear, Optimal' problem formulations (4.7A) and (4.7B) have similar spatial distributions and frequencies of possible monitoring-well positions. For long wall-clock times, distributions of the monitoring locations become very close as expected, since the 'Linear, Optimal' formulation in (4.7B) is based on the 'Optimal' EDO subset (4.7A). The main differences between the two problem formulations is that (4.7A) includes clustering spots of possible positions and (4.7B) can represent these spots by single positions because equally performing positions for the same possible contamination sources were thinned out with the 'Optimal'-reduction method. The position distributions for the 'Linear, Grid' (4.7C) and the 'Binary, Grid' (4.7D) formulations are less dense and include void spaces in the middle of the original EDO set. Obviously, the optimization was forced to follow a different strategy for the grid-based reduced EDO set, because the available positions in the middle area (void space) could not contribute as well as other positions. Therefore, the selected positions form two barriers, in front of the remote possible contamination sources and in front of the closer sources. As a conclusion, if using the 'Grid' problem formulations ((4.7C) and (4.7D)), decision-makers would unintentionally restrict the number of promising monitoring-well positions. In summary, neither the binary representation nor the grid-based reduction method can satisfy the complex requirements of this multi-objective optimization problem. They are limited in

the number and decision-relevant scope of the alternative solutions they yield.

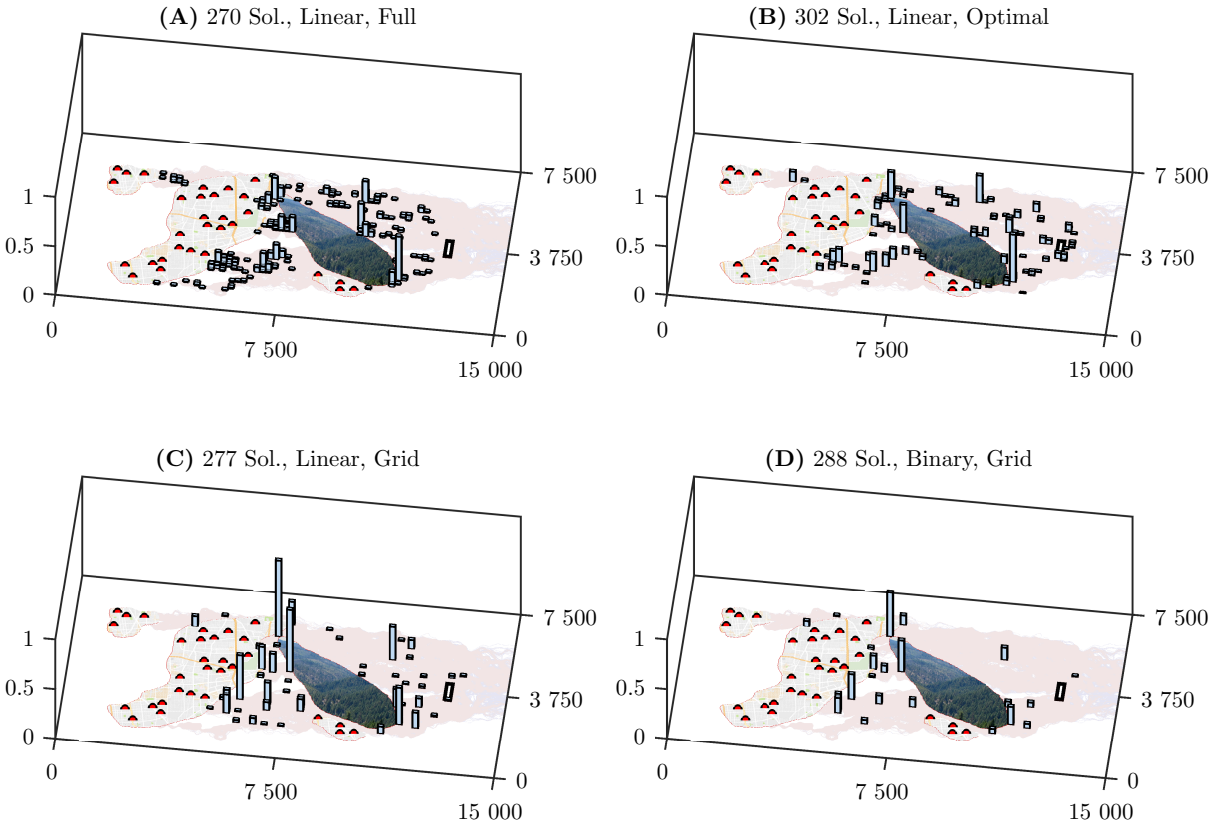


Figure 4.7: Distribution and frequency (blue bars) of monitoring-well positions from approximation sets on U_Protect for four problem formulations: (A) Scenario1: 'Full, Linear-Indexing Representations', (B) Scenario 2: 'Optimal, Linear-Indexing Representations', (C) Scenario 5: 'Grid, Linear-Indexing Representations', (D) Scenario 7: 'Grid, Binary Representations'.

Summarized, the new methods introduced here, i.e., the linear-indexing representation and the optimal reduction, produce superior results compared to other typically used methods. Their performance is more efficient, effective, and reliable.

4.5.6 Reduction Effects When Accounting for Uncertainty

So far, I did not consider uncertainty within the U_Protect test case. Considering Uncertainty, however, is important for robust optimization results and increases their reliability as shown in Figure 4.8. Figure 4.8 compares three sets of solutions for the U_Protect test case in the objective space with detection probability on the x-axis, early-warning time on the y-axis, and costs on the z-axis. The theoretical optimum in this il-

lustration is the backside lower-left corner. The orange set is a Pareto-front approximation for a single scenario, i.e., for U_Protect without considering uncertainty. The blue set is a Pareto-front approximation for multiple scenarios (around 300), i.e., with considering uncertainty in U_Protect. The red set illustrates the performance of the orange solutions (without uncertainty) on the U_Protect test case with uncertainty considered. For a better understanding of the three sets, all candidate solutions are projected on the two-dimensional planes. Figure 4.8 clearly shows that the optimization problem becomes more challenging when considering uncertainty (blue): the candidate solutions cannot reach the same quality level as the solutions without considering uncertainty (orange). It also illustrates the quality loss in performance when ignoring uncertainty (red vs. blue). Consequentially, uncertainty consideration helps prevent frustration due to high but unfulfilled expectations (in Figure 4.8: expected performance (orange) vs. effective performance (red)). Although the blue set totally outperforms the red set, the blue set is located in a critical objective-space range where any deviation from the real Pareto front should be as small as possible. The actual goal attainment neither exceeds 50% in early-warning time, nor in detection probability. That is, the goal attainment is low and the potential loss of, e.g., 1% detection probability is relatively larger compared to the case without considering uncertainty (orange). Concluding, with uncertainty consideration, an effective optimization is particularly desirable.

We already saw that the linear-indexing representation of the search space is preferred over the binary representation and, with the optimal reduction method, the optimization provides superior results. With uncertainty consideration, however, the base EDO set becomes more challenging for two reasons:

1. As I use the arithmetic mean to aggregate over all scenarios (cf. Section 3.4), the final EDO set enlarges compared with the single-scenario EDO set (in our case: 219,122 vs. 436,986 EDOs), because the possibly contaminated area increases with the number of possible flow paths for the possible contaminations.

2. Due to the averaging effect, the aggregated objective values of the single EDOs become similar (the EDO set becomes indifferent) and less EDOs are clearly dominated, hence can be removed.

While the linear-indexing representation is not affected by the more complex 'uncertainty' EDO set, the optimal reduction is presumed to be less effective and less efficient (less EDOs to remove, more EDOs to compare). A measure to increase both efficiency and effectiveness of the reduction algorithm is to use the boosted optimal reduction method introduced in Section 4.3. The boosted version homogenizes the objective space of the EDO set and more EDOs can be removed. Here, we use the round-function as homogenizer. Differently to the pure optimal reduction method, with the implemented homogenization, it cannot be guaranteed anymore that the EDO set contains all Pareto-optimal solutions. However, compared to the pure optimal reduction method, the boosted one totally outperformed (cf. Table 4.7): it is much faster and

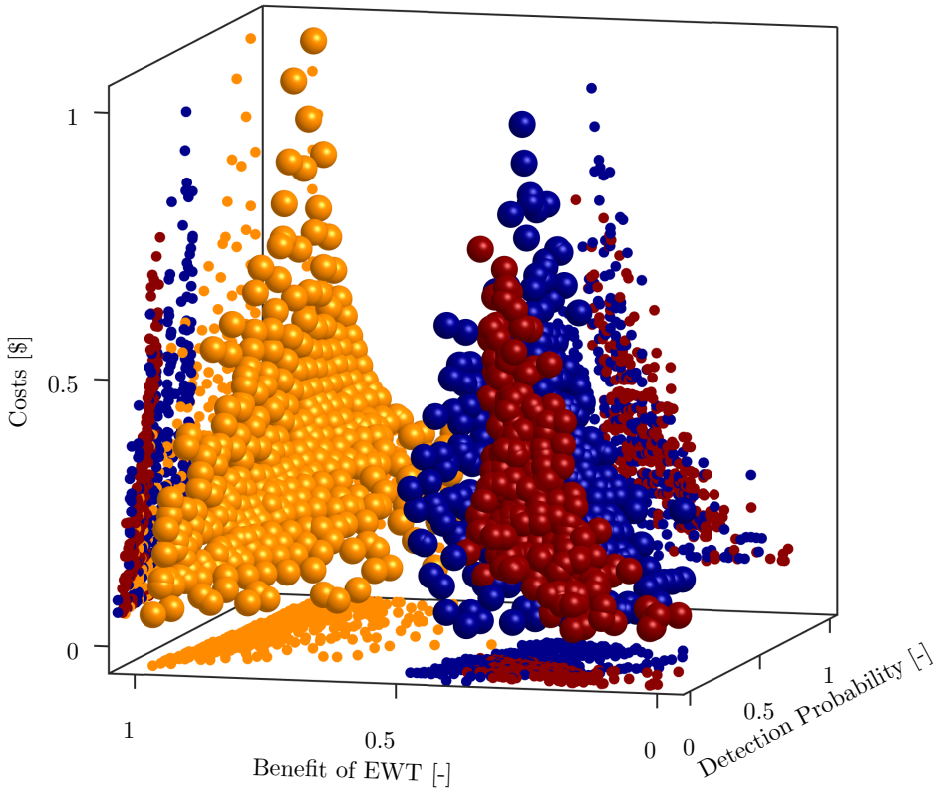


Figure 4.8: Pareto-approximations sets in the objective space. Orange spheres: approximation set without considering uncertainty. Blue spheres: approximation set with considering uncertainty. Red spheres: performance of the orange solutions (without uncertainty) when confronted with uncertainty. Dots: projection of the sets on the respective two-dimensional planes.

more effective. The question is, whether the missing guarantee of optimality has a significant drawback on the optimization.

Therefore, Figure 4.9 shows again the hypervolume dynamics as a function of NFE for the three alternative problem formulations with uncertainty consideration (cf. Table 4.3). All are based on the linear-indexing representation. The 'Full' problem formulation (Scenario 8) uses the base EDO set, the 'Optimal' formulation (Scenario 9) is based on the optimal-reduced EDO set, and the 'Optimal + Round' EDO set (Scenario

Table 4.7: Efficiency and Effectiveness of the Pure and Boosted Optimal Reduction Method

Method	Wall-Clock Time	Total Reduction	Reduction Rate
Pure Optimal	$\sim 10,800$ s	367,784	$\sim 84\%$
Boosted Optimal	~ 150 s	436,185	$\sim 99.8\%$

10) is based on the optimal reduced EDO set by the round-function boosted optimal-reduction method (cf. Section 4.3). Both reduced problem formulations are more efficient compared with the 'Full' problem formulation. The latter needs around 10^7 NFE to reach the same hypervolume as the 'Optimal' formulation after 6×10^5 NFE. After 5×10^7 NFE, it reaches the same hypervolume as the 'Optimal + Round' formulation after 5×10^5 NFE. Obviously, the size of the different EDO sets is the main player for the optimization performance. Although the 'Optimal + Round' formulation cannot guarantee to provide all Pareto-optimal solutions, it clearly outperformed the other two formulations.

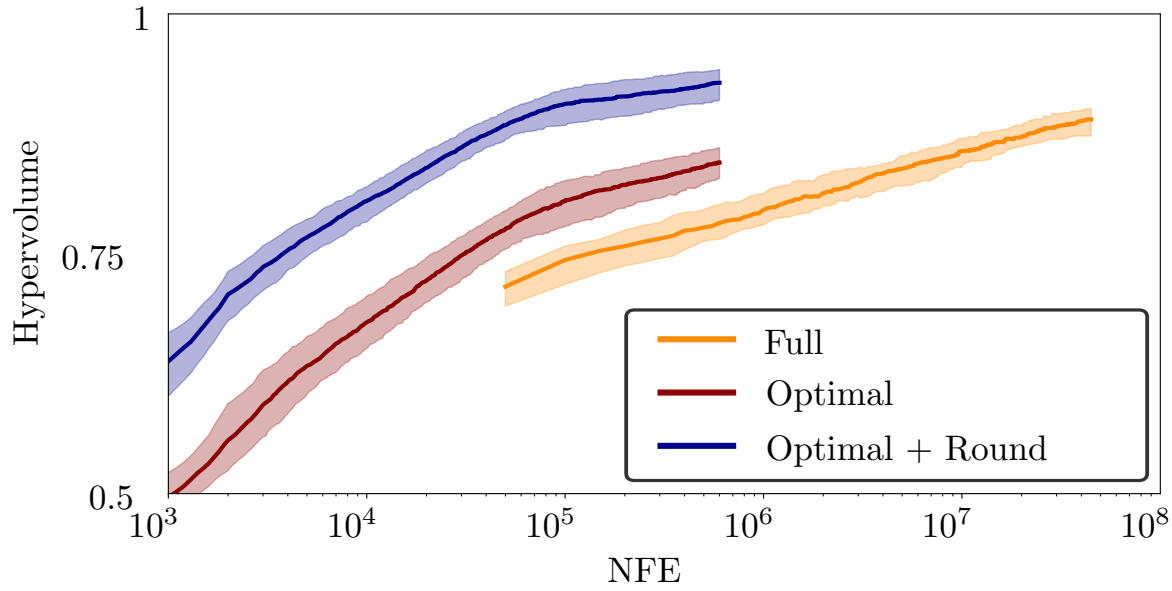


Figure 4.9: Runtime dynamics: Hypervolume over function evaluations (NFE) for three problem formulations (Scenarios 8-10) that differ in the used EDO set. All problem formulations are based on the linear-indexing representation. As in Figure 4.4, the shaded areas represent variations in hypervolume and are bounded by the 10% and 90% percentiles. The solid lines represent the median for the respective problem formulation. Due to the large number of function evaluation for the 'Full' problem formulation (please observe the logarithmic scale), data are available every 50,000 NFE.

Summarized, the 'Optimal + Round' formulation showed the best optimization performance and the EDO-set reduction is highly efficient and effective. Especially for short optimization wall-clock times, the rigorous but slow optimal reduction method is not advisable. The optimal reduction method works best for (1) heterogeneous EDO sets, i.e., with EDOs that can clearly be distinguished in their objective values, and (2) for homogeneous EDO sets, i.e., with EDOs that have the same objective values. Then, the optimal reduction would provide small EDO sets that are ideal for producing superior results with MOEAs. When considering uncertainty, however, the aggregation effects make the EDO set indifferent, i.e., with EDOs that are similar to each other, and the

reduction is less effective. In this case, the reduction can be supported by a homogenization method that equalizes similar EDOs.

4.6 Summary and Conclusions

This chapter approaches challenges that emerge when addressing discrete multi-objective optimization problems with large potentially severe combinatorial scaling in their decision spaces. Specifically, I (1) developed an efficient and reliable search-space representation, the linear-indexing representation, that can handle large decision spaces, and (2) I contribute a search space reduction method, the optimal reduction, that provides speed search while ensuring the attained multi-objective trade-offs are improved high-quality approximations to the true Pareto-optimal solutions. As benchmark, I use the U_Protect test case together with the general optimization problem (formulated in Section 3.2) with three objectives: (1) detection probability (to be maximized), (2) early-warning time (to be maximized), and (3) costs (to be minimized).

I compare the linear-indexing representation with the common binary representation. My analysis shows that:

1. Linear-indexing representation enables reliable and efficient optimization runs that clearly outperform the binary representation over time.
2. Unlike binary representation, the linear-indexing representation can handle large problems.

I also compare our optimal search space reduction method with optimization-based and grid-based reduction methods. The grid-based reduction methods can be applied to continuous and discrete problems and are independent of the objective functions, the optimal and optimization-based reduction methods are limited to discrete problems and monotonous objective functions. In the context of the U_Protect test case, I show that:

1. Optimal reduction strategies produce the best results but need an additional algorithm.
2. The grid-based reduction is highly effective but a brute-force approach with the potential of losing candidate Pareto-optimal monitoring alternatives.
3. Optimal reduction gets performance problems for highly indifferent decision spaces, but can be boosted with methods that homogenize them.

The overarching conclusion is that, for discrete problems, linear-indexing representation is the key for a reliable and efficient optimization performance. Also, whenever feasible, my proposed optimal reduction helps to improve quality and endpoint

value of multi-objective analyses for the design of groundwater monitoring systems for source water protection.

These insights can be transferred to a broad class of optimization problems. The following four conditions must be fulfilled to apply linear-indexing representation and optimal reduction:

1. The search space must be finite, as part of a discrete and finite optimization problem (for representation and reduction).
2. The problem must be a combinatorial problem with a search space that can be represented by a set of single decision elements (for representation and reduction).
3. The optimization problem must be multi-objective (for reduction).
4. The objective functions must be monotonic (for reduction).

The results enable the feasibility of multi-objective optimization might be possible on standard desktop computers in operationally valid runtimes.

Chapter 5

Problems and Solutions: Robust and Reliable Early-Warning Monitoring in Practice

Parts of this chapter have been published in the Journal TRANSPORT IN POROUS MEDIA under the title *Optimization for Early-Warning Monitoring Networks in Well Catchments Should Be Multi-objective, Risk-Prioritized and Robust Against Uncertainty* [19] and in the Journal WATER RESOURCES RESEARCH under the title *Reconnecting Stochastic Methods with Hydrogeological Applications: A Utilitarian Uncertainty Analysis and Risk Assessment Approach for the Design of Optimal Monitoring Networks* Bode et al. (2018b) [15]. I am reusing parts of the text and figures from these publications by the kind permissions of the publishers Springer and American Geophysical Union.

In Chapter 3, I introduced the optimization problem and showed that good trade-off solutions can be found. In Chapter 4, I specified the problem formulation with a focus on search-space reduction and representation. I showed that the linear-indexing representation is significantly more reliable and efficient compared to the classical binary representation. I also introduced possible reduction methods for the search space that can lead to better optimization results and to an easier handling of a large data volume. Although I could increase efficiency and effectiveness of the optimization with the methods introduced in Chapter 4, in practice, there are more fundamental challenges that may prohibit a straightforward application of such optimization approaches.

These challenges already occur in the preparation phase before the optimization. Many academic methods are fundamental and general, but not yet optimized for practical applications. This holds, in particular, for advanced methods in stochastic hydro(geo)logy, uncertainty quantification, and risk analysis. Many academic risk estimation methods emphasize rigorously quantitative results [e.g., 28, 159], which requires input data that are rarely available in practice. For water suppliers, however, it is especially difficult to collect the required data. For instance, the effort is very high to explore the catchment for possible contamination sources and to collect all required characteristics of these sources, and there are difficulties in evaluating their respective risk. Additional difficulties are caused by limitations in time and computing power for transport simulations, especially when uncertainty in hydro(geo)logy and/or location uncertainty of possible contamination sources is considered.

In this chapter, I present methods and strategies for tackling the following specific challenges:

- Water suppliers usually do not have access to high-performance computing clusters, while transport simulations and uncertainty quantification are often computationally demanding. Thus, water suppliers need strategies that are frugal enough to perform such tasks on standard desktop computers (Sections 5.1 and 5.3).
- Especially in urban areas, too many possible contamination sources could exist to be screened effectively. Therefore, water suppliers need to distinguish relevant from irrelevant sources to reduce the number of considered sources (Section 5.2). Still, the number of relevant sources might be large, hence they also need effective methods to downsize the number of relevant sources during the transport simulation (Section 5.5).
- Possible contamination sources must be treated differently according to the risk they pose for the supply wells. However, it is difficult or even impossible to determine their risk quantitatively. That is, water suppliers need a non-quantitative possibility to prioritize their inventory of possible contamination sources during the optimization (Section 5.4).
- It is very likely that many possible contamination sources exist that are unknown. A coordinated monitoring of these sources is not possible, but in fact, unknown sources are a big issue. Hence, water suppliers need a strategy to account for the need to monitor unknown sources (Section 5.6).

In the following sections, I will briefly explain the specific problems and introduce approaches how I tackle them. For clarification and/or conformation of the used approaches, I will present and discuss supporting results, if it is helpful. At the end of this chapter, I will summarize the main conclusions in Section 5.7.

5.1 Simplifying Breakthrough Curve Approximation

As mentioned in Section 2.3, I use the PTRW method in all transport simulations to approximate the spatial and temporal concentration distributions of contaminants. The straightforward way to simulate the related breakthrough curves in a control volume (CV) is to track the number and residence time of the particles that pass the respective CV. Knowing the volume $V\ \left[\mathrm{L}^3\right]$ of the ith CV, and the mass $m_\mathrm{p}\ [\mathrm{M}]$ of each particle j, one can easily approximate the breakthrough curve by counting the particles over time at volume i:

$$c_i\left(t\right) \approx \frac{1}{n_{\mathrm{e},i} V_i} \sum_{j=1}^{n_\mathrm{p}} \delta_{ij} m_{\mathrm{p},j}\,, \tag{5.1}$$

with the effective porosity n_e in V_i, the number of all particles n_p, and the Kronecker delta δ_{ij} that indicates whether particle j occupied volume i or not (1 or 0). However, this counting technique leads to discontinuous breakthrough curves. For continuous breakthrough curves, more advanced methods must be applied, e.g., kernel density estimation [129].

For PTRW, the approximation error of breakthrough curves decreases proportionally to the square root of the number of particles in each CV [e.g., 84]. That is, for large model domains, millions of particles are needed, and this leads to computationally demanding simulation runs. A common technique for reducing computational costs is to approximate the temporal moments of breakthrough curves, instead of approximating the entire curve [e.g., 77, 104]. Then, the breakthrough curve can be reconstructed with parametric shapes from these moments. The advantage is that temporal moments are more robust against low particle numbers so that the total number of particles can be reduced dramatically. A requirement of this technique is that an analytical parametric distribution function is available that is defined by the temporal moments and meets the general behavior of the real curves (i.e., the shape).

In Section 2.2, I already introduced the inverse Gaussian distribution (IG). The IG is the analytical solution for breakthrough curves of an instantaneous contamination release in uniform and parallel flow with uniform dispersion coefficient. The scaling factor m_0^{IG} $[\mathrm{MT/L^3}]$ of the IG is the zeroth moment of the breakthrough curve, i.e., it matches the time integral over the entire curve. It can be approximated as:

$$m_{0,i}^{\mathrm{IG}} \approx \frac{1}{n_{e,i} V_i} \sum_{j=1}^{n_p} \delta_{ij} m_{p,j} \Delta t_j \,, \tag{5.2}$$

with the residence time Δt_j [T] of particle j in CV V_i. Then, the breakthrough curve can be reconstructed as

$$c_i(t) \approx m_{0,i}^{\mathrm{IG}} IG \,. \tag{5.3}$$

For continuous mass release, the breakthrough curve is the temporal convolution of Equation 5.3:

$$C_i(t) \approx \int_0^t m_{0,i}^{\mathrm{IG}} IG d\tau \,, \tag{5.4}$$

with the contaminant release time τ. Equation 5.4 can be rearranged when substituting the mass $m_{p,j}$ in Equation 5.2 by the constant value $\dot{m}_{p,j} = \int \frac{m_{p,j}}{\tau} d\tau$:

$$M_{0,i}^{\mathrm{IGC}} \approx \frac{1}{n_{e,i} V_i} \sum_{j=1}^{n_p} \delta_{ij} \dot{m}_{p,j} \Delta t_j \,. \tag{5.5}$$

Then, continuous breakthrough curves can be expressed using the cdf of the IG (cf. Equation 2.13):

$$C_i(t) \approx M_{0,i}^{\mathrm{IGC}} \int_0^t IG d\tau = M_{0,i}^{\mathrm{IGC}} \cdot IGC(\tau_m) \,. \tag{5.6}$$

In order to show that the reconstructed breakthrough curves need significantly fewer particles compared to the straightforward density estimation via particle counting, I use a quasi-three-dimensional uniform velocity field ($15\text{m} \times 10\text{m} \times 10\text{m}$), instantaneously release particles at $(x = 5, y = 5)$, and measure the concentration over time at $(x = 15, y = 5)$. The time discretization is 10s and the dimension of the CV is $(0.7\text{m} \times 0.7\text{m} \times 10\text{m})$. Transport-relevant parameters can be found in Table 5.1. I also compare the reconstructed curves with an analytical solution for this problem to better show how the reconstructed curves are affected by particle reduction. Therefore, I use the depth-integrated analytical solution of the concentration distribution over time at location (x, y) that is given by Equation 5.7 [e.g., 8]:

$$c\left(t, x, y\right) = \frac{m}{B} \frac{1}{4\pi t \sqrt{D_{\mathrm{L}} D_{\mathrm{T}}}} \cdot \exp\left(\frac{\left(-x + vt\right)^2}{4 D_{\mathrm{L}} t} - \frac{y^2}{4 D_{\mathrm{T}} t} \right) , \tag{5.7}$$

with the released mass m, aquifer depth B, time t, velocity v, and longitudinal and transverse dispersion coefficients D_{L} and D_{T}.

Figure 5.1 shows four different approximations of the breakthrough curve with different numbers of particles. It clearly indicates that the moment-based breakthrough curves (green line) are less affected by small particle numbers than the particle-counting-based approximation. The deviations between the moment-based and the analytical curves are small for each case, while the particle-counting-based curve is not reliable with 10000 particles or less.

Although the scaled IG does not generally hold for realistic cases (e.g., with non-uniform flow), I chose this reconstruction method in general over the particle-counting method for two reasons:

1. Computational power is limited in practice, hence practitioners need robust and efficient methods.
2. Many groundwater models are built with standard software suites like ModFlow and use relatively simple zonations to represent heterogeneities. Within the zonations the assumption of uniform flow is acceptable.

Table 5.1: Transport-relevant parameters.

Initial mass m [kg]	Velocity v [m/s]	Molecular diffusion coefficient D_{m} [m^2/s]	Longitudinal dispersivity α_ℓ [m]	Transverse dispersivity α_{t} [m]
0.01	0.01	1×10^{-9}	1	0.1

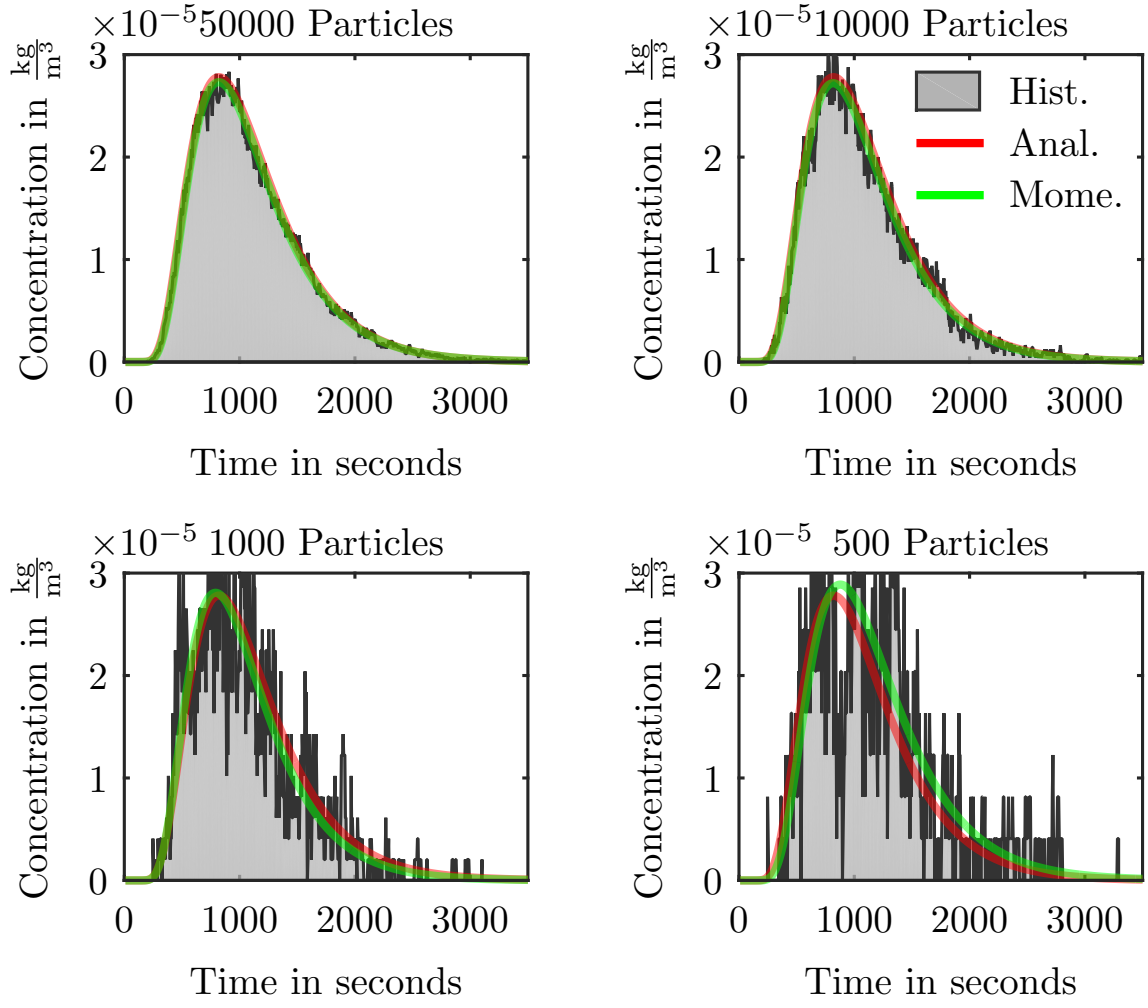


Figure 5.1: Comparison of particle-counting-based breakthrough curves via histogram (gray) with reconstructed curves (green) and the analytical solution (red).

5.2 Defining the Catchment

In many countries, groundwater protection zones restrict activities and storage of pollutants that are hazardous to water, but they cannot eliminate all water-hazardous pollutants, such that an inventory of possible groundwater contamination sources remains (see discussion in Chapter 1). These possible sources need to be monitored effectively to protect the water supply wells from contaminations if they fall into by the well catchment. Protection zones and actual well catchment often differ starkly in extent and location. Although both are based on the hydrogeological conditions [e.g., 46], the definition of protection zones is often practically and/or politically constrained. An al-

ready existing highway, for example, would not be rerouted if it goes straight through a currently defined catchment.

To provide clarity, as first and most important step, I propose to approximate the real well catchment (using backward transport simulation [122]) for four reasons:

1. The catchment could be just partly covered by the protection zone. Contamination sources outside the protection zone, but inside the catchment, still put the production wells and the produced water at risk.
2. Parts of the protection zone might not coincide with the actual catchment. Then, there may be contamination sources within the protection zone that do not affect the production wells, because they are not within the catchment. They are irrelevant for the water supplier.
3. Using backward transport simulation, one can estimate how much contaminant must be released at a certain location to still affect the production wells. Possible sources that cannot release the required mass of a contaminant are irrelevant for the water supplier, even if they fall into the catchment.
4. Using backward transport simulation, one can estimate the travel times from each location towards the production wells. Possible contamination sources that need longer than a certain time threshold to affect the production wells might be again irrelevant for the water supplier.

In summary, defining the well catchment, critical masses and travel times can guide water suppliers in their search for relevant possible contamination sources. Focusing only on the relevant sources identified in a reverse transport simulation reduces the computational demand of the subsequent forward transport simulations from all possible contamination sources to the supply wells. From the economic point of view, identifying irrelevant sources prevents the water supplier from installing monitoring wells that are useless.

To implement backward simulation, I invert the velocity field and solve a reverse transport problem (cf. Section 2.3) that starts at the production well. The resulting backward plume identifies the transport-relevant catchment of the well that needs to be monitored. Figure 5.7 in Section 5.6 exemplary shows relevant (green, yellow, and red dots) and irrelevant (gray dots) possible contamination sources as identified by the backward transport calculation. The optimized monitoring network (black crosses) only focuses on the relevant sources.

5.3 Uncertainty and Robust Optimization

When looking at the relatively thin predicted plumes occurring in Z_Based in Figure 5.3, it is apparent that the "real-life" performance of optimized early-warning mon-

itoring networks (compared to the performance predicted during the optimization) is highly sensitive to predictive uncertainty. Therefore, in Section 3.4, I suggested to include a corresponding representation of uncertainty during the optimization in order to achieve robustness. I call a monitoring network robust against a given uncertainty if its performance does not fall below expectation under the corresponding range of possibilities. In the following, I will discuss representations of uncertainty, effects of uncertainty, and the benefits of robust optimization.

Representation of Uncertainty

Geostatistically-based Monte-Carlo (MC) simulation is the standard academic approach to account for statistical uncertainty of aquifer parameters in groundwater models [e.g., 59, 65, 138]. However, in practice, MC simulations are impractical for four reasons. First, they are computationally too expensive to apply to large-scale or complex models typically needed for practical problems. Second, the underlying probability distributions of the uncertain parameters are not easy to define such that they would represent a realistic level of uncertainty. Third, numerical models used by water suppliers are usually built with standard software like ModFlow [74] or FeFlow [160] that lack a convenient interface for running MC simulations. Fourth, it is much easier to operate that type of software with zonation-based parametrizations or with pilot points [e.g., 132, 42] than with geostatistical random fields [e.g., 40, 88].

Despite these limitations, hydrogeologic uncertainty must be considered in the optimization to propose reliable network designs. Discussions with water suppliers led me to represent uncertainty by using well-selected scenarios instead of formal MC analyses. These scenarios have to satisfy two requirements:

1. They have to represent the dominant uncertainties that most influence the location and concentration of possible contaminant plumes.
2. The ensemble of scenarios should be chosen such that they cover the entire virtual contaminated area between the two enveloping plumes of this ensemble (an example of well-selected scenarios is shown in Figure 5.2 for uncertainty in ambient flow direction).

