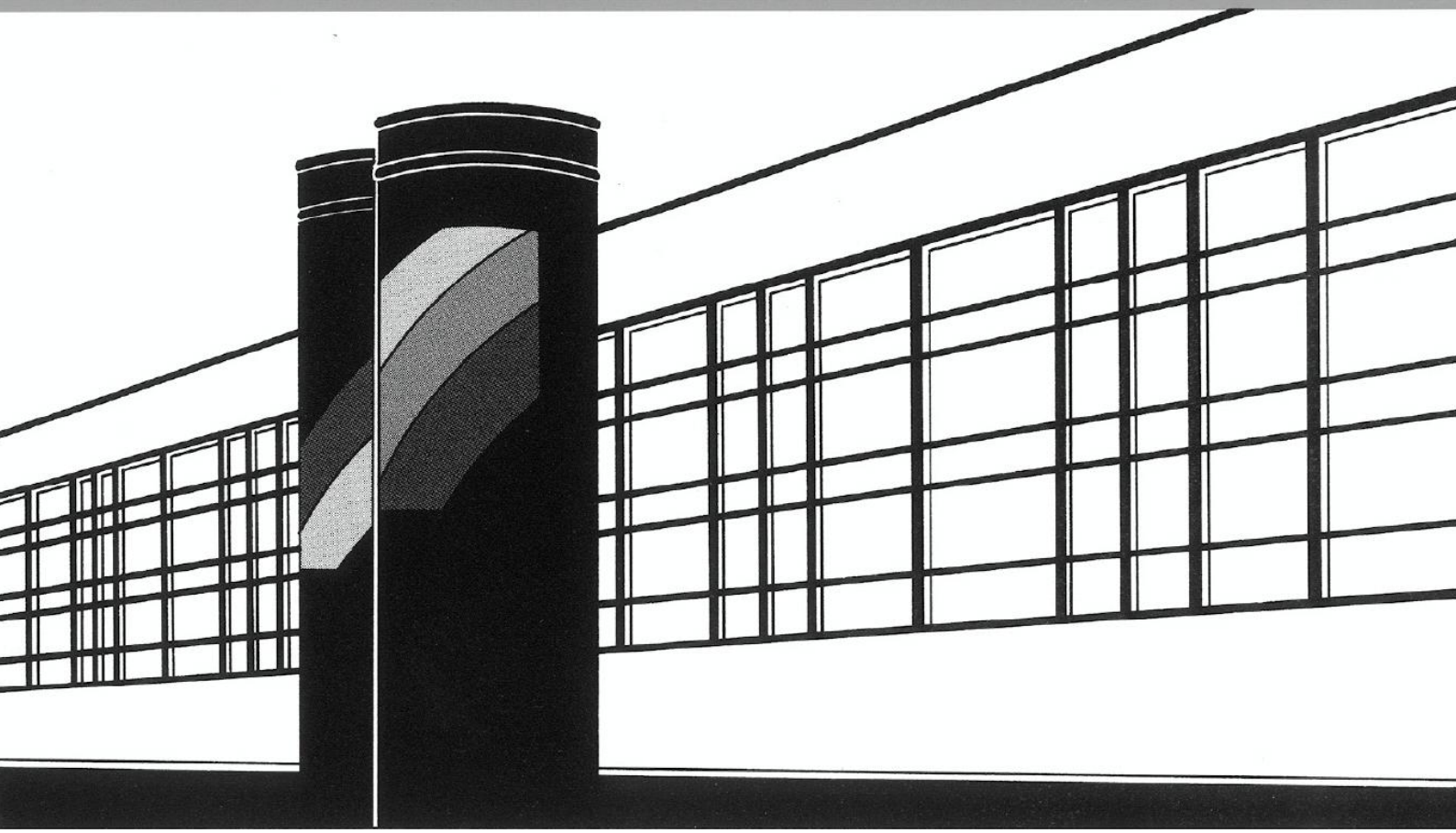
Universität Stuttgart

Institut für Wasser- und Umweltsystemmodellierung

Mitteilungen



Heft 229 Rainer Enzenhöfer

Risk Quantification and Management in Water Production and Supply Systems

Risk Quantification and Management in Water Production and Supply Systems

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

vorgelegt von

Rainer Enzenhöfer

aus Roth

Hauptberichter:	Jun.-Prof. Dr.-Ing. Wolfgang Nowak
Mitberichter:	Prof. Dr.-Ing. Rainer Helmig
Mitberichter:	Prof. Dr. Peter Grathwohl
Mitberichter:	Prof. Philip J. Binning.

Tag der mündlichen Prüfung: 17.12.2013

Institut für Wasser- und Umweltsystemmodellierung
der Universität Stuttgart
2014

Heft 229

Risk Quantification and Management in Water Production and Supply Systems

von
Dr.-Ing. Rainer Enzenhöfer

Eigenverlag des Instituts für Wasser- und Umweltsystemmodellierung
der Universität Stuttgart

D93 Risk Quantification and Management in Water Production and Supply Systems

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

Enzenhöfer, Rainer:
Risk Quantification and Management in Water Production and Supply Systems. Institut für Wasser- und Umweltsystemmodellierung, Universität Stuttgart. - Stuttgart: Institut für Wasser- und Umweltsystemmodellierung, 2014

(Mitteilungen Institut für Wasser- und Umweltsystemmodellierung, Universität Stuttgart: H. 229)
Zugl.: Stuttgart, Univ., Diss., 2014
ISBN 978-3-942036-33-7

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

Herausgegeben 2014 vom Eigenverlag des Instituts für Wasser- und Umweltsystem-modellierung

Druck: Document Center S. Kästl, Ostfildern

Acknowledgments

It has been a great pleasure for me to work under the supervision of Wolfgang. I always enjoyed the fruitful discussions about work and life, thus making it to a great place for working. Still, he knew how to challenge me, especially in terms of understanding all mathematical aspects of my work. Thanks for the patience, time and liberty to explore the field of risk quantification and management from a stochastic point of view. Furthermore, I want to thank Rainer, who opened up the opportunity for me to start my dissertation in the field of risk management in collaboration with the Landeswasserversorgung. I often think back at our KoWa times, which has been a great privilege for me to be part of supporting water research on a (inter-)national level. My gratitude is also to Philip, who has invited me to Denmark and challenged me with critical thoughts in our discussions. They improved my research and also shaped my personality and the way how to approach things in a more structured way. Last but not least, I have to thank Peter Grathwohl, who stepped into the breach for me to write up an expert opinion within the last minute and for being on my examination board.

I also want to thank my funding sources, which have been over the past years the German Research Foundation (DFG-KoWa), the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart and the Baden-Württemberg Stiftung.

Furthermore, I want to thank Sebastian Most, Michael Sinsbeck, Felix Bode, Julian Mehne, Andreas Geiges, Jonas Koch and Philip Leube for cross reading parts of my thesis or helping me by any upcoming questions. I am also indebted to the Landeswasserversorgung for their productive discussions and in specific Martin Emmert and Tobias Bunk.

I also want to thank my parents, who enabled me to study in Stuttgart and gave me the freedom to cut my own path through life. A special thank is rewarded to my wife, Kerstin, for always supporting me over the past six years, cutting back on her own wishes, taking our lovely two kids, Elias and Mattea, in order to free up time for my writing and so forth. Thank you very much. Here, I also want to thank my parents in law, who always opened their doors for Kerstin and our kids, when I have been traveling. My last gratitude belongs to God, the author of life and the ultimate risk manager to guarantee safety (Psalm 121:2-4).

Stuttgart, January 2014 Rainer Enzenhöfer

Contents

List of Figures

List of Tables

Abbreviations

Abbreviation	Denotation
ADI	Allowable Daily Intake
ALARP	As Low As Reasonable Practicable
ASTM	American Society of Testing and Materials
AWWA	American Water Works Association
BTEX	Benzene, Toluene, Ethylbenzene, Xylen
CDIs	Chronic Daily Intakes
CLEA	Contaminated Land and Exposure Assessment
CML	Customer Minutes Lost
DALY	Disability-Adjusted Life Years
DNAPL	Dense Non-Aqueous Phase Liquid
DVGW	German Technical and Scientific Association for Gas and Water
EPA	Environmental Protection Agency
GIS	Geographic Information System
GLUE	Generalized Likelihood Uncertainty Estimation
HACCP	Hazard Analysis and Critical Control Point
HI	Hazard Index
HQ	Hazard Quotient
IELCR	Individual Excess Lifetime Cancer Risk
IGD	Inverse Gaussian Distribution
IRIS	Integrated Risk Information System
IWA	International Water Association
MCL	Maximum Concentration Level
MODFLOW	Groundwater Flow Simulation Software by the US Geological Survey
MODPATH	Lagrangian Transport Simulation Software adapted to MODFLOW
NAPL	Non-Aqueous Phase Liquid
NOAEL	No-Observed-Adverse-Effect-Level
PEST	Model-Independent Parameter Estimation & Uncertainty Analysis Software
PREDUNC	Utility Program of PEST (Reduction of Prediction Uncertainty by Data Acquisition)

PTRW	Particle Tracking Random Walk
RAGS	Risk Assessment Guidance for Superfund
RAMCAP	Risk Analysis and Management for Critical Asset Protection
RBCA	Risk-based Corrective Action
REV	Representative Elementary Volume
RfC	Reference Concentration
RfD	Reference Dose
RI	Risk Index
RSF	Random Space Function
SEQ-ESO	Systme d'Evaluation de la Qualit des Eaux Souterraines (Quality index)
SF	Cancer Risk-based Slope Factor
STORM	Stakeholder-Objective Risk Measure
TDI	Tolerable Daily Intake
TU	Toxic Units
UN	United Nations
US	United States
VIP	Vulnerability Isopercentile
WHO	World Health Organization
WSP	Water Safety Plan
WVC	Well Vulnerability Criteria

Notation

The following table shows the significant symbols used in this work. Local notations, such as sub- and superscripts are explained in the text.

Symbol	Definition	Dimension
Greek Letters:		
α	Non-compliance level	$[\%]$
$\alpha_{L,T}$	Dispersivity (longitudinal, transversal)	$[L]$
β	Reliability	$[\%]$
γ	Water prizing function	$[\text{EUR}/L^3]$
γ_z	Semi-variogram	
ζ	Critical concentration level	$[M/L^3]$
η	Shape parameter	$[L]$
κ	Matérn shape parameter	$[-]$
λ	Failure rate per Time	$[1/T]$
$\lambda_{x,y}$	Correlation length	$[L]$
λ_ℓ	Lagranian parameter	
μ	Ensemble mean	
ν	Inverse Gaussian shape parameter	
ξ	Degradation factor	$[1/T]$
ξ	Stakeholder-objective risk measure	
ρ	Correlation coefficient	$[-]$
σ	Standard deviation	
τ	Critical time level	$[T]$
ϕ	Hydraulic pressure head	$[L]$
ψ	Contaminant-specific toxicological risk characteristic	
ω	Weighting factor	$[-]$
Ξ	Model function	
Σ	Model boundary	
Φ	Objective function	
Ω	Modeling domain	

Latin letters for scalars:

c	Concentration	$[M/L^3]$
d	Water demand	$[L^3/T]$
g	Discount rate	$[\%/T]$
h	Hazard	$[-]$
j	Contaminant type index	$[-]$
m	Mass	$[M]$
m_k	k-th temporal moment	$[MT^k/L^3]$
$\dot{m}$	Mass discharge	$[M/T]$
n	Number of unknown parameter values	$[-]$
n_e	Porosity	$[\%]$
q_r	Recharge rate	$\left[L/(TM^2)\right]$
s_i	Gauss-Integration Points	$[L]$
t	Time component	$[T]$
u	Utility value	
x	Spatial coordinate	$[L]$
y	Spatial coordinate	$[L]$
z	Random variable / spatial coordinate	$[-]/[L]$
A	Area	$\left[L^2\right]$
B	Benefit	
C	Costs	
CLE	Contaminant load exposure	$[M/T]$
C_z	Spatial covariance function of z	
D_m	Molecular diffusion coefficient	$[L^2/T]$
E	Entropy	$[-]$
F	Failure	$[-]$
G_k	Integral function of k-th moment	$[MT^k/L^3]$
H	Hydraulic head	$[L]$
H_n	Hermite polynomial of order n	
I	Indicator function	$[-]$
J	Annuality	$[1/T]$
K	Hydraulic conductivity	$[L/M]$
L	Likelihood	
MBR	Maximum blending ratio	$[-]$
MCR	Maximum concentration ratio	$[-]$
P	Probability	
Q_p	Pumping rate	$\left[L^3/T\right]$
R	Risk	
R_d	Retardation	$[-]$
RT	Response time	$[T]$
T_{Tot}	Total simulation time	$[L^2/T]$
T	Transmissivity	$[T]$
S	Spill location	$[L]$

U	Uncertainty measure	$[L^2]$
V	Volume	$[L^3]$
WET	Well exposure time	$[T]$
Y	Log-transmissivity	$[-]$
Z	Random variable	$[-]$

Symbols for vectors and matrices:

$\boldsymbol{\beta}$	Trend vector	
$\boldsymbol{\Gamma}$	Thikonov regularization matrix	
$\boldsymbol{\epsilon}$	Measurement error vector	
$\boldsymbol{\theta}$	Structural parameter vector	
$\boldsymbol{\xi}$	Normal distributed random vector	
$\boldsymbol{\omega}$	Weighting vector	
$\boldsymbol{d}$	Data measurement vector	
$\boldsymbol{h}$	Separation (distance) vector	$[L]$
$\boldsymbol{n}$	Normal vector	
$\boldsymbol{q}$	Darcy flux	$[L/T]$
$\boldsymbol{s}$	Random space vector	
$\boldsymbol{v}$	Velocity	$[L/T]$
$\boldsymbol{x}$	Spatial coordinates	$[L]$
$\boldsymbol{B}$	Displacement matrix	
$\boldsymbol{D}$	Hydromechanic dispersion tensor	$[L^2/T]$
$\boldsymbol{J}$	Jacobian (sensitivity) matrix	
$\boldsymbol{K}$	Hydraulic conductivity tensor	$[L/M]$
$\boldsymbol{R}$	Error covariance matrix	
$\boldsymbol{X}_p$	Particle position	$[L]$

Operators:

∇	Nabla operator
Δ	Difference operator
$\partial()$	Partial differential operator
$\cdot$	Scalar product
δ	Dirac delta function
$()^T$	Matrix transpose
$()^{-1}$	Matrix inverse

Kurzfassung

Motivation

97% des weltweit verfügbaren Frischwassers liegt als Grundwasser vor. Grundwasser ist eine begrenzte Ressource. Daher sind dessen Schutz und Management ein vorrangiges Ziel von gesellschaftlicher, ökologischer und ökonomischer Relevanz. Die konstante Nachfrage nach sicherem und sauberem Trinkwasser steht im direkten Konflikt mit gesellschaftlichen und wirtschaftlichen Landnutzungsansprüchen.
Wassergebietsmanager sind daher stetig mit der Aufgabe konfrontiert, alle potenziellen Gefahren in einem Wasserschutzgebiet bestmöglich zu kennen, zu kontrollieren und zu beherrschen, so dass ein Wechsel vom Schutzgebietsmanagement hin zum aktiven Risikomanagement stattfindet. Trotz dieses angestrebten Wechsels fehlt in vielen Bereichen eine klare Definition zur Handhabung von Unsicherheiten. Unsicherheiten im Risikomanagementsystem treten beispielsweise bei der Parametrisierung von Modellvariablen, der Wahl des Simulationsmodells und Diskretisierungsansatzes, der mangelnden Kenntnis über die geologische Beschaffenheit des Untergrundes und der geringen Datengrundlage sowie dem menschlichen Handeln auf. Durch eine genauere Kenntnis des Systems, z.B. des Untergrunds, können die physikalischen Prozesse modell-technisch und mathematisch besser approximiert und damit die Unsicherheit reduziert werden. Aven (2011) zeigt, dass eine Unsicherheitsanalyse ein integraler Bestandteil einer Risikoabschätzung ist. Diese versucht nach Kaplan and Garrick (1981) die drei Fragen (1) *Was kann passieren?*, (2) *Wie wahrscheinlich ist es?* und (3) *Wie hoch ist das Schadensmaß, wenn es eintritt?* zu beantworten.
Leider existieren kaum umfassende systembezogene Konzepte für Brunneneinzugsgebiete, die Risiken von der Gefahrenquelle bis zum Schutzgut simulationsgestützt, probabilistisch und physikalisch basiert quantifizieren, und diese mit entscheidungstheoretischen Ansätzen koppeln, um die menschliche Entscheidungsvariabilität mit zu berücksichtigen. Somit ist bislang eine rationale und optimale Entscheidungsunterstützung unter Berücksichtigung der Risikowahrnehmung im Risikomanagement von Trinkwasser nicht möglich.

Zielformulierung

Es soll daher ein robustes Risikomanagementsystem erarbeitet werden, welches eine **probabilistisch-quantitative** und **physikalisch basierte** Analyse des Rohwassers und damit eine Begrenzung von Gefährdungen in der Trinkwasserwirtschaft erlaubt. Das globale Ziel dieser Arbeit ist die Einführung eines innovativen Bayesschen Risikokonzepts, um das Risiko in der Trinkwassergewinnung aus Grundwasser zu quantifizieren und mit Hilfe

von robusten und neuen Entscheidungskonzepten in Abhängigkeit der unterschiedlichen Interessensziele zu steuern und zu minimieren. **Quantitative Risikoanalyseansätze** sind den qualitativen überlegen, da diese zum Beispiel eine Aggregierung von Schäden am Empfänger auf einer Kardinalskala ermöglichen und somit den Gesamtschaden im Vergleich zu qualitativen Methoden (Ordinalskala) genauer beziffern. Darüber hinaus erlauben quantitative Methoden ein objektives, transparentes Risikomanagement und bieten in Verbindung mit ökonomischen Konzepten Risikomanagern ein wertvolles Entscheidungssystem. Das Risikomanagementsystem soll **flexibel** gestaltet sein, so dass einzelne Module im Risikomodell verhältnismäig leicht ausgetauscht werden können (vergleiche Elemente und Knoten im Fehlerbaum). Dies ermöglicht die Verwendung des Risikosystems mit beliebig vorhandener Software und wenigen Daten. **Probabilistische** Methoden ermöglichen die Quantifizierung der Modell- und Paramaterunsicherheit, so dass Risikomanager die aktuelle Risikosituation besser einschätzen können. Aufgrund der komplexen hydrogeologischen Struktur im Untergrund und dem Unwissen über deren genaue Verteilung (mangelnde Datenlage) unterliegt die Modellierung von Transportflüssen in Grundwassersystemen groen Unsicherheiten. Mit Hilfe von datengetriebenen Bayesschen Kalibrierungstechniken kann die vorhandene Unsicherheit reduziert werden, so dass die Risikowerte näher an der Realität liegen. Massenflüsse drfen nur mit **physikalisch basierten** Transportmodellen ber die Zeit, dem Ort und der Schadenshäufigkeit (Frequenz) aggregiert werden. Dies ermöglicht eine genauere Bestimmung des Gesamtschadens und damit eine verbesserte Risikoabschätzung, da einzelne Gefährdungen häufig keine Gefahr für den Empfänger darstellen, aber in ihrer Summe diesen beeinträchtigen. Vor diesem Hintergrund muss eine verlässliche Risikoanalyse und Priorisierung von Gefährdungen kumulative Aspekte mit berücksichtigen. Insgesamt erlaubt die Arbeit eine genauere Analyse, Quantifizierung und bessere Begrenzung von kumulativen Risiken, eine zielgerichtete Einrichtung von Monitoringsystemen, und ein szenario-basiertes optimales Risikomanagement. Im Vordergrund steht die methodisch-konzeptionelle Entwicklung des Risikoquantifizierung- und -managementsystems, flankiert von zwei Fallstudien. Als Risiko wird das erwartete Nichteinhalten eines a-priori gesetzten Schwellenwertes definiert, das eine sichere und saubere Trinkwasserversorgung verhindert.

Ansatz

In dieser Arbeit werden zwei quantitative Risikomodelle vorgestellt. Ein Modell, VIP, weist mit probabilistischer Information schwellenwertüberschreitende Grenzlinien zum Schutz des Brunnenwassers durch Schutzgebietsausweisung aus. Das andere Risikomodell, STORM, berechnet das Gesamtrisiko basierend auf interessenspezifischen Risikomaßzahlen und einer kumulativen, aggregierenden Betrachtung von Risiken. Das daraus resultierende Gesamtrisiko setzt sich aus der Überlagerung von stoffungleichen, räumlich, zeitlich, und periodisch versetzten Gefährdungen zusammen. Die flexible, robuste und risikobasierte Entscheidungsanalyse basiert in beiden Methoden auf vier Brunnenvulnerabilitätskriterien (Frind et al., 2006). Diese Vulnerabilitätskriterien werden mit Hilfe von Monte Carlo Simulationen probabilistisch ausgewertet und damit epistemische Unsicherheit quantifiziert.

Dadurch sind für jedes Kriterium und an jedem Ort im Einzugsgebiet probabilistische Informationen in Form von kumulativen Verteilungsfunktionen vorhanden. Die probabilistische Information zu den vier Kriterien ist:

1. Die Wahrscheinlichkeitsverteilung der Spitzenankunftszeit für mögliche Einträge von Schadstoffen an jedem Raumpunkt des Einzugsgebietes.
2. Die Wahrscheinlichkeitsverteilung der Spitzenkonzentration im Brunnen und damit ein Maß für das Verdünnungspotential des Aquifers aufgrund von Diffusion und Dispersion.
3. Die Wahrscheinlichkeitsverteilung der vorhandenen Reaktionszeit für Gegenmaßnahmen vom Tag des Schadeneintritts bis zum dem Tag, an dem die Stoffkonzentration im Brunnen über einem bestimmten Grenzwert liegt.
4. Die Wahrscheinlichkeit der Dauer dieser Grenzwertüberschreitung, gemessen als Brunnenausfallzeit.

Sie dienen als Grundlage zur Ableitung jeglichen transport-basierten Risikomaßes, z.B. „customer minutes lost“, „toxic unit“, „hazard quotient“ und andere.
Als Gegenmaßnahme zu teuren Monte Carlo Simulationen wird das Strömungsfeld umgekehrt und der Transport rückwärts vom Brunnen bis zu allen möglichen Punkten im Einzugsgebiet berechnet. Dies umgeht die Notwendigkeit für jeden einzelnen Diskretisierungspunkt des Einzugsgebietes ein Vorwärtsmodell laufen zu lassen. Als weitere Modellreduktion wird der Ansatz der zeitlichen Momente (e.g., Harvey and Gorelick, 1995) verwendet. Zur Auswertung der Vulnerabilitätskriterien muss die komplette Information zum Konzentrationsverlauf am Brunnen vorliegen. Die Kombination aus (partikel-basierten) zeitlichen Momenten (hier: Lagrange (Salamon et al., 2006) und Euleransatz (Harvey and Gorelick, 1995)) und der analytischen oder numerischen Kurvenrekonstruktion (hier: inverse Gauß Verteilung (Folks and Chhikara, 1978) und nicht-linearer max-entropischer Ansatz in log-Zeit (Harvey and Gorelick, 1995)) liefern alle transportrelevanten Informationen für ein flexibles und robustes Risikomanagement. Mangelnde Daten über die Materialeigenschaften des Untergrunds und fehlende Kenntnis der Materialheterogenität führen zu unterschiedlichen Parametersätzen gleicher Modellgüte, so dass eine Bandbreite von möglichen Ergbenissen existiert. Mit Hilfe der Geostatistik können für diese Parametersätze unterschiedliche Heterogenitätsfelder generiert werden, die den Materialeigenschaften des Systems entsprechen. Zur Reduzierung der epistemischen Unsicherheit wurde das geostatistische Modell mit Hilfe des Bayes'schen Ansatzes stochastisch (formelles Bayes'sche GLUE, Feyen et al., 2003) auf direkte und indirekte Daten kalibriert und damit aktualisiert. Bei deterministischer Kalibrierung von Strömungsmodellen kann die Kovarianzmatrix der Parameter nach der Kalibrierung berechnet und für die konditionelle Simulation verwendet werden. Die Ergebnisse aus der Transportsimulation dienen als Grundlage für beide Risikomodelle (STORM und VIP), um einerseits das Gesamtrisiko aus dem Versagen mehrerer verschiedener Gefährdungen am Empfänger, andererseits eine risikobewusste Schutzgebietsausweisung zu quantifizieren. Das VIP und das STORM-Konzept sind modell- und schutzgebietsunabhängig. Dies wird durch zwei methodisch getrennte VIP Ansätze an einem synthetischen und realem Brunneneinzugsgebiet illustriert. Die quantitativ erfassten Risikowerte werden durch Zuweisung von Nutzwerten für

ein transparentes, rationales Entscheidungssystem verwendet, um Zielformulierungen unterschiedlicher Interessengruppen optimal bedienen zu können.

Schlussfolgerung

Diese Arbeit zeigt, dass in allen Belangen quantitative Risikoansätze den qualitativen überlegen sind. Ebenso weist sie darauf hin, dass nur physikalisch-basierte probabilistische Informationen innerhalb einer Risikoabschätzung und eines robusten risikobasierten Entscheidungshilfesystems essentiell sind. Die Aggregation von Risiken, die sich zeitlich am Empfänger überlagern, führt nur durch Aggregation auf der Ebene von Massenflüssen zu plausiblen und korrekten Risikomaßen. Es dürfen nur Gefährdungen mit gleicher Schadensausprägung (z.B. Nitrat, chlorierte Kohlenwasserstoffe) auf Basis der Massenflüsse zu einem kumulativen Risiko aggregiert werden. Das Gesamtrisiko ergibt sich in Abhängigkeit der Risikomazahl z.B. durch Summation oder durch ein anderes Rechenverfahren. Des Weiteren erlaubt die Bayes'sche Modellkalibrierung sowohl a-priori Wissen zur Reduzierung von Unsicherheit einzubinden, als auch eine realitätsnähere Einschätzung der Modellergebnisse. Es konnte gezeigt werden, dass die Unsicherheitsreduzierung nicht per se Risiko minimiert und von dem gewählten Nutzwert abhängt. Risiko selbst ist von der Risikomaßzahl, welche das Ziel der Risikoabschätzung definiert, dem Risikoansatz und dem Risikoverhalten abhängig. Bestimmte Gefährdungen zeigten einen höheren Einfluss auf häufig wiederkehrende Ereignisse als andere und umgekehrt. Mit Hilfe dieser Differenzierung und der Berücksichtigung der Versagenswahrscheinlichkeit eines Gefährdungsträgers kann eine klare Priorisierung der Gefährdungen basierend auf dem ausgehenden Risiko stattfinden, so dass der Vergleich zwischen Gefährdungen mit niedrigem, häufigem und hohem, seltenen Schaden gelingt. Des Weiteren führt die Summe der Einzelrisiken zu einem Risikowert, welches die Priorisierung und die gesamte Risikosituation fehlerhaft, gegebenenfalls mit groen Folgen fr das Risikomanagement und der Risikominimierung, einschätzt. Selbst als Worst-Case Szenario unterschätzt die Summation der Einzelschäden den maximal beobachteten Gesamtschaden um mehr als $50\,\%$. Nichtsdestotrotz sind $98\,\%$ aller Gesamtschäden kleiner als die Summe der Einzelschäden. Insgesamt wurden unterschiedliche robuste und risikobasierte Entscheidungsunterstützungssysteme vorgestellt, die eine interessensspezifische, optimale Entscheidung in Abhängigkeit der jeweiligen Risikowahrnehmung erlauben.

Summary

Motivation

97% percent of the world's usable freshwater is stored as groundwater, which is a limited resource. Thus, its protection and management is a world-wide major societal, health-related, ecologic and economic concern. The constant demand for clean and safe drinking water is in direct conflict with social and economic land-use claims.
Therefore, water managers are challenged to know (1) what kind of hazards exist within the water catchment, (2) how these hazards can be controlled and (3) knowing that they are controlled. Thus, water management shifts from fixed and thus passive wellhead delineation zones to active risk management. Despite this desired change, a clear definition on dealing with uncertainties in risk assessment and management for drinking water supply systems is still missing. Aven (2011) shows that uncertainty analysis is an integral part of risk assessment. Also, the US EPA (2001) promulgates cumulative probability distribution functions to assess confidence bounds, regarding the risk prediction. These uncertainties are, for example, a result of measurement error, model conceptualization and parameterization. Therefore, it is necessary to quantify uncertainty as part of risk assessment. According to Kaplan and Garrick (1981), risk assessment addresses three questions (1) What can happen?, (2) What is the probability that it happens? and (3) What is the damage after it happens? Thus, in general risk is the „*uncertainty about and severity of the consequences of an activity with respect to something that humans value*" (Aven and Renn, 2009).
Unfortunately, only few comprehensive risk concepts exist for drinking water supply systems that address risk from source to receptor, while considering uncertainty and physically-based modeling aspects. Modularized, transport-based and probabilistic risk quantification models coupled with a rational, and stakeholder-objective decision analysis framework for groundwater supply systems do not yet exist. Only with this type of comprehensive risk model, stakeholders are able to estimate risk at the receptor level most accurately. This supports stakeholders to take risk-informed, implementable, transparent, and evidence-based decisions (Pollard et al., 2008) in an uncertain environmental framework. Furthermore, it pushes water governance to the next higher level.

Goal

The main purpose of this work is to present a new methodological risk concept within a Bayesian framework to quantify and manage risk within groundwater resources for drink-

ing water supply, utilizing smart decision analysis concepts based on multiple stakeholder-objectives. The risk concept has to be **quantitative**, **flexible**, **probabilistic** and **physically-based**.
Quantitative risk assessment approaches are superior to qualitative ones. For example, they allow the aggregation of hazard impacts, provide transparency due objectivity, and enable risk-informed management that is based on cardinal scale and economic concepts. The **flexibility** in the risk model allows stakeholders to easily exchange single modules (compare fault-tree: nodes or events) with ready available software and modeling techniques in a plug and play mode. **Probabilistic** methods that quantify uncertainty provide a prediction space of many possible outcomes, such that stakeholders can better evaluate the current risk situation. Especially in case of subsurface heterogeneity and the lack of knowledge about the structural distribution, it is indispensable to quantify uncertainty. It is possible to reduce uncertainty by Bayesian-based conditioning techniques, moving risk estimates closer to reality. State-of-the-art transport-based models are able to calculate the cumulative hazard impact at the target objective as required by European Commission (2003). Only **physically-based** transport models allow aggregation of mass discharges across space, time and frequency. This allows risk managers to evaluate hazards more precisely as individual hazards are often deemed to be no risk, although contributing to the overall expected impact at the well. Therefore, hazard ranking across the catchment has to be performed in a cumulative environmental setting. Thus, risk quantification concepts are in demand to provide valuable and indispensable information for water stakeholders that are quantitative, flexible, probabilistic and physically-based. Only admitting uncertainty and utilizing this type of risk framework stakeholders are able to take transparent, robust, rational, and risk-informed decisions.

Approach

To fulfill the stated goal, I present two quantitative risk modeling frameworks in this thesis, where the first approach is an integral part of the other. The first model, VIP, provides probabilistic information to delineate vulnerability-based wellhead protection zones for a given non-compliance probability level. The second risk model, STORM, quantifies the overall risk based on the transport results of VIP, taking cumulative risk aspects and the stakeholder-objective views into account. The overall risk is deduced from the total concentration breakthrough curve at the well, where mass discharges from different release locations of the same contaminant type are aggregated across space, time and frequency. The accumulation over hazard type is available by statistical analysis. In both cases, a flexible, robust and risk-informed decision analysis is based on the four well vulnerability criteria by Frind et al. (2006). These four intrinsic vulnerability criteria are cast into a Monte Carlo framework, resolving epistemic uncertainty. They deliver pixel-wise probabilistic information for each location across the catchment. The additional probabilistic information is:

1. The probability of **peak arrival time** from all potential spill locations to be faster than a required minimum time.
2. The probability of **peak concentrations** in the well to be larger than some maximum allowed level (e.g., an MCL).