Further criteria for the quality of scenarios are discussed, e.g., by Kosow and Gaßner (2008) [97].

Intense discussions with water providers showed that, for most of the well catchments, the ambient flow direction is the most influential factor. It is controlled by seasonal changes in groundwater recharge and by the pumping rate of the production wells. Catchments for other supply wells were highly influenced by the stages of nearby

rivers. The key challenge is to investigate how these factors control risks to the supply wells so that the water suppliers could select scenarios that best represent their concerns.

For demonstrating and discussing the effects of ignoring uncertainty, I consider three different uncertainty scenarios (S1-S3) on Z_Based for the optimization:

S1: We have an idealistic situation without any uncertainty (Equation 3.3).

S2: We have uncertainty in modeling our system represented by the four hydraulic scenarios listed in Table 3.2, and consider it during the optimization procedure (Equation 3.10).

S3: We assume that we have an idealistic situation, but in reality, the used model is subject to uncertainty (represented by the four hydraulic scenarios listed in Table 3.2).

The Effects of Uncertainty

Figure 5.3 shows the uncertainty scenarios S1 and S2 introduced above on two selected monitoring networks, one optimized without considering uncertainty (scenario S1, black cross mark), and one optimized considering uncertainty (scenario S2, white plus mark). The black-crosses network is the so-called best compromise solution [e.g., 157], which is defined as the one Pareto-optimal solution with the smallest Euclidean

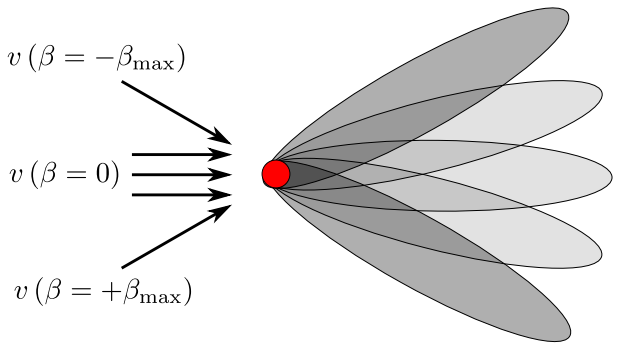


Figure 5.2: Example of well-selected scenarios for uncertainty in ambient flow direction $v\,(\beta)$. Assumed main flow direction is from the left to the right $(\beta = 0)$. **Red dot:** Possible contamination source. **Dark-gray shaded plumes:** Possible contamination plumes that would occur for the two selected extreme scenarios $\beta = \pm\beta_{\max}$. **Light-gray shaded plumes:** Possible contamination plumes that would occur for scenarios with smaller deviations in the flow direction from the main flow direction. The entire scenario ensemble represents the possible contaminated area assuming an expected maximum uncertainty.

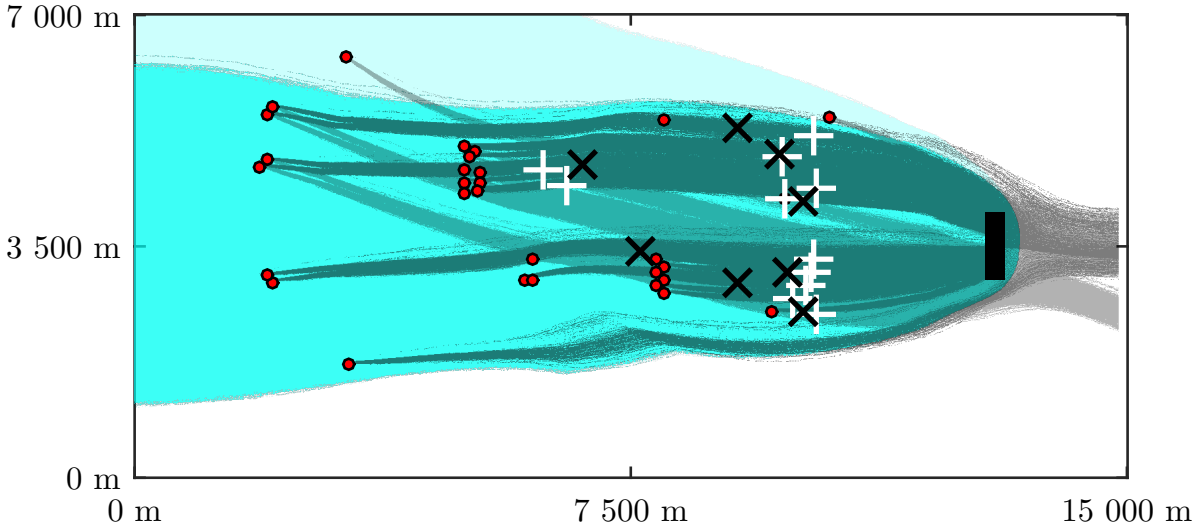


Figure 5.3: Two selected optimized monitoring networks from the basic multi-objective optimization without uncertainty (black cross mark) and with uncertainty (white plus mark) (see Equations 3.3 and 3.10 in Sections 3.2 and 3.4). Entire blue area: well catchment from all hydraulic scenarios; Dark blue area: well catchment from hydraulic scenario 1. Black rectangle: position of the well gallery; Red points: possible contamination sources. Gray dot clouds: visualization of the simulated plumes from each possible contamination source for all hydraulic scenarios; Dark gray dot clouds: simulated plumes for hydraulic scenario 1. The different hydraulic scenarios can be found in Table 3.2

.

distance to the idealistic unattainable optimum in normalized axes (cf. Chapter 3). Through optimized positioning of only eight monitoring wells, it attains a detection probability of 91% and an early-warning utility of 0.73. This relatively high performance at relatively low costs can be achieved by two strategies:

1. monitoring the well catchment at positions where several predicted contaminant plumes coincide and neglecting possible contamination sources that have clearly separated plumes;
2. by avoiding to attain early warning functionality by covering the most remote sources.

The network considering uncertainty has a similar performance in detection probability and early-warning time, but for reaching these goal attainments it needs eleven monitoring wells (i.e., three additional ones). It follows the same strategies as the network discussed above, but the density of each monitoring-well group increased. The relatively thin plumes only overlap in the close vicinity of the production wells. That is, demanding the same number of monitoring wells when considering different hydraulic scenarios would push the monitoring network towards the production well and early-warning time would suffer.

The Benefit of Robust Optimization

Figure 5.4 shows the obtained fronts for the three uncertainty scenarios S1-S3. A suitable measure of robustness is the difference between the optimized predicted performance and the particular performance obtained after actual installation. For better comparison within a single figure, it shows only projections onto the cost/time plane. Scenario S1 (black) performs the best, compared to the other fronts. However, in practice, a scenario without uncertainty does not exist. Considering uncertainty (scenario S2), the results (red) are more modest in their claims of performance during the optimization compared to those without uncertainty (black). The results are less optimistic because it is harder to guarantee a reliable and early detection of contaminant plumes when admitting that one does not know their exact locations and mutual overlaps. Ignoring uncertainty (scenario S3) leads to differences between the expected performance (black) and the actual performance in simulated virtual reality (green), i.e., evaluated for and averaged over all considered hydro(geo)logic scenarios. In theory, the most expensive solution (from the black front) has a close-to-perfect benefit in early-warning time. Compared to that, the same solution performs poorly in virtual reality (green).

Compared to the green front, one could either find monitoring networks with the same performance in early-warning time but with fewer costs, or monitoring networks with

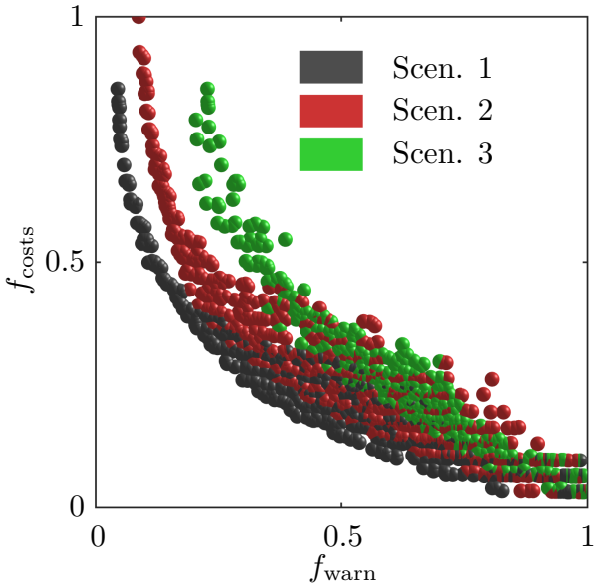


Figure 5.4: Projection of Pareto fronts for three multi-objective optimization scenarios. Scenario 1: Black front from an optimization that does not consider uncertainty. Scenario 2: Red front obtained when optimizing the performance on average over four hydraulic scenarios. Scenario 3: Green front illustrates the performance of the non-robust optimization results on average over the four hydraulic scenarios.

a much better performance with identical costs. Obviously, the black solutions are only optimal for one hydraulic scenario, but not for the considered uncertainties.

The difference in early-warning time between the black front and the red front at given cost values can be interpreted as the costs of having (and considering) uncertainty. The difference between the black front and the green front at given cost values can be interpreted as costs of not considering uncertainty, and these costs are larger than the costs of considering uncertainty. Without any surprise, all scenarios agree that one cannot expect any performance at the limit of no costs.

As a conclusion, considering uncertainty in the optimization is in fact required, because otherwise, one will obtain overly optimistic model-based results that do not perform as desired when installed in reality. The hydraulic scenario analysis performed here is very crude and serves mainly for illustration. In practice, it would be advisable to also account for the other sources of uncertainty mentioned in Section 2.4 and to work with a much larger number than the four hydraulic scenarios from Table 3.2, following the rules of well-selected scenarios discussed above. Using well-selected scenarios, one can adopt more efficient approaches that explicitly consider high-risk scenarios. This can build on the years or even decades of experience of water suppliers to propose a set of representative scenarios.

5.4 Risk Prioritization of Possible Contamination Sources

The optimization problems formulated in Chapter 3 and in the previous sections of Chapter 5 do not yet prioritize according to the severity of possible contamination sources. A possible contamination source can be more or less severe depending on the type of the hazardous activity, stored or handled contaminants and contaminant masses, compliance with applying regulations and so forth. A common approach to prioritize possible contamination sources is to assign a (statistically motivated) weight to each source that reflects factors like the frequency of accidents, the stored contaminant mass, mobility, and toxicity. Provided that all required data are available to calculate these weights through quantitative and probabilistic risk estimations [e.g., 50, 52, 60], this approach would be statistically rigorous. Unfortunately, it is often impractical, or sometimes even impossible to get all relevant data concerning the possible contamination sources (e.g., the probability of failure of a nuclear power plant). Hence it becomes impossible to conduct the probabilistic analyses required to define quantitative weights.

Therefore, following concepts of qualitative risk assessment and to provide a practical possible contamination-source prioritization that also supports the individual concerns of water suppliers, I categorize all possible contamination sources into the classes *severe*, *medium* and *almost tolerable* [e.g., 26, 78]. This categorization can be visualized through the color scheme of traffic lights, i.e., as red, yellow and green. It expresses the prioritization preferences for monitoring in the sense that "red" possible contamination sources should be monitored with first priority, "yellow" ones with secondary priority, and "green" ones with the smallest priority.

In order to accommodate this categorization in the optimization problem, I introduce separate objective functions for the detection probability and for the early-warning time of each risk category. Because individual monitoring wells can cover possible contamination sources from several risk categories, the cost function cannot be separated. Thus, the original problem with three objectives

$$\mathbf{d}_{\text{opt}} = \underset{\mathbf{d}\in\mathbf{D}}{\arg\min} \left(f_{\text{detect}}, f_{\text{warn}}, f_{\text{cost}}\right) \tag{5.8}$$

obtains a total of seven objectives:

$$\mathbf{d}_{\text{opt}} = \underset{\mathbf{d}\in\mathbf{D}}{\arg\min} \left[f_{\text{det}}^{(\text{red})}, f_{\text{warn}}^{(\text{red})}, f_{\text{det}}^{(\text{yellow})}, f_{\text{warn}}^{(\text{yellow})}, f_{\text{det}}^{(\text{green})}, f_{\text{warn}}^{(\text{green})}, f_{\text{cost}}\right]. \tag{5.9}$$

Although the optimization became more complex, the benefit of this approach is that water suppliers can prioritize the possible contamination sources following their own individual assessment, expert opinion, or public concerns. This is possible without the need for a difficult and costly acquisition of data to calculate quantitative risk weights.

For demonstrating and discussing the qualitative risk prioritization, I present optimization results of the optimization problem according to a robust formulation of Equation 5.9, i.e., working with a risk categorization and uncertainty. In this optimization, I categorize the inventory of possible contamination sources of Z_Based into severe, medium and almost tolerable sources, and compare the results to those obtained without risk categorization according to Equation 3.10 (see Section 3.4).

In Figure 5.5, I compare a selected Pareto-optimal monitoring network from the risk-prioritized formulation (black cross marks) with a selected one obtained from Section 5.3, i.e., without risk prioritization. The prioritized solution was selected to achieve a 100% detection probability and a large early-warning time utility (here: 86%) for all severe possible contamination sources (red ones) while maintaining the number of monitoring wells at eleven, i.e., using the same number as the non-prioritized solution. This results in a reduced detection probability and early-warning utility across the remaining risk inventory. The key performance data of both networks can be found in Table 5.2. By comparison of these two networks, one can see that two additional severe possible contamination sources (red) are now monitored (one at [4km, 1.5km] in domain coordinates and one at [7km, 3km]). One monitoring well roughly at [7km, 4km] has been removed, because it mainly served to monitor medium and almost tolerable (yellow, green) possible contamination sources. Other significant changes appear around [11km, 4.5km], where a group of three monitoring wells was replaced by

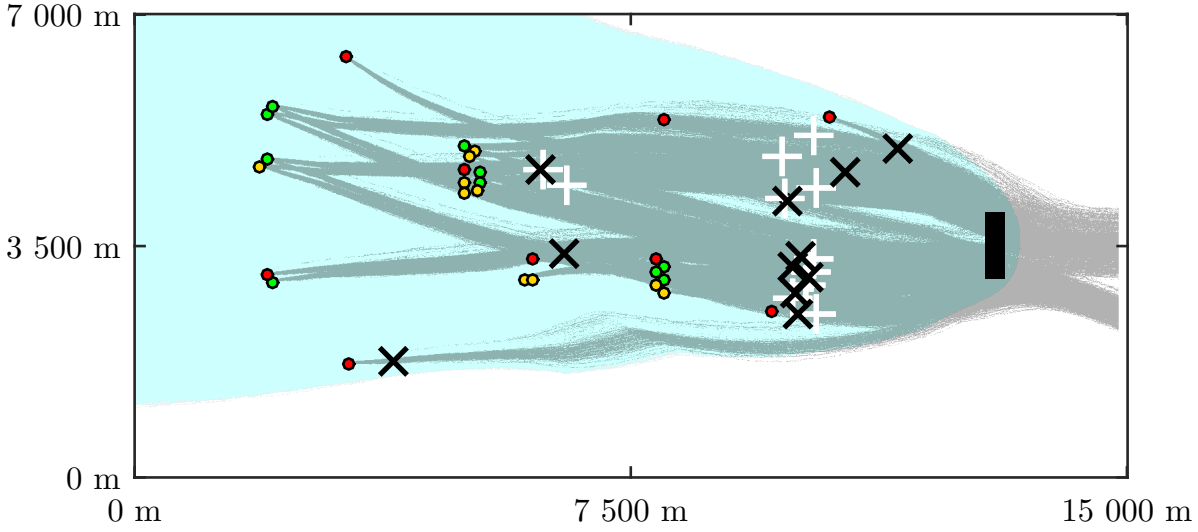


Figure 5.5: Comparison of a risk-prioritized solution from the robust multi-objective optimization with a solution without risk categorization. Black cross marks: prioritized monitoring network; White plus marks: network without prioritization; Red, yellow, green dots: prioritized possible contamination sources; All other symbols and colors: same as in Figure 5.3.

two monitoring wells that coincide better with the possible plume paths of only severe possible contamination sources (red).

There are several strategies available for decision makers to screen through the list of Pareto-optimal results. One option is to screen through the possible compromises between overall coverage and prioritized coverage at fixed costs. A second option is to start with perfect coverage of severe sources, and then to investigate the additional costs and willingness to pay for additional coverage of medium and almost tolerable sources. With this, decision-makers are able to make the necessary compromises between the performance for different risk categories under budgetary constraints in an informed and transparent manner. They do not have to quantify weights before the optimization but can see the consequences of their preferences directly at the Pareto-approximation front.

5.5 Areas with Many Possible Contamination Sources

Some possible contamination patterns can be difficult to characterize because they include multiple possible sources distributed in a relatively large area (e.g., cities, landfills, industrial parks, ...). Practically, it can be too expensive to produce a detailed map of each source. Furthermore, in many cases, water suppliers do not have the right to collect all of the information necessary to describe the separate sources. Often, even if the data could be collected, the computing time and memory requirements to treat each source individually in the transport simulations may be prohibitive. For practical application, this type of source area (area with uncertainty in point-source locations) should be handled using a simpler, representative model.

To accommodate this preference, I introduced the concept of virtual representative sources. Instead of accounting for each possible source in, e.g., an urban area, artificial sources are used, closely spaced along a line at the downstream edge of the source area. The underlying idea is that the plumes emanating from any source within the area will cross this downstream boundary. Furthermore, placing the virtual sources as close as

Table 5.2: Key data of the different monitoring networks illustrated in Figure 5.5

Monitoring network	Risk categories	Detection Probability [−]	Early-warning time utility [−]
Non-prioritized	All	0.97	0.79
Non-prioritized	Severe	0.88	0.72
Prioritized	All	0.80	0.66
Prioritized	Severe	1	0.86

possible to the edges of the source area, with reference to the flow lines, provides a conservative estimate of impact. That is, the virtual contaminants represented by the virtual contaminant sources will arrive sooner and at higher concentrations than the actual contaminants from the real possible contamination sources within the source area. In most cases, the difference in estimated warning time is small because the size of the source area is small relative to the distance of the downstream boundary to the production wells. therefore, this simplification provides at least as much protection as would be provided if every real source was modeled individually; however, it offers three practical advantages:

1. water suppliers do not need to undertake time-consuming and expensive data acquisition in urban areas;
2. they can account for possible contamination sources that are uncertain in location within urban areas;
3. the transport simulation requires less computational effort and time to consider the smaller number of virtual representative sources.

5.6 Unknown Possible Contamination Sources

When designing a monitoring network, the most difficult possible contamination sources to consider are the (potentially numerous) sources within well catchments that are not even known to exist. Reasons for this include the limited existence of information on land use in the private or industrial sector, restricted access to existing information due to data privacy policies, known possible sources with unknown locations (from vague historical records) or unforeseeable hazards that fall under the classical category of black swans [158]. These unknown possible contamination sources are especially problematic for water suppliers with production wells in large, densely populated areas and in regions with intensive farming. The only way to achieve robustness in the face of unknown sources is to represent the possible existence of them through an additional virtual component in the risk model. While it is important to guard against unknown sources, the virtual risk model must be as pragmatic and simple as possible. For example, using a dense grid of virtual sources over the entire well catchment would lead to unacceptably high computational costs. I chose to model the unknown risks using a densely spaced set of virtual point sources that surround the water supply wells (cf. Figure 5.6, the virtual blue possible contamination sources at the pumping well). In the following, I call this concept *line of attack*. This risk model is plausible because any contaminant from (unknown) possible contamination sources has to pass the line of attack before affecting the pumping wells. Thus, a resulting monitoring network that can monitor the virtual sources on the line of attack is also able to monitor any remaining risk that is farther away from the pumping wells. For all

unknown spills, the line of attack is a conservative (worst-case) representation in the following aspects (compared to the actual properties of the unknown possible source represented by this model):

1. the modeled travel time from the line of attack to the well is always shorter;
2. the detectable width of the modeled plumes is always smaller;
3. the contaminant impact on the well is always larger.

The line of attack is located at a distance in travel time coordinates that is defined by the user. Essentially, this represents the minimum travel time that a user is comfortable allowing themselves to install countermeasures for water quality protection. All residual unknown sources between the line and the pumping wells cannot be controlled with the resulting monitoring networks. Hence, the choice of travel time to the well is a compromise between residual-risk coverage and achievable early-warning time. Known possible sources between the line of attack and the pumping wells need to be monitored with source-targeted wells.

The line of attack approach potentially offers large savings in monitoring costs and in design optimization. This is because there will be no further advantage in placing any monitoring wells farther away from the well than the line of attack if only a robust detection probability is essential and the user-defined minimal early-warning time is sufficient. In this case, there is no need to collect data about sources beyond this distance, either. The corresponding monitoring networks usually exhibit a line of defense formed by monitoring wells just downgradient of the line of attack. However, outside of the line of attack, there might exist known possible contamination sources toxic enough such that a reliable monitoring in detection probability and early-warning time is required. Then, additional monitoring wells close to these known possible contamination sources are appropriate.

To facilitate the line-of-attack concept, I followed the approach described in Section 5.4 and suggest to classify the unknown possible contamination sources into a risk class. However, they are unknown, and hence they cannot be classified into one of the three already existing risk classes in a meaningful manner. To overcome this problem, I introduce an additional risk class (blue) that represents unknown possible contamination sources in distributed areas and unknown sources placed on a line of attack:

$$\mathbf{d}_{\text{opt}} = \underset{\mathbf{d}\in\mathbf{D}}{\arg\min} \left[f_{\text{det}}^{(\text{red})}, f_{\text{warn}}^{(\text{red})}, f_{\text{det}}^{(\text{yellow})}, f_{\text{warn}}^{(\text{yellow})}, f_{\text{det}}^{(\text{green})}, f_{\text{warn}}^{(\text{green})}, f_{\text{det}}^{(\text{blue})}, f_{\text{warn}}^{(\text{blue})}, f_{\text{cost}} \right] \tag{5.10}$$

With that, the number of objective functions increases again (from seven to nine). But, this further increase in complexity has the benefit of accommodating the subjective concerns and expert judgments of stakeholders, allowing them to alter the weights on the virtual sources to represent their relative risk tolerance of known and unknown sources.

Figure 5.6 shows one selected optimized network, which can be separated in two parts with different tasks: the first part (white pluses) specializes only on good coverage of the residual risks (*line of defense* close to the line of attack), and the second part (white crosses) is an augmentation of the line of defense in order to improve the performance in monitoring all the known possible contamination sources. The model for residual risks is the representation by the line of attack (blue dots) that lies between the well gallery and the bulk area of the catchment. Here, the line of attack is a line of hypothetical point sources placed at a travel time of two years away from the well gallery. For simplicity of the analysis, here I drop the classification of the known possible sources into different risk classes.

The performance data for both parts are shown in Table 5.3. The strategy for covering the residual risk is the line of defense that directly matches the line of attack. In the solution shown, this line is composed of 17 monitoring wells. This strategy achieves a detection probability of 81% and claims an early warning utility of 0.68. The detection probability number is accurate for all residual risks (behind the line of attack). The early warning time, however, is strongly suboptimal, because most residual risks could be detected earlier when knowing their location.

The line of defense provides a good coverage for almost all possible residual risk sources anywhere in the catchment. Only 19% of the entire catchment area cannot

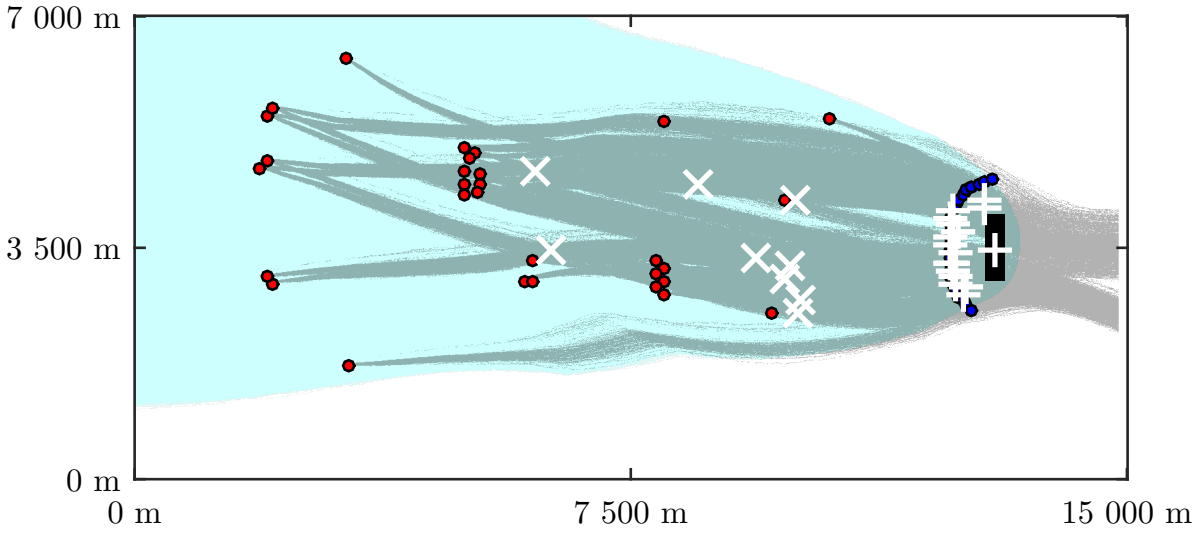


Figure 5.6: Comparison of two early-warning networks from the optimization with residual risks (see Equation 5.10). White plus mark: monitoring network specialized only on residual risks; White cross mark: augmentation to improve the performance for the known possible contamination sources; Blue circles: model of residual risk ("line of attack"), partly covered by the white cross marks; all other symbols and colors: same as in Figure 5.3.

Table 5.3: Key data of the different monitoring networks illustrated in Figure 5.6

Monitoring network	Sources	Detection probability	Early-warning utility	Monitoring wells
Only residual	Residual	0.81	0.68	17
Only residual	All known	0.65	0.07	17
Augmented	Residual	0.81	0.68	26
Augmented	All known	0.96	0.66	26

be controlled by the line of defense. Hence, it is clear that it should also provide a good detection probability for the known possible contamination sources. However, the detection probability of the line of defense for all known sources is only 65%. This means that 35% of the known sources lay inside the area which is not controlled by the line of defense, which is more than the 19% predicted during the optimization. This can be explained by considering the known source inventory as only a *small* sample out of all possible source positions (the remaining risk). The early-warning utility for the known sources is unsatisfactory (0.07). In order to improve this performance, additional monitoring wells in the more remote parts of the catchment are required. The cross mark show a selected corresponding augmentation of the current network by 9 monitoring wells. These additional wells increase the detection probability for the known sources from 65% to 96% and the corresponding early-warning utility from 0.07 to 0.66.

Depending on their risk aversion and risk perception, a decision maker could either start with coverage of residual risks and then buy more early-warning time for known sources (possibly for the severe ones first), or start with network solutions that offer a satisfactory performance for known possible contamination sources at affordable costs, and then investigate the cost-performance trade-offs for additional coverage of residual risks. The provided information is valuable decision support in situations where decision makers are highly risk-averse, or where the residual risks are known to exist at unknown locations. In situations penetrated with deep uncertainty in flow conditions, I expect the residual-risk solution to provide a large degree of robustness.

To show the robustness potential of the line of defense, I compare the performance of a line of defense (white plus mark in Figure 5.7) to a "regular" monitoring network (black cross mark in Figure 5.7) and show that the line of attack is sufficient to represent unknown possible contamination sources. Therefore, I use the first hydraulic scenario of Z_Based (cf. Table 3.1, here with the longitudinal dispersivity of $\alpha_\ell = 1$m and the transverse dispersivity of $\alpha_t = 0.1$m (cf. Figure 5.7)). I consider 100 scenarios of unknown possible contamination sources. Each of them contains 15 randomly located unknown and the 15 known possible contamination sources. The ratio of known and unknown sources is arbitrary and could favor known or unknown sources. The considered domain of all possible locations is restricted to the well catchment.

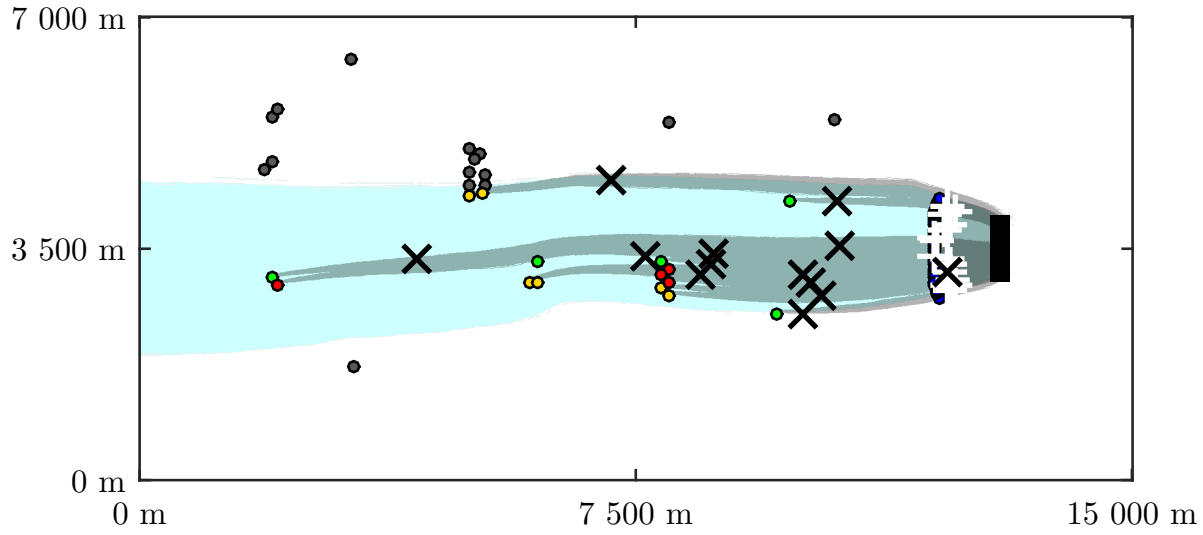


Figure 5.7: Model domain of the test case. Gray dots: possible contamination sources outside the catchment; Black cross mark: monitoring network with no prioritization; White plus mark: line of defense; all other symbols and colors: same as in Figure 5.3

Figure 5.8 illustrates the performance of both networks for each scenario (blue dots for the black-crosses network and blue diamonds for the line of defense). They cluster into two clouds (illustrated as shaded areas). For both networks, the objective values of the known possible contamination sources are the same for each considered scenario because they are not affected by the varying unknown sources (cf. the red, yellow and green dots/diamonds in Figure 5.8). The line of defense is significantly more robust (against uncertainty in location of possible contamination sources) in terms of detection probability. The variations of the blue diamonds in Figure 5.8 are much smaller than the variations of the blue dots, and the entire cloud of blue diamonds is close to optimal detection. The center of the diamonds cloud is approximately at 85% detection probability, while the center of the cloud of dots is only near 50% detection probability. There are two reasons for the discrepancy between the expected and actual detection probability (100% vs. ∼85%) of the line of defense:

1. The breakthrough curves of a close contamination source might have a short duration of visibility at the line of defense so that the probability of missing the contamination plume due to low sample frequencies increases.

2. The line of defense cannot monitor contamination sources that are within the line of defense, i.e., between the line of defense and the well gallery.

As expected, the line of defense performs worse in early-warning time compared to the black-crosses network due to its short distance to the well gallery. However, budgetary limitations usually make it impossible to install a monitoring network that can

provide a high detection probability and a high benefit in early-warning time for all unknown possible contamination sources. From this point of view, it is not necessary to model all possible unknown sources (e.g., using a dense grid of sources over the entire domain) because under cost constraints the optimization algorithm would find a line of defense as Pareto-optimal solution. Relaxing the computationally expensive approach of a dense source grid to the line of attack, however, reduces the number of Pareto-optimal solutions. The trade-off between loss in early-warning time and detection probability due to the positioning of the line of defense has to be preselected by the water supplier. But, especially for well catchments with a difficult inventory of possible contamination sources, i.e., with a high expected number of unknown sources, the line of defense is a strategy worth considering.

In summary, the main benefit of the line of attack is that the corresponding line of defense is robust in detection probability against the occurrence of unknown possible contamination sources. The main drawback is the loss in early-warning time, because of its short distance to the production wells.

5.7 Summary and Conclusions

In this chapter, I presented possible strategies to apply academic concepts in practice such that they are frugal enough to run on standard personal computers. For the transport simulation, I specifically recommended

- to approximate the temporal moments of the breakthrough curves and then to reconstruct them from those moments, instead of approximating them directly. This way, significantly fewer particles are needed to reach stable breakthrough curves.
- to determine the catchment of the production wells via backward transport simulation. This step helps to identify and remove possible contamination sources that will be irrelevant in all further analysis steps, because (1) they are outside the catchment, (2) they do not contain enough contaminant mass to threaten the wells, (3) the travel time distance of their contaminant is too long.
- to consider uncertainty for robust optimization results, because the actual performance of optimized early-warning networks is highly susceptible to predictive uncertainty of contaminant transport towards the well.

Further, I introduced a qualitative risk-prioritization approach that enables water suppliers to account for their own (subjective) risk prioritization of each possible contamination source instead of difficult (or even impossible) risk-quantification approaches. I also presented a strategy to minimize the effort for screening areas with many possible contamination sources by representing them with virtual sources. Finally, I introduced

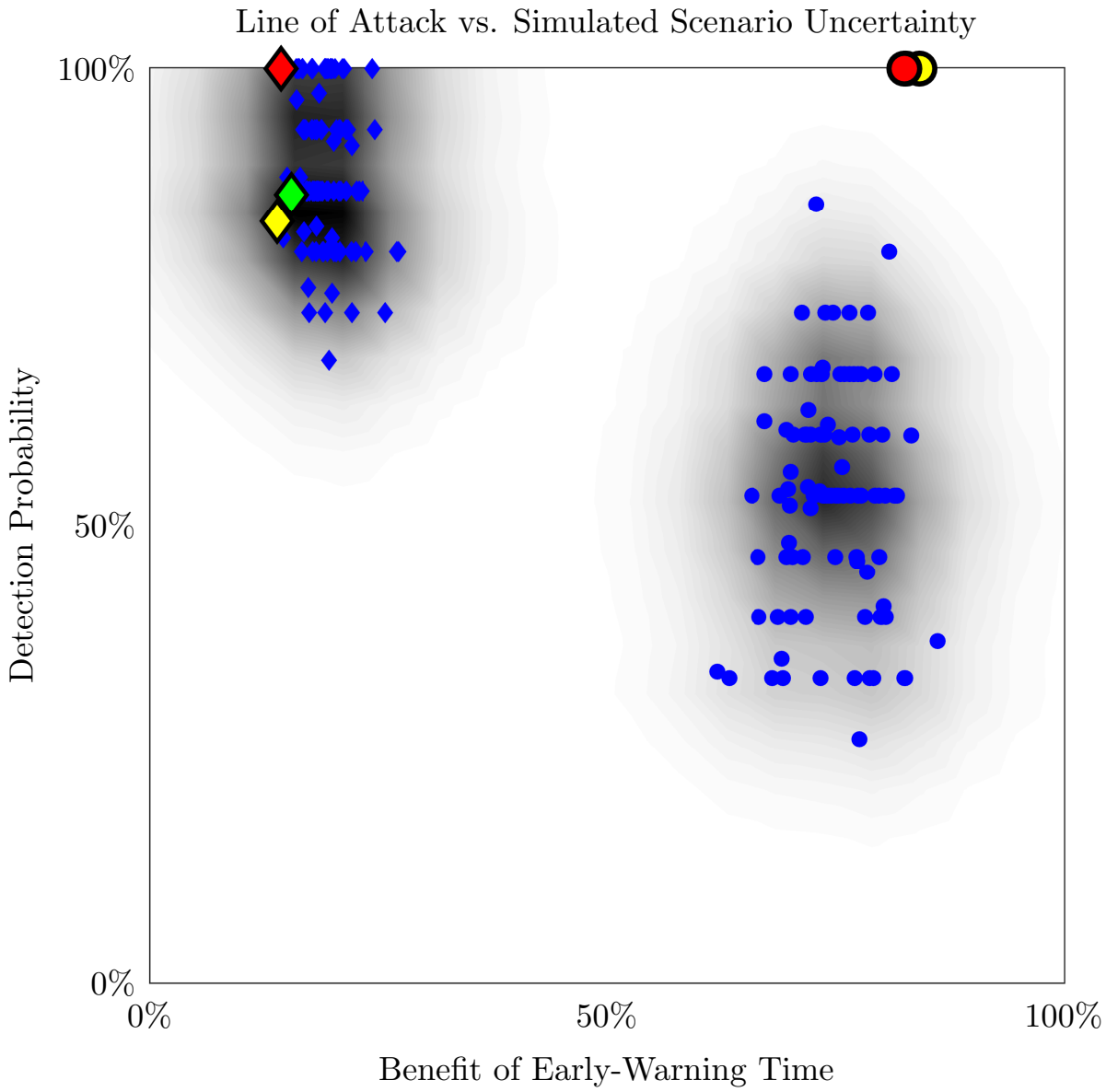


Figure 5.8: Trade-off between detection probability and early-warning time for two different monitoring networks. Red and yellow dot (according to the risk class): performance of the monitoring network optimized for all known possible contamination sources (black crosses in Figure 5.7). The green dot is overlain by the red and yellow dot; Red, yellow and green diamonds: performance of the line of defense regarding the known possible contamination sources; Blue dots: performance of the black-crosses monitoring network regarding 100 scenarios with 15 randomly located contamination sources; Blue diamonds: performance of the line of defense confronted with these scenarios. Black shaded areas: density of the point clouds for a better visibility.

a risk model that represents unknown possible contamination sources and enable an effective and robust monitoring.

I find the following conclusions most noteworthy:

- In practice, Monte-Carlo simulations are not feasible to represent uncertainty. Instead, uncertainty can be considered by using representative scenarios. It is important to follow the rules of well-selected scenarios and to involve decision makers directly and actively in defining the scenarios.
- A prioritization of possible contamination sources and a corresponding augmentation of the multi-objective optimization offers options for cost-efficient focus on only the most severe sources and provides valuable insights into the trade-offs for additional coverage of medium or almost tolerable sources.
- Early-warning monitoring networks can even be optimized to cover possible contamination sources that are unknown in existence or location through an adequately chosen residual risk model. The step from quantitative to qualitative risk assessment still requires that possible contamination sources are known in their location and existence. Often, there are unknown possible contamination sources in well catchments. In such cases, a useful strategy is to move away from reliable solutions (in terms of solutions with a quantified probability of success) to robust solutions (in terms of still serving the purpose sufficiently well and being sufficiently reliable even in the face of unknown possible contamination sources). Robustness is typically paid for by a small decrease in classical reliability, or by additional costs.

Overall, the presented utilitarian approaches serve the goal to adopt computationally demanding and data-hungry, complex statistical methods for practical application. Hence, it is one step towards a reconnection of academia and practice.

Chapter 6

An Analytical Approach

This chapter is based on the manuscript *An Analytical Approach for Positioning a Single Monitoring Well to Reliably Detect Point-Source Contaminant Plumes* [13]. I am reusing parts of the text and figures from this manuscript.

In the previous chapter, I presented methods and approaches that enable the use of the overall optimization approach in practice. Some of these methods dealt with uncertainty in aquifer parameters, or virtual contamination sources that represent, e.g., possible contamination sources that are uncertain in location. When uncertainty is considered, the uncertain parameters become random variables and influence the reliability and optimality of possible monitoring-well locations.

Using a proper uncertainty representation (e.g., MC simulation or scenario analysis) for flow and transport simulation and the robust optimization-problem formulation from Equation 3.10, an ensemble of varying outcomes can be generated and analyzed. As already shown in Figure 5.4, optimization results for the individual scenarios differ from the results for the entire ensemble. For instance, without considering location uncertainty of a contamination plume, a single monitoring well can reliably detect the contamination everywhere within the plume. With considering location uncertainty of the plume, a single monitoring well can reliably detect the contamination only where all possible plumes overlap. In simulation-based modeling, many more sources of uncertainty exist.

In this chapter, I specifically investigate the effects of physical mechanisms and parameter uncertainty on reliable monitoring-well positions. I provide analytical solutions that disclose the factors that influence the optimal positioning of monitoring wells with a special emphasis on uncertainty in hydro(geo)logy and spill location. Developing a fundamental understanding is crucial for being able to evaluate and understand optimal monitoring-network designs. Otherwise, inverse thinking is not possible. In this context, inverse thinking means that decision-makers try to understand the positioning rules behind an optimization outcome for the sake of validation or learning. In other words, they might not be able answering the following question:

> *Why does [some optimization algorithm] suggest these monitoring-well locations although they do not seem to be optimal to my current understanding?*

With the fundamental findings of this chapter, optimization results can roughly be predicted and finally justified.

6.1 Goal, Benefits, and Approach

Goals

The goal of this chapter is to obtain a fundamental understanding of the factors and processes that influence the optimal position of a single monitoring well. The monitoring well should ensure a guaranteed detection of continuous spills that could occur from a single possible contamination source. The chance of a guaranteed detection is mainly evaluated by varying flow directions and uncertain spill locations (contaminant plume could by-pass the monitoring well), as well as by varied dilution of the contaminant due to transverse dispersion (concentration of the contaminant at the monitoring well could fall below the detection limit). For simplicity, the approach pursued here is designed for detecting long-term contaminant sources. Thus, a steady-state approach is employed and monitoring frequency is not considered.

Expected Benefits

The main outcome of this chapter is a fundamental process understanding for finding good monitoring well positions when conditions are uncertain. Detection monitoring networks involving many monitoring points and many contaminant sources are subject to the same underlying processes, albeit with more complexity due to additional wells and sources. Hence, the results obtained here can be used to better understand and guide formal optimization procedures for monitoring networks and to facilitate the inverse thinking mentioned above.

Approach

I use an analytical model to determine the optimal monitoring-well locations for a single contaminant source. A major benefit of employing an analytical approach is that closed-form expressions for monitoring-well locations can be obtained, directly explaining the role of transport parameters. The main interest is on detecting plumes that emerge from continuous (steady state) point sources of contaminants. Contaminants are assumed to be conservative, i.e., no sorption or reactions occur. Since the point source is continuous, longitudinal dispersion plays only a minor role and can be neglected [e.g., 73].

The point source is placed at the origin of the coordinate system, and the two-dimensional steady-state analytical solution of the advection-dispersion equation (ADE) for a depth-averaged infinite aquifer in 2D with constant coefficients (i.e., homogeneous aquifer with uniform thickness and uniform flow) is employed. Here, the Gaussian plume is used as solution for the ADE (cf. Equation 2.11). It is mapped to a shifted and rotated coordinate system, such that the contamination source lies at the origin and the velocity v is aligned with the positive x-axis. We shall call this coordinate system the *standard space*, and denote the coordinates and the concentration distribution in the standard space with a tilde: $\widetilde{C}(\widetilde{x}, \widetilde{y})$.

To compute the concentration at a point (x, y) in physical space, first, it has to be transformed to the standard space. We define a transformation T that maps physical coordinates into standard coordinates: $(\widetilde{x}, \widetilde{y}) = T(x, y)$. If the spill location is at (x_0, y_0) in physical space and the velocity is rotated by the angle β, then we obtain the following transformation T:

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}. \tag{6.1}$$

It follows:

$$C(x, y) = \widetilde{C}(T(x, y)). \tag{6.2}$$

When considering all parameter values (x_0, y_0, β) of this transformation to be fixed, the observable plume concentration C depends only on the spatial coordinates x and y.

Next, a detection limit C_{thres} is introduced, which depends on the measurement equipment accuracy. If the concentration at any point exceeds C_{thres}, the plume is detectable at that point. Otherwise, it is not. Hence, the isoline $C = C_{\text{thres}}$ from Equation 6.2 outlines the detectable portion of the plume. In the following, I will only work with the geometry and outline of the detectable portion, since I am only interested in whether and where the contaminant spill can be found by detecting its plume. Thus we define:

$$d(x, y) = \begin{cases} 1 & \text{if } C(x, y) \geq C_{\text{thres}} \\ 0 & \text{otherwise} \end{cases} \tag{6.3}$$

where $d(x, y)$ is an indicator field for detectability marking the detectable portion with $d = 1$.

When considering parametric dependence and uncertainty, the location and shape of the detectable plume portion depend on a number of factors:

1. the released mass per time of the contamination $\dot{m}_{\text{in}}$,
2. the thickness of the aquifer B,
3. the seepage velocity in longitudinal direction v,

4. the rotation angle β and offset (x_0, y_0) of the flow-aligned coordinate system,
5. the transverse dispersion D_T,
6. and the value of the detection limit C_{thres}.

In this chapter, I investigate how these factors and their uncertainties affect the geometry, size and position of the detectable portion of the plume, and how these factors define possible positions for the monitoring well, such that a spill is guaranteed to be detected.

To analyze the dependence of the detection indicator function $d(x, y)$ (Equation 6.3) on these factors, the investigated parameters are explicitly included in the definition of d: $d(x, y, \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a vector containing the parameters. If any of the parameters $\boldsymbol{\theta}$ is uncertain, then $d(x, y, \boldsymbol{\theta})$ becomes a random boolean function. The expected value of d is then the detection probability P:

$$P(x, y) = \mathbb{E}_\theta [d(x, y, \boldsymbol{\theta})] = \int_\Omega d(x, y, \boldsymbol{\theta})\, p(\boldsymbol{\theta})\, d\boldsymbol{\theta}\,. \tag{6.4}$$

Here, Ω denotes the parameter space of the uncertain parameters $\boldsymbol{\theta}$ considered, and $p(\boldsymbol{\theta})$ is their assumed probability density function. $P(x, y)$ is the resulting map of detection probability. In Sections 6.2 and 6.3, I will derive analytical solutions for the geometric outline of the portion with a detection probability $P(x, y) = 1$, i.e., 100% detection probability.

In more general terms, the optimal monitoring position is defined by the location of highest achievable detection probability. Whether a 100% detection probability exists or not depends on the parameters and the level of uncertainty in the system. I will analyze how the probability map $P(x, y)$ and its maximum value changes under variations of source strength $\dot{m}_{in}$, absolute velocity v, detection limit C_{thres}, and transverse dispersion α_t in Section 6.4.

6.2 Characteristics of the Plume

In this section, I introduce the isoline of a concentration distribution based on the ADE solution from Equation 2.11. I derive the characteristics of this isoline shown in Figure 6.1. I will use these characteristics in the following sections for investigating effects of physical mechanisms and their uncertainties.

Isoline

By inserting $C = C_{thres}$ into Equation 2.11 and solving for y, we obtain the isoline as a function of $x \in [0, x_{tip}]$, with x_{tip} as the end point of the detectable portion of the

plume (cf. Equation 6.7):

$$y_{\text{thres}}(x) = \pm\sqrt{\frac{-4D_{\text{T}}x}{v} \cdot \ln\left(\frac{C_{\text{thres}}B}{\dot{m}_{\text{in}}} \cdot \sqrt{4\pi v D_{\text{T}} x}\right)}, \tag{6.5}$$

where y_{thres} is the y-position of the isoline obtained by inserting $C = C_{\text{thres}}$.

Characteristics

If we consider point-sources and flow fields without any uncertainties, the detection probability $P(x, y)$ is binary (cf. Equations 6.3 and 6.4): we can achieve a guaranteed detection anywhere within the detectable portion of the plume defined by Equation 6.5. In the following, I call the area of 100% detection probability the *area of guaranteed detection* (Ω_{GD}) and use its size A_{GD} as a measure of robustness for the monitoring well placement problem. A large A_{GD} means that monitoring wells can be positioned to withstand unforeseen shifts or changes of the plume due to unknown and unconsidered uncertainties in the system description. In the case without uncertainty, Ω_{GD} assumes its largest possible size. It can be calculated by integrating the detectable width over the length of the detectable plume in Equation 6.5:

$$A_{\text{GD}} = 2 \cdot \int_{x_0}^{x_{\text{tip}}} y_{\text{thres}}(x)\, dx\,, \tag{6.6}$$

with x_0 and x_{tip} as the starting and the ending point of the detectable plume portion along the x-axis (cf. Figure 6.1). The starting point is $x_0 = 0$, and the end point x_{tip} follows from Equation 6.5 by inserting $y_{\text{thres}} = 0$:

$$x_{\text{tip}} = \left(\frac{\dot{m}_{\text{in}}}{C_{\text{thres}}B}\frac{1}{\sqrt{4\pi v D_{\text{T}}}}\right)^2. \tag{6.7}$$

Solving the integral in Equation 6.6 yields

$$A_{\text{GD}} = \left|\left(\frac{\dot{m}_{\text{in}}}{C_{\text{thres}}B\sqrt{3\pi}}\right)^3 \frac{1}{v^2 D_{\text{T}}} \cdot \sqrt{\pi}\right|. \tag{6.8}$$

It is useful to describe the geometry of A_{GD}, for example, the x coordinate where Ω_{GD} has its largest width and the corresponding value of that width. This position can be found by searching for the point where the first derivative assumes a value of zero in one branch of the detection outline:

$$\frac{\partial y_{\text{thres}}}{\partial x} = 0 \Longleftrightarrow x_{\text{opt}} = \frac{1}{e}\left(\frac{\dot{m}_{\text{in}}}{BC_{\text{thres}}}\frac{1}{\sqrt{4\pi v D_{\text{T}}}}\right)^2, \tag{6.9}$$

where the constant e is *Euler's number*. Inserting x_{opt} in Equation 6.5 yields the maximum width w_{max}

$$w_{\text{max}} = 2 \cdot \sqrt{\frac{1}{2e\pi}} \cdot \frac{\dot{m}_{\text{in}}}{C_{\text{thres}} B v} \,. \tag{6.10}$$

As can be seen in Figure 6.1, the shape of the detectable portion of the contaminant plume is roughly elliptical, following from the Gaussian function. For x coordinates less than the maximum width of the detectable plume, its shape is mainly influenced by velocity and transverse dispersion. From upstream of the spill location, only clean water is coming into the mix. Downstream, beyond the maximum width, the shape is dictated by the detection limit and dilution due to transverse dispersion. Note that the detectable plume is bounded by a concentration isoline that is defined by the detection limit. Contaminants do occur outside the detectable plume, particularly downstream and transverse to the flow, but are not measurable.

6.3 Analytical Solutions

Uncertainty in Spill Location

In practice, the location of point sources is often uncertain. For example, a tank farm at an oil refinery includes many gas tanks, and each gas tank could fail. Uncertainty varies from case to case: the possible area affected by a contaminant spill from a large oil refinery containing many possible sources is very uncertain, while there is less uncertainty for an individual gas tank. When placing monitoring wells, we wish to ensure that detection is guaranteed and two questions are of particular interest:

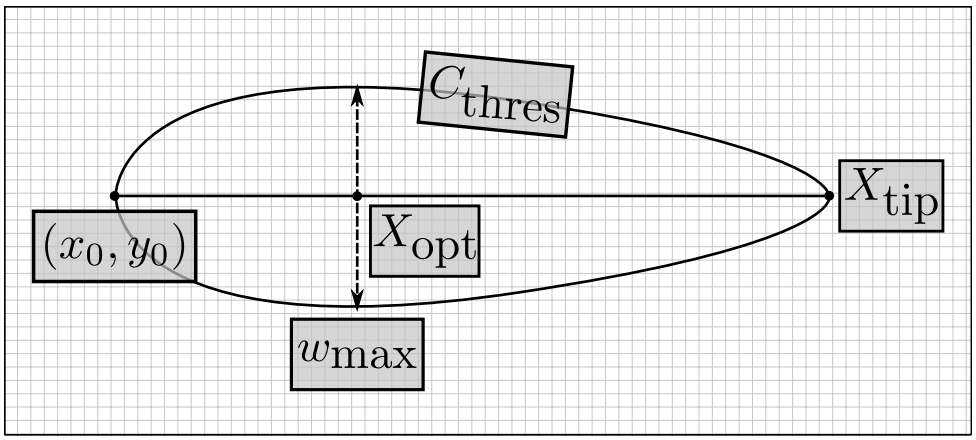


Figure 6.1: Geometry aspects: Isoline of a contaminant plume starting in (x_0, y_0). The outline of the plume is obtained by a concentration threshold C_{thres}. The maximum length of the isoline plume is given by X_{tip} and its maximum width by w_{max} at X_{opt}.

1. **Backward transport:** What is the maximum allowed uncertainty in spill location for a given monitoring well position and a requirement of 100% detection probability?

2. **Forward transport:** Given a level of uncertainty in contaminant-source locations, what are possible locations for a 100% detection probability monitoring well?

First, we consider the backward transport problem. For a given monitoring-well location we define a *permitted uncertainty area* Ω_{PU} in which a detectable point source could occur. Then, for the forward transport problem, we can find an extended Ω_{GD} for possible monitoring-well locations if the actual *uncertainty area* Ω_{U} of the point source is a subset of the Ω_{PU}.

For the forward problem, we will identify the Ω_{PU} of a monitoring well placed at $(x_{\mathrm{m}}, y_{\mathrm{m}})$. Whenever multiple possible source locations are involved, the concept of reverse transport [e.g., 60, 100, 121] offers smart solution approaches. We use it here to identify the visual range of the monitoring well: we invert the flow field and swap the role of point source (now at $(x_{\mathrm{m}}, y_{\mathrm{m}})$) and monitoring well in Equation 6.5. Hence, the outline of Ω_{PU} is the mirrored outline of the detectable portion of the forward plume.

If the shape of Ω_{U} is exactly known, Ω_{PU} can be used for controlling whether a sin-

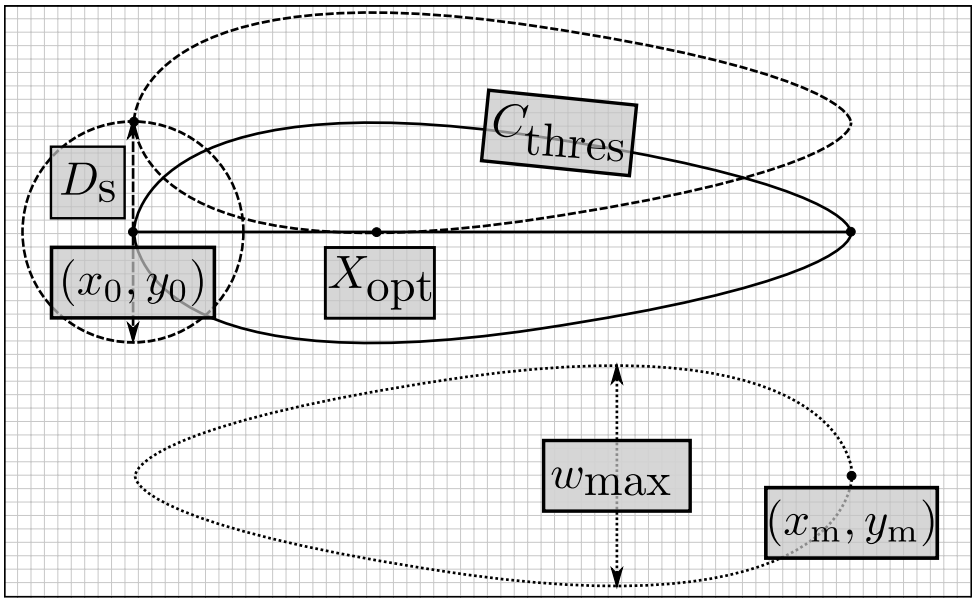


Figure 6.2: Uncertainty in spill location (x_0, y_0): Ω_{PU} can be calculated via reverse transport (dotted plume below) starting from the monitoring-well position $(x_{\mathrm{m}}, y_{\mathrm{m}})$. From a simplified parameterization of Ω_{U} by a circle (dashed circle top) an uncertainty diameter (D_{s}) is retrieved. Its maximum length is equal to the maximum width $D_{\mathrm{s}} \leq w_{\mathrm{max}}$ of Ω_{PU}. All point sources along the dashed outline of the circle with the centerpoint (x_0, y_0) can be monitored at X_{opt}.

gle monitoring well at (x_m, y_m) is sufficient to monitor the point source or additional monitoring wells are required. Additional monitoring wells are required if Ω_U is not completely covered by Ω_{PU}. If the shape of Ω_U is unknown, a most conservative approximation of its shape it is a circle (i.e., no preferred direction of uncertainty). Then, we are interested in its maximum permitted diameter D that still leads to a guaranteed detection at a single point. As the limiting case, we only have to consider the two plumes that start at the top-most and bottom-most points of the Ω_U-circle in transverse transport setting. These two plumes intersect at the downstream distance X_{opt} where the maximum plume width occurs, as identified in Section 6.2. The corresponding maximum width w_{max} is the same as the maximum uncertainty diameter D_s of the Ω_U-circle (see Figure 6.2, top).

For an uncertainty diameter $D_s < w_{max}$, there exists an Ω_{GD}, and its size depends on the difference between D_s and w_{max}. That is, Ω_{GD} is a function of D_s. A closed-form expression for Ω_{GD} cannot be found analytically, so I made a simplification. I conducted a parametric study over different uncertainty diameters D_s (see Figure B.1 and B.2 in Appendix B.4). I determined that we can simplify the problem to only consider uncertainty in transverse direction to the flow (fixed x-position of the point source, y-position uncertain). Ω_{GD} is then shown by the gray shaded portion in Figure 6.3. As can be seen, its perimeter is defined by the intersection of the contaminant plumes

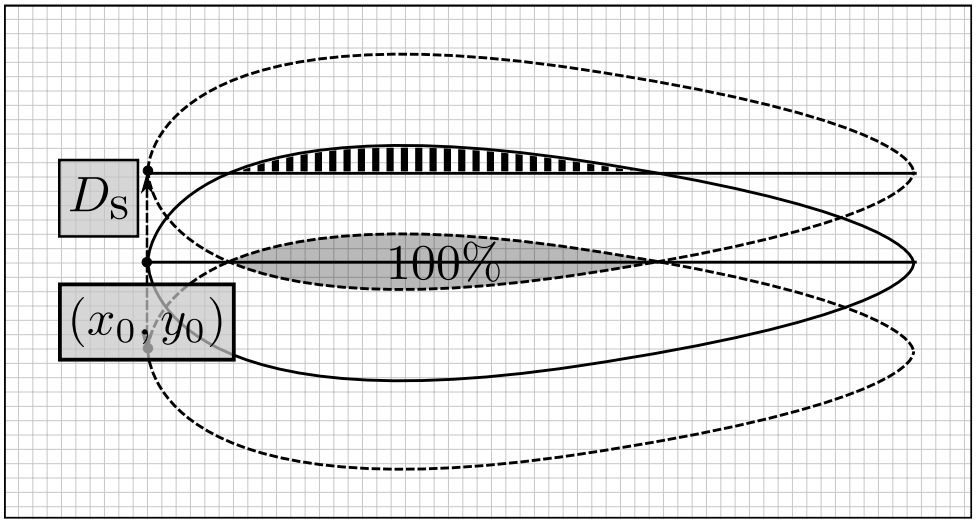


Figure 6.3: Ω_U-circle is simplified by only considering uncertainty in transverse direction of the flow. This uncertainty is represented by a vertical line with the length of D_s. For $D_s < w_{max}$, an Ω_{GD} occurs (gray portion). To calculate its size A_{GD}, formally we have to integrate over the overlapping portion of the two shifted plumes (dashed plumes). Instead, we shift the x-axis and integrate over the portion between the x-axis and the solid-line plume (checked portion) that starts in (x_0, y_0). Due to symmetry, this is half of the gray shaded Ω_{GD}.

originating at $\left(x_0, \pm\frac{D_\mathrm{s}}{2}\right)$. Its size A_GD can be found by integrating the plume between the limits x_1 and x_2 defining its upstream and downstream ends:

$$A_\mathrm{GD} = 2 \cdot \int_{x_1}^{x_2} y_\mathrm{thres}\,(x) - \frac{1}{2} D_\mathrm{s}\, dx\,. \tag{6.11}$$

To find the integration limits we must find the roots of Equation 6.5 shifted vertically by a distance $\frac{D_\mathrm{s}}{2}$:

$$y_\mathrm{thres}\,(x) - \frac{1}{2} D_\mathrm{s} = 0\,. \tag{6.12}$$

They are shown in Appendix B.2 to be:

$$x_{1,2} = \frac{v D_\mathrm{s}^2}{-8 D_\mathrm{T}\, \mathrm{W}_j\left(\frac{v D_\mathrm{s}^2 \left(\frac{C_\mathrm{thres} B}{\dot{m}_\mathrm{in}} \cdot \sqrt{4\pi v D_\mathrm{T}}\right)^2}{-8 D_\mathrm{T}}\right)}\,, \tag{6.13}$$

where $\mathrm{W}\,(\cdot)$ is the Lambert W function [25]. Since there are two different branches of the Lambert W function for $\arg \mathrm{W}\,(\cdot) = \kappa$, $\kappa \in \left[-\frac{1}{e}, 0\right)$, we distinguish them by the index j in Equation 6.13:

$$j = \begin{cases} -1 & \text{if } \left(\frac{v D_\mathrm{s}^2 \left(\frac{C_\mathrm{thres} B}{\dot{m}_\mathrm{in}} \cdot \sqrt{4\pi v D_\mathrm{T}}\right)^2}{-8 D_\mathrm{T}}\right) < -\frac{1}{e} \\ 0 & \text{if } \left(\frac{v D_\mathrm{s}^2 \left(\frac{C_\mathrm{thres} B}{\dot{m}_\mathrm{in}} \cdot \sqrt{4\pi v D_\mathrm{T}}\right)^2}{-8 D_\mathrm{T}}\right) \geq -\frac{1}{e} \end{cases}\,.$$

Uncertainty in Ambient Flow Direction

Uncertainty in ambient flow direction is one of the most challenging problems in hydrologic modeling. Different hydrological scenarios (e.g., different pumping rates in groundwater-extraction wells) greatly influence the main direction of groundwater flow. Without detailed hydrologic monitoring data, the uncertainty in flow direction can be very large.

I describe uncertainty in ambient flow direction by an uncertainty interval of the rotation angle $\beta \in [-\beta_0, +\beta_0]$ from the centerline of our coordinate system (cf. Figure 6.4). Without any uncertainty, no deviations to the centerline would occur ($\beta_0 = \pm 0^\circ$), whereas $\beta_0 = \beta_\mathrm{max} = \pm 90^\circ$ is the maximum meaningful uncertainty to consider (see below). In practical applications, the two limiting rotation angles $\pm\beta_0$ can be estimated from simulations of hydro(geo)logical extreme cases, or by interpretation of hydrogeological maps.

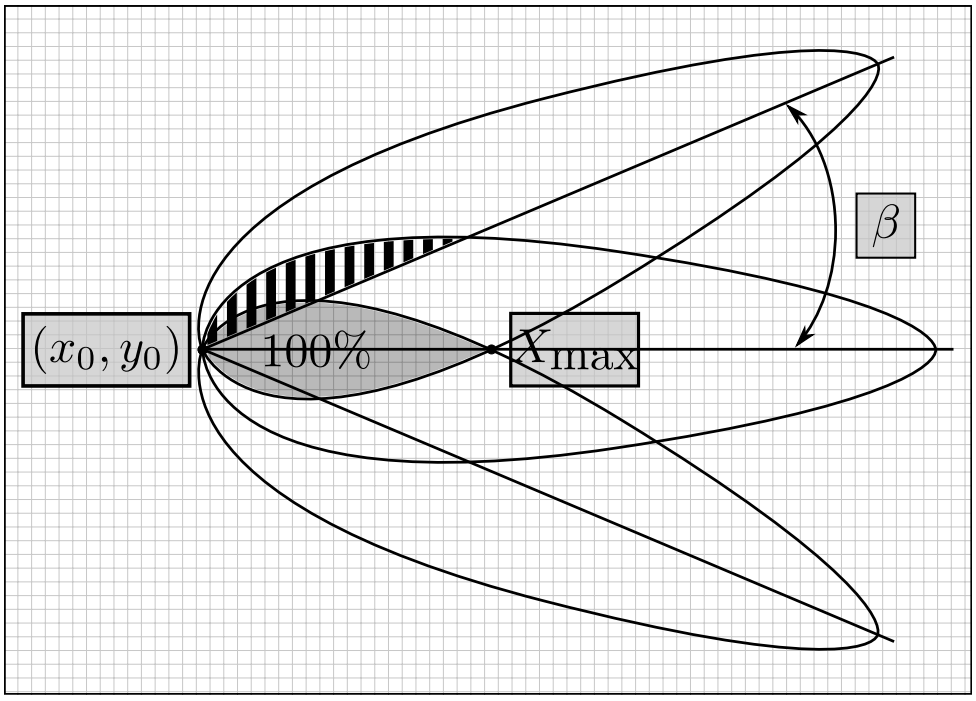


Figure 6.4: Uncertainty in ambient flow direction: All possible plumes that start in (x_0, y_0) with a maximum rotation angle $\pm\beta_0$ overlap in the gray shaded portion, the A_{GD}. The two limit plumes rotated by $\pm\beta_0$ intersect in $X_{\max}$. To calculate A_{GD}, formally one has to integrate over the overlapping portion of the two limit plumes from zero to $x_{\max}$. For easier calculations, instead of rotating the plume we rotate the x-axis and integrate over the unrotated plume from zero to the intersection with the rotated x-axis (checked portion). Due to symmetry, this is half A_{GD}.

For the extreme case of $\beta = \pm 90°$, we can find exactly one monitoring point that achieves a 100% detection probability, and this point is the source position at (x_0, y_0) itself. For smaller angles, the overlapping portion of all possible plumes (i.e., Ω_{GD}) is defined by intersecting the plumes for the extreme cases with $\beta = \pm\beta_0$. Ω_{GD} collapses to the single point at the source with an increasing angle $\left(\lim_{\beta \to \pm 90°} A_{\mathrm{GD}}(\beta) = 0 \right)$.

Formally, we use the transformation T from Equation 6.1 to calculate the point of intersection $(x = x_{\max}, y = 0)$ between the detection outline $(C = C_{\mathrm{thres}})$ of the two rotated plumes at the end-points and the unrotated x-axis with $x_0 = y_0 = 0$:

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} x_{\max} \\ 0 \end{pmatrix}. \tag{6.14}$$

Due to the symmetry of the plume, the point of intersection lies on the x-axis. To compute its position $(x_{\max}, 0)$, we calculate the intersection point $(x_{\mathrm{int}}, y_{\mathrm{int}})$ of the unrotated outline of the plume starting in (x_0, y_0) and a line $(y = m \cdot x)$ with the slope $m = \tan(\beta_0)$. Afterwards, we use the transformation T (Equation 6.14) to obtain $x_{\max}$. Employing this approach, the intersection value x_{int} is:

$$x_{\mathrm{int}} = \frac{2D_{\mathrm{T}}}{v \tan^2(|\beta_0|)} \cdot \mathrm{W}_0 \left(\frac{1}{8\pi} \left(\frac{\tan(|\beta_0|)\, \dot{m}_{\mathrm{in}}}{C_{\mathrm{thres}} B D_{\mathrm{T}}} \right)^2 \right), \tag{6.15}$$

with $\mathrm{W}(\cdot)$ as Lambert W function and uncertainty angle β_0. Employing the transformation T yields:

$$x_{\max} = \frac{x_{\mathrm{int}}}{\cos(\beta_0)}. \tag{6.16}$$

As mentioned above, an Ω_{GD} can be found whenever the flow direction varies in an angular range of $\mathrm{abs}(\beta) < 90°$. To compute its size A_{GD}, we need to integrate the gray-shaded portion in Figure 6.4, which is once again composed of two identical half-portions because of symmetry. To make the integration easier, we evaluate that half-portion in the unrotated coordinate system, see the checked portion in Figure 6.4:

$$A_{\mathrm{GD}} = 2 \cdot \left(\int_0^{x_{\mathrm{int}}} y_{\mathrm{thres}}(x)\, dx - \int_0^{x_{\mathrm{int}}} \tan(|\beta_0|) \cdot x\, dx \right), \tag{6.17}$$

with $x_{\mathrm{int}} = \cos(\beta_0) \cdot x_{\max}$. The second term in Equation 6.17 is the orthogonal triangle between the checked portion and the x-axis. Using the formulas from Appendix B.1, A_{GD} can be calculated.

6.4 Analysis of the Coordinating Factors

Transverse dispersion, source strength, quality of the monitoring equipment, and velocity are factors appearing in the analytical solution given in Equation 2.11 and all

subsequent results in Chapter 6. The following sections describe (1) each factor and (2) their influence on the optimal monitoring-well positions derived in Section 6.3. Finally, the effects of uncertainty in spill location and ambient flow direction are discussed.

Investigated Parameters

Transverse Dispersion The width and the length of the detectable plume portion (Equation 6.5) are strongly dependent on the transverse dispersion coefficient D_T. The maximum width w_{max} of the detectable plume, however, is independent of D_T (cf. Equation 6.10). A larger transverse dispersion leads to a generally shorter detectable area, since peak concentrations along the plume centerline spread faster towards the limit width w_{max}, and then dissipate faster to values below the detection threshold. Vice versa, a small transverse dispersion leads to long-stretched detectable area.