3. The probability of the time window available to react (=**reaction time**) after a spill event until a critical concentration level is exceeded in the well (e.g., drinking water standard) and the probability to be smaller than a minimum critical value required to take adequate counter measures.

4. The probability that the well has to be shut down or is exposed to a non-compliance contamination level for more than a given critical duration (= **exposure time**).

These four measures serve as intermediate risk level to deduce any known risk measure, such as Customer Minutes Lost, Toxic Units, Hazard Quotient and so forth.
Due to the probabilistic approach via Monte Carlo simulations, the featured concept would lead to high computational costs, especially in light of stochastic conditioning schemes. As a counter measure, model reduction techniques such as the reverse flow formulation (e.g., Neupauer and Wilson, 2001) and temporal moment-based transport calculations are employed. STORM relies on the transport results of the VIP approach. In order to demonstrate the flexibility of the risk concept, two novel model combinations are presented. It also shows that probabilistic risk assessment is already available with common software. One set of tools uses a Eulerian transport description resolving heterogeneity on and above a fine grid scale in order to distinguish between dilution and plume location. The backward temporal moment approach (e.g., Harvey and Gorelick, 1995) is combined with a maximum entropy-based reconstruction technique in log-time, utilizing Gauss-Hermite integration. The second model suite combines particle-based moments with analytical breakthrough curve reconstruction. The particle-based moments are obtained by MODPATH (e.g., Pollock, 1994) extended with a random walk step to account for dispersion. The breakthrough curve reconstruction technique is based on the inverse Gaussian distribution (e.g., Folks and Chhikara, 1978). For both models, conditional geostatistical fields have been generated, in order to quantify and reduce epistemic uncertainty. The first model is conditioned on direct and indirect data by the formal Bayesian Generalized Likelihood Uncertainty Estimator (e.g., Feyen et al., 2003). The other model provides conditional flow simulations based on the post-calibration covariance matrix which is available by automated deterministic calibration models, such as PEST (e.g., Doherty and Hunt, 2010). The presented VIP concept is catchment- and model-independent. The transport-based vulnerability results serve as a basis for robust and transparent decision making. The concept improves the accuracy of current and possible future wellhead protection outlines by risk-aware delineation (VIP). The VIP approach is demonstrated for a synthetic and a real test catchment.
The transport results are used for STORM to quantify the overall risk of water supply failure, when water is abstracted from groundwater. The STORM model is a forward approach and assigns to each individual hazard across the catchment a probability of failure. These hazards fail independently of each other, such that mass discharges of the same contaminant type are aggregated at the well, although hazards have failed at different times and locations with varying frequency. The expected impact per year or the impact for a given return period can be estimated per hazard, contaminant type and across all risk sources. Thus, STORM is a valuable tool for active risk management to install most effective mitigation measures and to find the most severe hazards through prioritization.

Conclusion

The thesis demonstrates that quantitative, probabilistic and physically-based risk models are superior to qualitative or deterministic risk assessment approaches in all aspects to support transparent and robust decisions. The overall risk allows stakeholders to monitor risk over time, in order to detect risk trends. Only physically-based models are capable to aggregate mass discharges at the well that overlap in space, time and frequency. Therefore, only mass-discharge-based risk measures are accurate enough to base decisions upon, as individual hazards alone often pose no risk. Hazards with varying units (across chemical type) are aggregated based on statistical analysis and not per summation. This is due to the fact that risk estimation based on critical levels are highly non-linear. Furthermore, the Bayesian framework allows to reduce epistemic uncertainty. Within a synthetic case study, the worth of data in light of accuracy, quantity and data type has been investigated based on pre-defined sampling scenarios. The study shows that measurement accuracy plays an important role in uncertainty reduction and that the sampling campaign with nine measurements delivers the cost-optimal situation for delineating wellhead protection zones with a $90\,\%$ safety level. Overall, the four vulnerability criteria set into a probabilistic framework provide the intermediate basis, to obtain any known transport-based risk measure. It has also been shown that prioritization depends on the risk model, risk estimate, risk objective and risk-attitude. There exist risk sources that influence more frequent impacts than others and vice versa. Due to this discernment and the consideration of hazard failure frequency it is possible to prioritize hazards and compare hazards with low, frequent impacts to high, rare events. Furthermore, the sum of hazard-wise risk levels lead to a bias in total risk with consequences in misjudging the current risk situation and thus the risk management. Even in case of a worst-case scenario this summation underestimates the maximum observed impact level by more than $50\,\%$ by covering $98\,\%$ of all possible severity levels. By coupling a decision analysis framework to the risk quantification approaches robust, transparent and economic-based risk management is available for water risk managers.

1. Introduction

„A ship is safe in harbor, but that's not what ships are for."
by William G.T. Shedd (1820 – 1894)

1.1. Motivation

In 2010, the UN declared supply with clean and safe drinking water to be a new human right (A/64/L.63/Rev.1, 2010). To date many national regulations exist worldwide to protect drinking water resources (e.g. DVGW, 2006; US EPA, 1993). Nevertheless, the World Health Organization (WHO) finds that poor water quality still poses a severe risk to human health, both in developing and industrialized countries (WHO, 2004). They recommend using a risk-based approach in order to ensure fresh, clean and safe drinking water via implementing water safety plans into national legislation from catchment to tap (Davison et al., 2005).
To date, approximately 74% of the drinking water in Germany is supplied from groundwater (UBA, 2013). Denmark is almost fully supplied by groundwater. Many other countries face a similar situation. These groundwater resources are vulnerable to human activities and natural phenomena. Industrial, agricultural, transport and geogenic hazards are typically located across the drinking water catchments and pose a severe risk, varying in magnitude and toxicity, to the supplied drinking water. Thus, stakeholders may ask themselves questions, such as:

1. How safe is the current supply situation?
2. Which hazards pose the greatest risk to the supply well?
3. How and to what extent can the supply situation (reliability) be improved?
4. What are the most effective risk mitigation measures?

Risk assessment and management schemes help to answer the above questions. In fact, by adopting a risk framework, stakeholders will start to determine all risk sources in the catchment, estimate the risk at the receptor level and take risk-informed, implementable, transparent and evidence-based decisions (Pollard et al., 2008).

1.2. Problem Setting and Goals

As soon as water stakeholders start the challenging task of risk assessment, they are confronted with additional questions that are more detailed and related to the risk analysis and

management approach:

i. What are suitable and reliable risk analysis methods that avoid risky (hence deterministic) and too conservative (hence simplistic models) decisions? (I-1 to I-6, II-1)

ii. How valuable are qualitative (semi-quantitative) and easy to use risk assessment tools in spite of their approximate character? (I-2)

iii. How can epistemic uncertainty aspects be incorporated into risk analysis and what are the benefits? (II-3)

iv. What are relevant risk measures and how can risk be measured in a quantitative sense? (III-2, III-3)

v. How to get an overall risk level across all hazards in the catchment despite its variability? (III-1)

vi. To what extend can cumulative risk assessment lead to alternative decisions? (III-1)

vii. To what extend can risk measures be used to manage or mitigate risk? (II-3, III-3)

viii. Do different risk measures lead to different decisions? (III-2, III-3)

In light of these questions, the overall aim for this thesis is:

> *to assess the reliability of safe water supply by developing smart risk quantification and management concepts and tools that support risk-informed and robust decision making in the light of uncertainty, using state-of-the art transport-based models and statistical methods.*

Therefore, I will combine the conceptual ideas of physically-based risk assessment of contaminated sites with geostatistical and probabilistic methods within a Bayesian framework. This will improve the risk management process to enable risk-informed decisions in an uncertain framework. By doing so, I will present three major contributions to the community.

1. The first contribution is an overview and in-depth **introduction in risk assessment** and management in the field of environmental engineering. Advantages, disadvantages and limitations of qualitative and quantitative risk models are discussed that will allow risk managers to judge the reliability and suitability of risk analysis methods.

2. The second contribution (**VIP**) casts a deterministic macro-dispersive vulnerability-based risk model into a probabilistic framework in order to account for uncertainties in the model used. VIP stands for *V*ulnerability *I*so*P*ercentile that show spatially mapped lines of non-compliance with pre-defined safety levels. VIP is a physically- and catchment-based backward risk assessment concept. It introduces four well vulnerability values as intermediate risk levels.

3. The third major contribution (**STORM**) utilizes the vulnerability isopercentile concept (VIP) to establish a forward cumulative risk model that is able to account for failure probabilities for each hazard and to adapt to the objectives of water stakeholders by implementing stakeholder-view-dependent risk measures. STORM is an acronym for stakeholder-objective risk measures.

Along these three major contributions, there exist several sub-aims. These sub-aims will help to answer the above stated questions. Please note, the numbering in the brackets above relates the questions to the proposed sub-aims.

1.3. Specific Goals and Contributions

Introduction in Risk Assessment (Chapter 3)

I-1) Risk guidance: In this thesis, I aim to provide general guidance for water stakeholders and scientists to perform risk assessment and management for drinking water catchments the best possible way in line with current regulations and scientific standards (Chapter 3). Furthermore, I aim at clarifying the use of risk terminology, as risk is clearly defined in the international standard (ISO, 2009), but often used more vaguely in the water-related literature.

I-2) Choice of risk model: In order to avoid pitfalls and misconception, while using qualitative and quantitative risk models, risk managers and scientists need to be critical with available methods and software. With Chapter 3 and Chapter 10, I want to demonstrate the strength and weaknesses of qualitative, deterministic and probabilistic risk approaches.

I-3) Physically-based risk models: The delineation of wellhead protection zones is classically based on deterministic physically-based models. Also risk assessment of contaminated sites is based on deterministic transport-based models. In this thesis, I want to demonstrate that meaningful risk analysis is only available with mass-discharge-based risk models, as only these models will be able to adequately aggregate the severity of impact at the target level (Chapters 3, 7, and 10).

I-4) The role of dispersion in risk assessment: Dispersion lowers peak concentration of contaminants at the well. The presented risk concepts are able to distinguish between these dilution effects and actual uncertainty in plume location, as they avoid upscaled Fickian macrodispersion. Nevertheless, deterministic transport models often use for simplicity advective-only assumption. A detailed discussion on why dispersion matters is given in Sections 3.6.3 and 8.3.

I-5) Probabilistic risk models: These models enable risk managers to take decisions in light of cumulative probability distributions of the estimated risk measures. Within this thesis, I want to demonstrate that additional information is available by quantifying epistemic uncertainty, supporting risk-informed decision analysis (Chapters 4 to 6).

I-6) Modularity: Risk models need to be flexible to adapt to the local situation and available software. The aim is to introduce software- and catchment-independent risk frameworks, where each step can be performed by another model. Therefore, I will show the modularity of the step-wise procedures of VIP (Chapter 4) and STORM (Chapter 7).

VIP (Chapters 4-6, 8, 9)

II-1) Applicability to practitioners: In this thesis, I want to show that risk-based wellhead delineation (here: the VIP framework) is easy to understand, conceptually simple and that VIP maps can be intuitively interpreted without special training. For this reason, I will apply the VIP framework with readily available software (Chapter 9) and demonstrate it on an actual catchment.

II-2) Risk quantification and management with VIP :

- The resulting risk levels are commonly based on advective or upscaled macrodispersive transport calculations. With VIP I will show that *peak arrival time* is the more conservative and better risk measure for wellhead delineation in comparison to bulk arrival time (e.g., Section 8.4).
- VIP will introduce a decision analysis framework that allows *risk-aware wellhead delineation*. Questions on the actual well safety or the increase in safety at no (or only little) additional costs in comparison to the current delineated area are addressed (e.g., Section 6.2).
- How much can uncertainty be reduced by collecting more and better data? One of my goals is to demonstrate the *worth of data* in wellhead protection zone delineation and to what extent it is worth to add additional data in light of economic constraints (e.g., Section 6.1).

II-3) Reduction opportunities in computational time: Combining physically-based and probabilistic risk models within a Bayesian geostatistcial framework is one of the major goals as previously stated. This directly leads to dramatically high computational costs due to the required Monte Carlo framework. Therefore, the aim is to reduce computational time by adequate model reduction techniques (e.g., Section 5.5, Chapter A).

STORM (Chapters 7, 10)

III-1) Cumulative risk assessment: Risk estimation is challenged by the variability of hazards. They differ in contaminant properties, spatial location across the entire catchment and temporal occurrence of hazard failures. With STORM, I want to demonstrate how to assess cumulative risk levels, despite the complexity in aggregating the severity levels of different types of hazards (e.g., Section 7.5). Furthermore, the validity of risk summation is discussed (e.g., Sections 3.6.5 and 10.2.2).

III-2) Stakeholder-objective views: Within a participatory process different interest groups exist that value different risk objectives. This makes risk assessment close to an impossible task to satisfy all required needs. It is therefore my aim, to introduce intermediate risk levels obtained from one model that are based on mass discharge information. Any risk measure should be quickly and easy available through these intermediate risk levels (e.g., Section 7.5). The influence of varying risk objectives on the resulting decisions is also investigated (e.g., Section 10.3.3).

III-3) Risk estimation and management with STORM :

- In environmental systems, *risk estimates* are classically calculated by probability times damage. In this conception, risk is conditional to failure events that have in fact occurred. STORM estimates the statistically expected damage within a given return period, also considering the time between failures for assessing a risk estimate. Chapter 10 will demonstrate the difference between the two risk estimation approaches.
- STORM is capable to investigate the effectiveness of scenario-based *risk mitigation* options. How risk mitigation is influenced by several factors, such as risk attitude, time-dependent mitigation strategy, or options available, is further investigated in Section 10.3.
- A major advantage of quantitative risk models is *hazard prioritization*. In Section 10.3.1, I will discuss challenges in hazard prioritization and compare hazard ranking lists based on different risk models, risk measures and risk preference.

1.4. Framing the Thesis

This section provides a brief overview on assessing risk in the field of water resource protection. The aim of this section is to frame the approach pursued in this thesis (see Chapter 2). Relevant sections are highlighted, where an in-depth discussion exists.
As previously mentioned, the WHO (Davison et al., 2005) proposed water safety concepts to determine and control all known hazards from catchment to tap. The multi-barrier concept (O'Connor, 2002) intensifies this requirement by stating that each barrier from source to tap should guarantee safe drinking water. Within a drinking water catchment, there exist several threats that may alter the quality of groundwater resources, such as agricultural activities and substances (e.g., fertilizer, pesticides), industrial areas (e.g., dry cleaners, gas stations) or hazards from settlements (e.g., sewage system, waste site). Thus, risk assessment and management for drinking water supply begins on the catchment level to protect the used water resources (Section 3.4) and should follow the source-pathway-receptor concept (e.g., US EPA, 1989). This is similar to risk assessment for contaminated sites (Section 3.3.2).

Qualitative and Quantitative Risk Assessment

The most classical form is wellhead protection by delineating time-related protection zones (e.g., Stauffer et al., 2005, Section 3.6). Despite this, there exist multiple risk assessment models to estimate the risk of individual hazards. On international and national level, the WHO (Davison et al., 2005) or the DVGW (DVGW, 2009), respectively, propose to use qualitative risk estimation methods in order to rank and prioritize hazards and thus to identify potential threats to the drinking water source. In fact, many stakeholders follow these guidelines and rely on qualitative risk methods. Nevertheless, qualitative methods are of limited use to

support stakeholders in taking risk-informed and informative decisions (Neukum and Azzam, 2009). Especially in the context of cumulative impact assessment as requested by the European Commission (2003), qualitative methods fail to accurately predict the cumulative severity level. In addition, Cox (2008) argues that ranking of individual hazards that fall within the same risk category on a categorical scale used in qualitative methods is impossible. A detailed discussion on qualitative risk models and their usefulness is provided in Section 3.5.
Only quantitative risk assessment methods can actually improve effective, rational, transparent, and honest risk management decisions regarding prioritization and risk reduction (Section 3.6). These methods quantify risk levels of individual hazards on cardinal scales. Thus, prioritization of all hazards within one hazard category is available. Here, I distinguish between two communities, one using physically-based deterministic transport models (Section 3.6.1), the other using probabilistic risk models (Section 3.6.2).