Therefore, simulation of detection areas requires careful selection of the transverse dispersion coefficient α_t. Often, approaches employing macroscopic dispersion coefficients are used to represent the uncertainty of the actual plume path due to the lack of knowledge of aquifer heterogeneities. Such approaches simulate more dispersed plumes to reflect uncertainty in plume position, but cannot be used to calculate contaminant concentrations [e.g., 61]. Instead, parameters like the flow rotation angle and uncertainty in spill location should be used in my current analyses to determine the uncertainty in position of the plume.

The homogeneous model employed in this chapter only considers hydromechanical dispersion. Hydromechanical dispersion simulates the actual dilution occurring to dissipate the plume concentration below the detection limit. However, the transverse dispersion can be assigned higher values than a fraction of the grain diameter of the porous medium, because the used ADE (Equation 2.11) is depth-integrated [131]. I use the Scheidegger parameterization [143] $D_T = \alpha_t v + D_e$ with D_e as the effective diffusion coefficient and analyze the impact of α_t. The influence of transverse dispersivity on optimal monitoring is represented by a set of curves in each of the following figures.

Source Strength In general, the source strength of a contaminated site or source is unknown and can only be estimated using data acquired in extensive field campaigns. Through techniques of contaminant source identification, the volume of contaminated soil can be inferred for spills that have already occurred [e.g., 150, 156]. Here, I consider point sources with a given mass discharge $\dot{m}_{in}$ as the parameter describing source strength.

Measurement Equipment The quality of measurement equipment has a major effect on the detectable portion of the plume. High accuracy implies low values for C_{thres}.

Since the outline of the detectable plume is defined by $C = C_{\text{thres}}$, the detectable portion expands as C_{thres} becomes smaller. The three quantities C_{thres}, B and $\dot{m}_{\text{in}}$ occur generally together as the term $\frac{C_{\text{thres}}B}{\dot{m}_{\text{in}}}$, so we do not need to investigate them individually in Equations 2.11 and 6.5 and subsequent derivations. Within this term, C_{thres} and $\dot{m}_{\text{in}}$ have opposing impacts: a high mass discharge has the same effect as a low detection limit.

Ambient Velocity The ambient flow velocity changes for different hydro(geo)logical scenarios. In Equation 2.11, the velocity is not a simple linear scaling factor for the plume, because it also influences the strength of transverse dispersion ($D_{\text{T}} = \alpha_{\text{t}} v + D_{\text{e}}$) [143]. As discussed previously, transverse dispersion has a high impact on the shape of the plume and, therefore, the impact of ambient velocity also needs to be investigated.

A_{GD} is Highly Sensitive to Stronger Release, Lower Detection Limit, and Transverse Dispersion

A_{GD} is a function of the scaling factor $\frac{C_{\text{thres}}B}{\dot{m}_{\text{in}}}$ which contains the detection limit C_{thres}, and the depth-specific mass discharge $\frac{\dot{m}_{\text{in}}}{B}$. It also depends inversely on the transverse dispersion D_{T} and velocities. The dependency on the scaling factor and α_{t} is shown in Figure 6.5.

The smaller transverse dispersion or the scaling factor, the larger A_{GD}. These dependencies are very pronounced. It can be seen in Figure 6.5 that a smaller transverse dispersion leads to a smaller sensitivity to the detection limit (please observe the logarithmic scale of the vertical axis). The reason for the reduced sensitivity is that a small α_{t} leads to a narrow plume width over a long distance. Narrow plumes have very high concentration gradients so that changes in C_{thres} hardly show any effect.

Slow Velocities Increase A_{GD} by Increasing the Maximum Width and the Length of the Plume

As discussed above, A_{GD} is sensitive to C_{thres}, $\dot{m}_{\text{in}}$, and D_{T}. However, it is also sensitive to the ambient velocity (cf. Figure 6.6) for two reasons: First, transverse dispersion depends on the velocity. Small values in velocity decrease D_{T} and lead to the effect discussed above. Second, the width and the length of the detectable portion of the plume strongly depend on the velocity in a direct fashion: the maximum width grows linearly with decreasing velocities (due to the smaller dilution of $\dot{m}_{\text{in}}$ into a slower ambient flow, cf. Equation 6.10) and the length increases quadratically (due to the additional effect of slower dispersion, cf. Equation 6.7). Figure 6.6 schematically illustrates the impact of the velocity and source strength on A_{GD}.

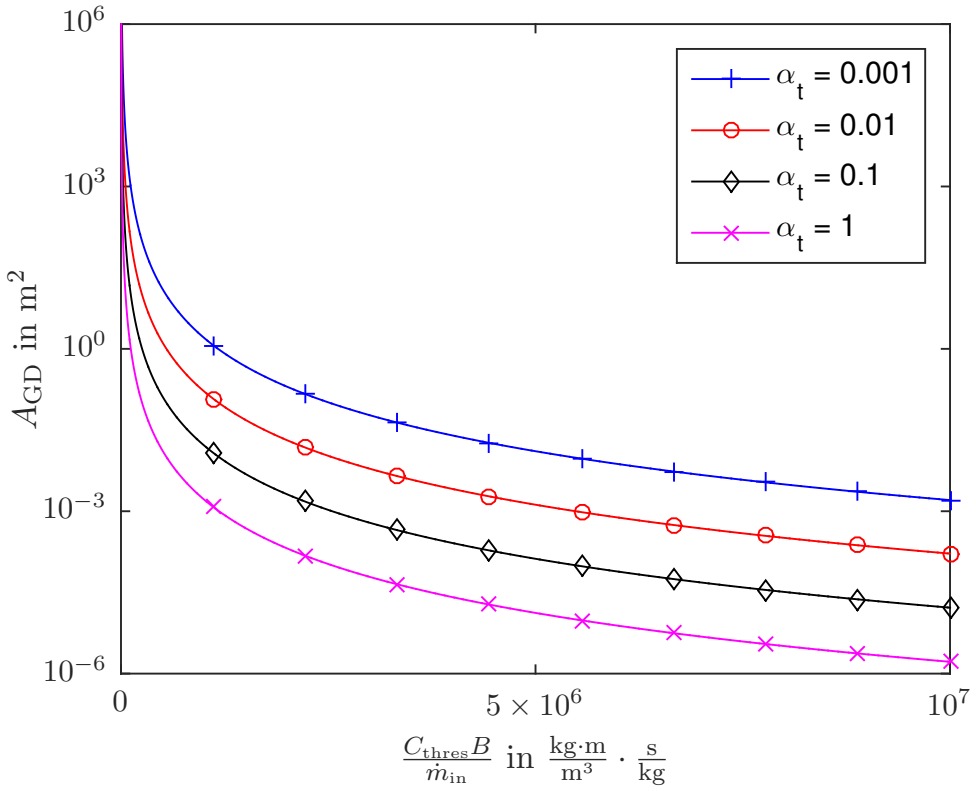


Figure 6.5: A_{GD} as a function of the factor of detection limit, aquifer depth and mass flux for a constant velocity of 1×10^{-5} m/s. The set of curves is defined by four transverse dispersion coefficients (α_t in m).

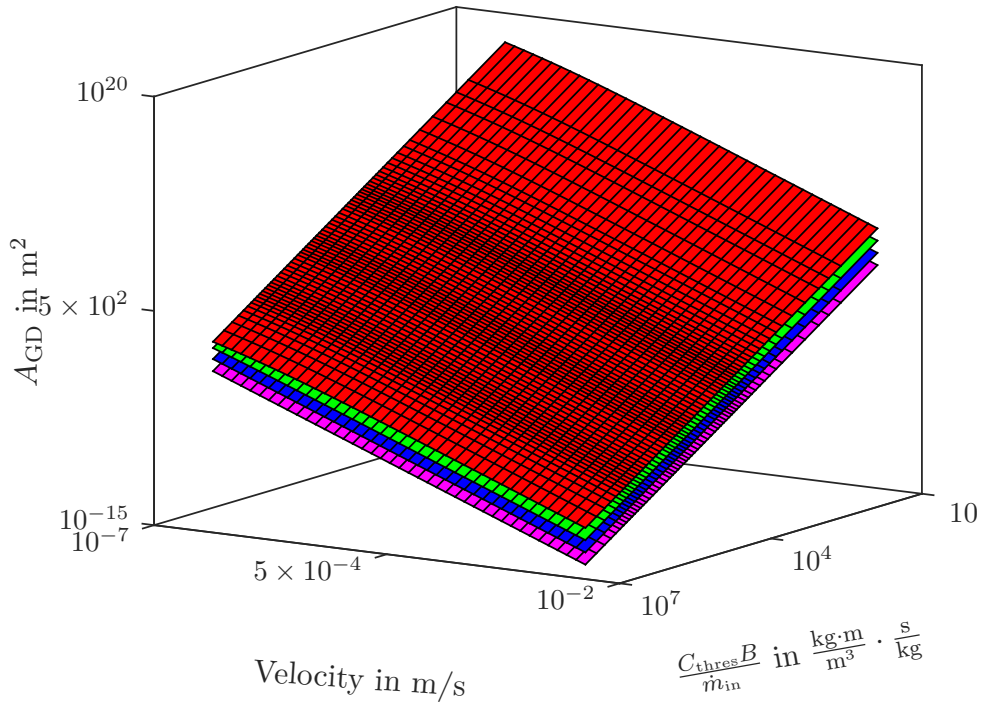


Figure 6.6: A_{GD} as a function of the scaling factor (detection limit times aquifer depth divided by mass flux) and velocities, for four different transverse dispersion coefficients α_{t}. The four layers are defined by the transverse dispersion coefficients: red ($\alpha_{\mathrm{t}} = 0.001\,\mathrm{m}$), green ($\alpha_{\mathrm{t}} = 0.01\,\mathrm{m}$), blue ($\alpha_{\mathrm{t}} = 0.1\,\mathrm{m}$), magenta ($\alpha_{\mathrm{t}} = 1\,\mathrm{m}$).

Low Detection Limit and Low Transverse Dispersion Pushes the Optimal Monitoring Well Away from the Source

The optimal position for a monitoring well is the position where the plume has its largest width $\left(X_{\text{opt}} = \left(x_{\text{opt}}, y_0\right)\right.$, Equation 6.9$)$, because that position is most robust against estimation errors in any of the parameters. As shown in Figure 6.7, the position of this point is highly influenced by the detection limit. Low values in C_{thres} increase the distance between X_{opt} and the source because the plume has more time to spread laterally until it is dissipated below the detection limit. The same effect can be seen for small values in α_t: a small transverse dilution leads to a very stretched plume that reaches its largest width far away from the source.

Large Uncertainty in Spill Location Sharply Reduces A_{GD}

The influence of the source location uncertainty diameter D_s on A_{GD} is small until D_s approaches w_{max}. The effect of D_s is much smaller than the influence of the transverse

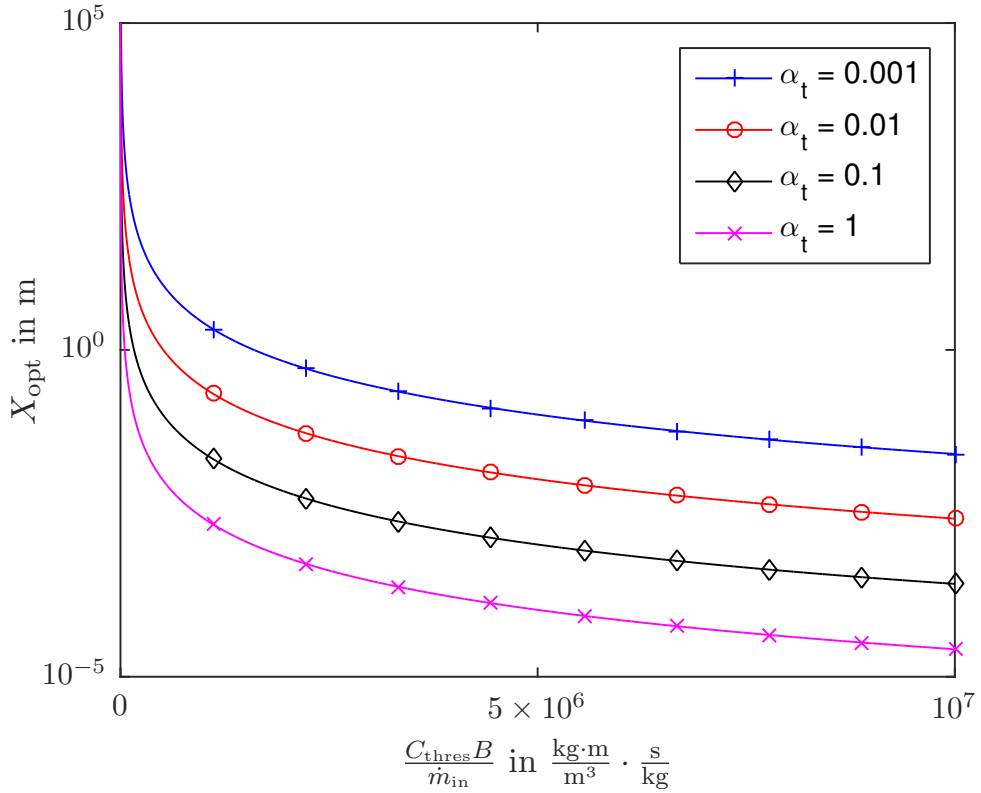


Figure 6.7: X-position of the maximum width of the plume X_{opt} as a function of the factor of detection limit, aquifer depth and mass flux for a constant velocity of 10^{-5} m/s. The set of curves is defined by four transverse dispersion coefficients (α_t in m).

dispersion or the detection limit (cf. Figure 6.8). However, for values of D_s close to w_{max}, A_{GD} strongly decreases. In the limiting case of $D_s = w_{max}$, just one point remains where 100% detection probability can be achieved and A_{GD} shrinks to zero. The remaining point is located at the point of largest plume width (X_{opt}, see Section 6.4 and Figure 6.7). Note that the maximum width of the plume w_{max} is independent of uncertainty in transverse dispersion and only the position of the maximum width changes in longitudinal direction with different values in α_t (cf. Equation 6.10 that is independent of D_T).

Angular Uncertainties Require Close Monitoring

In the scenarios with angular uncertainty, we define X_{max} to be the maximum distance from the source where one can still achieve a 100% detection probability. For infinitesimal values of the limiting uncertainty angle β_0, the limit of x_{max} is the tip of the plume $\left(\lim_{\beta_0 \to 0^\circ} x_{max} = x_{tip}\right)$ (cf. Figure 6.9). For high values of β_0, the limit

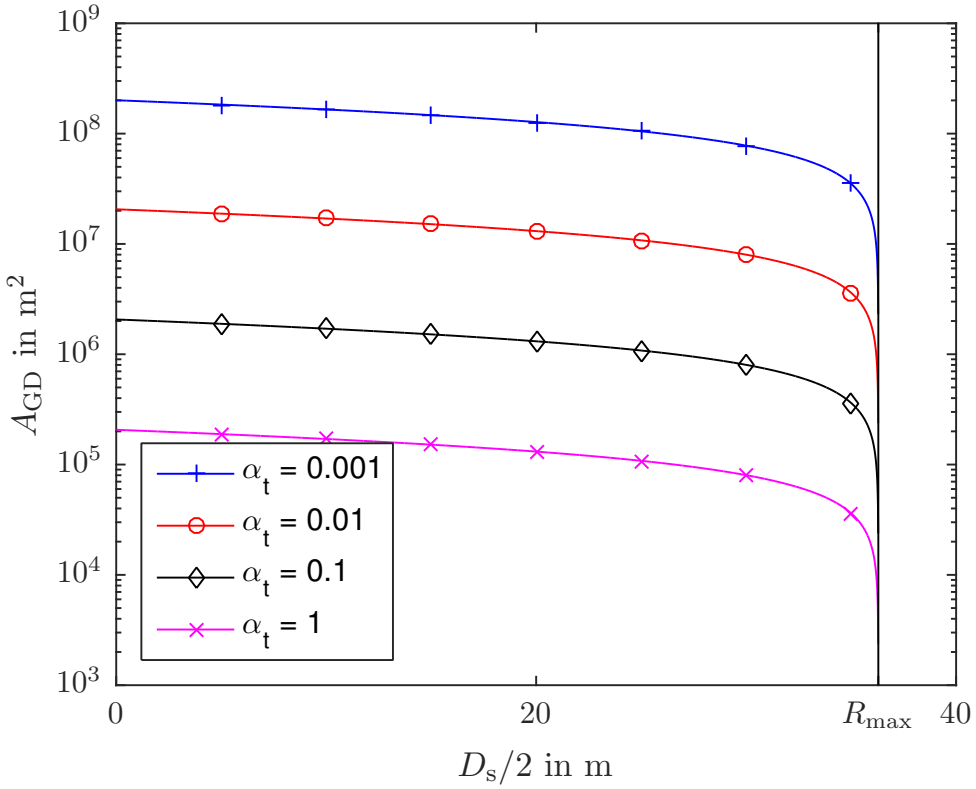


Figure 6.8: A_{GD} as a function of the uncertainty diameter D_s. In the limit $D_s/2 = R_{max} = w_{max}/2$, A_{GD} shrinks to zero. The set of curves is defined by four transverse dispersion coefficients (α_t in m).

of $x_{\max}$ is zero $\left(\lim_{\beta_0 \to \pm 90^\circ} x_{\max} = 0\right)$. The first ten degrees of uncertainty angle β_0 ($\text{abs}\,(\beta_0) < 10^\circ$) cause 99% of the total effect (please observe the logarithmic scale for $x_{\max}$). For $\text{abs}\,(\beta_0) > 10^\circ$, the changes in $x_{\max}$ are merely in the range of centimeters. $x_{\max}$ is very sensitive to changes in β_0 for small values in α_t because the maximum width of these plumes is far from the source. Increasing β_0 just slightly already leads to a strongly increased vertical shift of the plumes along the x-axis. Thus, A_{GD} decreases significantly.

Angular Uncertainties Lead to the Failure of Monitoring

Figure 6.10 clearly shows that A_{GD} for uncertainty in ambient flow direction dramatically decreases with increasing angular uncertainty β_0. This affects especially plumes with small values in α_t because the distances between the origin of the sources and the positions where these plumes have their maximum width $\left(X_{opt}\right)$ increase with decreasing α_t (cf. Figure 6.7). These plumes are relatively thin, such that the sizes of A_{GD} are highly sensitive to changes in β_0. With large values in β_0, however, A_{GD} becomes infinitesimal for all considered plumes and exists only theoretically. Such small A_{GD} have no practical benefits, and a reliable monitoring is impossible.

6.5 Summary and Conclusions

In this chapter, I investigated the physical mechanisms and their uncertainties that control the optimal placement of monitoring wells. I considered the most simple case: A single monitoring well is positioned in order to observe whether or not a single possible point-source has emitted a contaminant plume or not. I developed analytical solutions for the optimal monitoring position and its robustness as functions of several parameters of interest: the radius of uncertainty in contaminant release location (x_0, y_0), the value of the depth-specific contaminant mass discharge $\frac{\dot{m}_{in}}{B}$, the value of ambient flow velocity v and its uncertainty interval in angle β, the value of the detection limit C_{thres}, and the value of transverse dispersivity α_t. The derived analytical solutions showed that

1. the area of guaranteed detection greatly increases with decreasing detection limit and transverse dispersion,
2. a low detection limit and low transverse dispersion pushes the optimal monitoring well away from the source,
3. a large uncertainty in spill location drastically reduces the area of guaranteed detection,

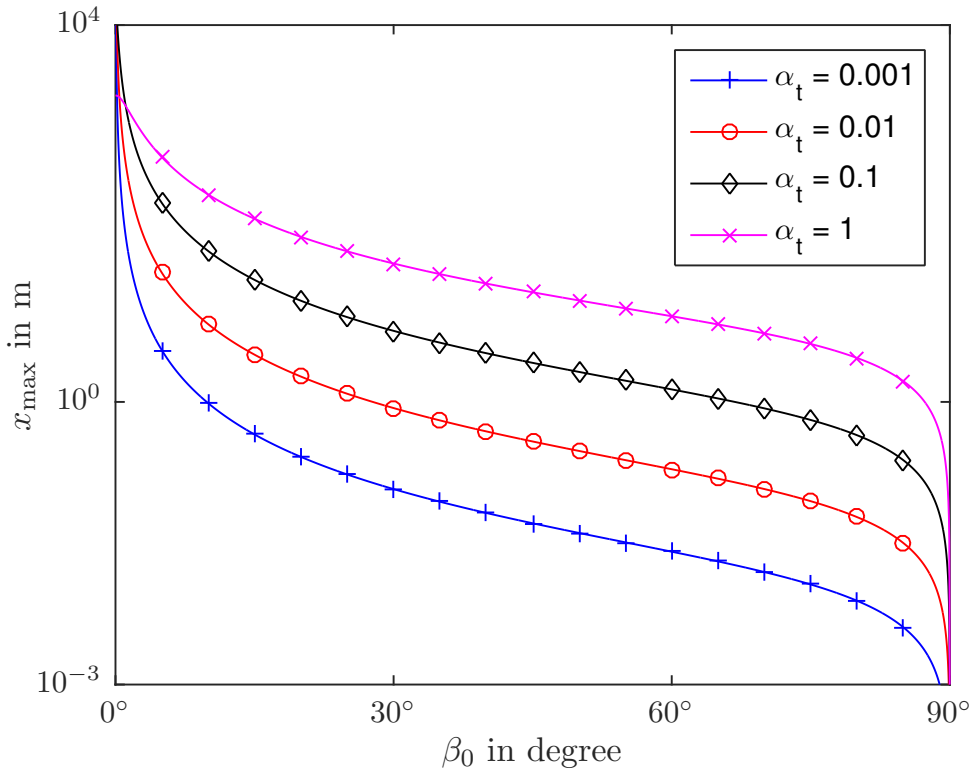


Figure 6.9: Intersection point x_{max} as a function of angular uncertainty $\pm\beta_0$ in ambient flow direction. For $\beta_0 = 0$ the intersection point is at $x_{max} = x_{tip}$. The set of curves is defined by four different transverse dispersion coefficients (α_t in m).

4. angular uncertainties require spatially close monitoring and spoil the robustness of monitoring.

Because the solutions are provided closed form, it is easy to analyze the effects of the controlling factors and their uncertainty. That is, the solutions can be used

- to improve the fundamental understanding of controlling factors that influence optimal monitoring-well position and support (or enable) the inverse thinking mentioned in the very beginning of this chapter.
- to provide decision support for evaluating and upgrading existing monitoring networks, acknowledging the inherent uncertainty through the evolvement of detection probabilities.

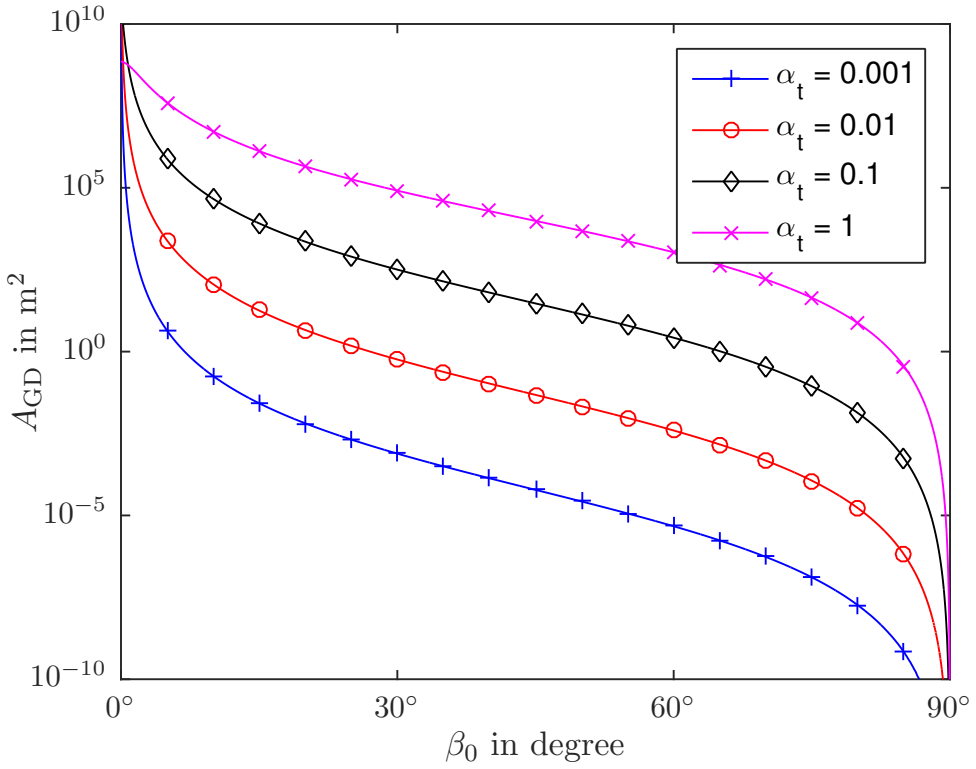


Figure 6.10: A_{GD} as a function of angular uncertainty $\pm\beta_0$ in ambient flow direction. The set of curves is defined by four different transverse dispersion coefficients (α_t in m).

Chapter 7

Insights from Practical Application

Parts of this chapter have been published in the Journal WATER RESOURCES RESEARCH under the title *Reconnecting Stochastic Methods with Hydrogeological Applications: A Utilitarian Uncertainty Analysis and Risk Assessment Approach for the Design of Optimal Monitoring Networks* Bode et al. (2018b) [15]. I am reusing parts of the text and figures from this publications by the kind permission of the publisher American Geophysical Union.

As already mentioned in Chapter 1, this work aims to develop simple-to-use methods for hydrogeological problems in practice. The previous chapters all work towards this goal: I adjusted and simplified the optimization problem (Chapter 4), I developed feasible concepts regarding the groundwater contaminant simulation for practical applications (Chapter 5), and I developed simple analytical tools to evaluate optimization results under uncertainty (Chapter 6).

This chapter demonstrates that these methods and concepts can successfully be transferred to practice by presenting select numerical results from the cooperation project introduced in Chapter 1. These results show that efficient, effective, and reliable monitoring-design optimization is possible on real-world cases under computational limitations and time constraints.

However, the transfer from academia to practitioners requires careful and effective communication to ensure that data acquisition, data evaluation, and invested project resources support the interests of both parties. The fact that communication is crucial in co-productions and collaborations between academics and practitioners has already been identified in literature related to stakeholder engagement [e.g., 2, 47]. Some common problems in such collaborations are: the use of jargon; discipline-specific software with steep learning curves; and other group-specific, implicit mindsets used by both parties.

Therefore, in the second part of this chapter, I highlight important aspects of project management learned that led to a mutually successful project. The lessons learned are much more general than the specific project. I believe they will be useful for future collaborations between academic and industrial partners in the face of uncertainty for

many hydrogeologic, hydrologic, and other environmental problems. Further literature regarding challenges, benefits, and drawbacks related to stakeholder engagement in environmental modeling can be found in the reviews of Reed (2008) [133] and Krueger et al. (2012) [98]. The interested reader is also referred to the special issue *Water Governance, Stakeholder Engagement, and Sustainable Water Resources Management* edited by Megdal et al. (2017) [114]. It gives a broad overview of recently published studies on stakeholder engagement related to the field of water resources management.

7.1 Real Case

The case study investigated here is based on the drinking-water well catchment of the *Zweckverband Landeswasserversorgung (LW)* in Stuttgart, Germany. The catchment is located south-east of Stuttgart at the Swabian Jura and the Donauried. It has an area of around 300 km^2. The geological composition of the catchment in vertical direction is characterized by five layers. A large area is dominated by a jura stratum (from north-west to south-east) that drops below tertiary molasse layers close to the production wells. Above the molasse layers, there are gravel layers with a thickness of around $8 - 10$ m and a covering soil layer. Large parts of the jura stratum are karst aquifers. For water production, more than 200 pumping wells are grouped in eight well galleries (see purple squares in Figure 7.1), producing a maximum of 2.5 m^3/s fresh water. The catchment produces about 40×10^6 m^3 fresh water per year.

Within the catchment, more than 1,500 possible contamination sources could be identified by the LW from which 143 were labeled as "relevant". These relevant sources were divided into three groups according to their subjective risk (cf. Section 5.4): "almost tolerable", "medium", and "severe" sources (cf. green, yellow, and red dots in Figure 7.1).

7.1.1 Transport Simulation

The transport simulation is based on a three-dimensional flow field generated with ModFlow and provided by the LW. It includes the five different geological layers described above. The different layer-specific and transport-relevant parameters can be found in Table 7.1. For transport simulation, the model domain is discretized with a regular grid $(50\text{m} \times 50\text{m} \times 5\text{m})$ and, in total, includes 77,589,720 cells. The catchment outline is approximated using more than 6,000,000 particles in the backward transport simulation with RPTRW (cf. Section 2.3). Once the catchment was found, some more possible contamination sources could be labeled as irrelevant (cf. Section 5.2). These were either not within the calculated catchment, or could not hazard the production wells, because of subcritical contaminant mass or too long travel times. In Figure 7.2,

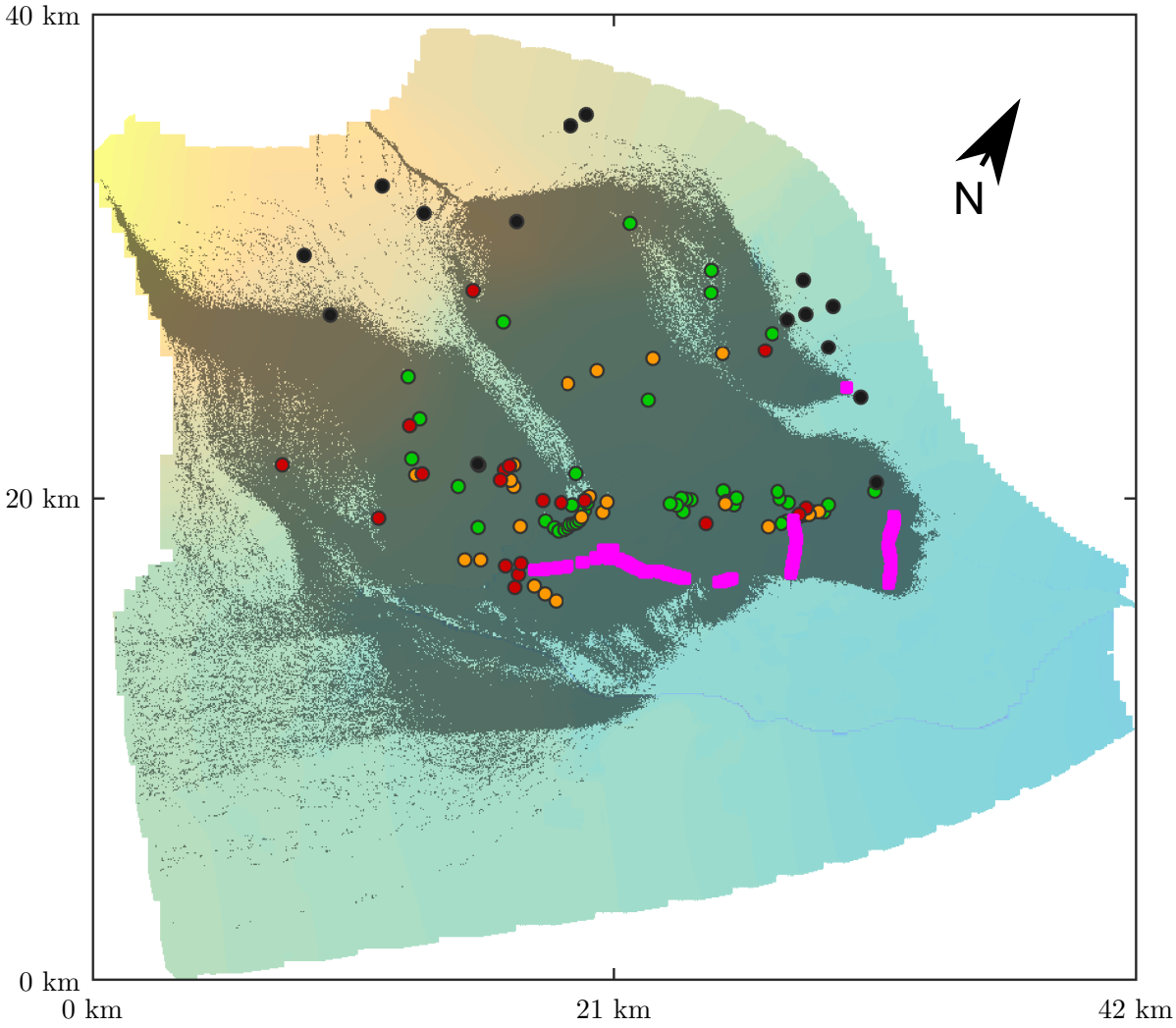


Figure 7.1: Top view of the real-case model domain. The colored background area represents the groundwater level (downward gradient from yellow to turquoise). Purple squares: production wells; Gray overlay: simulated well catchment; Colored dots: 143 possible contamination sources that are irrelevant (gray), almost tolerable (green), medium (yellow), severe (red).

these irrelevant possible sources are marked as gray dots. The number of relevant sources could be reduced to 120.

The forward transport simulation from the possible contamination sources was performed with PTRW (cf. Section 2.3) and 5000 particles per source. This simulation is the basis of the multi-objective optimization (cf. Section 3.1) and defines the search area of meaningful monitoring-well positions, i.e., positions where, in principle, a possible contamination source can be detected.

Altogether, the transport simulations needed around 50 hours on a single Intel Core i5-4590 CPU with a maximum frequency of 3.3GHz, and 32 GB RAM. While most of the simulation time was required for defining the catchment (41 hours), the forward transport simulation only needed nine hours.