Structured Risk Assessment

As early as in 1975, the U.S. Nuclear Regulatory Commission (1974) performed probabilistic risk assessment studies for nuclear power plants by fault tree analysis. The applied probabilistic methods quantify uncertainty and risk by assigning failure probabilities to single elements of the tree. Only recently, fault tree analysis models or graph-based systems (e.g., Nilsen and Aven, 2003) gained importance in environmental systems and subsequently in drinking water supply studies (e.g., Lindhe et al., 2009; Rodak and Silliman, 2012; Sadiq et al., 2007). These graph-based probabilistic models enable stakeholders to consider hazards across location and contaminant types within the entire catchment, while considering uncertainty aspects. The system is modularized into sub-components, containing probability information on possible failure events. Furthermore, graph-based models allow to flexibly exchange single tree elements (nodes) with more sophisticated models or with software and knowledge available. For example, Lindhe et al. (2009) performed a quantitative risk analysis via fault tree analysis, considering only expert knowledge to assign probabilities to their tree elements without using analytical or numerical software tools. Although being a promising instrument, up to date, these models cannot aggregate mass at the receptor level to adequately predict the overall and combined severity of impact (e.g., Bolster et al., 2009; Fernàndez-Garcia et al., 2012).

Physically-based Risk Assessment

A decade later, the US Environmental Protection Agency (EPA) introduced the risk assessment guidance for Superfund sites in groundwater engineering (US EPA, 1989), which marks the beginning of deterministic mass-discharge-based risk quantification (e.g., Cushman et al., 2001, Section 3.3.2). Hereupon, many studies followed and adopted the proposed source-pathway-receptor concept, estimating the level of severity at the receptor by using numerical or analytical transport models (e.g., Jamin et al., 2012, Section 3.3.2). Due to a conservative approach in estimating risk, many authors assume advective-only transport calculations (e.g., Tait et al., 2004). This assumption may lead to estimates that misjudge the actual

risk situation, because dispersion lowers concentration levels in the pumped raw water, as the contaminant plume mixes and spreads along the pathway to the well (e.g., Kitanidis, 1994). A detailed discussion on dispersion is given in Section 3.6.3. Overall, physically-based models are often scenario-based, deterministic and focused on known contaminated sites, neglecting uncertainty.

Probabilistic Risk Assessment

As early as 1998, Evers and Lerner (1998) asked the question of how uncertain wellhead protection delineation is. Upscaled parameters such as hydraulic conductivity introduce model and thus prediction errors. Geostatistical models resolve the unknown heterogeneity of the subsurface. Therefore, setting these geostatistical fields into a probabilistic context helps to quantify parameter uncertainty related to the subsurface (e.g., de Barros et al., 2012; Rubin, 2003). For example, Feyen et al. (2003) used a Bayesian modeling framework to further reduce the parameter uncertainty by conditioning the geostatistical model to measured hydraulic head and conductivity values. Many more studies exist that use a Bayesian approach to update their model accuracy by more and better data (e.g., de Barros et al., 2009, Chapter 6). Further aspects of uncertainty in risk assessment are provided in Section 3.6.4.

Cumulative Risk Assessment

Furthermore, physically-based models allow aggregation of impacts across space and time of contaminants with identical contaminant properties (e.g., Troldborg et al., 2008). Cumulative impact assessment is required by the European Commission (2003) that directly leads to the task of cumulative risk assessment (e.g., US EPA, 2007). In cumulative risk assessment there exists many challenges, such as aggregation across different contaminant types, accounting for different failure times, temporal arrival of contaminants at the well, spatial distribution across the catchment, and so forth. Some problems, such as aggregation across space or different contaminant types are already solved by using utility functions (e.g., Fishburn, 1970) within a physically-based transport model (e.g., Jamin et al., 2012; Troldborg et al., 2008). Nevertheless, all studies related to cumulative physically-based risk assessment neglect uncertainty and failure frequency of individual hazards. In addition, cumulative or total risk assessment advances the identification of risk trends, if assessed and monitored over a longer time period. Therefore, well safety concepts are in demand of a risk assessment framework that admits and quantifies uncertainty, while considering state-of-the-art mass-discharge-based transport models (Section 3.6.5).

Stakeholder-objective Risk Measures

Frind et al. (2006) were among the first to consider dispersion in impact assessment by introducing an upscaled Fickian macrodispersion transport model to assess well vulnerability

at the drinking water well within a backward deterministic risk model (e.g., Cushman et al., 2001). The information obtained from a contaminant breakthrough curve is summarized by four intrinsic well vulnerability criteria. These well vulnerability criteria determine the impact of possible contamination load from source to receptor, exactly as desired by Einarson and Mackay (2001). The concept of well vulnerability is a fundamental part of this thesis, as it provides the necessary information to support stakeholders in transparent and risk-informed decision making (Chapter 6). One of the major contributions of this thesis will be to add the uncertainty aspect to the well vulnerability criteria, as probabilistic and risk-related information are missing. Therefore, the concept is explained in more detail in Chapter 4. Just recently, de Barros et al. (2013) used well vulnerability criteria in the context of environmental performance metrics. Still, the focus in probabilistic risk assessment is on environmental and human-health risk assessment, as stated by Öberg and Bergbäck (2005), and not on technical or economic-related issues.
Nevertheless, various stakeholder groups such as policy makers, water utility managers, environmental scientists or consumers are concerned about different risk objectives. Each stakeholder group measures impact on their own relevant severity scale, leading to multiple risk measures with as many possibilities for competing decisions. Example are the impact on the environment and on human-health (e.g., Öberg and Bergbäck, 2005) with risk measures such as toxic units (e.g., McKnight et al., 2012), hazard quotients (e.g., Hodgson, 2012) and daily intake rates (e.g., Rodak and Silliman, 2012; US EPA, 1989). Lindhe et al. (2009) used a risk measure taken from the energy sector, which is called customer minutes lost to assess the technical down-time of the Gothenburg water supply system.
Up to presence, there exists no study that considers economical, technical, environmental or health-related risk aspects all at once, thus satisfying the information needs of stakeholders from different disciplines with one single risk concept. MacGillivray et al. (2006) state that "there are potential tensions between managing the risks of a commercial water business and the overarching public health", especially in the light of financial pressure. Therefore, it is indispensable to provide information that deliver stakeholder-objective risk measures (STORM) that calculate risk measures as mentioned above and beyond. These and other meaningful stakeholder-objective risk measures are available by the four well vulnerability criteria. Providing a STORM concept will be the third major contribution of this thesis, next to the probabilistic well vulnerability concept and the introductory overview to risk assessment and management. A detailed discussion on STORM is given in Chapter 7.

Risk Management and Decision Analysis

The choice of the best management option can easily get very complex, when considering the spatial, temporal and cumulative hazard impact across several risk sources with different impact dimensions (hazard type). Risk quantification is one part of risk management. The risk estimates are evaluated in light of critical (pre-defined, regulatory-based) levels. In case, risk is unacceptable hazard prioritization supports to target effective mitigation options (e.g., Troldborg et al., 2008), such as land-use changes (e.g., Rodak and Silliman, 2012). None of these studies considered risk of non-compliance in light of uncertainty reduction neither through conditioning schemes (e.g., Schöniger et al., 2012) nor improving wellhead

protection zone delineation to capture more accurately the well catchment. Hazard prioritization (e.g., Troldborg et al., 2008), choice of mitigation alternatives (e.g., Lindhe et al., 2009; Rodak and Silliman, 2012) or improving the system reliability are risk management options, that are available all at once within the VIP and STORM framework. The concept of decision analysis (e.g., Freeze et al., 1990) is used to find the best available scenario-choice. A detailed review on risk management and decision analysis is given in Chapters 3 and 6.

Most Relevant Quantitative Risk Studies

Overall, in the light of implementable and still reliable risk management decisions, stakeholders need models and scenarios, that should neither be overly expensive (e.g., simplistic models) nor too risky (e.g., deterministic models). There exist only few risk assessment approaches that fulfill these requirements to ensure and quantify water supply safety by combining both probabilistic and physically-based aspects with state-of-the-art models (e.g., de Barros et al., 2009; Rodak and Silliman, 2012; Tait et al., 2004). Rodak and Silliman (2012) use a fault tree analysis framework, thus failing to aggregate risk levels. Tait et al. (2004) use among others an advective-only assumption, zonation-based approach and neglect the fact of impact aggregation due to failure frequency. de Barros et al. (2009) focuses on improving human health risk measures through additional hydraulic conductivity data. The latest work by de Barros et al. (2013) assesses the influence of relevant length scales to environmental performance metrics and the benefit of uncertainty reduction for more accurate environmental performance metrics without considering cumulative risk aspects, model flexibility or risk measures beyond human-health.

1.5. Thesis Description - A Starting Point

Due to the shortcomings sketched in the previous section and the goals of Section 1.2, I define the thesis as:

> *Risk assessment and management for drinking water supply systems has to be probabilistic, quantitative and physically-based. Only physically-based models enable the aggregation of mass-discharges at the well, which is the only valid way to account for cumulative effects across hazards in space, temporal arrival and frequency. Aggregation across different contamination types is only available by statistical analysis on the level of risk estimation. The flow and transport calculations should avoid upscaled models that already include averaging procedures, because they would hinder a clean statistical analysis. Also, they should be set into a Bayesian and/or geostatistical probabilistic modeling framework to quantify and reduce epistemic uncertainty. That being said on a conceptual level, adopting the flexibility and modularity of graph-based risk methods will be a valuable step on the way from theory to practice. For practical application, it is also relevant to develop a unique concept that offers all desirable risk measures for multiple stakeholder views out of one method. Thus, stakeholder decisions can be supported by stakeholder-objective risk measures and by intuitive decisions analysis frameworks.*

2. Global Approach and Structure

The overall aim for this thesis is:

> *to assess the reliability of safe water supply by developing smart risk quantification and management concepts and tools that support risk-informed and robust decision making in the light of uncertainty, using state-of-the art transport-based models and statistical methods.*

Water supply safety has two dimensions, first providing enough water (quantitative aspect) and second water that is of sufficient quality. The thesis focuses on the last aspect, protecting groundwater resources that are used for drinking water supply, to avoid system failure due to exceedance of critical threshold levels.
In order to set the stage for the hypothesized thesis in Section 1.5, the presented study will provide in **Chapter 3** an overview on current risk regulations and available risk concepts in the field of environmental engineering and drinking water supply to set the frame for quantitative risk assessment. The starting point of Chapter 3 is to clarify risk terminology across disciplines and for this thesis. With the help of Chapter 3, I will demonstrate the advantage of *quantitative* risk assessment methods in comparison to *qualitative* ones, and why quantitative methods should be *probabilistic* and *physically-based*. Furthermore, I will also address the need for *cumulative risk assessment*. These findings are applicable to any research field, such as automotive, financial markets, medicine, IT services and so forth. Therefore, Chapter 3 provides an excellent overview on risk assessment and management that should enable risk managers to judge the suitability and validity of a given risk model.

Two Risk Modeling Concepts (VIP and STORM)

Starting from **Chapter 4**, I will introduce two quantitative, physically-based and probabilistic risk modeling concepts. These two concepts will improve risk quantification and management by quantifying parameter uncertainty within a Bayesian framework and by providing rational and robust decision analysis based on the stakeholder-dependent risk objectives. Both concepts follow the source-pathway-receptor concept, as shown in Fig. 2.1.
The first approach follows the idea of *backward* risk assessment, delineating wellhead protection zones probabilistically via *V*ulnerability *I*so*P*ercentiles (**VIP** concept). The VIP approach is based on nine methodological steps that cover modules # 5, 6, 8, 9 and 10 as shown in Fig. 2.1. The VIP concept provides probabilistic risk information on contaminant well exposure.
The second approach is a *forward* risk analysis, assessing *ST*akeholder-*O*bjective *R*isk *M*easures (**STORM** framework) within a modularized framework. The transport results

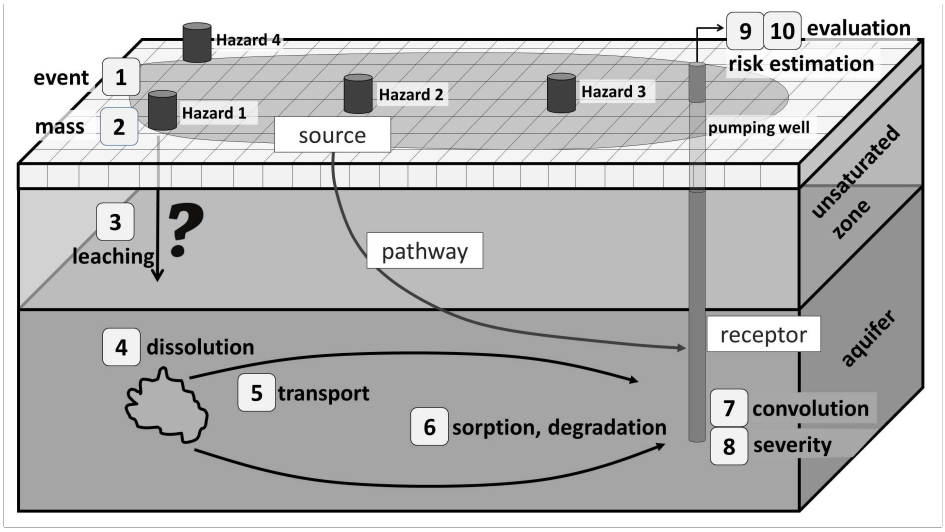


Figure 2.1.: Conceptual regional risk model (STORM) to characterize from source (modules 1, 2) via pathway (modules 3-6) to receptor (modules 7, 8) the level of severity, supporting risk-informed decisions (modules 9, 10).

of the VIP approach are used within the STORM concept. STORM is based on ten modules as shown in Fig. 2.1 and extends the VIP approach by modules 1 to 4 and 7, thus relating exposure risk to a temporal frequency component.

Global VIP Approach and Structure

The VIP concept aims to protect drinking water resources via delineation of probabilistic capture or protection zones. It provides the possible exposure risk to contaminants solely based on the location of hazards. The flow and transport model requires geostatistical methods and parameterization of hydraulic conductivity to account for the spatial subsurface heterogeneity. The calibrated flow model is set into a Bayesian modeling framework to condition the model to observation data, thus reducing epistemic uncertainty. In order to distinguish between the uncertainty in plume location and actual dilution, Monte Carlo simulations are used to resolve spatial heterogeneity while using local-scale dispersivities. Monte Carlo methods lead to cumulative distribution functions instead of just statistical moments, providing the probabilistic risk-based information, as proposed by the US EPA (1997). Nevertheless, Monte Carlo simulations require a large computational demand, especially in the light of model conditioning. As a counter measure, the flow field is reversed and transport-based time-related moments are calculated to efficiently reconstruct the full history of contaminant breakthrough curves for each point in the domain, either by analytical or numerical solution schemes. From these breakthrough curves, intrinsic or post-processed specific vulnerability criteria can be deduced such as peak arrival time and maximum peak concentration. When compared to critical levels in each realization within a Monte Carlo framework (e.g., reflecting regulatory standards or personal risk acceptance levels), isopercentile lines of non-compliance can be mapped to visualize safety margins across the catch-

ment. Then isopercentile lines serve as input for decision analysis, such as risk-aware delineation or the worth of data to reduce epistemic uncertainty.
A detailed step-by-step explanation of the risk quantification concept on a conceptual level is provided in **Chapter 4**. The key idea is to set the four well vulnerability criteria by Frind et al. (2006) into a probabilistic and risk context.
Chapter 5 provides the mathematical description of the VIP concept. The individual model components are explained, following a step-wise procedure. The appendix provides a smart and fast breakthrough curve reconstruction technique based on Gauss-Hermite integration (Appendix: Chapter A).
Chapter 6 describes the use of probabilistic well vulnerability criteria in a risk management context, in order to support stakeholders in taking risk-informed decisions based on rational decision analysis concepts. These VIP-based decision concepts include among others, risk-aware delineation, the concept of data worth, and the benefit of using additional data for robust decisions.
The VIP concept is applied to two test catchments: a synthetic (**Chapter 8**) and a real test site (**Chapter 9**). The two test sites are chosen to demonstrate two different toolboxes that simulate flow and transport processes for assessing the vulnerability isopercentile lines independently.