7.1.2 Optimization

The optimization is based on the risk-prioritized optimization formulation introduced in Equation 5.9, with detection probability f_{det} and early-warning time utility f_{warn} for each of the three risk classes. For the early-warning time utility function f_{warn} the three user-defined values (cf. Section 3.2.2) are set to: the minimum desirable early-warning time $\hat{t}^{\min} = 2$ years, the corresponding early-warning time utility $\hat{u}^{\min} = 0.75$, and the maximum desirable early-warning time $\hat{t}^{\max} = 30$ years. The cost function f_{cost} is split up in installation costs $f_{\text{cost}}^{(\text{i})}$ and operational cost $f_{\text{cost}}^{(\text{o})}$. The installation costs $f_{\text{cost}}^{(\text{i})}$ requires a detailed drilling-cost map that was provided by the LW. Besides the drilling costs, it also includes fixed installation costs (e.g., the planning-costs of a monitoring well). The operational cost function $f_{\text{cost}}^{(\text{o})}$ is a summation of all operational costs, i.e., sample drawing and sample analysis, averaged maintenance and repair costs of the planned monitoring wells, etc.. That is, the problem formulation includes eight objective functions:

$$\mathbf{d}_{\text{opt}} = \underset{\mathbf{d}\in\mathbf{D}}{\arg\min} \left[f_{\text{det}}^{(\text{r})}, f_{\text{warn}}^{(\text{r})}, f_{\text{det}}^{(\text{y})}, f_{\text{warn}}^{(\text{y})}, f_{\text{det}}^{(\text{g})}, f_{\text{warn}}^{(\text{g})}, f_{\text{cost}}^{(\text{i})}, f_{\text{cost}}^{(\text{o})} \right]. \tag{7.1}$$

Table 7.1: Transport relevant parameters

Layer	Longitudinal dispersivity α_ℓ	Transversal dispersivity α_{t}	Volumetric porosity Φ_{v}	Effective porosity Φ_{e}
1	20	1	0.15	0.15
2	20	1	0.035	0.035
3	150	7.5	0.01	0.01
4	150	7.5	0.01	0.01
5	150	7.5	0.0003	0.0003

From the transport simulation and its spatial resolution, we obtain a set of 924,520 possible monitoring-well positions (cf. gray overlay in Figure 7.2). The optimal reduction method introduced in Section 4.3.3, reduces these to 7,660 possible monitoring-well positions (white pluses in Figure 7.2). Please note that possible monitoring-wells also have a vertical component, not shown in Figure 7.2. Figure 7.2 is the top view of the model domain and possible monitoring-well positions in vertical direction are projected to the top. Hence, they cannot be visually distinguished in vertical direction.

The optimization was performed with the Borg MOEA (cf. Section 4.1.1). The default settings for the genetic operators can be found in Table 4.4. The ϵ-box precision was set to €1000 for installation costs, to €100 for operational costs, and to 0.01 for all detection-probability and early-warning time utility functions. I used the linear-indexing representation introduced in Section 4.2.2. Each solution is defined by 121 decision variables: 60 binary decision variables that define the number of actual monitoring wells, 60 integer variables that define the position of the monitoring wells, and a single decision variable that specifies the sampling frequency (one, four, 12, or 48 times per year for all planned wells). So, overall, the actually reduced search space still contains $\sim 4.3 \times 10^{151}$ possible monitoring-network designs.

On the hardware mentioned above, the reduction took approximately one hour and the optimization took approximately four hours.

7.1.3 Results and Discussion

Figure 7.3 shows a selection of the obtained approximation set (black lines). All objective values are normalized and their ideal value is one. While the early-warning time utility and detection probability functions are between zero and one per definition, the two cost functions are normalized with the maximum values over all solutions. Here, the figure only shows 10% of the entire solutions, i.e., 5,591 out of 55,909 solutions, selected after the criterion of best compromise for detection probability for the three different risk classes.

Figure 7.3 shows that the optimization algorithm was unable to find solutions with satisfying detection probabilities for almost tolerable (max. $\sim$ 45%) and for the medium (max. $\sim$ 60%) possible contamination sources. The maximum detection probability for severe sources is at around 90%. However, analysis of the transport simulation shows that without any limitation in the number of possible monitoring wells the maximum reachable detection probability values for this problem are 69% for almost tolerable, 66% for medium, and 97% for severe possible contamination sources. That is, the differences between the actual goal attainment of the optimization and the maximum possible values express the difficulty of the optimization problem.

The red line in Figure 7.3 represents the objective values of the monitoring network that is illustrated in Figure 7.2 as blue crosses. It is defined by 29 monitoring wells

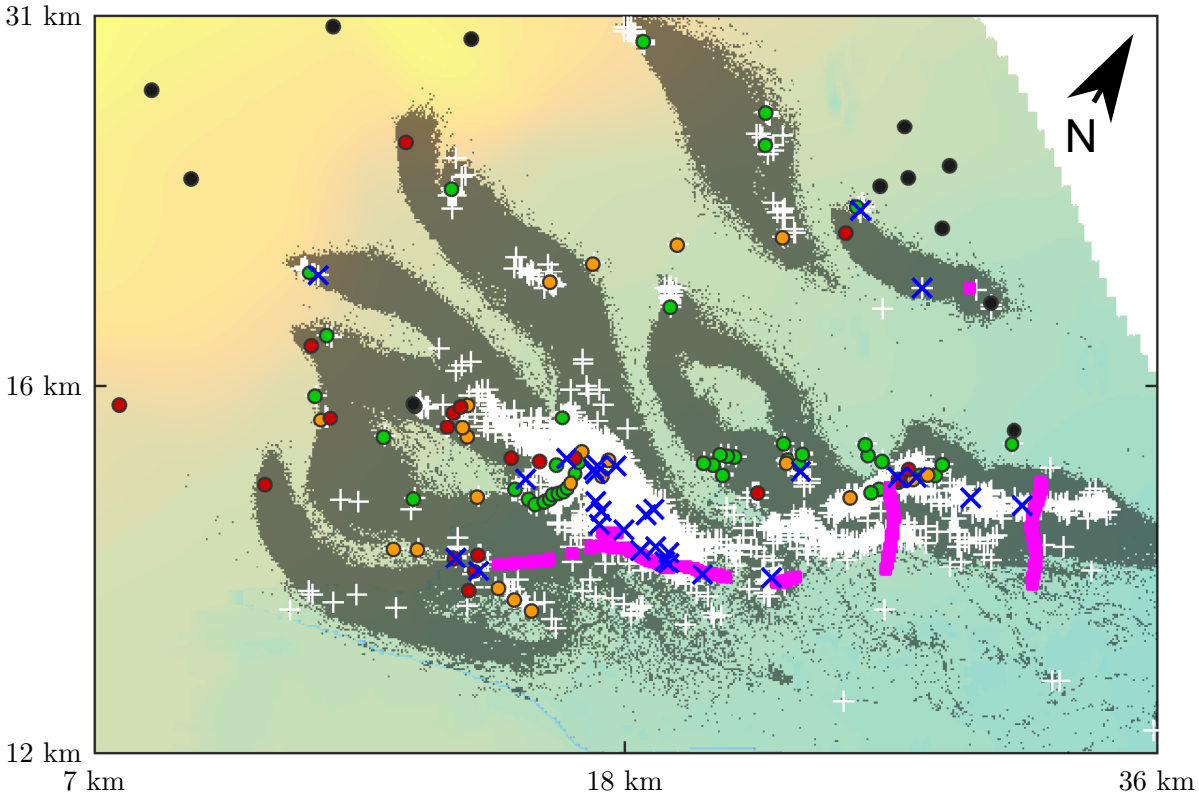


Figure 7.2: Detail of the top view from Figure 7.1. Gray overlay: possible monitoring-well positions; White pluses: remaining possible monitoring-well positions after optimal reduction (cf. Section Section 4.3.3); Blue crosses: exemplary optimized monitoring network; all other mark/colors: same as in Figure 7.1.

and provides a detection probability of 90% for the severe, 38% for the medium, and 26% for the almost tolerable possible contamination sources. The early-warning time utilities are between 0.01 and 0.04. The installation costs are around €$2.2 \times 10^6$and the operational costs are €3.2×10^5 per year. The high operational costs are caused by the high sampling frequency of 48 samples per year and monitoring well. This sampling frequency is necessary because there are several possible contamination sources with a short visibility at the monitoring-well positions. With lower sampling frequencies, detection probability would fall for all sources.

Most of the monitoring wells form a barrier close to the production wells, where many possible contamination sources overlap. Especially, the group of sources at $[18\text{km}, 15\text{km}]$ needs heavy monitoring by $\frac{2}{3}$ of all monitoring wells. Some of the remaining monitoring wells protect the production wells in a selective way, i.e., they monitor in a one-to-one situation (e.g., at $[12\text{km}, 21\text{km}]$ or the monitoring wells at $[27\text{km}, 21\text{km}]$).

Altogether, in absolute values, the presented monitoring network is expensive in installation and operation. However, this problem formulation ignores already existing monitoring wells that could be easily included by adapting the drilling-cost map (no costs at positions with already existing monitoring wells). The detection probability of severe and medium sources can only be increased by increasing the sample frequency. Focusing only on the severe possible contamination sources could heavily reduce operational and installation costs because the number of relevant sources would drop from 120 to 22. The LW has run its own various versions of this optimization problem with an earlier version of the methods developed here, released as a software package called PROMETEUS. PROMETEUS was developed within the project together with the DVGW introduced in Section 1.1. The results of the LW are published in the final report of this project [67].

As a conclusion, in this section, I exemplary showed that the methods introduced in the previous chapters are serving the goal to enable water suppliers the use of academic concepts regarding the optimization of monitoring networks under uncertainty. Results show, that efficient and effective monitoring networks can be found in reasonable computing time (in total, 55 hours were needed for the entire process).

7.2 Communication Strategies

In this section, I present suggestions for how collaboration between the practical and the academic sides can be improved on projects that include uncertainty as challenge. Three key strategies that emerged during the project (introduced in Chapter1) are emphasized in the following subsections. All of them required frequent communication

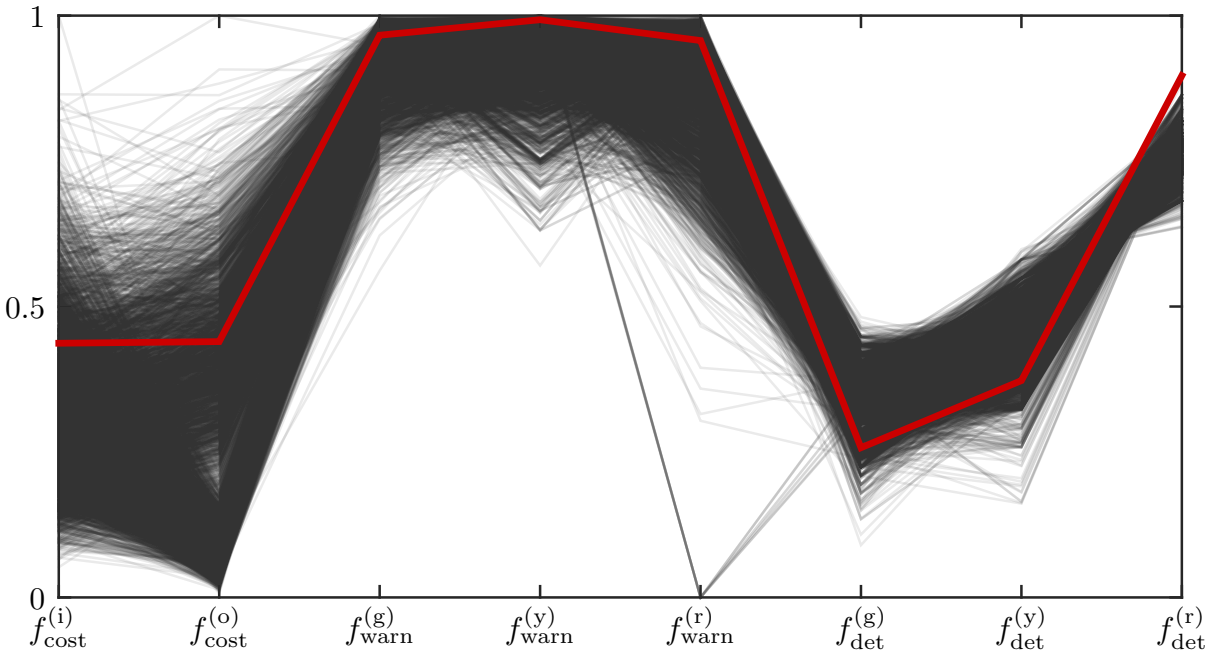


Figure 7.3: Selection of the Pareto-approximation set. X-axis: objective functions; Y-axis: normalized objective values (cost-values were normalized with the costs of the most expensive network); Black lines: non-dominated solutions; Red line: solution illustrated in Figure 7.2 as blue crosses.

through meetings involving the entire project group or sub-groups comprised of, e.g., only the operating engineers or only single water suppliers.

Speaking the Same Language Is Important

The biggest challenge and most important element in the collaboration was to develop a common language. This language has to be developed from the beginning of the project and has to be continually evolved as new challenges and opportunities arise. Clear communication is especially important for concepts related to uncertainty, risk, and optimality because common conceptions and intuitions often differ starkly from scientific definitions in these realms. As a result, it can be important to state all definitions, even those that may seem obvious, to avoid miscommunication. Our strategy during the project was to maintain a close connection between the academic team and one water supply company that represented the practical side. The process involved long discussions, exacting explanations, translations, and interpretations; but, it resulted in glossaries that could be shared across the entire project group. The results found rapid acceptance within the entire group because both practitioners and academics were involved in their generation.

Good Visualization of Methods and Concepts Is Very Helpful

Whenever the project group developed or introduced a new method or concept, it was important that it was explained clearly to every participant as quickly as possible. While a large body of literature focuses on visualization techniques of scientific data [e.g., 55, 81, 82], conceptual illustrations are strongly related to the problem; hence it is difficult to develop general rules for designing them. However, simple and illustrative figures (as simple as possible, as complex as necessary [e.g., 151]) were the most impactful medium for communicating new methods and concepts. An example of simple and story-telling illustration is given in Figure 7.4 that shows the line-of-attack concept introduced in Section 5.6. A combined task force of one representative water supplier and some academic participants together created story-telling visualizations and illustrative flow charts to capture the key concept. The practitioners usually led this effort because they had direct knowledge of what would be accessible and informative to other practitioners. The academic side helped to prioritize the most important aspects of each idea that kept the analyses on solid scientific grounding. In the best case, these illustrations were intuitively understandable in such a way that they worked as anchors for the represented concepts after they were introduced and immensely simplified communication during the project.

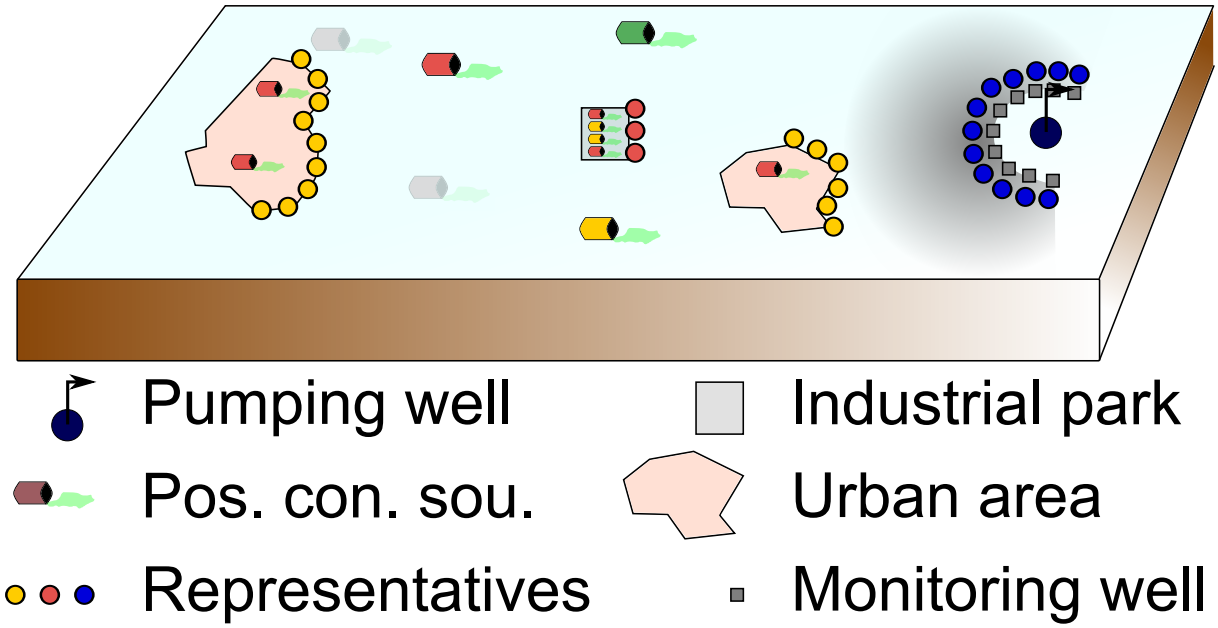


Figure 7.4: Example of a story-telling illustration. The figure illustrates the line-of-attack concept to tackle the problem of unknown possible contamination sources. The blue dots represent virtual contamination sources and the gray squares are the respective monitoring network. The gray shadow illustrates the monitoring direction of the network.

Finding the Same Objectives Supports the Collective

A strong collaboration results in a whole that is greater than the sum of its parts. To achieve this, each group has to have its own motivations and objectives that drive their contributions to the team effort. This statement, which applies at a technical level, also holds at the higher level of the overall purpose of the project. In the extreme, the practical participants might be focused on developing less expensive, innovative tools to design better monitoring networks. Meanwhile, the academic participants may be most interested in gaining access to a large pool of real data to test their methods. As with all collaborations, much of the effort lies in aligning these differing motivations to leverage the strengths of each partner. We found that it was useful to acknowledge the separate motivations of the partners but to work openly to define a joint objective of developing a useful and practical software with state-of-the-art methods and perspectives. This required relaxation of expectations on the academic side because purely academic methods could not always be applied feasibly. However, in some cases, this required flexibility from the practitioners, who sometimes assume that seemingly complex methods were not worth the time or effort. Often, the academic partners made use of jointly-developed illustrations to communicate the methods under consideration. Interestingly, this tension led to balanced but innovative approaches, as described in Section 5.4. Ultimately, the shared overall objective was well served by the practical fil-

ter placed on the academic ideas and by the requirement that both sides work to find a common language and compelling illustrations to understand all aspects of the project. This process further benefited from the diversity of the practical side, which included decision makers, field engineers, modelers and administrative employees for different water suppliers working in very different well catchments. Additionally, employees from the ministry and representatives from the head organization (DVGW) participated. Each of them brought different perspectives and placed different emphases on individual technical problems, needs, and concerns.

In summary, during the project, communication was critical. I found the following three strategies to be most useful:

1. **Terminology:** The project group developed glossaries for uncertainty and risk terminology to ensure that we had a common language. This basic vocabulary minimized unnecessary misunderstandings and formed the foundation for building common understanding as the group chose and developed more involved analyses later in the project.

2. **Illustrations:** The project group worked as small, collaborative, multidisciplinary groups to create clear illustrations to visualize and understand concepts related to scientific analysis, uncertainties, and risks. These made it easy for everyone involved to understand these concepts. The illustrations were critical for building solid understanding of complex concepts across all partners. Importantly, the practitioners led in the development of these illustrations because they were in the best position to know whether a figure would communicate concepts clearly to those who would use the software in practice.

3. **Joint objectives:** The project group recognized that each participant had different main goals when forming the team, but we committed to developing a joint objective that we could all use to guide constructive collaboration.

The following three general conclusions may extend beyond this project:

1. **Flexibility:** Practitioners are aware of the *uncertainty problem* and have developed strategies to conduct their operations under uncertainty. Academics do not need to explain uncertainty or convince practitioners to care about it. But, they must clearly explain why their methods add value to practitioners by better defining critical uncertainties. However, academics must understand that many of their methods to quantify uncertainty are not feasible in practice because of data limitations or computational demands. Therefore, academics must be flexible in modifying and simplifying their approaches to provide the best possible insights within budgetary limits.

2. **Direction:** In some cases, the practitioners' desired level of simplicity would require unacceptable (from the viewpoint of academics) modifications of academic

methods. Then, it is incumbent upon academics to work with practitioners to develop clear illustrations that can explain the need for the more advanced analyses. Both sides must approach this process openly, nudging toward a common solution and avoiding the temptation to demand simplicity or rigor.

3. **Communication:** The importance of communication cannot be overstated: from initial efforts to develop common terminology, to the development of a joint objective, to continually developing visual aides to ensure integrated understanding. This willingness to spend the time required to share ideas is the only way to demonstrate the honesty, transparency, and mutual appreciation required to leverage each partner's contributions toward a common goal. Each partner will bring unique strengths to the process, but communication is the quintessence of a successful collaboration.

Chapter 8

Summary, Conclusions, and Outlook

In this chapter, I will summarize my contributions presented in this thesis (Section 8.1), and give some concluding remarks on what I have learned during the transfer of the developed methods to practice (Section 8.2). In the final Section 8.3, I present some ideas for further research that are based upon this thesis.

8.1 Summary

The overarching goal of this thesis was to develop appropriate concepts and computationally efficient methods for the search of optimal monitoring networks in drinking-water well catchments. In most well catchments, many possible contamination sources endanger fresh-water aquifers and water supply wells. To control the emerging risk, monitoring networks are typically used, which track the groundwater quality prior to production. Finding optimal monitoring networks, however, is challenging because their objectives are often manifold and competing. The three main objectives I considered in this thesis were (1) to maximize the detection probability of spilled contamination, (2) to maximize the early-warning time for water suppliers to provide sufficient reaction time for installing countermeasures, and (3) to minimize the overall costs of the monitoring network. Beyond the conceptualization for the optimization problem, objectives for the involved computational methods were (4) to be efficient, and (5) robust against either statistical uncertainty, scenario uncertainty, or unknown unknowns. The work was strongly influenced by a project called *Risk-Based Groundwater Monitoring for Well Head Protection Areas* together with the *German Technical and Scientific Association for Gas and Water* (DVGW) and national water-supply companies.

Within the four Chapters 3 to 6 that contains my original contributions, I approached the overall problem from different directions:

1. In Chapter 3, I formulated the multi-objective optimization problem and developed the corresponding objective functions: detection probability, early-warning time, and costs. I presented aggregation rules of the objective values, (1) if a monitoring network monitors more than a single possible contamination source, and (2) if the transport model is subject to predictive uncertainty and multiple scenarios or realizations are considered. In Chapter 3, I also developed two test cases: a zonation-based model, called Z_Based, that abstracts typically used models

from water-supply companies, and U_Protect, an urban source water protection test case that captures key complexities for real-world source water protection in groundwater-based water supplies. In an exemplary discussion on U_Protect, I showed that the three considered objectives are in fact competing, and hence that multi-objective optimization is an appropriate approach to solve the optimization problem.

2. In Chapter 4, I investigated methods in detail for search improvements for multi-objective and discrete combinatorial optimization problems. I developed (1) an efficient and reliable search-space representation, the linear-indexing representation, that can handle large decision spaces without significant computational losses, and (2) I contributed a search space reduction method, the optimal reduction, that provides speedy search while ensuring that the attained multi-objective trade-offs are high-quality approximations to the true Pareto-optimal solutions.

   In the context of the U_Protect test case, I showed that both the linear-indexing representation and the optimal reduction method clearly improved optimization results, compared to selected other representation and reduction methods, e.g., binary representation and grid-based reduction.

   The overarching conclusion is that, for discrete problems, linear-indexing representation is the key for a reliable and efficient optimization performance. Additionally, when feasible, the proposed optimal reduction helps to improve quality and end-point value of multi-objective analyses for the design of groundwater monitoring systems for source water protection. These insights can be transferred to a broad class of optimization problems. My results enable the feasibility of multi-objective optimization on standard desktop computers in operationally valid runtimes.

3. In Chapter 5, I presented possible strategies to apply academic concepts in practice such that they are frugal enough to run on standard personal computers, and that they can be operated with realistically available data. These included strategies regarding the transport simulation, a qualitative risk-prioritization approach that enables water suppliers to account for their own (subjective) fears against possible contamination sources, a strategy to minimize the effort for screening areas with many possible contamination sources, and a risk model that represents unknown possible contamination sources and enable an effective and robust monitoring.

   The main points I concluded in this chapter were (1) that Monte-Carlo simulations are not feasible to represent uncertainty in practice, and well-selected scenarios should be preferred, (2) that a qualitative risk prioritization within the optimization helps find cost-efficient monitoring networks, and (3) that unknown possible contamination sources can be considered by a proper risk model. Overall, the presented utilitarian approaches serve the goal to adopt computationally

demanding and data-hungry complex statistical methods for practical application.

4. In Chapter 6, I investigated the physical mechanisms and their uncertainties that control the optimal placement of monitoring wells. I considered the most simple case: A single monitoring well is positioned in order to observe whether or not a single possible point-source has emitted a contaminant plume or not. I developed analytical solutions for the optimal monitoring position and its robustness as functions of several parameters of interest. Because the solutions are provided closed form, it is easy to analyze the effects of the controlling factors and their uncertainty. The solutions can be used to provide decision support for evaluating and upgrading existing monitoring networks, acknowledging the inherent uncertainty through the evolvement of detection probabilities.

In Chapter 7, I applied the developed methods and strategies on a real case and presented communication strategies for a successful collaboration between academia and industry. The three main strategies found during the project with the DVGW and national water suppliers were (1) that a clear terminology is key for a mutual understanding, (2) that clear and story-telling illustrations help understand complex scientific concepts and methods regarding uncertainty analysis and risk assessment, and (3) that a joint objective between the collaboration partners guides the right direction of the project.

From the project, I found three general conclusions regarding collaborations between academia and practice. First, practitioners are aware of the *uncertainty problem*, but they do not have instruments and methods to properly tackle it. The academic methods, however, are often not feasible and need modifications in data hungriness and computational efficiency. Second, too much simplification on complex scientific methods is not constructive and academics and practitioners need to find a trade-off solution between simplicity and rigor. Finally and most importantly, communication is the quintessence of a successful collaboration.

8.2 Conclusions From Transfer to Practice

In this section, I will make some general conclusions I have learned during the collaboration within the DVGW project. Parts of this section have been published in Bode et al. (2018b) [15].

Collaboration It can be useful to recognize that the process of jointly defining an academic/practical partnership is a multi-objective optimization exercise between scientific priorities and practical utility. There is a useful creative tension that arises when

practitioners demand simple approaches and academics are required to explain and advocate for more advanced analyses. The process winnows potentially expensive approaches with little practical benefit while forcing all partners to clearly understand the relevance of all aspects of the study. The key point is that academics should listen to the needs of the users, design the analyses to provide them with useful information and provide tools to help them to define their preferences and to integrate them with the science.

Optimization Multi-objective optimization is a powerful concept, but it needs to be communicated clearly as a strength for decision makers to achieve practical buy-in.

Risk From the practical side, approaches to risk assessment that stressed clarity, simplicity, and applicability were preferred; in general, qualitative risk assessment was preferred to rigorous, but less practical, quantitative approaches. In some cases, the only feasible risk analysis will be fully qualitative, based on subjective public concerns and expert judgments of stakeholders. This will require thoughtful collaboration to define the appropriate model ensemble and to decide how to apply it to decision making.

8.3 Outlook

In the following, I present some ideas for further research that are built upon my thesis.

- In Chapter 3, I introduced the objective-function formulation of the cost function that is a simple summation of arising expenses and does not consider consequential costs. I argued that the consequential costs are one of the main motivation for water suppliers to invest in optimal monitoring networks, hence there is no need for considering these costs. In practice, however, politicians, stakeholders, and decision-makers often must be convinced to invest in such networks who are not directly involved in the daily business of water-providing companies. Then, considering and modeling consequential costs can be very persuading and could be a necessary step to help protect fresh-water aquifers.
- The three main objectives of the monitoring networks to be fulfilled are detection probability, early-warning time, and costs. The optimized networks should control the risk coming from the identified possible contamination sources. In case of a groundwater contamination and a detection by a monitoring well, the water provider is warned but has no information about the contamination source. In future work, the formulated optimization problem could be extended by an objective function that evaluates the ability of monitoring-well locations and monitoring networks to identify the actual contamination source. Monitoring networks

that are designed to also consider source identification would minimize the contaminated area, hence would minimize remediation costs of the contaminated soil.

- In Chapters 3 and Chapter 4, I investigated optimization under uncertainty. When considering uncertainty, the stability of the optimization strongly depends on the aggregation function over the ensemble of scenarios/realizations. Here, it would be interesting to determine, how much uncertainty can be tolerated by the optimization problem and the different aggregation functions until the optimization collapses. The goal would be to find optimal trade-offs between robust optimization and uncertainty reduction, and the question is: when should the decision maker invest in uncertainty reduction and how long can the optimization still find satisfying results under increasing uncertainty?

- In Chapter 4, I presented the linear-indexing representation of the search space. The corresponding index vector emulates the actual search space and simplifies the sampling and mutation/recombination step during the optimization. The random sampling is based on a uniform distribution. The idea for further research is to sort the elements of the index vector by characteristics of their objective values via, e.g., k-means clustering, and to provide multiple additional sampling distributions that are related to these different clusters. The sampling distribution used for a solution could be assigned by an additional decision variable and must be the same in the mutation/recombination step for a defined number of function evaluations. Ongoing research already showed that this idea can increase the diversity of the Pareto-front approximation by better exploring the edges of the front.

- In Chapter 5, I stated that well-selected scenarios are better suited than Monte-Carlo simulation in practice for uncertainty representation. Although I clearly defined what exactly is meant by well-selected, it might be difficult for water suppliers to find well-selected scenarios that satisfy the definition. Therefore, ongoing research should be invested in developing guidelines for smart and easy scenario selections.

- Chapter 6 provides analytical solutions for optimal monitoring location under uncertainty for single possible contamination sources. These solutions could be extended to multiple possible contamination sources.

- Overall, for practical application, it would be exceptional to have a tool that evaluates the quality of sequentially built monitoring networks to better guide the effective installation. Water providers would clearly benefit from it since they could investigate, which monitoring well brings the most gain in quality. Thus, they could prioritize the sequence of installation.

Appendix A

Analytical Solutions

A.1 Derivation of the 2D Steady-State Advection-Dispersion Equation

We consider the advection-dispersion equation for uniform and parallel flow in x-direction and a uniform dispersion tensor **D**:

$$\frac{\partial C}{\partial t} + v_{\mathrm{x}} \cdot \nabla C - \nabla \cdot (\mathbf{D} \nabla C) = 0\,, \tag{A.1}$$

with concentration C, time t and velocity v_{x} (v in the following). For a two-dimensional, steady-state case, Equation A.1 simplifies to

$$v\frac{\partial C}{\partial x} - D_{\mathrm{xx}}\frac{\partial^2 C}{\partial x^2} - D_{\mathrm{yy}}\frac{\partial^2 C}{\partial y^2} = 0\,. \tag{A.2}$$

As explained in Section 2.2, we can neglect longitudinal dispersion and the term $D_{\mathrm{xx}}\frac{\partial^2 C}{\partial x^2}$ in Equation A.2 vanishes. Then, we can apply a transformation of coordinates, substituting x by travel time τ: $x = v\tau$ and x turns into a time coordinate. The changed Equation A.2

$$\frac{\partial C}{\partial \tau} - D_{\mathrm{yy}}\frac{\partial^2 C}{\partial y^2} = 0 \tag{A.3}$$

is the one-dimensional diffusion equation, Fick's second law, respectively. For an infinite domain, the solution for Equation A.3 is

$$C\,(y,\tau) = \frac{c_1}{\sqrt{\tau}} \exp\left(-\frac{y^2}{4D_{\mathrm{yy}}\tau}\right)\,, \tag{A.4}$$

with the boundary condition

$$\lim_{y \to \pm\infty} C\,(\tau, y) = 0$$

and c_1 as an arbitrary constant [27]. With the initial condition $C\,(\tau = 0, y) = \frac{\dot{m}_{\mathrm{in}}}{v}\delta\,(y)$ with the Dirac-Delta Function [41], in Equation A.4 the constant c_1 equals $\frac{\dot{m}_{\mathrm{in}}}{v\sqrt{4\pi D_{\mathrm{yy}}\tau}}$. With the inverse transformation from τ to x, the final solution of Equation A.2 is

$$C\,(x,y) = \frac{\dot{m}_{\mathrm{in}}}{B\sqrt{4\pi D_{\mathrm{yy}} v x}} \exp\left(-\frac{y^2 v}{4D_{\mathrm{yy}} x}\right)\,. \tag{A.5}$$

The division by the aquifer depth B highlights that this solution is depth-averaged. In the following, the transverse dispersion D_{yy} is denoted as D_{T}.

Appendix B

Derivations for Chapter 6

B.1 Derivation of A_{GD}

In this section, I derive the antiderivative of y_{thres} with respect to x in Equation 6.6. As integration bounds, we use $x_1 = 0$ and $x_2 = x_{\mathrm{tip}}$. Since y_{thres} is symmetric, we can double the calculated area:

$$A_{\mathrm{GD}} = 2 \cdot \int_{x_1}^{x_2} y_{\mathrm{thres}}\,(x)\,dx \tag{B.1}$$

with

$$y_{\mathrm{thres}} = \sqrt{\frac{-4D_{\mathrm{T}}}{v} x \cdot \ln\left(\frac{C_{\mathrm{thres}} B}{\dot{m}_{\mathrm{in}}} \cdot \sqrt{4\pi v D_{\mathrm{T}} x}\right)}\,. \tag{B.2}$$

Using

$$\psi = \frac{-4D_{\mathrm{T}}}{v} \tag{B.3}$$

and

$$\rho = \frac{C_{\mathrm{thres}} B}{\dot{m}_{\mathrm{in}}} \cdot \sqrt{4\pi v D_{\mathrm{T}}} \tag{B.4}$$

Equation B.2 can be simplified and the antiderivative can be expressed as

$$A_{\mathrm{GD}} = 2\sqrt{\psi} \cdot \int_{x_1}^{x_2} \sqrt{x \cdot \left(\ln(\rho) + \frac{1}{2}\ln(x)\right)}\,dx\,. \tag{B.5}$$

This integral can be simplified by substituting

$$x = \exp\left(\frac{2}{3}z^2\right) \cdot \frac{1}{\rho^2}$$

and hence

$$dx = \exp\left(\frac{2}{3}z^2\right) z \cdot \frac{4}{3\rho^2} dz$$

to

$$A_{\mathrm{GD}} = 2\sqrt{\psi} \cdot \frac{4}{3^{\frac{3}{2}}} \frac{1}{\rho^3} \cdot \int_{z_1}^{z_2} \exp\left(z^2\right) \cdot z^2 dz\,. \tag{B.6}$$

Considering only the integral $\int \exp\left(z^2\right) \cdot z^2 dz$ we can use partial integration ($\int uv'dz = uv - \int u'vdz$) with $u = z^2$ and $v' = \exp\left(z^2\right)$. While the derivative of $u = z^2$

is simple to solve ($u' = 2z$), we can use the definition of the imaginary error function to solve the integral of $\exp\left(z^2\right)$. The imaginary error function [e.g., 123] is defined as

$$\operatorname{erfi}(z) = -i \cdot \operatorname{erf}(iz) = \frac{2}{\sqrt{\pi}} \int_0^z \exp\left(x^2\right) dx\,.$$

With that, $\int \exp\left(z^2\right) dz = \frac{\sqrt{\pi}}{2} \operatorname{erfi}(z) + const.$ and we can solve the integral from Equation B.6:

$$\int_{z_1}^{z_2} \exp\left(z^2\right) \cdot z^2 dz = \left[\frac{\sqrt{\pi}}{2} z^2 \operatorname{erfi}(z)\right]_{z_1}^{z_2} - \sqrt{\pi} \int_{z_1}^{z_2} z \operatorname{erfi}(z)\, dz\,. \tag{B.7}$$

To solve Equation B.7 we can again apply partial integration with $u = z$ ($u' = 1$) and $v' = \operatorname{erfi}(z)$. The antiderivative of the error function can be found in Abramowitz and Stegun (1965) [1]:

$$\int \operatorname{erf}(x)\, dx = x \operatorname{erf}(x) + \frac{1}{\sqrt{\pi}} \exp\left(-x^2\right) + const.$$

We still have to transform $v = \int \operatorname{erfi}(z)\, dz$ to $v = -\int \operatorname{erf}(t)\, dt$ with $t = iz$. With that, $v = z \operatorname{erfi}(z) - \frac{\exp\left(z^2\right)}{\sqrt{\pi}} + const.$ and

$$\int_{z_1}^{z_2} z \operatorname{erfi}(z)\, dz = \left[\frac{1}{2} z^2 \operatorname{erfi}(z) - \frac{1}{2\sqrt{\pi}} z \exp\left(z^2\right) + \frac{1}{4} \operatorname{erfi}(z)\right]_{z_1}^{z_2}\,. \tag{B.8}$$

Putting all together yields

$$\begin{aligned} A_{\mathrm{GD}} =& 2\sqrt{\psi} \cdot \frac{4}{3^{\frac{3}{2}}} \frac{1}{\rho^3} \int_{z_1}^{z_2} \exp\left(z^2\right) \cdot z^2 dz \\ =& 2\sqrt{\psi} \cdot \frac{1}{3^{\frac{3}{2}}} \frac{1}{\rho^3} \left[\left(2z \cdot \exp\left(z^2\right) - \sqrt{\pi} \cdot \operatorname{erfi}(z)\right)\right]_{z_1}^{z_2}\,, \end{aligned} \tag{B.9}$$

with

$$z_{1,2} = \sqrt{3 \ln(\rho) + 3 \ln\left(\sqrt{x_{0,1}}\right)}\,. \tag{B.10}$$

For solving the integral in Equation 6.6, we can summarize a constant factor γ:

$$\gamma = \left(\frac{\dot{m}_{\mathrm{in}}}{C_{\mathrm{thres}} B \sqrt{3\pi}}\right)^3 \frac{1}{v^2 D_{\mathrm{T}}}\,.$$

and calculate the integration bounds z_1 and z_2. Inserting $x_1 = 0$ and $x_2 = x_{\mathrm{tip}}$ in Equation B.10 leads to $z_1(x = 0) = \sqrt{-\infty} = i\infty$ and $z_2\left(x = x_{\mathrm{tip}}\right) = 0$, respectively. For the sake of readability, x_{tip} is defined as

$$x_{\mathrm{tip}} = \left(\frac{\dot{m}_{\mathrm{in}}}{C_{\mathrm{thres}} B} \frac{1}{\sqrt{4\pi v D_{\mathrm{T}}}}\right)^2\,.$$

With that, A_{GD} can be written as

$$A_{\mathrm{GD}} = \gamma i \cdot \left[2x \exp\left(x^2\right) - \sqrt{\pi} \cdot \operatorname{erfi}(x) \right]_{z_1}^{z_2} . \tag{B.11}$$

Inserting the integration bounds leads to the final results $A_{\mathrm{GD}} = A_{\mathrm{GD}}(z_2) - A_{\mathrm{GD}}(z_1)$:

$$A_{\mathrm{GD}} = \left| -\gamma \cdot \sqrt{\pi} \right| . \tag{B.12}$$

B.2 Derivation of the Integration Limits for A_{GD} Considering Uncertainty in Spill Location

To solve the integral in Equation 6.11, we need to find the roots of Equation 6.12. We are starting again with Equation B.2, simplified using ψ and ρ (Equations B.3 and B.4). Instead of using the uncertainty diameter D_{s} from Equation 6.12, now we are using a general shift y_0:

$$y_{\mathrm{thres}}(x) - y_0 = 0 . \tag{B.13}$$

Inserting y_{thres} into Equation B.13 yields

$$y_0^2 - \psi (x - x_0) \left(\ln(\rho) + \ln\left(\sqrt{x - x_0}\right) \right) = 0 , \tag{B.14}$$

with the spill location (x_0, y_0). Rearranging Equation B.14 yields

$$(x - x_0) \left(\ln\left(\rho^2\right) + \ln(x - x_0) \right) - \frac{2y_0^2}{\psi} = 0 . \tag{B.15}$$

Multiplying Equation B.15 with ρ^2 and rearranging yields

$$(x - x_0) \rho^2 \left(\ln\left(\rho^2\right) + \ln(x - x_0) \right) = \frac{2y_0^2 \rho^2}{\psi} . \tag{B.16}$$

Equation B.16 can be also expressed as:

$$\exp\left(\ln(x - x_0) + \ln\left(\rho^2\right) \right) \left(\ln\left(\rho^2\right) + \ln(x - x_0) \right) = \frac{2y_0^2 \rho^2}{\psi} , \tag{B.17}$$

and we can apply the Lambert W function $\mathrm{W}(\cdot)$ [e.g., 25]. It has the property that for $a = b \exp(b)$ the function is $\mathrm{W}(a) = b$. When using the Lambert W function, we get

$$\mathrm{W}\left(\frac{2y_0^2 \rho^2}{\psi} \right) = \ln(x - x_0) + \ln\left(\rho^2\right) . \tag{B.18}$$

Finally, in Equation B.16 the term $\ln\left(\rho^2\right) + \ln(x - x_0)$ can be substituted by Equation B.18 and the roots $x_{1,2}$ are:

$$x_{1,2} = \frac{2y_0^2}{\psi \, \mathrm{W}_j\left(\frac{2y_0^2 \rho^2}{\psi} \right)} + x_0 , \tag{B.19}$$

with $j \in \{-1, 0\}$.

B.3 Derivation of the Integration Limits for A_{GD} Considering Uncertainty in Ambient Flow Direction

In order to calculate the integral in Equation 6.17, we need to find the intersection point X_{int} in Figure 6.4. Again, we start with Equation B.2, simplified using ψ and ρ (Equations B.3 and B.4) and now look for the roots of

$$y_{thres}(x) - \tan(\beta)\, x = 0\,, \tag{B.20}$$

with the uncertainty angle β. Rearranging Equation B.20 yields

$$x^2 - \frac{\psi \ln(\rho)}{\tan^2(\beta)} x - \frac{\psi}{2\tan^2(\beta)} x \ln(x) = 0\,. \tag{B.21}$$

Since we are not interested in the trivial solution $x = 0$ we only consider the equation

$$x - \frac{\psi \ln(\rho)}{\tan^2(\beta)} - \frac{\psi}{2\tan^2(\beta)} \ln(x) = 0\,. \tag{B.22}$$

Equation B.22 can be rearranged to

$$-\frac{2\tan^2(\beta)}{\psi} \exp(\ln(x)) = -\ln\left(x\rho^2\right)\,. \tag{B.23}$$

Multiplying Equation B.23 with $\exp\left(-\ln\left(x\rho^2\right)\right)$ yields

$$\begin{aligned} &-\frac{2\tan^2(\beta)}{\psi} \exp(\ln(x)) \exp\left(-\ln\left(x\rho^2\right)\right) = \\ &-\ln\left(x\rho^2\right) \exp\left(-\ln\left(x\rho^2\right)\right)\,. \end{aligned} \tag{B.24}$$

Equation B.24 can be also expressed as:

$$-\frac{2\tan^2(\beta)}{\psi}\frac{1}{\rho^2} = -\ln\left(x\rho^2\right) \exp\left(-\ln\left(x\rho^2\right)\right)\,, \tag{B.25}$$

and we can again apply the Lambert W function $\mathrm{W}(\cdot)$:

$$\mathrm{W}\left(-\frac{2\tan^2(\beta)}{\psi}\frac{1}{\rho^2}\right) = -\ln\left(x\rho^2\right)\,. \tag{B.26}$$

Finally, in Equation B.23 the term $-\ln\left(x\rho^2\right)$ can be substituted by Equation B.26:

$$x = -\frac{\psi}{2\tan^2(\beta)} \mathrm{W}\left(-\frac{2\tan^2(\beta)}{\psi}\frac{1}{\rho^2}\right)\,. \tag{B.27}$$

B.4 Why it is Appropriate to Only Consider Uncertainty in Transverse Direction to the Flow of the Spill Location

In Section 6.3 we substitute the Ω_U-circle by uncertainty in transverse direction to the flow. Here, I show why this substitution is valid without losses. We are starting with Equation B.19 (this equation calculates the integration limits for A_{GD} considering uncertainty in spill location (cf. Section 6.3)) and substitute the spill location (x_0, y_0) in polar coordinates for the parameterized circle: $x_0 = r \cdot \cos(\alpha)$ and $y_0 = r \cdot \sin(\alpha)$. With that, the roots $x_{1,2}$ depend only on the angle α, assuming the radius r of the circle is known and constant:

$$x_{1,2}(\alpha) = \frac{2r^2 \sin^2(\alpha)}{\psi \, W_j\left(\frac{2r^2 \sin^2(\alpha)}{\psi}\rho^2\right)} + r \cdot \cos(\alpha) \,, \tag{B.28}$$

with $j \in \{-1, 0\}$. Next, we take the derivative of Equation B.28 with respect to α, using the quotient rule and the derivative for the Lambert W function $\left(\frac{dW(x)}{dx} = \frac{W(x)}{x(1+W(x))}, \text{ for } x > -\frac{1}{e} \Rightarrow j = 0\right)$ [cf. 25]:

$$\frac{\partial x_{1,2}(\alpha)}{\partial \alpha} = \frac{2r^2 \sin(2\alpha)}{\psi\left(1 + W_0\left(\frac{2r^2 \sin^2(\alpha)}{\psi}\rho^2\right)\right)} - r \sin(\alpha) \,. \tag{B.29}$$

The next step is to find the roots of Equation B.29:

$$\frac{2r^2 \sin(2\alpha)}{\psi\left(1 + W_0\left(\frac{2r^2 \sin^2(\alpha)}{\psi}\rho^2\right)\right)} - r \sin(\alpha) = 0 \,. \tag{B.30}$$

This cannot be solved analytically because we still have a $\sin(\alpha)$ in the argument of the Lambert W function. However, a parametric study with Equations B.29 and B.30 over $r \in \left[0, \frac{D_s}{2}\right]$ is shown in Figure B.1 and B.2. Here, we can see that the combinations of $r = \frac{D_s}{2}$ and $\alpha = \pm 90°$ are the limiting cases of A_{GD} and the Ω_U-circle can be represented appropriately by uncertainty transverse to the direction of flow.

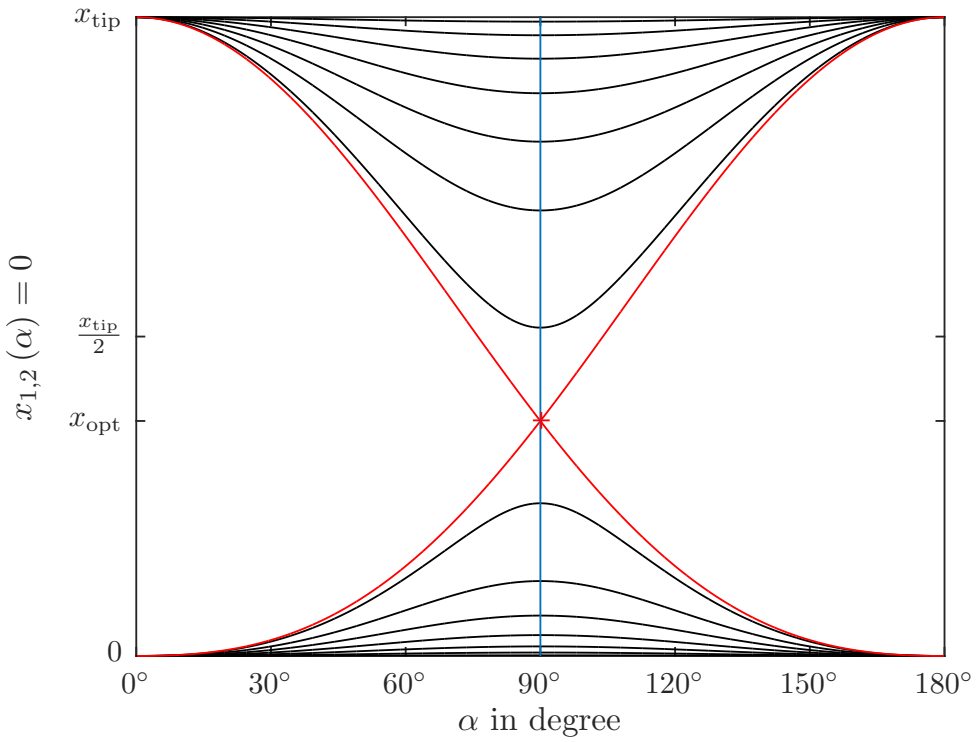


Figure B.1: Intersection of the plume outlines $(y_{\text{thres}}(x))$ and the x-axis over starting positions α of the plumes on the edges of Ω_U parametrized as a circle. The starting positions are limited to the outline of a circle that has a radius $r \leq \frac{D_s}{2}$. The set of curves (black) is defined by varying values of $r \in \left[0, \frac{D_s}{2}\right]$. The red line illustrates the result for $r = \frac{D_s}{2}$. The vertical blue line illustrates that the extreme values of all curves are independent of r and occur for $\alpha = 90°$.

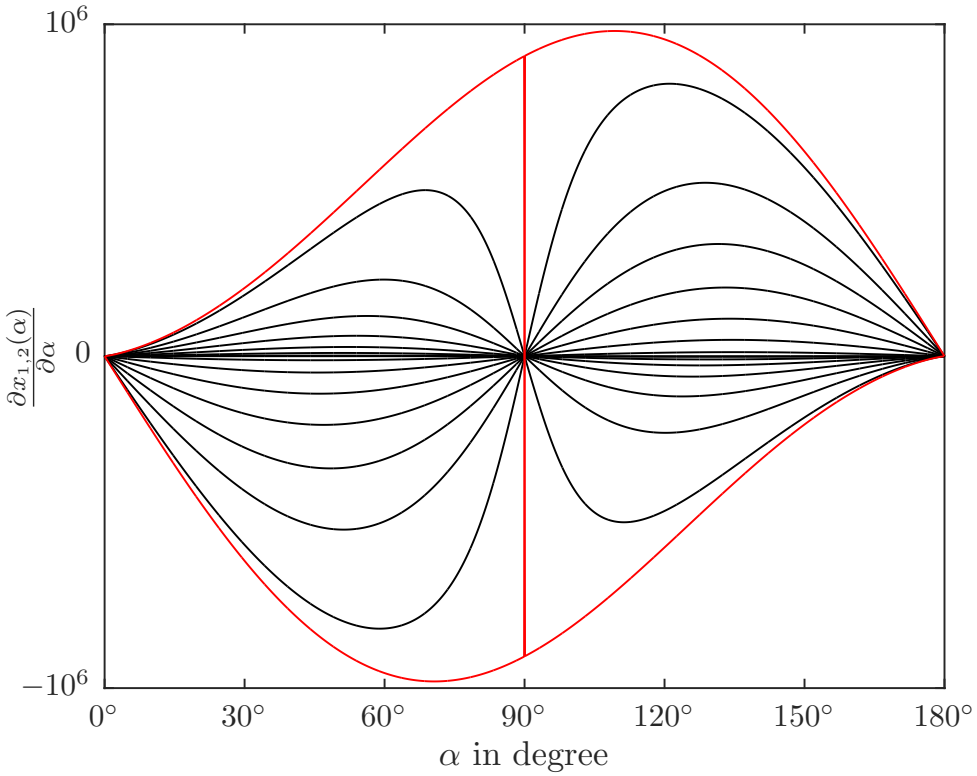


Figure B.2: The derivative of root-function shown in Figure B.1 over the angle α. The set of curves (black) is defined by varying values of $r \in \left[0, \frac{D_s}{2}\right]$. The red line illustrates the result for $r = \frac{D_s}{2}$, that is, the uncertainty in spill location where only one overlapping point of all possible plumes exists that evolve from the edges of the related Ω_U-circle.

Bibliography

[1] Abramowitz, M. and Stegun, I. A. (1965). *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*. Dover Books on Mathematics, New York, 9 edition.

[2] Allen, E., Kruger, C., Leung, F.-Y., and Stephens, J. C. (2013). Diverse Perceptions of Stakeholder Engagement within an Environmental Modeling Research Team. *Journal of Environmental Studies and Sciences*, 3(3):343–356.

[3] Allen, R. M. and Kanamori, H. (2003). The potential for earthquake early warning in southern California. *Science (New York, N.Y.)*, 300(5620):786–789.

[4] Aller, L., Bennet, T., Lher, J., and Petty, R. (1987). DRASTIC A standardized system for evaluating groundwater pollution potencial using hydrogeologic setting. *U.S. EPA Report 600/2-87-035 Ada Oklahoma.*, 29(1):23–37.

[5] Babbar, M. and Minsker, B. S. (2006). Groundwater Remediation Design Using Multiscale Genetic Algorithms. *Journal of Water Resources Planning and Management*, 132(5):341–350.

[6] Bean, J. C. (1994). Genetic Algorithms and Random Keys for Sequencing and Optimization. *ORSA Journal on Computing*, 6(2):154–160.

[7] Bear, J. (1961). Some experiments in dispersion. *Journal of Geophysical Research*, 66(8):2455–2467.

[8] Bear, J. (1972). *Dynamics of Fluids in Porous Media*. Dover Publications.

[9] Benson, A. K., Payne, K. L., and Stubben, M. A. (1997). Mapping groundwater contamination using dc resistivity and VLF geophysical methods - A case study. *Geophysics*, 62(1):80–86.

[10] Bode, F., Binning, P. J., and Nowak, W. (2013). What Factors Coordinate the Optimal Position of a Single Monitoring Well Down Gradient of a Hazardous Site. In *AGU Fall Meeting 2013, Abstract: H21I-1175*, San Francisco, CA, USA.

[11] Bode, F., Binning, P. J., and Nowak, W. (2014a). Enhanced Multi-objective Optimization of Groundwater Monitoring Networks. In *NGWA Groundwater Summit 2014*, Denver, CO, USA.

[12] Bode, F., Binning, P. J., and Nowak, W. (2014b). Multi-Objective Optimization of Groundwater Monitoring Networks: Computational Costs vs. Optimization Accuracy. In *XX. International Conference on Computational Methods in Water Resources (CMWR)*, Stuttgart, Germany.

[13] Bode, F., Binning, P. J., and Nowak, W. (2018a). An Analytical Approach for Positioning a Single Monitoring Well to Reliably Detect Point-Source Contaminant Plumes. *Manuscript*, In Prepara.

[14] Bode, F., Emmert, M., Zigelli, N., and Nowak, W. (2017). Optimale Grundwassermessstellennetze: Multi-kriterielle Optimierung als Entscheidungshilfe. *gwf Wasser+Abwasser*, 158:111–121.

[15] Bode, F., Ferré, T., Zigelli, N., Emmert, M., and Nowak, W. (2018b). Reconnecting Stochastic Methods with Hydrogeological Applications: A Utilitarian Uncertainty Analysis and Risk Assessment Approach for the Design of Optimal Monitoring Networks. *Water Resources Research*.

[16] Bode, F., Loschko, M., and Nowak, W. (2014c). A Risk-Based Multi-Objective Optimization Concept for Early-Warning Monitoring Networks. In *AGU Fall Meeting 2014, Abstract: H23K-1023*, San Francisco, CA, USA.

[17] Bode, F., Loschko, M., and Nowak, W. (2015). How to Decide? Multi-Objective Early-Warning Monitoring Networks for Water Suppliers. In *EGU General Assembly 2015, Geophysical Research Abstracts 17: EGU2015-9270*, Vienna, Austria.

[18] Bode, F. and Nowak, W. (2013). Optimizing early-warning monitoring systems for improved drinking water resource protection. In *NUPUS*, Bergen, Norway.

[19] Bode, F., Nowak, W., and Loschko, M. (2016a). Optimization for Early-Warning Monitoring Networks in Well Catchments Should Be Multi-objective, Risk-Prioritized and Robust Against Uncertainty. *Transport in Porous Media*, 114(2):261–281.

[20] Bode, F., Nowak, W., and Reed, Patrick. M. Reuschen, S. (2016b). Putting Man in the Machine: Exploiting Expertise to Enhance Multiobjective Design of Water Supply Monitoring Network. In *AGU Fall Meeting 2016, Abstract: H51A-1413*, San Francisco, CA, USA.

[21] Bode, F., Reed, P. M., Reuschen, S., and Nowak, W. (2018c). Search Space Representation and Reduction Methods to Enhance Multi-Objective Water Supply Monitoring Design. *Water Resources Research*, Submitted.

[22] Bode, F. and Reuschen, Sebastian Nowak, W. (2015). Never Use the Complete Search Space: A Concept to Enhance the Optimization Procedure for Monitoring Networks. In *AGU Fall Meeting 2015, Abstract: IN11B-1774*, San Francisco, CA, USA.

[23] Bosman, P. and Thierens, D. (2003). The balance between proximity and diversity in multiobjective evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 7(2):174–188.

[24] Cerny, V. (1985). Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *Journal of Optimization Theory and Applications*, 45(1):41–51.

[25] Corless, R. M., Gonnet, G. H., Hare, D. E. G., Jeffrey, D. J., and Knuth, D. E. (1996). On the LambertW function. *Advances in Computational Mathematics*, 5(1):329–359.

[26] Cox, L. A. T. (2008). What's Wrong with Risk Matrices? *Risk Analysis*, 28(2):497–512.

[27] Crank, J. (1979). *The mathematics of diffusion*. Oxford University Press.

[28] Cushman, D. J., Driver, K. S., and Ball, S. D. (2001). Risk assessment for environmental contamination: an overview of the fundamentals and application of risk assessment at contaminated sites. *Canadian Journal of Civil Engineering*, 28(S1):155–162.

[29] Dantzig, G., Orden, A., and Wolfe, P. (1955). The generalized simplex method for minimizing a linear form under linear inequality restraints. *Pacific Journal of Mathematics*, 5(2):183–195.

[30] Davison, A., Howard, G., Stevens, M., Callan, P., Fewtrell, L., Deere, D., Bartram, J., World Health Organization. Water, S., and Team, H. (2005). *Water safety plans: managing drinking-water quality from catchment to consumer*. Geneva: World Health Organization.

[31] Deb, K. (2001). *Multi-Objective Optimization using Evolutionary Algorithms*. John Wiley & Sons, Chichester, UK.

[32] Deb, K. and Agrawal, R. B. (1994). Simulated Binary Crossover for Continuous Search Space. *Complex Systems*, 9:1–34.

[33] Deb, K., Agrawal, S., Pratap, A., and Meyarivan, T. (2000). A Fast Elitist Non-dominated Sorting Genetic Algorithm for Multi-objective Optimization: NSGA-II. In *Parallel Problem Solving from Nature PPSN VI*, pages 849–858. Springer.

[34] Deb, K., Joshi, D., and Anand, A. (2002a). Real-coded evolutionary algorithms with parent-centric recombination. In *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No.02TH8600)*, volume 1, pages 61–66. IEEE.

[35] Deb, K., Mohan, M., and Mishra, S. (2003). A fast multi-objective evolutionary algorithm for finding well-spread pareto-optimal solutions. Technical report, KanGAL, Kanpur, India.

[36] Deb, K., Pratab, S., Agarwal, S., and Meyarivan, T. (2002b). A Fast and Elitist Multiobjective Genetic Algorithm: NGSA-II. *IEEE Transactions on Evolutionary Computing*, 6(2):182–197.

[37] Deb, K. and Sachin, J. (2002). Running Performance Metrics for Evolutionary Multi-Objective Optimization. Technical report, KanGAL, Kanpur, India.

[38] Deuflhard, P. (2011). *Newton Methods for Nonlinear Problems*, volume 35 of *Springer Series in Computational Mathematics*. Springer Berlin Heidelberg, Berlin, Heidelberg.

[39] Diersch, H.-J. G. (2014). *FEFLOW: Finite Element Modeling of Flow, Mass and Heat Transport in Porous and Fractured Media*. Springer Science & Business Media.

[40] Dietrich, C. R. and Newsam, G. N. (1993). A fast and exact method for multidimensional gaussian stochastic simulations. *Water Resources Research*, 29(8):2861–2869.

[41] Dirac, P. A. M. (1958). *The Principles of Quantum Mechanics*. Oxford Science Publications, 4 edition.

[42] Doherty, J. (2003). Ground Water Model Calibration Using Pilot Points and Regularization. *Groundwater*, 41(2):170–177.

[43] Dorigo, M. (1992). *Optimization, learning and natural algorithms*. Ph. d. thesis, Politecnico di Milano, Italy.

[44] Dossing Overheu, N., Tuxen, N., Flyvbjerg, J., Aabling, J., Andersen, J. A., Pedersen, J. K., Thyregod, T., Binning, P. J., and Bjerg, P. L. (2013). Risk-based prioritisation of ground water threatening point sources at catchment and regional scale. *AquaConSoil 2013 - Programme and Book of Abstracts*, 485:171–172.

[45] DVGW (2003). Messnetze zur Überwachung der Grundwasserbeschaffenheit in Wassergewinnungsgebieten. *Technische Regel - DVGW W 108*.

[46] DVGW (2006). Richtlinien für Trinkwasserschutzgebiete; Teil 1: Schutzgebiete für Grundwasser. *Technische Regel - DVGW W 101*.

[47] Eden, S., Megdal, S., Shamir, E., Chief, K., and Mott Lacroix, K. (2016). Opening the Black Box: Using a Hydrological Model to Link Stakeholder Engagement with Groundwater Management. *Water*, 8(5):216.

[48] Emmert, M., Zigelli, N., Haakh, F., Bode, F., and Nowak, W. (2016). Risikobasiertes Grundwassermonitoring für Wasserschutzgegiebte. *energie - wasser-praxis*, 8.

[49] Enzenhoefer, R. (2014). *Risk quantification and management in water production and supply systems*. Dissertation, University of Stuttgart.

[50] Enzenhoefer, R., Binning, P., and Nowak, W. (2015). STakeholder-Objective Risk Model (STORM): Determining the aggregated risk of multiple contaminant hazards in groundwater well catchments. *Advances in Water Resources*, 83:160–175.

[51] Enzenhoefer, R., Bunk, T., and Nowak, W. (2014). Nine Steps to Risk-Informed Wellhead Protection and Management: A Case Study. *Groundwater*, 52(S1):161–174.

[52] Enzenhoefer, R., Nowak, W., and Helmig, R. (2012). Probabilistic exposure risk assessment with advective-dispersive well vulnerability criteria. *Advances in Water Resources*, 36:121–132.

[53] EU (2018). Safer drinking water for all Europeans. *European Commission - Press release*, IP/18/429.

[54] Faber, M. H. and Maes, M. a. (2003). Modeling of Risk Perception in Engineering Decision Analysis. In *Proceedings 11th IFIP WG7.5 Working Conference on Reliability and Optimization of Structural Systems*, volume 7, pages 113–122.

[55] Fayyad, U., Grinstein, G. G., and Wierse, A. (2002). *Information Visualization in Data Mining and Knowledge Discovery*. Morgan Kaufmann Publishers, Academic Press.

[56] Fernàndez-Garcia, D. and Sanchez-Vila, X. (2011). Optimal reconstruction of concentrations, gradients and reaction rates from particle distributions. *Journal of Contaminant Hydrology*, 120-121:99–114.

[57] Folks, J. and Chhikara, R. (1978). The Inverse Gaussian Distribution and Its Statistical Application–A Review. *Journal of the Royal Statistical Society. Series B (Methodological)*, 40(3):263–289.

[58] Fonseca, C. M. and Fleming, P. J. (1993). Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization. *Icga*, 93(July):416–423.

[59] Freeze, R. A. (1975). A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media. *Water Resources Research*, 11(5):725–741.

[60] Frind, E., Molson, J., and Rudolph, D. (2006). Well Vulnerability: A Quantitative Approach for Source Water Protection. *Groundwater*, 44(5):732–742.

[61] Gelhar, L. W. and Axness, C. L. (1983). Three-dimensional stochastic analysis of macrodispersion in aquifers. *Water Resources Research*, 19(1).

[62] Glover, F. (1986). Future paths for integer programming and links to artificial intelligence. *Computers & Operations Research*, 13(5):533–549.

[63] Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison Wesley Pub Co Inc.

[64] Goldberg, D. E., Korb, B., and Deb, K. (1989). Messy Genetic Algorithms: Motivation, Analysis, and First Results. *Engineering*, 3:493–530.

[65] Gómez-Hernández, J. J. and Gorelick, S. M. (1989). Effective groundwater model parameter values: Influence of spatial variability of hydraulic conductivity, leakance, and recharge. *Water Resources Research*, 25(3):405–419.

[66] Gorelick, S. M., Evans, B., and Remson, I. (1983). Identifying sources of groundwater pollution: An optimization approach. *Water Resources Research*, 19(3):779–790.

[67] Haakh, F., Heß, C., Kaupe, M., Ries, J., and Nowak, W. (2016). Abschlussbericht - Risikobasiertes Grundwassermonitoring für Wasserschutzgebiete. Technical report, DVGW e.V.

[68] Hadka, D. and Reed, P. (2012). Diagnostic Assessment of Search Controls and Failure Modes in Many-Objective Evolutionary Optimization. *Evolutionary Computation*, 20(3):423–452.

[69] Hadka, D. and Reed, P. (2013). Borg: An Auto-Adaptive Many-Objective Evolutionary Computing Framework. *Evolutionary Computation*, 21(2):231–259.

[70] Hadka, D. and Reed, P. (2015). Large-scale parallelization of the Borg multiobjective evolutionary algorithm to enhance the management of complex environmental systems. *Environmental Modelling & Software*, 69:353–369.

[71] Hajela, P. and Lin, C. Y. (1992). Genetic search strategies in multicriterion optimal design. *Structural Optimization*, 4(2):99–107.

[72] Halton, J. H. (1970). A Retrospective and Prospective Survey of the Monte Carlo Method. *SIAM Review*, 12(1):1–63.

[73] Ham, P. A., Schotting, R. J., Prommer, H., and Davis, G. B. (2004). Effects of hydrodynamic dispersion on plume lengths for instantaneous bimolecular reactions. *Advances in Water Resources*, 27(8):803–813.

[74] Harbaugh, A. W. (2005). MODFLOW-2005, The U.S. Geological Survey Modular Ground-Water Model - the Ground-Water Flow Process. In *6. Modeling techniques*, chapter A16. US Geological Survey Reston, Reston, Virginia, USA.

[75] Harik, G. R. and Lobo, F. G. (1999). A parameter-less genetic algorithm. In *Proceedings of the 1st Annual Conference on Genetic and Evolutionary Computation-Volume 1*, pages 258–265. Morgan Kaufmann Publishers Inc.

[76] Hartigan, J. A. (1975). *Clustering Algorithms*. John Wiley & Sons.

[77] Harvey, C. F. and Gorelick, S. M. (1995). Temporal Moment-Generating Equations: Modeling Transport and Mass Transfer in Heterogeneous Aquifers. *Water Resources Research*, 31(8):1895–1911.

[78] Hokstad, P., Roestum, J., Sklet, S., Rosén, L., Pettersson, T. J. R., Linde, A., Sturm, S., Beuken, R., Kirchner, D., and Niewersch, C. (2009). Methods for risk analysis of drinking water systems from source to tap - Guidance report on risk analysis. *TECHNEAU*, D 4.2.4.

[79] Ishibuchi, H. and Murata, T. (1996). Multi-objective genetic local search algorithm. In *Proceedings of IEEE International Conference on Evolutionary Computation*, pages 119–124. IEEE.

[80] Jacobson, E., Andricevic, R., and Morrice, J. (2002). Probabilistic capture zone delineation based on an analytic solution. *Groundwater*, 40(1):85–95.

[81] Kehrer, J. and Hauser, H. (2013). Visualization and Visual Analysis of Multifaceted Scientific Data: A Survey. *IEEE Transactions on Visualization and Computer Graphics*, 19(3):495–513.

[82] Keim, D. (2002). Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):1–8.

[83] Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of ICNN'95 - International Conference on Neural Networks*, volume 4, pages 1942–1948. IEEE.

[84] Kinzelbach, W. (1988). The Random Walk Method in Pollutant Transport Simulation. In Custodio, E., Gurgui, A., and Ferreira, J. P. L., editors, *Groundwater Flow and Quality Modelling*, pages 227–245. Springer Netherlands, Dordrecht.

[85] Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). Optimization by Simulated Annealing. *Science*, 220(4598):671–680.

[86] Kita, H., Ono, I., and Kobayashi, S. (2000). Multi-Parental Extension of the Unimodal Normal Distribution Crossover for Real-Coded Genetic Algorithms. *Transactions of the Society of Instrument and Control Engineers*, 36(10):875–883.

[87] Kitanidis, P. K. (1986). Parameter Uncertainty in Estimation of Spatial Functions: Bayesian Analysis. *Water Resources Research*, 22(4):499–507.

[88] Kitanidis, P. K. (1997). *Introduction to Geostatistics: Applications to Hydrogeology*. Cambridge University Press.

[89] Knowles, J. and Corne, D. (2002). On metrics for comparing nondominated sets. In *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No.02TH8600)*, volume 1, pages 711–716. IEEE.

[90] Koch, J. and Nowak, W. (2014). A method for implementing Dirichlet and third-type boundary conditions in PTRW simulations. *Water Resources Research*, 50(2):1374–1395.

[91] Koch, J. and Nowak, W. (2015). Predicting DNAPL mass discharge and contaminated site longevity probabilities: Conceptual model and high-resolution stochastic simulation. *Water Resources Research*, 51(2):806–831.