Global STORM Approach and Structure

The STORM concept aims to protect drinking water resources via prioritization of hazards across the catchment in light of their individual and cumulative impact to the receptor. There are two key points that are featured by STORM. First, stakeholder views are multi-objective and related to human-health, environmental or technical risk aspects. The presented risk quantification framework STORM is flexible enough to adapt to these desired risk objectives by using the vulnerability criteria as intermediate risk levels. Secondly, it accounts for the spatial distribution and temporally recurring failure events of hazards across the catchment by mass-discharge aggregation at the receptor. The aggregation across different contaminant types is available via statistical analysis.
The STORM approach is modularized into ten sub-models (see Fig. 2.1), following the physical and temporal consecutive steps that a contaminant takes from source to receptor. To each possible risk source across the domain, a failure rate and a probability distribution of mass release are assigned. The occurrence of failure events per hazard are modeled via a binomial distribution (module 1, Fig. 2.1). The mass being released to the subsurface is also randomized (module 2, Fig. 2.1). These two modules are denoted as event model. The flow and contaminant transport through the unsaturated zone can be described by a time-lag model, and dissolution of contaminant mass to the aquifer is considered by means of a simple leaching model. Some contaminants may form pools in a separate immiscible phase, and thus dissolves only slowly into the ambient groundwater flow. This process attenuates the arrival, duration and severeness of peak concentration (modules 3 and 4, Fig. 2.1). These two modules are denoted as immission model. In a multi-hazard setting, the different hazard mass discharges of the same type may overlap at the well and form cumulative concentration breakthrough curves in the well. The transport and impact

model accounts for the pathway segment and thus estimates the damage at the receptor by calculating the well vulnerability information from these cumulative concentration breakthrough curves (modules 5-8, Fig. 2.1). Risk is expressed by the expected annual impact or the expected impact given a return period for each hazard, contaminant type and across type (modules 9, Fig. 2.1). With this information, stakeholders are able to evaluate risk, prioritize hazards and risk mitigation methods (module 10, Fig. 2.1). Risk mitigation is scenario-based and can lower risk by reducing failure probability (uncertainty), mass or both aspects. These different scenarios are set into an objective decision analysis, which risk mitigation alternative serves the objective best.
In summary, the STORM framework is a powerful tool to assess cumulative risk, pertained by spatial, temporal, frequency- and type-dependent hazard variability. It accounts for the multi-objective views of stakeholders and allows stakeholders to prioritize risk sources in a complex environment of interacting hazards at the receptor level (aggregation). The modularization into ten elements follows the idea of graph-based models to simply exchange individual modules with readily available knowledge and software at any time. The available information on each hazard is stored in a hazard database. A second database stores all contaminant-specific flow and transport properties.
A detailed explanation per module of the STORM framework is provided in **Chapter 7**. Chapter 7 focuses on accurate aggregation need and the advantage of using stakeholder-objective risk measures in risk management.
Chapter 10 demonstrates the application of the risk framework to a real test site, based on transport results from the VIP application shown in Chapter 9 and a synthetic hazard database. Furthermore, Chapter 10 presents possible risk quantification and management options. It focuses on and compares lists of hazard prioritization that are based on the chosen return period, risk objective, risk estimate and risk model. In addition to this, different scenario-based risk mitigation options are compared.

3. Introduction to Risk Assessment and Management

Aven and Renn (2009) and other authors unify risk terminology and concepts across disciplines in a consistent way. Despite these efforts, different scientific communities are still reluctant to set risk into a consistent and understandable way as defined by the international standard organization (ISO) and thus consistent across disciplines. For example, risk is often defined as probability times damage. In fact, risk is not a product per se, but rather a combination of both. Even zooming into one community, such as the environmental engineering one, shows that risk assessment is disunited, adding to the overall confusion. Although there exist clear guidelines, risk assessment and management is interpreted and used within each community differently due to historical reasons.
Therefore, it is one of my first aims to clarify the risk terminology in order to establish a meaningful and consistent terminology for risk assessment and management in this thesis (Section 3.1). Second, I will provide an overview on risk management as a general framework (see Section 3.2) and, in specific, for the field of environmental engineering (Section 3.3). Third, a short review on risk guidelines and regulations for water supply systems is presented in Section 3.4, with a special focus on German national guidelines. Finally and fourth, limitations and challenges in qualitative (Section 3.5) and quantitative (Section 3.6) risk assessment are discussed. Qualitative risk assessment includes index-based vulnerability methods, while the discussion on quantitative approaches focuses on graph-based and physically-based methods. Each of these three environmental modeling concepts (index-, graph- and physically-based) has historically built their own risk community and terminology.

3.1. Terminology

Risk assessment and management concepts have been widely praised and adopted to multiple research fields and applications. They support stakeholders and risk managers to take risk-informed decisions, regarding economical, social, environmental, technical, and health-related questions. This includes identification and prioritization of hazards in environmental systems or of project potentials, the preservation of system reliability, safety aspects and more. Literature shows (e.g., Aven, 2011), that there still exists a great confusion on how to use risk terms. In order to avoid misunderstanding, this section defines the most important terms for this thesis. The two terms risk and vulnerability are discussed in specific.

This is important, as I show in the remainder of this thesis (see Section 1.4) that many risk assessment studies are by definition vulnerability rather than risk assessment frameworks.

Risk is defined according to the AS/NZS 4360 Risk Management Standard (AS/NZS 4360:2004, 2004) as the chance of something happening that will have an impact upon predefined objectives. Thus, risk is measured in terms of consequences (impact) and likelihood (chance). This definition follows the commonly used definition by Kaplan and Garrick (1981), defining risk as a set of scenarios, where each scenario is defined by probability and consequence. Probability is here used in concurrence to likelihood or frequency. Aven (2011) argues that, within the risk definition, the term uncertainty should be used instead of probability, in order to account for all aspects of uncertainties within risk analysis, e.g., the lack of knowledge about outcomes or future event occurrences. Aven and Renn (2009) define risk as „uncertainty about and severity of the consequences of an activity with respect to something that humans value". To express this level of uncertainty, the term probability is commonly used, whereas other methodologies also exist to express uncertainty, e.g., possibility and evidence theory. Keeping probability instead of likelihood, one has to be aware that „probability is always conditional on a background knowledge" (Aven, 2011). Performing uncertainty analysis, e.g., through Monte Carlo simulations, also results in a probability density distribution of an outcome. This probability is a measure of variation and conditional to the degree of the resolved uncertainty. Furthermore, risk and all subsequent components, probability and damage, are estimated at the receptor level.

Risk analysis is a process that includes the definition of scope, hazard identification and risk estimation.

Risk assessment is risk analysis (see above) followed by risk evaluation. Risk assessment distinguishes between three possible risk estimation procedures, which are specified by the type of risk estimation (qualitative, semi-quantitative and quantitative).

Risk estimation is the process to obtain or calculate risk, respectively. There exist qualitative, semi-quantitative or quantitative risk estimation methods.

Qualitative risk estimation describes the level of uncertainty (resp. probability) and level of severity on ordinal scales, such as low, medium, high or on a point scale with a low maximum number of points, ranging, e.g., from 1 (very low) to 5 (very high), based on subjective and hard data.

Semi-quantitative risk estimation is formally identical to qualitative methods, just using a finer scale to categorize and rank hazards, regarding their probability and level of severity. Semi-quantitative methods often include intermediate damage and/or probability levels, which are based on aggregated case-specific criteria. They also employ cardinal operators (such as summation or averaging), although risk is measured on ordinal scale. This discrepancy to use cardinal operators on ordinal scales will receive more attention during the later parts of this thesis.

Quantitative risk estimation describes the level of severity and uncertainty on cardinal scales. Risk quantification methods include process- and statistically-based

approaches. Statistical or black-box methods allow to empirically calculate the damage probability from a large data sample. Process- or physically-based models calculate the damage probability function by mathematical expressions that represent nature at best.

Risk management can be understood as coordinated activities to direct and control a system with respect to risk (Aven, 2011). Risk management includes risk analysis (scope, hazard identification, risk estimation), risk evaluation, risk reduction (or risk mitigation, see below) and treatment ISO (2009).

Risk mitigation is an action to reduce a current risk level to an acceptable level.

Risk measure quantifies or measures the consequence (severity) that a receptor is exposed to, measured in a given dimension. Different choices of risk measures usually exist, each one possibly leading to different risk estimates and management options.

Risk source are all activities and objects that may potentially alter a desired system or organization state. Risk sources can have negative and positive effects. Here, risk sources with negative consequences are used interchangeably with the term hazard (see below).

Exposure is a process that a person or organism gets into contact with a hazard, thus bridging the gap between a hazard and a risk (Kolluru et al., 1996). Aven (2012) follows this definition by stating that exposure describes the fact that an object is being subject to an event that could possibly induce a negative impact.

Exposure risk assessment accounts for uncertainty in the pathway segment, thus describing the combination of exposure uncertainty and the level of severity at the receptor.

Hazard / Threat are terms used for human activities, objects or natural phenomena that may negatively deteriorate the receptor (see below), i.e., that can induce a negative impact on the protective good of interest.

Hazard failure is the event that a risk source fails with possible consequences to the system.

Receptor is the last step within the source-pathway-receptor risk concept by the US EPA (1989). The receptor depends on the scope of the risk analysis and may be, for example, a human population, an organization, system or the drinking water well.

Regional risk assessment defines the spatial extent of the hazard identification. Here, regional is defined as the catchment that contributes to the aquifer from which drinking water is drawn (see Troldborg et al., 2008).

Total risk or overall risk is the cumulative effect of uncertainty and consequences by all available risk sources at the receptor level.

Vulnerability of a receptor depends on the sensitivity to cope with a risk source. In terms of well vulnerability (Frind et al., 2006), the term vulnerability is distinguished into intrinsic and specific:

Intrinsic vulnerability depends only on the hydrogeological characteristics such as aquifer properties (US EPA, 1993), the overlying soil and geological material.

Specific vulnerability depends, in addition to intrinsic vulnerability, also on contaminant properties, such as retardation and degradation.

The International Standard Organization defines vulnerability as „intrinsic properties of something resulting in susceptibility to a risk source that can lead to an event with a consequence". Aven and Renn (2009) reformulates this statement in compliance with the above risk definition, that vulnerability is „uncertainty about and severity of consequences of an activity, given the occurrence of an (initiating) event.". Thus vulnerability and risk are based on the same concept, except that „vulnerability is always conditional on an event or risk source" (Aven, 2011). The proposed vulnerability definitions are in alignment with each other.

3.2. Risk Management - A General Framework

The literature shows clear disagreement on definitions and scope of risk management. Lindhe et al. (2009) defines risk management in the context of decision making, starting with the question whether the current risk is acceptable or not. As a next step, they compared several risk mitigation options within a decision analysis framework to find the best solution (Lindhe et al., 2011). Tartakovsky (2013) relates risk management to the task of finding the best decision either through optimization or by decision analysis. Furthermore, the author includes aspects of communicating risk and uncertainty to all participants involved in the risk assessment. Risk communication and reporting is often unintentionally ignored, especially within the scientific environmental modeling community, although being part of the risk management process. The objective of risk communication is to transparently convey information to all participants of the risk assessment process, such as local authorities, public and others.
In order to resolve these disagreements, national and international standards should be considered. By definition of the ISO (2009), risk management is more than merely taking a decision and communicating risk. Risk management actually involves all steps starting from risk analysis and risk assessment to the mitigation of risk and monitoring the system to ensure that risk is controlled (see Fig. 3.2). Nevertheless, most authors (e.g., Lindhe et al., 2011; Tartakovsky, 2013) relate risk management only to the additional working steps directly following the risk assessment. This point of view seems to be established and accepted community-wide. There exist several standards, which explain in detail the concept of risk management, such as DVGW (2009); ISO (2009) and the Australian Management Standard (AS/NZS 4360:2004, 2004). All these concepts share basic principles, although using a different risk terminology.
Thus, it makes good sense to adopt an even more global view, taken from management in general. One management framework that features the basic idea of risk management or any other management and decision process is the PDCA (Plan - Do - Check - Act) cycle. The PDCA cycle consists of four phases:

(1) In the „planning“ period, one has to identify the target and objectives of the risk assessment study and identify all potential hazards that my alter the defined goals. Furthermore, acceptable outcome levels are determined.

(2) The „Do“-phase includes the collection of all available information and data through monitoring or other means. With the available information, the outcome is estimated.

(3) The results are compared to the predefined acceptable outcome levels („check“) to define whether they are acceptable or not.

(4) The last phase adjusts the results to the point of acceptance through means of risk reduction methods („Act“), if necessary.

The PDCA management framework is continuous and helps to constantly and continuously improve the processes in a system, e.g., supply safety (DVGW, 2009). This is very close to risk management, which also aims to be a continuous developing process to improve the system state. Thus, risk management also incorporates the four presented phases, only naming and structuring the work flow differently.
A modified risk management framework according to ISO is illustrated in Fig. 3.2 and compared to the PDCA cycle. The ISO risk management framework is very detailed, thus being an excellent raw model for this thesis. The individual steps are explained next.

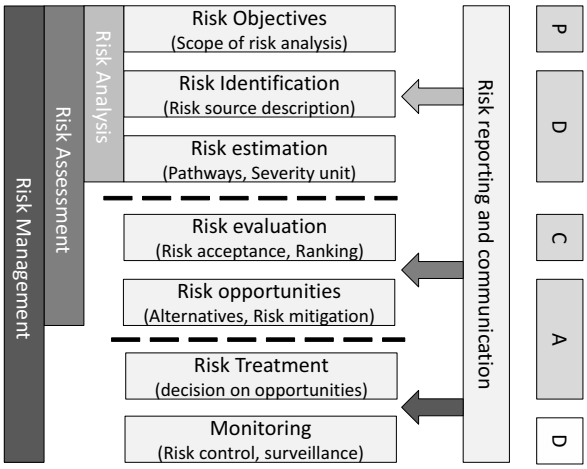


Figure 3.1.: Risk management framework modified after ISO (2009), compared to the PDCA cycle.

Risk Objective: The goal of this step is to define the objectives and targets of the risk assessment study. It includes the definition of the scenario to be analyzed. The following questions should be addressed: Which object is at risk? What is the unit to measure risk in, by which risk measure, and which risk level is acceptable and which one is not? Also, the type of risk assessment study is defined. This is either human-health or environmental risk assessment, following European Commission (2003) or US EPA (1989) standards. Nevertheless, the risk assessment objective could also be of technical or

economical nature, such as quantitative supply safety with a minimal water production down-time or supply with cost-minimal effort. Depending on the risk objective, the relevant risk estimation criteria and suitable risk models are chosen, respectively.

Risk identification: The risk identification phase includes, among others, the hazard identification, i.e., the collection of all potential risk sources in the system that might lead to a positive or negative effect on the target objective. In addition, all possible exposure routes from the risk source to the receptor are identified and reported. This challenges risk managers to know their system, including all relevant processes, in depth. There exist several hazard identification tools that help risk managers to identify all relevant hazards, such as using expert knowledge, performing a Delphi study or brainstorming of participants and experts. During the risk identification, process potential risk mitigation options may come up, which should be marked down as early as possible for a later use. All information gathered to this point should be systematically documented. This documentation is often done with the help of hazard databases, e.g., using GIS systems.