[92] Koch, J. and Nowak, W. (2016). Identification of contaminant source architectures-A statistical inversion that emulates multiphase physics in a computationally practicable manner. *Water Resources Research*, 52(2):1009–1025.

[93] Kollat, J. and Reed, P. (2006). Comparing state-of-the-art evolutionary multiobjective algorithms for long-term groundwater monitoring design. *Advances in Water Resources*, 29(6):792–807.

[94] Kollat, J. and Reed, P. (2007). A computational scaling analysis of multiobjective evolutionary algorithms in long-term groundwater monitoring applications. *Advances in Water Resources*, 30(3):408–419.

[95] Kollat, J., Reed, P., and Kasprzyk, J. (2008). A new epsilon-dominance hierarchical Bayesian optimization algorithm for large multiobjective monitoring network design problems. *Advances in Water Resources*, 31(5):828–845.

[96] Kollat, J. B., Reed, P. M., and Maxwell, R. M. (2011). Many-objective groundwater monitoring network design using bias-aware ensemble Kalman filtering, evolutionary optimization, and visual analytics. *Water Resources Research*, 47(2).

[97] Kosow, H. and Gaßner, R. (2008). *Methods of Future and Scenario Analysis: Overview, assessment, and selection criteria.* German Development Institute (DIE), Bonn.

[98] Krueger, T., Page, T., Hubacek, K., Smith, L., and Hiscock, K. (2012). The role of expert opinion in environmental modelling. *Environmental Modelling & Software*, 36:4–18.

[99] Kübert, M. and Finkel, M. (2006). Contaminant mass discharge estimation in groundwater based on multi-level point measurements: A numerical evaluation of expected errors. *Journal of Contaminant Hydrology*, 84(1-2):55–80.

[100] Kunstmann, H. and Kinzelbach, W. (2000). Computation of stochastic wellhead protection zones by combining the first-order second-moment method and Kolmogorov backward equation analysis. *Journal of Hydrology*, 237(3-4):127–146.

[101] Kursawe, F. (1991). A variant of evolution strategies for vector optimization. In *Parallel Problem Solving from Nature*, pages 193–197. Springer-Verlag, Berlin/Heidelberg.

[102] LaBolle, E. M., Fogg, G. E., and Tompson, A. F. B. (1996). Random-Walk Simulation of Transport in Heterogeneous Porous Media: Local Mass-Conservation Problem and Implementation Methods. *Water Resources Research*, 32(3):583–593.

[103] Laumanns, M., Thiele, L., Deb, K., and Zitzler, E. (2002). Combining Convergence and Diversity in Evolutionary Multiobjective Optimization. *Evolutionary Computation*, 10(3):263–282.

[104] Leube, P. C., Nowak, W., and Schneider, G. (2012). Temporal moments revisited: Why there is no better way for physically based model reduction in time. *Water Resources Research*, 48(11).

[105] Liu, Y. and Minsker, B. S. (2004). Full Multiscale Approach for Optimal Control of In Situ Bioremediation. *Journal of Water Resources Planning and Management*, 130(1):26–32.

[106] Loaiciga, H. A., Charbeneau, R. J., Everett, L. G., Fogg, G. E., Hobbs, B. F., and Rouhani, S. (1992). Review of groundwater quality monitoring network design. *Journal of Hydraulic Engineering-Asce*, 118(1):11–37.

[107] Luke, S. (2013). *Essentials of Metaheuristics*. Lulu Raleigh.

[108] MacFarlane, D., Cherry, J., Gillham, R., and Sudicky, E. (1983). Migration of contaminants in groundwater at a landfill: A case study. *Journal of Hydrology*, 63(1-2):1–29.

[109] Maier, H., Kapelan, Z., Kasprzyk, J., Kollat, J., Matott, L., Cunha, M., Dandy, G., Gibbs, M., Keedwell, E., Marchi, A., Ostfeld, A., Savic, D., Solomatine, D., Vrugt, J., Zecchin, A., Minsker, B., Barbour, E., Kuczera, G., Pasha, F., Castelletti, A., Giuliani, M., and Reed, P. (2014). Evolutionary algorithms and other metaheuristics in water resources: Current status, research challenges and future directions. *Environmental Modelling & Software*, 62:271–299.

[110] Marler, R. T. and Arora, J. S. (2004). Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization*, 26(6):369–395.

[111] Massmann, J. and Freeze, R. A. (1987a). Groundwater contamination from waste management sites: The interaction between risk-based engineering design and regulatory policy: 1. Methodology. *Water Resources Research*, 23(2):351–367.

[112] Massmann, J. and Freeze, R. A. (1987b). Groundwater contamination from waste management sites: The interaction between risk-based engineering design and regulatory policy: 2. Results. *Water Resources Research*, 23(2):368–380.

[113] Matérn, B. (1986). *Spatial Variation*, volume 36 of *Lecture Notes in Statistics*. Springer New York, New York, NY.

[114] Megdal, S., Eden, S., and Shamir, E. (2017). Water Governance, Stakeholder Engagement, and Sustainable Water Resources Management. *Water*, 9(3):190.

[115] Member, S., Member, S., Veeravalli, V. V., Veeravalli, V. V., Member, S., and Member, S. (2003). Decentralized Detection in Sensor Networks. *Ieee Transactions On Signal Processing*, 51(2):407–416.

[116] Metropolis, N. and Ulam, S. (1949). The Monte Carlo Method. *Journal of the American Statistical Association*, 44(247):335–341.

[117] Meyer, P. D. and Brill, E. D. (1988). A method for locating wells in a groundwater monitoring network under conditions of uncertainty. *Water Resources Research*, 24(8):1277–1282.

[118] Meyer, P. D., Valocchi, A. J., and Eheart, J. W. (1994). Monitoring network design to provide initial detection of groundwater contamination. *Water Resources Research*, 30(9):2647–2659.

[119] Michalak, A. M. and Kitanidis, P. K. (2004). Estimation of historical groundwater contaminant distribution using the adjoint state method applied to geostatistical inverse modeling. *Water Resources Research*, 40(8).

[120] Moutsopoulos, K. N., Gemitzi, A., and Tsihrintzis, V. a. (2008). Delineation of groundwater protection zones by the backward particle tracking method: Theoretical background and GIS-based stochastic analysis. *Environmental Geology*, 54(5):1081–1090.

[121] Neupauer, R. M. and Wilson, J. L. (2001). Adjoint-derived location and travel time probabilities for a multidimensional groundwater system. *Water Resources Research*, 37(6):1657–1668.

[122] Neupauer, R. M. and Wilson, J. L. (2002). Backward probabilistic model of groundwater contamination in non-uniform and transient flow. *Advances in Water Resources*, 25(7):733–746.

[123] Ng, E. W. and Geller, M. (1969). A table of integrals of the Error functions. *Journal of Research of the National Bureau of Standards, Section B: Mathematical Sciences*, 73B(1):1.

[124] Nicklow, J., Reed, P., Savic, D., Dessalegne, T., Harrell, L., Chan-Hilton, A., Karamouz, M., Minsker, B., Ostfeld, A., Singh, A., and Zechman, E. (2010). State of the Art for Genetic Algorithms and Beyond in Water Resources Planning and Management. *Journal of Water Resources Planning and Management*, 136(4):412–432.

[125] Nocedal, J. and Wright, S. J. (2006). *Numerical Optimization*. Springer Berlin Heidelberg.

[126] Nowak, W., Bode, F., and Loschko, M. (2015). A Multi-objective Optimization Concept for Risk-based Early-warning Monitoring Networks in Well Catchments. *Procedia Environmental Sciences*, 25:191–198.

[127] Nowak, W., Schwede, R. L., Cirpka, O. A., and Neuweiler, I. (2008). Probability density functions of hydraulic head and velocity in three-dimensional heterogeneous porous media. *Water Resources Research*, 44(8).

[128] Pareto, V. (1964). *Cours d'économie politique*. Librairie Droz.

[129] Parzen, E. (1962). On Estimation of a Probability Density Function and Mode. *The Annals of Mathematical Statistics*, 33(3):1065–1076.

[130] Pollock, D. W. (2016). User guide for MODPATH Version 7 - A particle-tracking model for MODFLOW. Technical report, US Geological Survey.

[131] Pool, M., Carrera, J., Dentz, M., Hidalgo, J. J., and Abarca, E. (2011). Vertical average for modeling seawater intrusion. *Water Resources Research*, 47(11).

[132] RamaRao, B. S., LaVenue, A. M., De Marsily, G., and Marietta, M. G. (1995). Pilot Point Methodology for Automated Calibration of an Ensemble of conditionally Simulated Transmissivity Fields: 1. Theory and Computational Experiments. *Water Resources Research*, 31(3):475–493.

[133] Reed, M. S. (2008). Stakeholder participation for environmental management: A literature review. *Biological Conservation*, 141(10):2417–2431.

[134] Reed, P., Hadka, D., Herman, J., Kasprzyk, J., and Kollat, J. (2013). Evolutionary multiobjective optimization in water resources: The past, present, and future. *Advances in Water Resources*, 51:438–456.

[135] Reed, P., Minsker, B., and Valocchi, A. (2000). A Multiobjective Approach to Long-Term Groundwater Monitoring Design. In *Building Partnerships*, pages 1–5, Reston, VA. American Society of Civil Engineers.

[136] Reed, P. M. and Minsker, B. S. (2004). Striking the Balance: Long-Term Groundwater Monitoring Design for Conflicting Objectives. *Journal of Water Resources Planning and Management*, 130(2):140–149.

[137] Robert, C. P. and Casella, G. (2004). *Monte Carlo Statistical Methods*. Springer Texts in Statistics. Springer New York, New York, NY.

[138] Rubin, Y. (2003). *Applied Stochastic Hydrogeology*. Oxford University Press, New York.

[139] Saenton, S. and Illangasekare, T. (2004). Determination of DNAPL entrapment architecture using experimentally validated numerical codes and inverse modeling. In *Computational Methods in Water Resources: Volume 1*, pages 767–778. Elsevier.

[140] Salamon, P., Fernàndez-Garcia, D., and Gómez-Hernández, J. J. (2006). A review and numerical assessment of the random walk particle tracking method. *Journal of Contaminant Hydrology*, 87(3-4):277–305.

[141] Sampat, P. (2000). Deep trouble: The hidden threat of groundwater pollution. *Worldwatch Paper*, (154):55.

[142] Schaffer, J. (1985). Multiple objective optimization with vector evaluated genetic algorithms. In *The 1st international Conference on Genetic Algorithms*, pages 93–100. Lawrence Erlbaum Associates. Inc., Publishers.

[143] Scheidegger, A. E. (1961). General theory of dispersion in porous media. *Journal of Geophysical Research*, 66(10):3273–3278.

[144] Schraudolph, N. N. and Belew, R. K. (1992). Dynamic Parameter Encoding for genetic algorithms. *Machine Learning*, 9(1):9–21.

[145] Schwede, R. L. and Cirpka, O. A. (2010). Stochastic evaluation of mass discharge from pointlike concentration measurements. *Journal of Contaminant Hydrology*, 111(1-4):36–47.

[146] Shah, R. and Reed, P. (2011). Comparative analysis of multiobjective evolutionary algorithms for random and correlated instances of multiobjective d-dimensional knapsack problems. *European Journal of Operational Research*, 211(3):466–479.

[147] Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Monographs on Statistics and Applied Probability, London, 1 edition.

[148] Sinha, E. and Minsker, B. S. (2007). Multiscale island injection genetic algorithms for groundwater remediation. *Advances in Water Resources*, 30(9):1933–1942.

[149] Smith, G. D. (1985). *Numerical Solution Of Partial Differential Equations: Finite Difference Methods*. Oxford University Press, 3 edition.

[150] Snodgrass, M. F. and Kitanidis, P. K. (1997). A geostatistical approach to contaminant source identification. *Water Resources Research*, 33(4):537–546.

[151] Sober, E. (1981). The Principle of Parsimony. *The British Journal for the Philosophy of Science*, 32(2):145–156.

[152] Srinivas, N. and Deb, K. (1994). Muiltiobjective Optimization Using Nondominated Sorting in Genetic Algorithms. *Evolutionary Computation*, 2(3):221–248.

[153] Stauffer, F., Guadagnini, A., Butler, A., Franssen, H. J. H., van de Wiel, N., Bakr, M., Riva, M., and Guadagnini, L. (2005). Delineation of source protection zones using statistical methods. *Water Resources Management*, 19(2):163–185.

[154] Storck, P., Eheart, J. W., and Valocchi, A. J. (1997). A method for the optimal location of monitoring wells for detection of groundwater contamination in three-dimensional heterogenous aquifers. *Water Resources Research*, 33(9):2081–2088.

[155] Storn, R. and Price, K. (1997). Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4):341–359.

[156] Sun, A. Y., Painter, S. L., and Wittmeyer, G. W. (2006). A robust approach for iterative contaminant source location and release history recovery. *Journal of Contaminant Hydrology*, 88(3-4):181–196.

[157] Talbi, E.-G. (2009). *Metaheuristics: from design to implementation*. John Wiley & Sons, Inc.

[158] Taleb, N. N. (2007). *The Black Swan: The Impact of the Highly Improbable*. Random House.

[159] Tartakovsky, D. M. (2013). Assessment and management of risk in subsurface hydrology: A review and perspective. *Advances in Water Resources*, 51:247–260.

[160] Trefry, M. G. and Muffels, C. (2007). FEFLOW: A Finite-Element Ground Water Flow and Transport Modeling Tool. *Groundwater*, 45(5):525–528.

[161] Troldborg, M., Lemming, G., Binning, P. J., Tuxen, N., and Bjerg, P. L. (2008). Risk assessment and prioritisation of contaminated sites on the catchment scale. *Journal of Contaminant Hydrology*, 101(1-4):14–28.

[162] Tsutsui, S., Yamamura, M., and Higuchi, T. (1999). Multi-parent recombination with simplex crossover in real coded genetic algorithms. *Proceedings of the 1999 Genetic and Evolutionary Computation Conference (GECCO'99)*, pages 657–664.

[163] UN (2010). The Human Right to Water and Sanitation. *Sixty-fourth General Assembly, GA/10967*, 108thMeeti.

[164] US EPA (1993). Guidelines for delineation of wellhead protection areas. *EPA-440/5-93-001*.

[165] Van Veldhuizen, D. A. and Lamont, G. B. (1998a). Evolutionary Computation and Convergence to a Pareto Front. *Late Breaking Papers at the Genetic Programming 1998 Conference*, pages 221–228.

[166] Van Veldhuizen, D. A. and Lamont, G. B. (1998b). Multiobjective evolutionary algorithm research: A history and analysis. Technical report, Air Force Institute of Technology, Ohio.

[167] Varljen, M. D. and Shafer, J. M. (1991). Assessment of uncertainty in time-related capture zones using conditional simulation of hydraulic conductivity. *Ground Water*, 29(5):737–748.

[168] Vrugt, J. A. and Robinson, B. A. (2007). Improved evolutionary optimization from genetically adaptive multimethod search. *Proceedings of the National Academy of Sciences*, 104(3):708–711.

[169] Walker, W., Harremoës, P., Rotmans, J., van der Sluijs, J., van Asselt, M., Janssen, P., and Krayer von Krauss, M. (2003). Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated Assessment*, 4(1):5–17.

[170] Walters, G. A. and Lohbeck, T. (1993). Optimal Layout of Tree Networks Using Genetic Algorithms. *Engineering Optimization*, 22(1):27–48.

[171] Walters, G. A. and Smith, D. K. (1995). Evolutionary Design ALgorithm for Optimal Layout of Tree Networks. *Engineering Optimization*, 24(4):261–281.

[172] Wen, X.-H. and Gómez-Hernández, J. J. (1996). The Constant Displacement Scheme for Tracking Particles in Heterogeneous Aquifers. *Ground Water*, 34(1):135–142.

[173] WHG (2009). Öffentliche Wasserversorgung, Wasserschutzgebiete, Heilquellenschutz: Festsetzung von Wasserschutzgebieten.

[174] Whitley, D. (1994). A genetic algorithm tutorial. *Statistics and Computing*, 4(2).

[175] Whitley, D., Mathias, K., and Fitzhorn, P. (1991). Delta coding: An iterative search strategy for genetic algorithms. *Proceedings of the Fourth International Conference on Genetic Algorithms*, pages 77–84.

[176] WHO (2004). *Guidelines for Drinking-Water Quality*. WHO Publications, Geneva, 3 edition.

[177] Wolpert, D. (1995). No free lunch theorems for search. Technical report, The Santa Fe Institute, Santa Fe.

[178] Wolpert, D. and Macready, W. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(1):67–82.

[179] Yenigül, N. B., Elfeki, A. M. M., van den Akker, C., and Dekking, F. M. (2006). A decision analysis approach for optimal groundwater monitoring system design under uncertainty. *Hydrology and Earth System Sciences Discussions*, 3(1):27–68.

[180] Zienkiewicz, O., Taylor, R., and Zhu, J. (2013). *The Finite Element Method: its Basis and Fundamentals*. Elsevier.

[181] Zitzler, E., Laumanns, M., and Bleuler, S. (2004). A Tutorial on Evolutionary Multiobjective Optimization. In Gandibleux, X., Sevaux, M., Sörensen, K., and T'kindt, V., editors, *Metaheuristics for Multiobjective Optimisation*, pages 3–37. Springer Berlin Heidelberg.

[182] Zitzler, E., Laumanns, M., and Thiele, L. (2001). SPEA2: Improving the Strength Pareto Evolutionary Algorithm. *TIK-Report*, 103:1–21.

[183] Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C., and da Fonseca, V. (2003). Performance assessment of multiobjective optimizers: an analysis and review. *IEEE Transactions on Evolutionary Computation*, 7(2):117–132.

[184] Zwahlen, F. (2003). Vulnerability and Risk Mapping for the Protection of Carbonate (Karst) Aquifers. Scope-Goals-Results. *Research and Development, Brussels*, 297:42.

Lebenslauf

Persönliche Daten

Name:	Felix Bode
Geburtsdatum:	18.02.1987
Geburtsort:	Diepholz

Bildungsweg

07/2013 - 12/2017	Promotion im Rahmen des Gradiuiertenkollegs *Nonlinearities and Upscaling in Porous Media* (NUPUS), Universität Stuttgart
10/2007 - 01/2013	Dipl.-Ing. Umweltschutztechnik, Universität Stuttgart
05/2007	Abitur am Gymnasium am Wall, Verden (Aller)

Arbeitserfahrung

07/2013 - 12/2017	Wissenschaftlicher Mitarbeiter am Institut für Wasser- und Umweltsystemmodellierung der Universität Stuttgart, Lehrstuhl für Stochastische Simulation und Sicherheitsforschung für Hydrosysteme
seit 04/2018	Projektleiter im Bereich der Optimierung und Planung bei DUALIS GmbH IT Solution, Dresden

Institut für Wasser- und Umweltsystemmodellierung
Universität Stuttgart

Pfaffenwaldring 61
70569 Stuttgart (Vaihingen)
Telefon (0711) 685 - 64717/64749/64752/64679
Telefax (0711) 685 - 67020 o. 64746 o. 64681
E-Mail: iws@iws.uni-stuttgart.de
http://www.iws.uni-stuttgart.de

Verzeichnis der Mitteilungshefte

5 Plate, Erich: *Beitrag zur Bestimmung der Windgeschwindigkeitsverteilung in der durch eine Wand gestörten bodennahen Luftschicht*, und
Röhnisch, Arthur; Marotz, Günter: *Neue Baustoffe und Bauausführungen für den Schutz der Böschungen und der Sohle von Kanälen, Flüssen und Häfen; Gestehungskosten und jeweilige Vorteile*, sowie
Unny, T.E.: *Schwingungsuntersuchungen am Kegelstrahlschieber*, 1967

6 Seiler, Erich: *Die Ermittlung des Anlagenwertes der bundeseigenen Binnenschiffahrtsstraßen und Talsperren und des Anteils der Binnenschiffahrt an diesem Wert*, 1967

7 *Sonderheft anläßlich des 65. Geburtstages von Prof. Arthur Röhnisch mit Beiträgen von* Benk, Dieter; Breitling, J.; Gurr, Siegfried; Haberhauer, Robert; Honekamp, Hermann; Kuz, Klaus Dieter; Marotz, Günter; Mayer-Vorfelder, Hans-Jörg; Miller, Rudolf; Plate, Erich J.; Radomski, Helge; Schwarz, Helmut; Vollmer, Ernst; Wildenhahn, Eberhard; 1967

8 Jumikis, Alfred: *Beitrag zur experimentellen Untersuchung des Wassernachschubs in einem gefrierenden Boden und die Beurteilung der Ergebnisse*, 1968

9 Marotz, Günter: *Technische Grundlagen einer Wasserspeicherung im natürlichen Untergrund*, 1968

10 Radomski, Helge: *Untersuchungen über den Einfluß der Querschnittsform wellenförmiger Spundwände auf die statischen und rammtechnischen Eigenschaften*, 1968

11 Schwarz, Helmut: *Die Grenztragfähigkeit des Baugrundes bei Einwirkung vertikal gezogener Ankerplatten als zweidimensionales Bruchproblem*, 1969

12 Erbel, Klaus: *Ein Beitrag zur Untersuchung der Metamorphose von Mittelgebirgsschneedecken unter besonderer Berücksichtigung eines Verfahrens zur Bestimmung der thermischen Schneequalität*, 1969

13 Westhaus, Karl-Heinz: *Der Strukturwandel in der Binnenschiffahrt und sein Einfluß auf den Ausbau der Binnenschiffskanäle*, 1969

14 Mayer-Vorfelder, Hans-Jörg: *Ein Beitrag zur Berechnung des Erdwiderstandes unter Ansatz der logarithmischen Spirale als Gleitflächenfunktion*, 1970

15 Schulz, Manfred: *Berechnung des räumlichen Erddruckes auf die Wandung kreiszylindrischer Körper*, 1970

16 Mobasseri, Manoutschehr: *Die Rippenstützmauer. Konstruktion und Grenzen ihrer Standsicherheit*, 1970

17 Benk, Dieter: *Ein Beitrag zum Betrieb und zur Bemessung von Hochwasserrückhaltebecken*, 1970

18 Gàl, Attila: *Bestimmung der mitschwingenden Wassermasse bei überströmten Fischbauchklappen mit kreiszylindrischem Staublech*, 1971, vergriffen

19 Kuz, Klaus Dieter: *Ein Beitrag zur Frage des Einsetzens von Kavitationserscheinungen in einer Düsenströmung bei Berücksichtigung der im Wasser gelösten Gase*, 1971, vergriffen

20 Schaak, Hartmut: *Verteilleitungen von Wasserkraftanlagen*, 1971

21 *Sonderheft zur Eröffnung der neuen Versuchsanstalt des Instituts für Wasserbau der Universität Stuttgart mit Beiträgen von* Brombach, Hansjörg; Dirksen, Wolfram; Gàl, Attila; Gerlach, Reinhard; Giesecke, Jürgen; Holthoff, Franz-Josef; Kuz, Klaus Dieter; Marotz, Günter; Minor, Hans-Erwin; Petrikat, Kurt; Röhnisch, Arthur; Rueff, Helge; Schwarz, Helmut; Vollmer, Ernst; Wildenhahn, Eberhard; 1972

22 Wang, Chung-su: *Ein Beitrag zur Berechnung der Schwingungen an Kegelstrahlschiebern*, 1972

23 Mayer-Vorfelder, Hans-Jörg: *Erdwiderstandsbeiwerte nach dem Ohde-Variationsverfahren*, 1972

24 Minor, Hans-Erwin: *Beitrag zur Bestimmung der Schwingungsanfachungsfunktionen überströmter Stauklappen*, 1972, vergriffen

25 Brombach, Hansjörg: *Untersuchung strömungsmechanischer Elemente (Fluidik) und die Möglichkeit der Anwendung von Wirbelkammerelementen im Wasserbau*, 1972, vergriffen

26 Wildenhahn, Eberhard: *Beitrag zur Berechnung von Horizontalfilterbrunnen*, 1972

27 Steinlein, Helmut: *Die Eliminierung der Schwebstoffe aus Flußwasser zum Zweck der unterirdischen Wasserspeicherung, gezeigt am Beispiel der Iller*, 1972

28 Holthoff, Franz Josef: *Die Überwindung großer Hubhöhen in der Binnenschiffahrt durch Schwimmerhebewerke*, 1973

29 Röder, Karl: *Einwirkungen aus Baugrundbewegungen auf trog- und kastenförmige Konstruktionen des Wasser- und Tunnelbaues*, 1973

30 Kretschmer, Heinz: *Die Bemessung von Bogenstaumauern in Abhängigkeit von der Talform*, 1973

31 Honekamp, Hermann: *Beitrag zur Berechnung der Montage von Unterwasserpipelines*, 1973

32 Giesecke, Jürgen: *Die Wirbelkammertriode als neuartiges Steuerorgan im Wasserbau*, und Brombach, Hansjörg: *Entwicklung, Bauformen, Wirkungsweise und Steuereigenschaften von Wirbelkammerverstärkern*, 1974

33 Rueff, Helge: *Untersuchung der schwingungserregenden Kräfte an zwei hintereinander angeordneten Tiefschützen unter besonderer Berücksichtigung von Kavitation*, 1974

34 Röhnisch, Arthur: *Einpreßversuche mit Zementmörtel für Spannbeton - Vergleich der Ergebnisse von Modellversuchen mit Ausführungen in Hüllwellrohren*, 1975

35 *Sonderheft anläßlich des 65. Geburtstages von Prof. Dr.-Ing. Kurt Petrikat mit Beiträgen von*: Brombach, Hansjörg; Erbel, Klaus; Flinspach, Dieter; Fischer jr., Richard; Gàl, Attila; Gerlach, Reinhard; Giesecke, Jürgen; Haberhauer, Robert; Hafner Edzard; Hausenblas, Bernhard; Horlacher, Hans-Burkhard; Hutarew, Andreas; Knoll, Manfred; Krummet, Ralph; Marotz, Günter; Merkle, Theodor; Miller, Christoph; Minor, Hans-Erwin; Neumayer, Hans; Rao, Syamala; Rath, Paul; Rueff, Helge; Ruppert, Jürgen; Schwarz, Wolfgang; Topal-Gökceli, Mehmet; Vollmer, Ernst; Wang, Chung-su; Weber, Hans-Georg; 1975

36 Berger, Jochum: *Beitrag zur Berechnung des Spannungszustandes in rotationssymmetrisch belasteten Kugelschalen veränderlicher Wandstärke unter Gas- und Flüssigkeitsdruck durch Integration schwach singulärer Differentialgleichungen*, 1975

37 Dirksen, Wolfram: *Berechnung instationärer Abflußvorgänge in gestauten Gerinnen mittels Differenzenverfahren und die Anwendung auf Hochwasserrückhaltebecken*, 1976

38 Horlacher, Hans-Burkhard: *Berechnung instationärer Temperatur- und Wärmespannungsfelder in langen mehrschichtigen Hohlzylindern*, 1976

39 Hafner, Edzard: *Untersuchung der hydrodynamischen Kräfte auf Baukörper im Tiefwasserbereich des Meeres*, 1977, ISBN 3-921694-39-6

40 Ruppert, Jürgen: *Über den Axialwirbelkammerverstärker für den Einsatz im Wasserbau*, 1977, ISBN 3-921694-40-X

41 Hutarew, Andreas: *Beitrag zur Beeinflußbarkeit des Sauerstoffgehalts in Fließgewässern an Abstürzen und Wehren*, 1977, ISBN 3-921694-41-8, *vergriffen*

42 Miller, Christoph: *Ein Beitrag zur Bestimmung der schwingungserregenden Kräfte an unterströmten Wehren*, 1977, ISBN 3-921694-42-6

43 Schwarz, Wolfgang: *Druckstoßberechnung unter Berücksichtigung der Radial- und Längsverschiebungen der Rohrwandung*, 1978, ISBN 3-921694-43-4

44 Kinzelbach, Wolfgang: *Numerische Untersuchungen über den optimalen Einsatz variabler Kühlsysteme einer Kraftwerkskette am Beispiel Oberrhein*, 1978, ISBN 3-921694-44-2

45 Barczewski, Baldur: *Neue Meßmethoden für Wasser-Luftgemische und deren Anwendung auf zweiphasige Auftriebsstrahlen*, 1979, ISBN 3-921694-45-0

46 Neumayer, Hans: *Untersuchung der Strömungsvorgänge in radialen Wirbelkammerverstärkern*, 1979, ISBN 3-921694-46-9

47 Elalfy, Youssef-Elhassan: *Untersuchung der Strömungsvorgänge in Wirbelkammerdioden und -drosseln*, 1979, ISBN 3-921694-47-7

48 Brombach, Hansjörg: *Automatisierung der Bewirtschaftung von Wasserspeichern*, 1981, ISBN 3-921694-48-5

49 Geldner, Peter: *Deterministische und stochastische Methoden zur Bestimmung der Selbstdichtung von Gewässern*, 1981, ISBN 3-921694-49-3, *vergriffen*

50 Mehlhorn, Hans: *Temperaturveränderungen im Grundwasser durch Brauchwassereinleitungen*, 1982, ISBN 3-921694-50-7, *vergriffen*

51 Hafner, Edzard: *Rohrleitungen und Behälter im Meer*, 1983, ISBN 3-921694-51-5

52 Rinnert, Bernd: *Hydrodynamische Dispersion in porösen Medien: Einfluß von Dichteunterschieden auf die Vertikalvermischung in horizontaler Strömung*, 1983, ISBN 3-921694-52-3, *vergriffen*

53 Lindner, Wulf: *Steuerung von Grundwasserentnahmen unter Einhaltung ökologischer Kriterien*, 1983, ISBN 3-921694-53-1, *vergriffen*

54 Herr, Michael; Herzer, Jörg; Kinzelbach, Wolfgang; Kobus, Helmut; Rinnert, Bernd: *Methoden zur rechnerischen Erfassung und hydraulischen Sanierung von Grundwasserkontaminationen*, 1983, ISBN 3-921694-54-X

55 Schmitt, Paul: *Wege zur Automatisierung der Niederschlagsermittlung*, 1984, ISBN 3-921694-55-8, *vergriffen*

56 Müller, Peter: *Transport und selektive Sedimentation von Schwebstoffen bei gestautem Abfluß*, 1985, ISBN 3-921694-56-6

57 El-Qawasmeh, Fuad: *Möglichkeiten und Grenzen der Tropfbewässerung unter besonderer Berücksichtigung der Verstopfungsanfälligkeit der Tropfelemente*, 1985, ISBN 3-921694-57-4, *vergriffen*

58 Kirchenbaur, Klaus: *Mikroprozessorgesteuerte Erfassung instationärer Druckfelder am Beispiel seegangsbelasteter Baukörper*, 1985, ISBN 3-921694-58-2

59 Kobus, Helmut (Hrsg.): *Modellierung des großräumigen Wärme- und Schadstofftransports im Grundwasser*, Tätigkeitsbericht 1984/85 (DFG-Forschergruppe an den Universitäten Hohenheim, Karlsruhe und Stuttgart), 1985, ISBN 3-921694-59-0, *vergriffen*

60 Spitz, Karlheinz: *Dispersion in porösen Medien: Einfluß von Inhomogenitäten und Dichteunterschieden*, 1985, ISBN 3-921694-60-4, *vergriffen*

61 Kobus, Helmut: *An Introduction to Air-Water Flows in Hydraulics*, 1985, ISBN 3-921694-61-2

62 Kaleris, Vassilios: *Erfassung des Austausches von Oberflächen- und Grundwasser in horizontalebenen Grundwassermodellen*, 1986, ISBN 3-921694-62-0

63 Herr, Michael: *Grundlagen der hydraulischen Sanierung verunreinigter Porengrundwasserleiter*, 1987, ISBN 3-921694-63-9

64 Marx, Walter: *Berechnung von Temperatur und Spannung in Massenbeton infolge Hydratation*, 1987, ISBN 3-921694-64-7

65 Koschitzky, Hans-Peter: *Dimensionierungskonzept für Sohlbelüfter in Schußrinnen zur Vermeidung von Kavitationsschäden*, 1987, ISBN 3-921694-65-5

66 Kobus, Helmut (Hrsg.): *Modellierung des großräumigen Wärme- und Schadstofftransports im Grundwasser*, Tätigkeitsbericht 1986/87 (DFG-Forschergruppe an den Universitäten Hohenheim, Karlsruhe und Stuttgart) 1987, ISBN 3-921694-66-3

67 Söll, Thomas: *Berechnungsverfahren zur Abschätzung anthropogener Temperaturanomalien im Grundwasser*, 1988, ISBN 3-921694-67-1

68 Dittrich, Andreas; Westrich, Bernd: *Bodenseeufererosion, Bestandsaufnahme und Bewertung*, 1988, ISBN 3-921694-68-X, *vergriffen*

69 Huwe, Bernd; van der Ploeg, Rienk R.: *Modelle zur Simulation des Stickstoffhaushaltes von Standorten mit unterschiedlicher landwirtschaftlicher Nutzung*, 1988, ISBN 3-921694-69-8, *vergriffen*

70 Stephan, Karl: *Integration elliptischer Funktionen*, 1988, ISBN 3-921694-70-1

71 Kobus, Helmut; Zilliox, Lothaire (Hrsg.): *Nitratbelastung des Grundwassers, Auswirkungen der Landwirtschaft auf die Grundwasser- und Rohwasserbeschaffenheit und Maßnahmen zum Schutz des Grundwassers*. Vorträge des deutsch-französischen Kolloquiums am 6. Oktober 1988, Universitäten Stuttgart und Louis Pasteur Strasbourg (Vorträge in deutsch oder französisch, Kurzfassungen zweisprachig), 1988, ISBN 3-921694-71-X

72 Soyeaux, Renald: *Unterströmung von Stauanlagen auf klüftigem Untergrund unter Berücksichtigung laminarer und turbulenter Fließzustände*,1991, ISBN 3-921694-72-8

73 Kohane, Roberto: *Berechnungsmethoden für Hochwasserabfluß in Fließgewässern mit überströmten Vorländern*, 1991, ISBN 3-921694-73-6

74 Hassinger, Reinhard: *Beitrag zur Hydraulik und Bemessung von Blocksteinrampen in flexibler Bauweise*, 1991, ISBN 3-921694-74-4, vergriffen

75 Schäfer, Gerhard: *Einfluß von Schichtenstrukturen und lokalen Einlagerungen auf die Längsdispersion in Porengrundwasserleitern*, 1991, ISBN 3-921694-75-2

76 Giesecke, Jürgen: *Vorträge, Wasserwirtschaft in stark besiedelten Regionen; Umweltforschung mit Schwerpunkt Wasserwirtschaft*, 1991, ISBN 3-921694-76-0