Risk estimation: Risk has an upside and downside element because the impact at the target can be positive or negative. Although being aware of upside risk, research communities commonly consider only downside risk. The estimation of risk is either measured by a monitoring network or calculated by qualitative or quantitative risk estimation methodologies. Most scientific work in risk assessment has been done for improved risk estimation, developing new tools and concepts to represent reality best and thus render the decision basis closer to reality.

Risk evaluation: The resulting risk is denoted as acceptable, ***a**s **l**ow **a**s **r**easonable **p**racticable* (ALARP), or as unacceptable. Assigning risk levels to the categorized impact and probability is by definition part of the risk evaluation process. This distinction is clear in qualitative risk assessment, as risk categories are assigned as a last step to the risk matrix (see explanation in Section 3.5.2). In quantitative risk estimation, risk evaluation is commonly only related to judging the risk estimate against pre-defined threshold levels, goals (see *risk objective*) or national standards, and evaluate whether the residual risk is acceptable or not. Risk acceptance depends on risk aversion (see *risk treatment*) and denotes the treatment of residual risk. In case risk is not acceptable, risk managers perform a hazard prioritization that depends on the hazard severity.

Risk opportunities: While risk evaluation focuses on the prioritization of risk sources according to their impact and probability at the receptor, the risk opportunities step focuses on the prioritization of mitigation measures in order to reduce the unacceptable risk the best way. The prioritization of risk mitigation options is based on the effectiveness of the reduction options to achieve the risk objectives defined in the first step. The decision by risk managers, which option to take (see *risk treatment*), can either be supported through optimization algorithms or by the concept of decision analysis. According to the ISO (2009), the establishment of risk opportunities is part of the risk evaluation process. The two steps are separated in this thesis, as determining the best possible risk mitigation measures is an additional step to hazard prioritization and can be independently performed of each other. Risk evaluation, risk opportunities and

risk estimation are closely related, but receive in this thesis a precise work description. Please note, that the individual step *risk opportunity* is not a separate working step in the ISO (2009) definition, but belongs to risk evaluation.

Risk treatment: According to the mitigation ranking and personal preferences, risk managers decide and implement one or more risk mitigation measures. Risk treatment is often denoted as the process of making and implementing a decision about the mitigation measure. As indicated before (see *risk evaluation*) the decision depends on the risk behavior or culture. There are three types of risk personalities or cultures:

Risk aversion leads to a decision, where the mitigation option is least uncertain but confident in its expected risk reduction, although it may be inferior to the other alternatives in the statistical average.

Risk neutrality leads to a decision based alone on the expected risk reduction, weighting effectiveness to reduce risk and uncertainty in the reduction potential the same.

Risk sympathy leads to a decision with highest risk reduction potential, although being uncertain.

Monitoring: In order to know, that risk mitigation measures were successfully implemented or that risk is below the pre-defined acceptable risk level (see *risk objective*), the system and all relevant processes or parameters are monitored. This step is continuous, as stakeholders should always know and re-confirm that all hazards and their impact at the receptor are under control (e.g., Davison et al., 2005).

Risk communication: In order to transparently reflect decisions and increase public awareness and confidence, risk communication should not be regarded as a separate process, but as being an integral part of risk assessment and management. It involves all interest groups and needs constant documentation of all findings.

Kaplan and Garrick (1981) summarize these steps by defining risk management as achieving an appropriate balance between realizing opportunities for gains while minimizing losses. Risk management is a continuous process, being constantly updated with new and most current information.

3.3. Risk Assessment and Management in the Field of Environmental Engineering

In this section, I want to first give a brief overview on how to possibly categorize risk models in the field of environmental risk assessment, at least into categories relevant for this thesis (Section 3.3.1). Secondly, the risk assessment of contaminated sites following the RAGS Manual (US EPA, 1989) is discussed (Section 3.3.2) and related to the definitions from Section 3.2. It is important to notice that this community around risk assessment of contaminated sites has dominated the risk discussion in the past decade in the field of environmental engineering.

3.3.1. Risk Model Categorization

Fig. 3.2 shows how to possibly categorize risk models. The focus of this overview is to support those that are new to this field of environmental risk assessment, in order to judge the reliability and suitability of risk models. The presented antipodes are chosen in such a way, that they are relevant for this thesis, providing literature examples for each category. It is obvious that any risk concept belongs to only one antipode, each. This table is not complete and more categories could be added, such as risk objectives (e.g., human-health, ecological, technical and so forth). The antipodes of potential risk models considered in the table are briefly explained next.

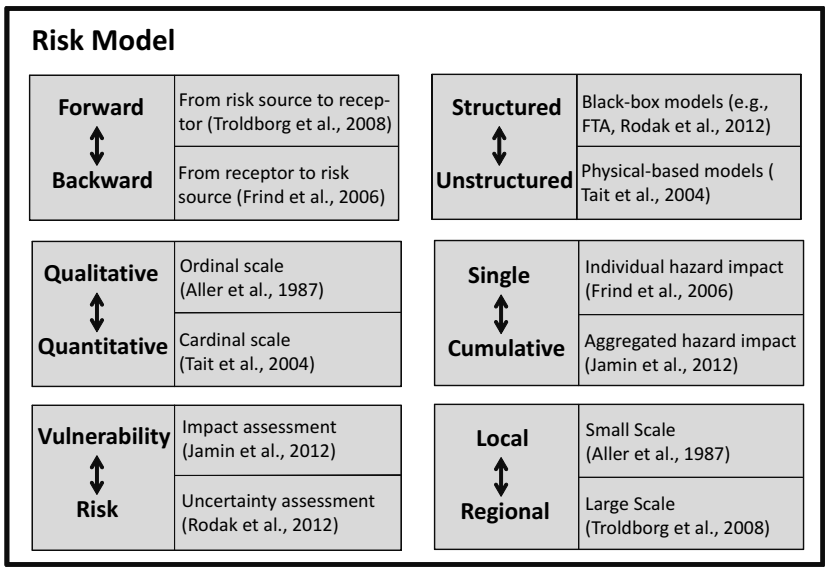


Figure 3.2.: Overview on risk types in risk assessment methods for contaminated sites.

Structured vs. Unstructured

According to Tartakovsky (2013), there exist two groups of probabilistic risk assessment methods. On the one hand, there are structured applications, such as risk models based on fault tree analysis (Bedford and Cooke, 2001). These fault tree analysis models have been predominantly used in the field of reliability engineering. On the other hand, Tartakovsky (2013) mentions unstructured models as a second group. Unstructured models are completely physically-based, avoiding a modularization and assignment of failure probabilities to single tree elements. These methods have been mainly used in the field of stochastic hydrogeology, resolving heterogeneity by geostatistical models (e.g., de Barros et al., 2012). Further explanation regarding physically-based unstructured and structured (graph-based) risk assessment approaches is given in Section 3.6.2.

Forward vs. Backward

Forward risk assessment calculates the level of risk from source to receptor, following the pathways between source and receptor in the forward direction (e.g., Tait et al., 2004), whereas backward or reverse risk assessment identifies isolines in the catchment, from which point a certain risk level at the well is exceeded (e.g., Frind et al., 2006) or how much mass has to be added in order to achieve an exceedance at a particular point in space (Cushman et al., 2001, see Fig. 10.1). These risk levels are commonly measured in concentration (e.g., maximum concentration level, MCL) or travel time (e.g., for wellhead protection delineation). The backward approach is centered around the receptor, and can be used together with the reverse transport formulations by Neupauer and Wilson (2001). More information regarding forward and reverse approaches is given by Cushman et al. (2001). In the later parts of this thesis, I will introduce two novel risk models. The first approach, called VIP, follows a backward risk assessment approach and the second one, called STORM, is a forward risk model.

Qualitative vs. Quantitative

Qualitative methods estimate risk on ordinal scales (e.g., risk matrices). A well-known field that uses qualitative methods is karstic aquifer vulnerability assessment (e.g., Aller et al., 1987; Goldscheider, 2005). Also, practitioners favor to adopt qualitative risk estimation techniques (e.g., Haakh et al., 2013), as these methods are intuitive and comparatively easy in application, thus leading to quick results. A detailed discussion on qualitative risk assessment is given in Section 3.5. Quantitative methods assess risk on cardinal scales and are further distinguished either in graph-based (e.g., Lindhe et al., 2009; Rodak and Silliman, 2012) or physically-based models (e.g., Tait et al., 2004). It is also possible that structured risk assessment models contain physically-based transport models (e.g., Nilsen and Aven, 2003). More details and limitations of quantitative methods is discussed in Section 3.6. Both concepts developed in this thesis (VIP and STORM) are quantitative and physically-based, yet copying the modular character from structural approaches.

Vulnerability vs. Risk

Many risk assessment models are rather vulnerability or exposure models (e.g., Chambon et al., 2011; Troldborg et al., 2008) than risk models, as they neglect natural variability and uncertainty. Exposure or vulnerability assessment can be called risk assessment, when the risk is conceptualized as being conditional to the fact that a contamination is present. Risk is always conditional on the assumptions that uncertain parameters are expressed as known variables. VIP is a probabilistic vulnerability method that considers uncertainty in physically-based model parameters, and thus could be perceived as risk estimation method, here defined as exposure risk assessment. STORM additionally considers the probabilities of hazard failures and uncertainty in mass load, providing clearly a risk assessment tool.

Single vs. Cumulative

The risk is calculated at the receptor by measuring the severity of impact in combination with likelihood. Often, the risk posed by hazards is assumed independently of each other, and thus hazards are treated as single and non-concurrent risk sources. This is common in qualitative risk assessment (e.g., Doessing Overheu et al., 2013; Haakh et al., 2013), but also in quantitative approaches (e.g., Einarson and Mackay, 2001; Frind et al., 2006). Troldborg et al. (2008) and Einarson and Mackay (2001) emphasized the importance to assess impact levels correctly by aggregating mass discharges at the well. Although assessing the cumulative impact at the well of one contaminant type, aggregation across hazard types is not possible, except using transfer functions such as DALY, micro-morts, groundwater quality index (Jamin et al., 2012) and so forth. While VIP follows the single-hazard mentality, STORM performs a quantitative aggregation across all hazards within a catchment, thus being a cumulative method.

Local vs. Regional

Risk is estimated at the receptor. Target receptors could be surface water, groundwater, indoor inhalation, plants, drinking water wells and so forth. Local risk assessment aims to estimate the contaminant impact on a receptor within its close vicinity. Nevertheless, drinking water is subject to all hazards across the catchment as the pumping well acts as an integral receptor. Jamin et al. (2012) refers to regional risk assessment, while calculating contaminant transport on the catchment-scale. Doessing Overheu et al. (2013) distinguishes between regional and catchment scale. Regional impact assessment might include several well catchments and is used in terms of a political boundary. In this thesis, both terms are used interchangeably. For a detailed comparison between local and catchment-scale processes, see Troldborg et al. (2008).

3.3.2. Risk Assessment of Contaminated Sites

According to Öberg and Bergbäck (2005), risk assessment of contaminated sites plays an important application in the field of environmental engineering. Ecological and human-health risk assessments, as part of the environmental risk assessment, are primarily influenced by the RAGS Manual, which has been published by the US EPA for Superfund sites (US EPA, 1989). RAGS is an acronym for *R*isk *A*ssessment *G*uidance for *S*uperfund. The introduction of risk assessment to Superfund sites has been motivated by the fact that clean-up of all contaminant sites would be too expensive and unfeasible, thus leading to remediation actions only of sites with the largest risk potential (Einarson and Mackay, 2001). Contaminant sites only pose a risk to the receptor, if a toxic substance reaches the target object. According to Einarson and Mackay (2001), this risk is measured in terms of concentration levels, derived from the mass discharge to the supply well. Cushman et al. (2001) provide an overview on the historical background of human-health risk assessment at contaminated sites.
Öberg and Bergbäck (2005) argue that probabilistic risk assessment helps to find a rational

and scientifically justifiable method to deal with uncertainty and variability in the context of toxicity assessment. The uncertain and varying input parameters are characterized by probability density distributions. Contaminated sites are challenged by natural variability. Hazards are spatially, temporal variable, and also the population that is exposed to contamination may change. Öberg and Bergbäck (2005) proposes two probabilistic risk assessment models in the context of contaminated sites, CalTox (CalTox, 1994) and CLEA (Jeffries, 2009). Both models are spreadsheet-based (EXCEL).
CalTox is sponsored by the State of California and calculates the emission of a chemical, the concentration of a chemical in *soil*, and the risk of an adverse health effect, following the dose model of the RAGS manual. It evaluates the distribution of a chemical among different environmental compartments (air, water, soil) and estimates human-health risk due to uptake of chemicals. CalTox relies for the toxicity assessment and risk characterization on measured contaminant concentrations in the root and vadose zone and in groundwater.
CLEA (=Contaminated Land and Exposure Assessment) is sponsored and developed by the British Environment Agency. It is also a site-specific risk assessment model that assesses the potential risk to human health, if a contamination is measured in the soil. The model can include further concentration measurements, such as concentration levels measured in soil air, ambient and indoor air or in fruits and vegetables (Jeffries, 2009).
These two example use measured or estimated concentration levels, performing a probabilistic toxicity assessment. Most models that provide concentration estimates by transport-based calculations do not yet incorporate probability distributions to account for uncertainty, especially not to resolve subsurface heterogeneity via geostatistical fields. Although these deterministic transport models claim to be risk models, they rather address exposure, impact or vulnerability than risk levels (see Fig. 3.2). This fact is also known and accepted (e.g., Troldborg, 2010). The risk aspect is then added by combining the exposure assessment with a toxicity assessment, leading to risk characterization, which is either probabilistic or deterministic. The individual steps to perform risk assessment of contaminated sites consist of four primary elements and is illustrated in Fig. 3.3.

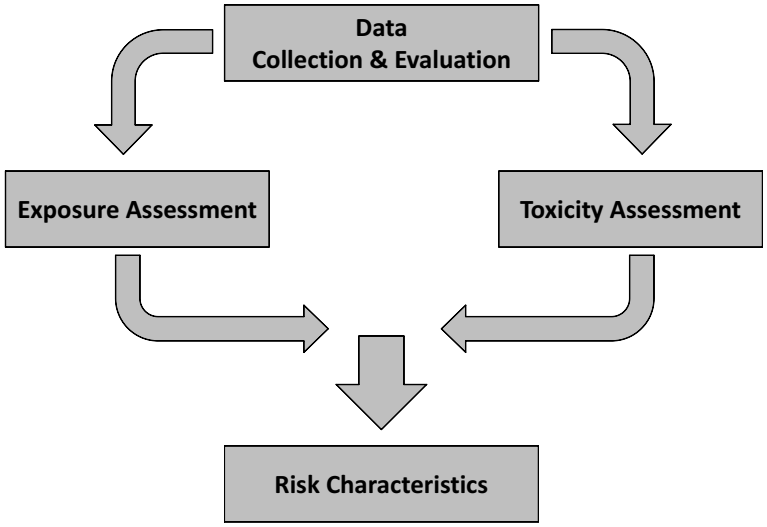


Figure 3.3.: Flow chart on risk assessment for contaminated sites.

Data Collection and Evaluation

As a first step, all available data about the contaminated site or area of investigation is collected. This includes the identification of substances that are of concern, evaluating the contaminant concentrations through sampling, examining the environmental setting and the characteristics of the contamination. These four elements provide data and information, e.g., on hydrogeological setting (aquifer properties) and contaminant properties, such as amount of mass being underground, factors controlling transport and fate and so forth. This information is necessary to perform transport simulation in the exposure assessment step, in order to estimate the impact of the contaminant plume at the receptor. Furthermore, potential exposure routes from source to receptor are sketched and inclosed within the study.
In the following, the attainment of the necessary information is split into three phases (Troldborg, 2010): desktop, field investigation and hazard identification phase. (1) The desktop phase includes the collection of all historical and actual information on resources and (manufactured) products, maps, photos and older studies about the site that help the risk manager to sketch the source and potential exposure pathways to the receptors. (2) The field investigation phase tries to justify the effort made in the desktop phase, proving that assumptions about the source and exposure routes are valid. This includes taking measurements such as core samples, concentration and head levels to confirm hydrogeological, hydrological maps and source assumptions. (3) The hazard identification phase determines all relevant information about relevant contaminants, reporting potential adverse health effects, biodegradability, physical and chemical properties and so forth.
These three phases result in a conceptual model for the risk assessment task that lead to the exposure assessment as a next step. The first element *data collection and evaluation* includes the similar tasks as the risk identification step of Fig. 3.2.

Exposure Assessment

The exposure assessment within human-health risk assessment identifies the population (receptor) that is possibly exposed to contamination. All potential exposure pathways are sketched, and for each exposure route quantitative estimates are developed through analytical or numerical transport models. The same is done for ecological risk assessment. The US EPA (1989) defines exposure pathway as the *course a compound takes from the source to the exposed individual (ecosystem).* Thus, an exposure pathway consists of a source of contamination, which has been released by any failure mechanism, the transport of the contaminant source to the receptor (see Fig. 2.1, modules 1 to 8) and a contact point with the contaminant of concern at the receptor through one or more exposure routes. As risk assessment of contaminated sites is commonly related to human-health risk, classical contact points (exposure routes) are (indoor) air inhalation, dermal contact (e.g., swimming, playing with sand) or any oral uptake through food or drinking water or children eating soil. A detailed list on potential exposure routes is given by Cushman et al. (2001). With this information for each contact point, quantitative estimates on magnitude, frequency and duration of an exposure are estimated. The magnitude of exposure is often measured as „chronic daily

intakes“ (CDIs).

$$CDI = \frac{c \times CR \times EF \times ED}{BW \times T_{AV}} \times \gamma_{ki}, \tag{3.1}$$

with c being the transport-based calculated average concentration over the study period, BW being body weight, ED the exposure duration and T_{AV} is the time period over which exposure is averaged. CR stands for the contact rate and γ_{ki} is a transfer function between groundwater and the contaminant concentration. The US EPA developed standard default values for typical exposure scenarios to estimate intake rates. There exist several other uptake formulations (e.g., Sander and Öberg, 2006) than the one proposed by the US EPA in Eq. (3.1). Nevertheless, Eq. (3.1) is the most prominent and widely used formulation.

Toxicity Assessment

Toxicity assessment characterizes the relationship between the exposure magnitude and adverse health or environmental effects. An exposure-response-relationship serves to identify the beginning of adverse toxic impact at the receptor and thus provides information on no-observed-adverse-effect-levels (NOAELs). The adverse health effects are distinguished between carcinogenic and non-carcinogenic effects. National standards and toxicological databases (e.g., Integrated Risk Information System, IRIS) provide information on toxicity properties. In these databases, information on „chronic reference dose“ (RfD) and „chronic reference concentration“ (RfC) values can be found, in order to evaluate the toxicity of non-carcinogenic substances through dermal contact or ingestion (RfD) and inhalation (RfC). RfD and RfC can be considered as the maximum daily uptake level with no effect (Cushman et al., 2001). Other parameters to value the maximum uptake are provided by acceptable or tolerable daily intake levels (ADI/TDI). To consider the exposure-response relationship, regarding cancer health risk, the US EPA introduced a contaminant-specific slope factor SF to value the relative toxicity of carcinogenic compounds (see Eq. 3.2).

Risk Characterization

As a final step, all information gathered in the previous steps are combined and the individual (qualitative or quantitative) risk values are developed. Non-carcinogenic risk estimates are obtained by the ratio between CDI and RfD/RfC, yielding a hazard index, which is aggregated across exposure pathway and substances. This hazard index proposed by the US EPA (1989) differs from the hazard indices developed within the COST program (Zwahlen, 2004). These two factors only share names, as the latter is a qualitative and the former a quantitative measure. For carcinogenic risk estimates, the individual excess cancer lifetime risk ($IELCR$) is calculated according to US EPA (1989):

$$IELCR = 1 - \exp\left(-CDI \times SF\right), \tag{3.2}$$

estimating the increased probability of cancer risk with CDI being the chronic daily intake rate (see Eq. 3.1). The slope factor SF is available via toxicity databases (e.g., IRIS). According to Rodak and Silliman (2012), the total cancer risk can be estimated as a sum of the risk

from each individual contaminant across all potential pathways. Acceptable risk levels are commonly given by national standards (see Section 3.4).

3.4. Water Safety Concepts - Standards and Guidelines

This section aims to review guidelines and standards regarding risk assessment and management that have the goal to guarantee safe drinking water supply. Safe drinking water supply requires clean aquifers and water sources, reliable treatment and safe distribution, which coincides with the three barriers of the multi-barrier approach (e.g., O'Connor, 2002). The WHO recommends implementing water safety plans into national regulation in order to ensure clean and fresh drinking water from catchment to tap (Davison et al., 2005). Drinking water is the most important consumable good and is by definition not exchangeable (DIN, 2000). Therefore, all drinking water safety concepts follow the call of the Bonn Charter to *ensure good safe drinking water that has the trust of consumers* (IWA et al., 2004). Good drinking water is characterized to be colorless, odorless, fresh and aesthetically appealing. These quality requirements are written into many national standards (e.g., DIN, 2000).
A classical risk control strategy for groundwater abstraction is to delineate time-related wellhead protection areas in a deterministic fashion. These areas are regulated in many countries by national standards or guidelines (e.g., DVGW, 2006; US EPA, 1993). Nevertheless, stakeholders need additional information to judge the degree of vulnerability of the source water to hazards that are located within the catchment, whether they are within the protection areas or outside. Hence, risk assessment and management initiates a paradigm shift from fixed capture zone areas and protection zones that restrict certain activities per se to actively check the risk level and thus the acceptance of hazards within the entire catchment. In order to support this paradigm shift, a more holistic approach is needed to ensure well safety and to provide more information to stakeholders, such as risk assessment and management.
Therefore, Section 3.4.1 provides an overview on the most fundamental guideline for risk assessment in drinking water supply, the water safety plan. Section 3.4.2 introduces risk standards required by the US EPA and Section 3.4.3 by the German Water Association. Both serve here as two illustrative examples how water safety plans are implemented on national level. Finally, a short review on wellhead delineation is provided, as this is relevant for the later parts of the thesis with a special focus on German regulations (see Section 3.4.4).

3.4.1. Water Safety Plans

The WHO (2004) states within their third drinking water guideline, that „drinking-water quality is an issue of concern for human health in developing and developed countries world-wide". The WHO proposes an integral and preventive risk-based framework from catchment to tap. Fig. 3.4 shows the process-oriented management scheme prepared by the WHO, called water safety plan. Risk models that follow the water safety plan comply with regulatory or national standards, e.g., health-based targets (intake rates). The water safety plan is based on the Hazard Analysis and Critical Control Point (HACCP) concept. HACCP was first used within the food industry and gives a clear framework to systematically assess

all hazards over all parts (=system analysis) and check through periodic monitoring, if critical points (=health-based targets) are exceeded or not. The HACCP aims to determine as early as possible a negative alteration of physical, chemical, technical or hygienic parameters within the complete chain of investigation. Thus, the water safety plan challenges water managers to know, control and monitor all possible hazards within the complete supply system from source to tap.
The water safety plan framework complies with the multi-barrier concept. Therefore, hazards across the three barriers, water resources, the water treatment plant and the distribution network, are identified (=system assessment). The monitoring of control measures (e.g., data collection related to the health-based targets) help to guarantee safe drinking water and a proper system operation. Management plans based on qualitative risk ranking should be used to mitigate risk, whether it does not comply with the health-based targets. The water safety plan is embedded into independent and periodic surveillance of the system. The desired result is that public health is not endangered, and the degree of goal fulfillment is communicated and published to the population. This strengthens the confidence of customers into the supplied drinking water as required by the Bonn Charter (IWA et al., 2004).

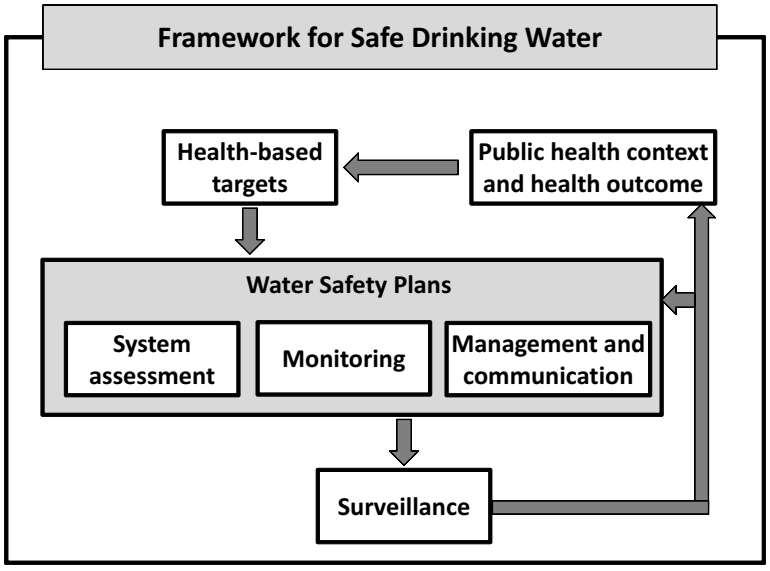


Figure 3.4.: Methodology of risk-based and process-oriented management according to Bonn Charter (IWA et al., 2004).

3.4.2. Risk Assessment and Management by US EPA Standards

In the United States of America the protection of drinking water is regulated by the safe drinking water act amendment (Tiemann, 2010) with the overarching goal of human-health protection. More than 70 regulations have been promulgated by the US EPA only between 1976 and 2002. These standards define Maximum Concentration Levels (MCL) for chemical and microbial contamination to avoid adverse health effects by consumption and exposure

of drinking water. The individual states are free to choose their own health standards, but have to be at least as strong as the ones proposed by the US EPA (2009).
Risk assessment in the field of environmental engineering started in the US with the RAGS Manual (US EPA, 1989). Risk assessment has been performed for Superfund sites. Later, several State regulatory agencies required risk assessment also for non-Superfund sites. In the 90'ies, the American Society for Testing and Materials (ASTM) published the Standard Guide for Risk-Based Corrective Action Applied at Petroleum Release Sites (ASTM RBCA Standard) that is based on the RAGS Manual. The goal of this standard was to remediate only those contaminated sites, which pose the greatest risk to the groundwater. As early as 2000, the US EPA considered the assessment of health effects due to chemical mixtures (US EPA, 2000). In 2007, the US EPA published a standard regarding cumulative health risk assessment of multiple chemicals, exposures and effects updating the 2003 standard. The evaluation of the cumulative impact involves all sources, contaminants and their interaction, media, exposure routes, vulnerability of individuals, and is measured via the hazard index. These multi-chemical exposures are ubiquitous, being present in soil, ambient air or drinking water.
Another framework for analyzing and managing risk, especially in the light of terrorist attacks, is RAMCAP: Risk Analysis and Management for Critical Asset Protection (ASME, 2006). The goal of this guideline and risk management methodology is to protect national infrastructures at risk, especially in the course of 9/11. A seven-step process provides the fundamental basis to understand and manage risk from terrorism. The American Water Works Association (AWWA) has adopted the RAMCAP framework to assess the risk of asset and system failure. The American Water Works Association emphasizes the need to consider the worst reasonable case scenarios in risk management, which makes risk assessment either risky or expensive.

3.4.3. German Risk Standards

In Germany, there exists a long tradition of water protection. Thus, many national standards to protect drinking water resources exist. Most regulations are promulgated by the German Association of Gas and Water (DVGW), such as the standards for technical safety management (DVGW, 2000) and wellhead delineation (DVGW, 2006). Maximum concentration levels are set by the federal ministry of health and are promulgated in the drinking water directive (TrinkwV, 2001). Regarding risk assessment and management, only two standards exist. These two standards address risk-based and process-oriented management of hazards during normal operations (DVGW standard W1001, 2009) and during critical situations (DVGW standard W1002, 2009), which are defined as crisis. In this context, a crisis is defined as situations where the water supply company alone cannot cope with the currently faced situation. The procedure in a crisis is strongly based on a military background, implying that one crisis leader has all the executive and legislative power.
In the remainder of this thesis, I only consider the *DVGW standard W1001* (2009) for risk management during normal operation. Supply safety under normal operation is guaranteed, if the water company is running properly. This is fulfilled by keeping health-related issues (e.g., TrinkwV, 2001), esthetic aspects (DIN, 2000) and technical requirements (e.g., work-

ing standards promulgated by the DVGW), such as sufficient pipe pressure and amount of water. These requirements are only achieved, if the involved stakeholders, government, water companies and consumers work jointly together. In order to guarantee supply safety, a methodology is proposed to deal with hazards and compare their relevance related to their potential to alter the drinking water quality. All parts of the supply chain (resource protection, water production, treatment, storage and distribution to the consumer) are considered to assess and evaluate hazards. Here, hazard is defined as a potential biological, chemical, physical or radiological detriment within the supply system. Thus, the guideline complements the required end-product control regulations by the technical safety management standard DVGW (2000) and is in line with the water safety plans of the WHO (Davison et al., 2005).

Fig. 3.5 shows the risk-based and process-oriented DVGW method for risk management in

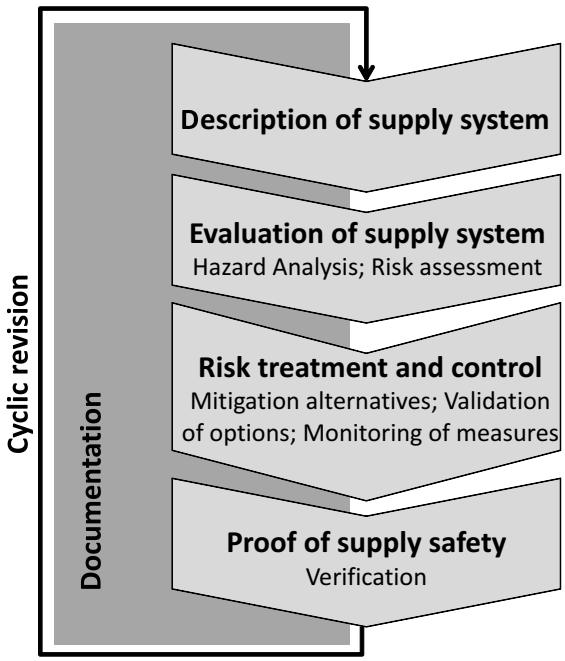


Figure 3.5.: Methodology of risk-based and process-oriented management according to DVGW (2009).

order to systematically identify, assess, evaluate and control risk sources. In the first step, the guideline suggests to describe the supply system with all potential hazards that have the potential to lead to detrimental water quality and quantity. In the next step, qualitative or semi-quantitative risk assessment approaches are proposed to assess the damage potential of these hazards, the probability of hazard failure, and to estimate the resulting risk. To illustrate the results and rank hazards, the international standard approach using a 3×3 or 5×5 risk matrix is suggested. The third step evaluates the suitability of risk mitigation alternatives, fulfilling the technical regulations published by the DVGW. Risk treatment methods that are not included in any of the technical guidelines of the DVGW are validated concerning their effectiveness, e.g., monitoring parameters before and after installation. The monitoring of water parameters and the frequency of sampling is essential to verify that risk

of known hazards is controlled. Finally, all observations, findings, data measured and methods used throughout the risk analysis need to be documented and periodically repeated to constantly improve supply safety. The presented outline of the DVGW W1001 is in line with the ISO standard.