77 Huwe, Bernd: *Deterministische und stochastische Ansätze zur Modellierung des Stickstoffhaushalts landwirtschaftlich genutzter Flächen auf unterschiedlichem Skalenniveau*, 1992, ISBN 3-921694-77-9, vergriffen

78 Rommel, Michael: *Verwendung von Kluftdaten zur realitätsnahen Generierung von Kluftnetzen mit anschließender laminar-turbulenter Strömungsberechnung*, 1993, ISBN 3-92 1694-78-7

79 Marschall, Paul: *Die Ermittlung lokaler Stofffrachten im Grundwasser mit Hilfe von Einbohrloch-Meßverfahren*, 1993, ISBN 3-921694-79-5, vergriffen

80 Ptak, Thomas: *Stofftransport in heterogenen Porenaquiferen: Felduntersuchungen und stochastische Modellierung*, 1993, ISBN 3-921694-80-9, vergriffen

81 Haakh, Frieder: *Transientes Strömungsverhalten in Wirbelkammern,* 1993, ISBN 3-921694-81-7

82 Kobus, Helmut; Cirpka, Olaf; Barczewski, Baldur; Koschitzky, Hans-Peter: *Versuchseinrichtung zur Grundwasser und Altlastensanierung VEGAS, Konzeption und Programmrahmen*, 1993, ISBN 3-921694-82-5

83 Zang, Weidong: *Optimaler Echtzeit-Betrieb eines Speichers mit aktueller Abflußregenerierung*, 1994, ISBN 3-921694-83-3, vergriffen

84 Franke, Hans-Jörg: *Stochastische Modellierung eines flächenhaften Stoffeintrages und Transports in Grundwasser am Beispiel der Pflanzenschutzmittelproblematik*, 1995, ISBN 3-921694-84-1

85 Lang, Ulrich: *Simulation regionaler Strömungs- und Transportvorgänge in Karstaquiferen mit Hilfe des Doppelkontinuum-Ansatzes: Methodenentwicklung und Parameteridentifikation*, 1995, ISBN 3-921694-85-X, vergriffen

86 Helmig, Rainer: *Einführung in die Numerischen Methoden der Hydromechanik*, 1996, ISBN 3-921694-86-8, vergriffen

87 Cirpka, Olaf: *CONTRACT: A Numerical Tool for Contaminant Transport and Chemical Transformations - Theory and Program Documentation -*, 1996, ISBN 3-921694-87-6

88 Haberlandt, Uwe: *Stochastische Synthese und Regionalisierung des Niederschlages für Schmutzfrachtberechnungen*, 1996, ISBN 3-921694-88-4

89 Croisé, Jean: *Extraktion von flüchtigen Chemikalien aus natürlichen Lockergesteinen mittels erzwungener Luftströmung*, 1996, ISBN 3-921694-89-2, vergriffen

90 Jorde, Klaus: *Ökologisch begründete, dynamische Mindestwasserregelungen bei Ausleitungskraftwerken*, 1997, ISBN 3-921694-90-6, vergriffen

91 Helmig, Rainer: *Gekoppelte Strömungs- und Transportprozesse im Untergrund - Ein Beitrag zur Hydrosystemmodellierung-*, 1998, ISBN 3-921694-91-4, vergriffen

92 Emmert, Martin: *Numerische Modellierung nichtisothermer Gas-Wasser Systeme in porösen Medien,* 1997, ISBN 3-921694-92-2

93 Kern, Ulrich: *Transport von Schweb- und Schadstoffen in staugeregelten Fließgewässern am Beispiel des Neckars*, 1997, ISBN 3-921694-93-0, vergriffen

94 Förster, Georg: *Druckstoßdämpfung durch große Luftblasen in Hochpunkten von Rohrleitungen 1997*, ISBN 3-921694-94-9

95 Cirpka, Olaf: *Numerische Methoden zur Simulation des reaktiven Mehrkomponententransports im Grundwasser*, 1997, ISBN 3-921694-95-7, *vergriffen*

96 Färber, Arne: *Wärmetransport in der ungesättigten Bodenzone: Entwicklung einer thermischen In-situ-Sanierungstechnologie*, 1997, ISBN 3-921694-96-5

97 Betz, Christoph: *Wasserdampfdestillation von Schadstoffen im porösen Medium: Entwicklung einer thermischen In-situ-Sanierungstechnologie*, 1998, SBN 3-921694-97-3

98 Xu, Yichun: *Numerical Modeling of Suspended Sediment Transport in Rivers*, 1998, ISBN 3-921694-98-1, *vergriffen*

99 Wüst, Wolfgang: *Geochemische Untersuchungen zur Sanierung CKW-kontaminierter Aquifere mit Fe(0)-Reaktionswänden,* 2000, ISBN 3-933761-02-2

100 Sheta, Hussam: *Simulation von Mehrphasenvorgängen in porösen Medien unter Einbeziehung von Hysterese-Effekten*, 2000, ISBN 3-933761-03-4

101 Ayros, Edwin: *Regionalisierung extremer Abflüsse auf der Grundlage statistischer Verfahren*, 2000, ISBN 3-933761-04-2, *vergriffen*

102 Huber, Ralf: *Compositional Multiphase Flow and Transport in Heterogeneous Porous Media*, 2000, ISBN 3-933761-05-0

103 Braun, Christopherus: *Ein Upscaling-Verfahren für Mehrphasenströmungen in porösen Medien*, 2000, ISBN 3-933761-06-9

104 Hofmann, Bernd: *Entwicklung eines rechnergestützten Managementsystems zur Beurteilung von Grundwasserschadensfällen*, 2000, ISBN 3-933761-07-7

105 Class, Holger: *Theorie und numerische Modellierung nichtisothermer Mehrphasenprozesse in NAPL-kontaminierten porösen Medien,* 2001, ISBN 3-933761-08-5

106 Schmidt, Reinhard: *Wasserdampf- und Heißluftinjektion zur thermischen Sanierung kontaminierter Standorte*, 2001, ISBN 3-933761-09-3

107 Josef, Reinhold:, *Schadstoffextraktion mit hydraulischen Sanierungsverfahren unter Anwendung von grenzflächenaktiven Stoffen*, 2001, ISBN 3-933761-10-7

108 Schneider, Matthias: *Habitat- und Abflussmodellierung für Fließgewässer mit unscharfen Berechnungsansätzen*, 2001, ISBN 3-933761-11-5

109 Rathgeb, Andreas: *Hydrodynamische Bemessungsgrundlagen für Lockerdeckwerke an überströmbaren Erddämmen*, 2001, ISBN 3-933761-12-3

110 Lang, Stefan: *Parallele numerische Simulation instätionärer Probleme mit adaptiven Methoden auf unstrukturierten Gittern*, 2001, ISBN 3-933761-13-1

111 Appt, Jochen; Stumpp Simone: *Die Bodensee-Messkampagne 2001, IWS/CWR Lake Constance Measurement Program 2001*, 2002, ISBN 3-933761-14-X

112 Heimerl, Stephan: *Systematische Beurteilung von Wasserkraftprojekten,* 2002, ISBN 3-933761-15-8, *vergriffen*

113 Iqbal, Amin: *On the Management and Salinity Control of Drip Irrigation,* 2002, ISBN 3-933761-16-6

114 Silberhorn-Hemminger, Annette: *Modellierung von Kluftaquifersystemen: Geostatistische Analyse und deterministisch-stochastische Kluftgenerierung,* 2002, ISBN 3-933761-17-4

115 Winkler, Angela: *Prozesse des Wärme- und Stofftransports bei der In-situ-Sanierung mit festen Wärmequellen,* 2003, ISBN 3-933761-18-2

116 Marx, Walter: *Wasserkraft, Bewässerung, Umwelt - Planungs- und Bewertungsschwerpunkte der Wasserbewirtschaftung,* 2003, ISBN 3-933761-19-0

117 Hinkelmann, Reinhard: *Efficient Numerical Methods and Information-Processing Techniques in Environment Water,* 2003, ISBN 3-933761-20-4

118 Samaniego-Eguiguren, Luis Eduardo: *Hydrological Consequences of Land Use / Land Cover and Climatic Changes in Mesoscale Catchments*, 2003, ISBN 3-933761-21-2

119 Neunhäuserer, Lina: *Diskretisierungsansätze zur Modellierung von Strömungs- und Transportprozessen in geklüftet-porösen Medien*, 2003, ISBN 3-933761-22-0

120 Paul, Maren: *Simulation of Two-Phase Flow in Heterogeneous Poros Media with Adaptive Methods*, 2003, ISBN 3-933761-23-9

121 Ehret, Uwe: *Rainfall and Flood Nowcasting in Small Catchments using Weather Radar*, 2003, ISBN 3-933761-24-7

122 Haag, Ingo: *Der Sauerstoffhaushalt staugeregelter Flüsse am Beispiel des Neckars - Analysen, Experimente, Simulationen -,* 2003, ISBN 3-933761-25-5

123 Appt, Jochen: *Analysis of Basin-Scale Internal Waves in Upper Lake Constance,* 2003, ISBN 3-933761-26-3

124 Hrsg.: Schrenk, Volker; Batereau, Katrin; Barczewski, Baldur; Weber, Karolin und Koschitzky, Hans-Peter: *Symposium Ressource Fläche und VEGAS - Statuskolloquium 2003, 30. September und 1. Oktober 2003,* 2003, ISBN 3-933761-27-1

125 Omar Khalil Ouda: *Optimisation of Agricultural Water Use: A Decision Support System for the Gaza Strip,* 2003, ISBN 3-933761-28-0

126 Batereau, Katrin: *Sensorbasierte Bodenluftmessung zur Vor-Ort-Erkundung von Schadensherden im Untergrund*, 2004, ISBN 3-933761-29-8

127 Witt, Oliver: *Erosionsstabilität von Gewässersedimenten mit Auswirkung auf den Stofftransport bei Hochwasser am Beispiel ausgewählter Stauhaltungen des Oberrheins*, 2004, ISBN 3-933761-30-1

128 Jakobs, Hartmut: *Simulation nicht-isothermer Gas-Wasser-Prozesse in komplexen Kluft-Matrix-Systemen*, 2004, ISBN 3-933761-31-X

129 Li, Chen-Chien: *Deterministisch-stochastisches Berechnungskonzept zur Beurteilung der Auswirkungen erosiver Hochwasserereignisse in Flussstauhaltungen,* 2004, ISBN 3-933761-32-8

130 Reichenberger, Volker; Helmig, Rainer; Jakobs, Hartmut; Bastian, Peter; Niessner, Jennifer: *Complex Gas-Water Processes in Discrete Fracture-Matrix Systems: Up-scaling, Mass-Conservative Discretization and Efficient Multilevel Solution,* 2004, ISBN 3-933761-33-6

131 Hrsg.: Barczewski, Baldur; Koschitzky, Hans-Peter; Weber, Karolin; Wege, Ralf: *VEGAS - Statuskolloquium 2004,* Tagungsband zur Veranstaltung am 05. Oktober 2004 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2004, ISBN 3-933761-34-4

132 Asie, Kemal Jabir: *Finite Volume Models for Multiphase Multicomponent Flow through Porous Media*. 2005, ISBN 3-933761-35-2

133 Jacoub, George: *Development of a 2-D Numerical Module for Particulate Contaminant Transport in Flood Retention Reservoirs and Impounded Rivers,* 2004, ISBN 3-933761-36-0

134 Nowak, Wolfgang: *Geostatistical Methods for the Identification of Flow and Transport Parameters in the Subsurface,* 2005, ISBN 3-933761-37-9

135 Süß, Mia: *Analysis of the influence of structures and boundaries on flow and transport processes in fractured porous media*, 2005, ISBN 3-933761-38-7

136 Jose, Surabhin Chackiath: *Experimental Investigations on Longitudinal Dispersive Mixing in Heterogeneous Aquifers,* 2005, ISBN: 3-933761-39-5

137 Filiz, Fulya: *Linking Large-Scale Meteorological Conditions to Floods in Mesoscale Catchments,* 2005, ISBN 3-933761-40-9

138 Qin, Minghao: *Wirklichkeitsnahe und recheneffiziente Ermittlung von Temperatur und Spannungen bei großen RCC-Staumauern*, 2005, ISBN 3-933761-41-7

139 Kobayashi, Kenichiro: *Optimization Methods for Multiphase Systems in the Subsurface - Application to Methane Migration in Coal Mining Areas*, 2005, ISBN 3-933761-42-5

140 Rahman, Md. Arifur: *Experimental Investigations on Transverse Dispersive Mixing in Heterogeneous Porous Media,* 2005, ISBN 3-933761-43-3

141 Schrenk, Volker: *Ökobilanzen zur Bewertung von Altlastensanierungsmaßnahmen,* 2005, ISBN 3-933761-44-1

142 Hundecha, Hirpa Yeshewatesfa: *Regionalization of Parameters of a Conceptual Rainfall-Runoff Model*, 2005, ISBN: 3-933761-45-X

143 Wege, Ralf: *Untersuchungs- und Überwachungsmethoden für die Beurteilung natürlicher Selbstreinigungsprozesse im Grundwasser,* 2005, ISBN 3-933761-46-8

144 Breiting, Thomas: *Techniken und Methoden der Hydroinformatik - Modellierung von komplexen Hydrosystemen im Untergrund*, 2006, ISBN 3-933761-47-6

145 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Müller, Martin: *Ressource Untergrund: 10 Jahre VEGAS: Forschung und Technologieentwicklung zum Schutz von Grundwasser und Boden,* Tagungsband zur Veranstaltung am 28. und 29. September 2005 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2005, ISBN 3-933761-48-4

146 Rojanschi, Vlad: *Abflusskonzentration in mesoskaligen Einzugsgebieten unter Berücksichtigung des Sickerraumes,* 2006, ISBN 3-933761-49-2

147 Winkler, Nina Simone: *Optimierung der Steuerung von Hochwasserrückhaltebeckensystemen,* 2006, ISBN 3-933761-50-6

148 Wolf, Jens: *Räumlich differenzierte Modellierung der Grundwasserströmung alluvialer Aquifere für mesoskalige Einzugsgebiete*, 2006, ISBN: 3-933761-51-4

149 Kohler, Beate: *Externe Effekte der Laufwasserkraftnutzung*, 2006, ISBN 3-933761-52-2

150 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhrmann, Matthias: *VEGAS-Statuskolloquium 2006,* Tagungsband zur Veranstaltung am 28. September 2006 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2006, ISBN 3-933761-53-0

151 Niessner, Jennifer: *Multi-Scale Modeling of Multi-Phase - Multi-Component Processes in Heterogeneous Porous Media*, 2006, ISBN 3-933761-54-9

152 Fischer, Markus: *Beanspruchung eingeerdeter Rohrleitungen infolge Austrocknung bindiger Böden,* 2006, ISBN 3-933761-55-7

153 Schneck, Alexander: *Optimierung der Grundwasserbewirtschaftung unter Berücksichtigung der Belange der Wasserversorgung, der Landwirtschaft und des Naturschutzes*, 2006, ISBN 3-933761-56-5

154 Das, Tapash: *The Impact of Spatial Variability of Precipitation on the Predictive Uncertainty of Hydrological Models,* 2006, ISBN 3-33761-57-3

155 Bielinski, Andreas: *Numerical Simulation of CO_2 sequestration in geological formations*, 2007, ISBN 3-933761-58-1

156 Mödinger, Jens: *Entwicklung eines Bewertungs- und Entscheidungsunterstützungssystems für eine nachhaltige regionale Grundwasserbewirtschaftung*, 2006, ISBN 3-933761-60-3

157 Manthey, Sabine: *Two-phase flow processes with dynamic effects in porous media - parameter estimation and simulation,* 2007, ISBN 3-933761-61-1

158 Pozos Estrada, Oscar: *Investigation on the Effects of Entrained Air in Pipelines,* 2007, ISBN 3-933761-62-X

159 Ochs, Steffen Oliver: *Steam injection into saturated porous media – process analysis including experimental and numerical investigations*, 2007, ISBN 3-933761-63-8

160 Marx, Andreas: *Einsatz gekoppelter Modelle und Wetterradar zur Abschätzung von Niederschlagsintensitäten und zur Abflussvorhersage, 2007,* ISBN 3-933761-64-6

161 Hartmann, Gabriele Maria: *Investigation of Evapotranspiration Concepts in Hydrological Modelling for Climate Change Impact Assessment*, 2007, ISBN 3-933761-65-4

162 Kebede Gurmessa, Tesfaye: *Numerical Investigation on Flow and Transport Characteristics to Improve Long-Term Simulation of Reservoir Sedimentation*, 2007, ISBN 3-933761-66-2

163 Trifković, Aleksandar: *Multi-objective and Risk-based Modelling Methodology for Planning, Design and Operation of Water Supply Systems*, 2007, ISBN 3-933761-67-0

164 Götzinger, Jens: *Distributed Conceptual Hydrological Modelling - Simulation of Climate, Land Use Change Impact and Uncertainty Analysis*, 2007, ISBN 3-933761-68-9

165 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhrmann, Matthias: *VEGAS – Kolloquium 2007,* Tagungsband zur Veranstaltung am 26. September 2007 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2007, ISBN 3-933761-69-7

166 Freeman, Beau: *Modernization Criteria Assessment for Water Resources Planning; Klamath Irrigation Project, U.S.*, 2008, ISBN 3-933761-70-0

167 Dreher, Thomas: *Selektive Sedimentation von Feinstschwebstoffen in Wechselwirkung mit wandnahen turbulenten Strömungsbedingungen*, 2008, ISBN 3-933761-71-9

168 Yang, Wei: *Discrete-Continuous Downscaling Model for Generating Daily Precipitation Time Series*, 2008, ISBN 3-933761-72-7

169 Kopecki, Ianina: *Calculational Approach to FST-Hemispheres for Multiparametrical Benthos Habitat Modelling,* 2008*,* ISBN 3-933761-73-5

170 Brommundt, Jürgen: *Stochastische Generierung räumlich zusammenhängender Niederschlagszeitreihen*, 2008, ISBN 3-933761-74-3

171 Papafotiou, *Alexandros: Numerical Investigations of the Role of Hysteresis in Heterogeneous Two-Phase Flow Systems*, 2008, ISBN 3-933761-75-1

172 He, Yi: *Application of a Non-Parametric Classification Scheme to Catchment Hydrology*, 2008, ISBN 978-3-933761-76-7

173 Wagner, Sven: *Water Balance in a Poorly Gauged Basin in West Africa Using Atmospheric Modelling and Remote Sensing Information*, 2008, ISBN 978-3-933761-77-4

174 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhrmann, Matthias; Schrenk, Volker: *VEGAS-Kolloquium 2008 Ressource Fläche III*, Tagungsband zur Veranstaltung am 01. Oktober 2008 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2008, ISBN 978-3-933761-78-1

175 Patil, Sachin: *Regionalization of an Event Based Nash Cascade Model for Flood Predictions in Ungauged Basins*, 2008, ISBN 978-3-933761-79-8

176 Assteerawatt, Anongnart: *Flow and Transport Modelling of Fractured Aquifers based on a Geostatistical Approach,* 2008, ISBN 978-3-933761-80-4

177 Karnahl, Joachim Alexander: *2D numerische Modellierung von multifraktionalem Schwebstoff- und Schadstofftransport in Flüssen*, 2008, ISBN 978-3-933761-81-1

178 Hiester, Uwe: *Technologieentwicklung zur In-situ-Sanierung der ungesättigten Bodenzone mit festen Wärmequellen,* 2009, ISBN 978-3-933761-82-8

179 Laux, Patrick: *Statistical Modeling of Precipitation for Agricultural Planning in the Volta Basin of West Africa,* 2009, ISBN 978-3-933761-83-5

180 Ehsan, Saqib: *Evaluation of Life Safety Risks Related to Severe Flooding,* 2009, ISBN 978-3-933761-84-2

181 Prohaska, Sandra: *Development and Application of a 1D Multi-Strip Fine Sediment Transport Model for Regulated Rivers,* 2009, ISBN 978-3-933761-85-9

182 Kopp, Andreas: *Evaluation of CO_2 Injection Processes in Geological Formations for Site Screening,* 2009, ISBN 978-3-933761-86-6

183 Ebigbo, Anozie: *Modelling of biofilm growth and its influence on CO_2 and water (two-phase) flow in porous media,* 2009, ISBN 978-3-933761-87-3

184 Freiboth, Sandra: *A phenomenological model for the numerical simulation of multiphase multicomponent processes considering structural alterations of porous media*, 2009, ISBN 978-3-933761-88-0

185 Zöllner, Frank: *Implementierung und Anwendung netzfreier Methoden im Konstruktiven Wasserbau und in der Hydromechanik*, 2009, ISBN 978-3-933761-89-7

186 Vasin, Milos: *Influence of the soil structure and property contrast on flow and transport in the unsaturated zone*, 2010, ISBN 978-3-933761-90-3

187 Li, Jing: *Application of Copulas as a New Geostatistical Tool*, 2010, ISBN 978-3-933761-91-0

188 AghaKouchak, Amir: *Simulation of Remotely Sensed Rainfall Fields Using Copulas*, 2010, ISBN 978-3-933761-92-7

189 Thapa, Pawan Kumar: *Physically-based spatially distributed rainfall runoff modelling for soil erosion estimation*, 2010, ISBN 978-3-933761-93-4

190 Wurms, Sven: *Numerische Modellierung der Sedimentationsprozesse in Retentionsanlagen zur Steuerung von Stoffströmen bei extremen Hochwasserabflussereignissen*, 2011, ISBN 978-3-933761-94-1

191 Merkel, Uwe: *Unsicherheitsanalyse hydraulischer Einwirkungen auf Hochwasserschutzdeiche und Steigerung der Leistungsfähigkeit durch adaptive Strömungsmodellierung*, 2011, ISBN 978-3-933761-95-8

192 Fritz, Jochen: *A Decoupled Model for Compositional Non-Isothermal Multiphase Flow in Porous Media and Multiphysics Approaches for Two-Phase Flow*, 2010, ISBN 978-3-933761-96-5

193 Weber, Karolin (Hrsg.): *12. Treffen junger WissenschaftlerInnen an Wasserbauinstituten*, 2010, ISBN 978-3-933761-97-2

194 Bliefernicht, Jan-Geert: *Probability Forecasts of Daily Areal Precipitation for Small River Basins,* 2011, ISBN 978-3-933761-98-9

195 Hrsg.: Koschitzky, Hans-Peter; Braun, Jürgen: *VEGAS-Kolloquium 2010 In-situ-Sanierung - Stand und Entwicklung Nano und ISCO -,* Tagungsband zur Veranstaltung am 07. Oktober 2010 an der Universität Stuttgart, Campus Stuttgart-Vaihingen*,* 2010*,* ISBN 978-3-933761-99-6

196 Gafurov, Abror: *Water Balance Modeling Using Remote Sensing Information - Focus on Central Asia,* 2010, ISBN 978-3-942036-00-9

197 Mackenberg, Sylvia: *Die Quellstärke in der Sickerwasserprognose: Möglichkeiten und Grenzen von Labor- und Freilanduntersuchungen,* 2010, ISBN 978-3-942036-01-6

198 Singh, Shailesh Kumar: *Robust Parameter Estimation in Gauged and Ungauged Basins,* 2010, ISBN 978-3-942036-02-3

199 Doğan, Mehmet Onur: *Coupling of porous media flow with pipe flow,* 2011, ISBN 978-3-942036-03-0

200 Liu, Min: *Study of Topographic Effects on Hydrological Patterns and the Implication on Hydrological Modeling and Data Interpolation,* 2011, ISBN 978-3-942036-04-7

201 Geleta, Habtamu Itefa: *Watershed Sediment Yield Modeling for Data Scarce Areas,* 2011, ISBN 978-3-942036-05-4

202 Franke, Jörg: *Einfluss der Überwachung auf die Versagenswahrscheinlichkeit von Staustufen,* 2011, ISBN 978-3-942036-06-1

203 Bakimchandra, Oinam: *Integrated Fuzzy-GIS approach for assessing regional soil erosion risks,* 2011, ISBN 978-3-942036-07-8

204 Alam, Muhammad Mahboob: *Statistical Downscaling of Extremes of Precipitation in Mesoscale Catchments from Different RCMs and Their Effects on Local Hydrology,* 2011, ISBN 978-3-942036-08-5

205 Hrsg.: Koschitzky, Hans-Peter; Braun, Jürgen: *VEGAS-Kolloquium 2011 Flache Geothermie - Perspektiven und Risiken,* Tagungsband zur Veranstaltung am 06. Oktober 2011 an der Universität Stuttgart, Campus Stuttgart-Vaihingen*,* 2011*,* ISBN 978-3-933761-09-2

206 Haslauer, Claus: *Analysis of Real-World Spatial Dependence of Subsurface Hydraulic Properties Using Copulas with a Focus on Solute Transport Behaviour,* 2011, ISBN 978-3-942036-10-8

207 Dung, Nguyen Viet: *Multi-objective automatic calibration of hydrodynamic models – development of the concept and an application in the Mekong Delta,* 2011, ISBN 978-3-942036-11-5

208 Hung, Nguyen Nghia: *Sediment dynamics in the floodplain of the Mekong Delta, Vietnam,* 2011, ISBN 978-3-942036-12-2

209 Kuhlmann, Anna: *Influence of soil structure and root water uptake on flow in the unsaturated zone,* 2012, ISBN 978-3-942036-13-9

210 Tuhtan, Jeffrey Andrew: *Including the Second Law Inequality in Aquatic Ecodynamics: A Modeling Approach for Alpine Rivers Impacted by Hydropeaking,* 2012, ISBN 978-3-942036-14-6

211 Tolossa, Habtamu: *Sediment Transport Computation Using a Data-Driven Adaptive Neuro-Fuzzy Modelling Approach,* 2012, ISBN 978-3-942036-15-3

212 Tatomir, Alexandru-Bodgan: *From Discrete to Continuum Concepts of Flow in Fractured Porous Media,* 2012, ISBN 978-3-942036-16-0

213 Erbertseder, Karin: *A Multi-Scale Model for Describing Cancer-Therapeutic Transport in the Human Lung,* 2012, ISBN 978-3-942036-17-7

214 Noack, Markus: *Modelling Approach for Interstitial Sediment Dynamics and Reproduction of Gravel Spawning Fish,* 2012, ISBN 978-3-942036-18-4

215 De Boer, Cjestmir Volkert: *Transport of Nano Sized Zero Valent Iron Colloids during Injection into the Subsurface*, 2012, ISBN 978-3-942036-19-1

216 Pfaff, Thomas: *Processing and Analysis of Weather Radar Data for Use in Hydrology*, 2013, ISBN 978-3-942036-20-7

217 Lebrenz, Hans-Henning: *Addressing the Input Uncertainty for Hydrological Modeling by a New Geostatistical Method*, 2013, ISBN 978-3-942036-21-4

218 Darcis, Melanie Yvonne: *Coupling Models of Different Complexity for the Simulation of CO_2 Storage in Deep Saline Aquifers*, 2013, ISBN 978-3-942036-22-1

219 Beck, Ferdinand: *Generation of Spatially Correlated Synthetic Rainfall Time Series in High Temporal Resolution - A Data Driven Approach*, 2013, ISBN 978-3-942036-23-8

220 Guthke, Philipp: *Non-multi-Gaussian spatial structures: Process-driven natural genesis, manifestation, modeling approaches, and influences on dependent processes*, 2013, ISBN 978-3-942036-24-5

221 Walter, Lena: *Uncertainty studies and risk assessment for CO_2 storage in geological formations*, 2013, ISBN 978-3-942036-25-2

222 Wolff, Markus: *Multi-scale modeling of two-phase flow in porous media including capillary pressure effects,* 2013, ISBN 978-3-942036-26-9

223 Mosthaf, Klaus Roland: *Modeling and analysis of coupled porous-medium and free flow with application to evaporation processes,* 2014, ISBN 978-3-942036-27-6

224 Leube, Philipp Christoph: *Methods for Physically-Based Model Reduction in Time: Analysis, Comparison of Methods and Application*, 2013, ISBN 978-3-942036-28-3

225 Rodríguez Fernández, Jhan Ignacio: *High Order Interactions among environmental variables: Diagnostics and initial steps towards modeling*, 2013, ISBN 978-3-942036-29-0

226 Eder, Maria Magdalena: *Climate Sensitivity of a Large Lake*, 2013, ISBN 978-3-942036-30-6

227 Greiner, Philipp: *Alkoholinjektion zur In-situ-Sanierung von CKW Schadensherden in Grundwasserleitern: Charakterisierung der relevanten Prozesse auf unterschiedlichen Skalen*, 2014, ISBN 978-3-942036-31-3

228 Lauser, Andreas: *Theory and Numerical Applications of Compositional Multi-Phase Flow in Porous Media*, 2014, ISBN 978-3-942036-32-0

229 Enzenhöfer, Rainer: *Risk Quantification and Management in Water Production and Supply Systems*, 2014, ISBN 978-3-942036-33-7

230 Faigle, Benjamin: *Adaptive modelling of compositional multi-phase flow with capillary pressure*, 2014, ISBN 978-3-942036-34-4

231 Oladyshkin, Sergey: *Efficient modeling of environmental systems in the face of complexity and uncertainty*, 2014, ISBN 978-3-942036-35-1

232 Sugimoto, Takayuki: *Copula based Stochastic Analysis of Discharge Time Series*, 2014, ISBN 978-3-942036-36-8

233 Koch, Jonas: *Simulation, Identification and Characterization of Contaminant Source Architectures in the Subsurface*, 2014, ISBN 978-3-942036-37-5

234 Zhang, Jin: *Investigations on Urban River Regulation and Ecological Rehabilitation Measures, Case of Shenzhen in China*, 2014, ISBN 978-3-942036-38-2

235 Siebel, Rüdiger: *Experimentelle Untersuchungen zur hydrodynamischen Belastung und Standsicherheit von Deckwerken an überströmbaren Erddämmen,* 2014, ISBN 978-3-942036-39-9

236 Baber, Katherina: *Coupling free flow and flow in porous media in biological and technical applications: From a simple to a complex interface description,* 2014, ISBN 978-3-942036-40-5

237 Nuske, Klaus Philipp: *Beyond Local Equilibrium — Relaxing local equilibrium assumptions in multiphase flow in porous media,* 2014, ISBN 978-3-942036-41-2

238 Geiges, Andreas: *Efficient concepts for optimal experimental design in nonlinear environmental systems*, 2014, ISBN 978-3-942036-42-9

239 Schwenck, Nicolas: *An XFEM-Based Model for Fluid Flow in Fractured Porous Media*, 2014, ISBN 978-3-942036-43-6

240 Chamorro Chávez, Alejandro: *Stochastic and hydrological modelling for climate change prediction in the Lima region, Peru*, 2015, ISBN 978-3-942036-44-3

241 Yulizar: *Investigation of Changes in Hydro-Meteorological Time Series Using a Depth-Based Approach*, 2015, ISBN 978-3-942036-45-0

242 Kretschmer, Nicole: *Impacts of the existing water allocation scheme on the Limarí watershed – Chile, an integrative approach*, 2015, ISBN 978-3-942036-46-7

243 Kramer, Matthias: *Luftbedarf von Freistrahlturbinen im Gegendruckbetrieb*, 2015, ISBN 978-3-942036-47-4

244 Hommel, Johannes: *Modeling biogeochemical and mass transport processes in the subsurface: Investigation of microbially induced calcite precipitation*, 2016, ISBN 978-3-942036-48-1

245 Germer, Kai: *Wasserinfiltration in die ungesättigte Zone eines makroporösen Hanges und deren Einfluss auf die Hangstabilität*, 2016, ISBN 978-3-942036-49-8

246 Hörning, Sebastian: *Process-oriented modeling of spatial random fields using copulas*, 2016, ISBN 978-3-942036-50-4

247 Jambhekar, Vishal: *Numerical modeling and analysis of evaporative salinization in a coupled free-flow porous-media system,* 2016, ISBN 978-3-942036-51-1

248 Huang, Yingchun: *Study on the spatial and temporal transferability of conceptual hydrological models*, 2016, ISBN 978-3-942036-52-8

249 Kleinknecht, Simon Matthias: *Migration and retention of a heavy NAPL vapor and remediation of the unsaturated zone*, 2016, ISBN 978-3-942036-53-5

250 Kwakye, Stephen Oppong: *Study on the effects of climate change on the hydrology of the West African sub-region*, 2016, ISBN 978-3-942036-54-2

251 Kissinger, Alexander: *Basin-Scale Site Screening and Investigation of Possible Impacts of CO_2 Storage on Subsurface Hydrosystems*, 2016, ISBN 978-3-942036-55-9

252 Müller, Thomas: *Generation of a Realistic Temporal Structure of Synthetic Precipitation Time Series for Sewer Applications*, 2017, ISBN 978-3-942036-56-6

253 Grüninger, Christoph: *Numerical Coupling of Navier-Stokes and Darcy Flow for Soil-Water Evaporation*, 2017, ISBN 978-3-942036-57-3

254 Suroso: *Asymmetric Dependence Based Spatial Copula Models: Empirical Investigations and Consequences on Precipitation Fields*, 2017, ISBN 978-3-942036-58-0

255 Müller, Thomas; Mosthaf, Tobias; Gunzenhauser, Sarah; Seidel, Jochen; Bárdossy, András: *Grundlagenbericht Niederschlags-Simulator (NiedSim3)*, 2017, ISBN 978-3-942036-59-7

256 Mosthaf, Tobias: *New Concepts for Regionalizing Temporal Distributions of Precipitation and for its Application in Spatial Rainfall Simulation,* 2017, ISBN 978-3-942036-60-3

257 Fenrich, Eva Katrin: *Entwicklung eines ökologisch-ökonomischen Vernetzungsmodells für Wasserkraftanlagen und Mehrzweckspeicher,* 2018, ISBN 978-3-942036-61-0

258 Schmidt, Holger: *Microbial stabilization of lotic fine sediments*, 2018, ISBN 978-3-942036-62-7

259 Fetzer, Thomas: *Coupled Free and Porous-Medium Flow Processes Affected by Turbulence and Roughness – Models, Concepts and Analysis*, 2018, ISBN 978-3-942036-63-4

260 Schröder, Hans Christoph: *Large-scale High Head Pico Hydropower Potential Assessment*, 2018, ISBN 978-3-942036-64-1

261 Bode, Felix: *Early-Warning Monitoring Systems for Improved Drinking Water Resource Protection*, 2018, ISBN 978-3-942036-65-8

Die Mitteilungshefte ab der Nr. 134 (Jg. 2005) stehen als pdf-Datei über die Homepage des Instituts: www.iws.uni-stuttgart.de zur Verfügung.