3.4.4. Wellhead Protection by Delineation

The delineation of wellhead protection zones is commonly performed in many countries world-wide. The common goal is to prevent drinking water from being contaminated by bacteria or other contaminants such as chlorinated solvent, pesticides and petroleum products. One example are the US EPA regulations (e.g., US EPA, 1993) that are, in turn, in accordance with the safe drinking water act amendment (Tiemann, 2010). The safe drinking water act requires from the individual States to develop source protection programs and in specific focus on two major hazards: microbial or chemical/radiological contaminants. The US EPA regulation suggests the delineation of time-related capture zones in line with the multi-barrier approach. The Wakerton, Ontario, Canada, tragedy in 2000 with several dead people and thousands of people getting ill from bacteriologically contaminated drinking water was a wake-up call across Canada to enforce regulations and improve province-wide time-related capture zones, following the multi-barrier concept (O'Connor, 2002). Here, I will focus on the German national standard DVGW (2006) as the relevant standard in this thesis for wellhead delineation.
The national standard DVGW (2006) splits the drinking water catchment into three zones. The inner zone closest to the production wells is called zone 1. The restriction level within this area is very high, abandoning almost all activities besides mowing the grass. In general, the land of zone 1 is owned by the supply company. The second zone is defined by an intrinsic 50-day advective travel time to the well, considering peak arrival time. The standard provides several methods to assess the travel-time-based delineation. Intensive agriculture and industrial activities are not allowed per se within these areas, to avoid microbial arrival at the production well. The third zone is split into two categories with only minor restrictions to land-use activities. The regulation of activities within wellhead protection zones is regulated alone by local authorities and, in more severe cases, in conjunction with the provincial (regional) ministry. Also, the delineation of wellhead protection zones is enforced by the ministry and not by the water company. It is also the ministry that decides, if a well capture zone deserves protection or not.
Furthermore, the standard (DVGW, 2006) provides a list of potential hazards split into seven categories, such as traffic and transport, residential area, agriculture, industry, waste water and others. In accordance to these categories, Haakh et al. (2013) developed a semi-quantitative risk assessment approach in collaboration with stakeholders of the DVGW research project „Risk assessment in drinking water catchments". This approach is further discussed in Section 3.5.

3.5. Qualitative Risk Assessment and Limitations

Qualitative risk methods are widely used across disciplines, such as civil defense, quality engineering, and project management just to name a few. Most likely, this is due to the simple and intuitive use of qualitative methods, even for non-trained risk managers. It is easy to implement expert knowledge and their results support risk management as a first step to better understand the system, to possibly pre-screen hazards and to identify a bundle of non-acceptable risk sources.
This section is motivated to guide early-stage risk managers to properly use qualitative risk assessment methods, while being aware of the limited predictive power of qualitative tools and while understanding several points of criticism. Actually, there is no need to distinguish between qualitative and semi-quantitative methods, as semi-quantitative methods often simply describe risk in finer resolved categories (see discussion below, Fig. 3.7).
In the following, I provide a a short introduction to qualitative risk assessment studies (Section 3.5.1) and risk matrices, as being most commonly the final stel in qualitative risk assessment (Section 3.5.2). Semi-quantitative vulnerability-based risk estimation tools are discussed next (Section 3.5.3), including risk intensity calculations according to Zwahlen (2004) and a newly published approach in the field of drinking water supply safety according to Haakh et al. (2013). Finally, recognized limitations using qualitative and semi-quantitative risk assessment approaches are summarized (Section 3.5.4).

3.5.1. Short Overview on Qualitative Methods

Aller et al. (1987) were one of the first authors to predict and map the vulnerability of aquifers to hazardous surface events within a catchment, using qualitative measures. Their goal was to map and prioritize locations where the aquifer is more vulnerable to hazardous surface events than at others. More studies followed, especially in the light of semi-quantitative vulnerability assessment for karstic aquifer systems, such as the European Cost Action 620 program (Zwahlen, 2004). In that program, risk intensity maps were estimated by combining vulnerability indices and pre-defined hazard indices of risk sources. Another European-funded project TECHNEAU („Technology Enabled Universal Access to Safe Water“) introduced several methods for risk estimation in the field of water supply safety. One of them is the so-called „coarse risk analysis“. It identifies the most severe threats within the system and helps setting the right priorities when implementing risk mitigation options (Hokstad et al., 2009). The coarse risk analysis identifies the severity and the probability of an event by categorization (e.g., 1 =very low to 6 =very high). The results are then presented via a risk matrix. All these index-based vulnerability methods and other qualitative risk approaches, such as risk matrices, work on ordinal scales (e.g., risk is low, medium, high; hazard points).
It has been argued that such qualitative methods are subjective in their results (e.g., Cox, 2008; Neukum and Azzam, 2009) and thus of limited use to support stakeholders in taking robust management decisions. A much larger problem to my opinion is that ordinal scales offer no mathematical operators, such as summation or multiplication, to aggregate individual risk values to total risk levels or for evaluating the risk product (damage times

likelihood). This limits qualitative risk estimation when trying to prioritize hazards in a consecutive order, or when trying to support risk-informed management decisions. A detailed discussion on cumulative risk aspects is given in Section 3.6.5.

3.5.2. Qualitative Risk Approach - Risk Matrix

A risk matrix is a table that consists of consequence and probability categories on each axis (rows and columns). Consequences are deemed to capture the severity, impact or damage, and probability is exchangeable with likelihood, uncertainty or frequency. Each cell in the table represents a combination of probability and consequence, leading to a risk, urgency, priority or management action level. An example of a simple 5×5 risk matrix is shown in Fig. 3.6.

Risk matrices are widely accepted and used across many disciplines (e.g., insurance, bank-

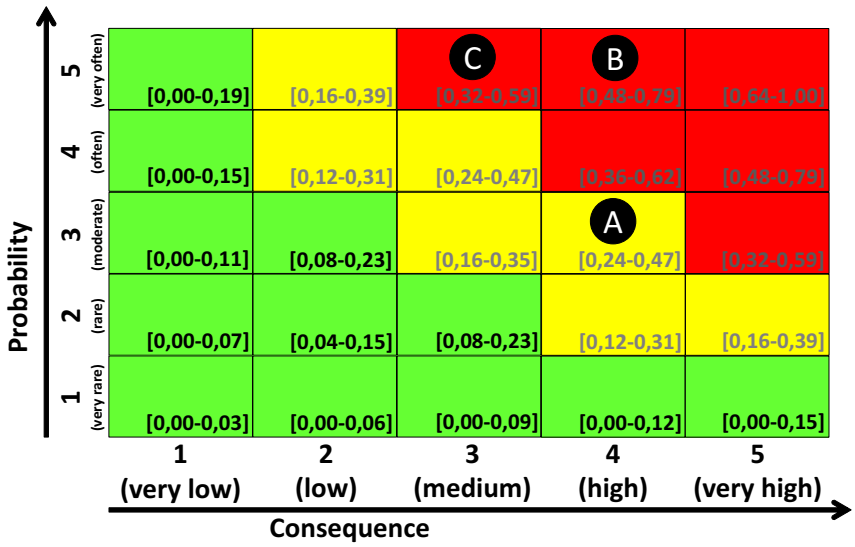


Figure 3.6.: A 5×5 risk matrix, showing three risk categories (acceptable (green), as-low-as-reasonable-practicable (yellow), not acceptable (red)) and three categorized hazards A to C. The squared brackets show the range of risk value per cell, when working on cardinal scale.

ing, transport). They provide a clear and understandable framework to perform risk assessment and management without intensive training in the risk sector. As a consequence, several parties can be quickly incorporated into the risk assessment process, leading to joint prioritization of management actions, hazards and mitigation options among others.
This simple accessibility to risk assessment and management can give rise to several dangerous situations, as risk matrices have several limitations. In the following, relevant questions are discussed that should help risk managers to apply risk matrices with care.

- *How to choose risk categories?*
  The three colors in Fig. 3.6 demonstrate the rating of each column-row-pair. The choice

of risk category that a column-row-pair receives is stakeholder-dependent and thus subjective to some extent. The US EPA (1989) proposes three categories, where risk is acceptable, as low as reasonable practicable (ALARP) or unacceptable. Nevertheless, there still exists the problem on how to distribute the risk categories within the cells of the matrix. It is therefore common to assume at first the definition by Kaplan and Garrick (1981), taking the product of the pair and thus categorize risk levels according to the product score. This leads to hyperbolic risk isolines and thus to *symmetric risk matrices*, if the range and the class width on the consequence and probability axis are constant. The assumption to classify risk according to the cardinal product score is obsolete, as risk is estimated on ordinal and not on cardinal scales within risk matrices. To my opinion, it is very likely, that the class interval widths between severity or probability levels are not constant. This is especially true for ordinal ranking. For example, Doessing Overheu et al. (2013) propose for mass discharge rates to the well a log-scale categorization to assign severity points. Therefore, *asymmetric risk matrices* are possible and risk is merely „some combination" and not the strict product of probability and consequences. Hazards categorized with very low probability and very high damage might get a different risk level than hazards that occur very often, but with very low damage potential. Hazards with higher risk category receive higher priority for risk mitigation and treatment, leading to a hazard ranking of $B(red) = C(red) > A(yellow)$ (see Fig. 3.6).

- *But how many risk categories are necessary?*
  As stated above, the US EPA (1989) claims that three risk categories are sufficient for risk assessment. The probability to distinguish two hazards, by falling in two separate risk classes given the 5×5-matrix in Fig. 3.6, is $\approx 63\%(= 1 - 1/25 \times [12 \times 12/25 + 7 \times 7/25 + 6 \times 6/25])$. Thus, the probability that two events are within one risk category is for this example $\approx 37\%$. This probability changes by choosing a different number of risk categories. It is obvious that, with increasing category resolution, the probability increases to be in separate risk categories, leading to possibly as many risk categories as column-row-pairs are available. From a stakeholder perspective, the more risk categories are available the better and more precise the risk ranking could be performed. Nevertheless, this is contrary to the argumentation of US EPA (1989) to rank hazards according to three risk categories. Also, the qualitative steps that lead to classification of any given hazard in any of the matrix cells are not overly accurate. Cox (2008) supports this statement and even enforces the ranking issue by claiming that risk ranking is only possible between the high and low risk category. Following this argumentation, the probability in the 5×5 example that two risk sources fall in the high and low risk category, is reduced to $\approx 12\%(= 6/25 \times 12/25)$. Thus, in $\approx 88\%$ of all cases, a clear ranking between two events is not possible. Increasing the amount of risk categories and assuming that hazards fall in two separate and significant risk categories will even reduce the ranking probability. An explanation on significant ranking is given next.

- *Is consecutive hazard ranking possible?*
  Ranking of hazards that fall in different risk categories is available within ordinal scales. A system with aligned risk categories shows that higher ranked risk levels are superior to the latter. Thus, by definition on ordinal scales, hazards assigned to

the same risk category cannot be distinguished between each other and thus ranked. Increasing the amount of risk categories tacitly leads to the assumption of cardinality, i.e., implying that intervals between levels (probability and consequence axis) are known and consistent. Let us assume that interval widths are known and thus risk can be mathematically expressed by the product of consequences and probability. For each column-row-pair, a minimum and maximum risk level is calculated through multiplication of the smallest and largest class value. The normalized range per cell is shown in Fig. 3.6 within the squared brackets. As stated above, the higher risk category is superior to the lower category, assuming that the largest risk value of the lower risk category is smaller than the smallest value of the higher risk category. This is true for all categories that have a positive inclination (see hazard A and B, lower left to upper right in Fig. 3.6) and false for negative ones (see hazard A and C. In this case, hazard A might has a higher risk value than the unacceptable hazard C). This is due to the hyperbolic character of risk levels. Therefore, Cox (2008) argues ranking of two hazards independent of positive and negative inclination is only possible, if the connected straight line between these two hazards passes through an intermediate risk level. There is no risk value within the acceptable column-row pairs that is larger than within the unacceptable classes. In Fig. 3.6 this would mean, that hazards can only be ranked according to acceptable and unacceptable risk and that only a pool of hazards within the unacceptable risk level is identified instead of a consecutive ranking of the top „five" events. An introduction on how to construct risk matrices in order to reduce the problem of false hazard ranking is presented by Cox (2008).

- *How valuable are risk matrices for decision making?*
  Risk matrices are often used for risk-informed management decisions, such as identifying the most severe hazards that pose an unacceptable risk. Let us assume three risk sources (e.g., A($C = 0.88; P = 0.51$), B($C = 0.83; P = 0.91$), C($C = 0.42; P = 0.82$)), as shown in Fig. 3.6, with P representing the probability scale and C the consequence axis. In this example, hazards B (red) and C (red) are the ones with the highest concerns that need special treatment by the risk manager. The validity of decision making with risk matrices is tested next. Let us assume that a different risk manager assigns to the three previous hazards a consistently lowered consequence level (e.g., $C = C - 0.15$). One would expect that the ranking of hazards is robust and would not change compared to the previous ranking. In fact, however the order changes. B (red) is still identified as the most severe hazard, but hazard C is now categorized to the ALARP risk category. With A and C now being in the same category, the choice between A and C is now more difficult. Thinking on cardinal scales as often unintentionally done, hazard A is identified as the more severe hazard than C, as the class product is $R_A = 0.12$ and thus larger than $R_C = 0.10$. This suggests a change in order compared to before. The same is true, if any parameter P or C is constantly changed, e.g., through a uniform effective risk mitigation measure. This problem is known as the principle of translation invariance (see further explanation to this effect by Cox (2008)). To my opinion this effect shows the non-cardinal behavior of risk matrices and that risk managers need to be critical, when using cardinal operations (e.g., multiplication and summation) in qualitative risk assessment approaches.

- *How to consider cumulative risk aspects?*
  Aggregation of risk is often followed by summation or building the product of risk levels (e.g., Jamin et al., 2012). As previously discussed, cardinal operations such as summation need a critical view, when using qualitative risk assessment. Not only risk matrices suffer that problem of translation invariance but also semi-quantitative methods (see Fig. 3.7). In order to consider cumulative aspects in a risk matrix, one has to accept the ordinal behavior of risk matrices, and then assign aggregated severity categories according to pre-defined combination rules. These rules assist risk managers and stakeholders in how to categorize two low risk values or one low and one high risk value. However, the ranking between these cumulative situations will almost surely be even less rigorous as the individual ranking.

A good reading on the limited use of risk matrices, such as decision making based on a ranking, is Cox (2008). The choice of risk categories is common knowledge. All ideas regarding cardinal and ordinal scales, such as cumulative risk aspects, and providing a critical review on risk matrices for the use in the environmental (hydrology) engineering community are new and so far to my knowledge first published in this context. There has been a paper on semi-quantitative vulnerability approaches by Neukum and Azzam (2009), moving from qualitative to quantitative approaches in order to make the comparison between karstic vulnerability models more transparent.

3.5.3. Semi-quantitative Risk Methods

Semi-quantitative risk assessment methods are also denoted as being DRASTIC-based (e.g., Doessing Overheu et al., 2013). As introduced before, DRASTIC is a methodology to assess an index-based vulnerability value for the groundwater body (Aller et al., 1987). It was the first of its kind and assigns for each factor a point value (severity of impact or susceptibility) within a given range (e.g., $1 - 10$, with 10 being most severe or susceptible). The six key factors of DRASTIC are **D**epth to water, net **R**echarge, **A**quifer media, **S**oil media, **T**opography, **I**mpact of vadose zone and hydraulic **C**onductivity of the aquifer. Each factor is independently weighted by a weight ω_i, leading to the following equation for the pollution potential π:

$$\pi = D \cdot \omega_D + R \cdot \omega_R + A \cdot \omega_A + S \cdot \omega_S + T \cdot \omega_T + I \cdot \omega_I + C \cdot \omega_C. \tag{3.3}$$

Also qualitative risk estimation methods adopted this intuitive concept of knowledge factorization by assigning points to individual factors (e.g., contaminant toxicity, amount of mass being stored) in order to estimate impact and the adjacent risk. Fig. 3.7 shows a schematic concept for factorization and thus the next step to semi-quantitative approaches. Risk is estimated as the combination of damage and probability, here represented by a product of both (Kaplan and Garrick, 1981). Semi-quantitative approaches use sub-factors to account for more detailed and available background information, leading to vulnerability values and intermediate damage levels. Coding this background knowledge is motivated by the idea to make risk assessment more transparent and objective. The combination is based on multiplication and summation, here building the product of sub-factors and the sum of basic

knowledge factors (e.g., age, toxicity). In the following, two semi-quantitative approaches are illustrated (Haakh et al., 2013; Zwahlen, 2004).

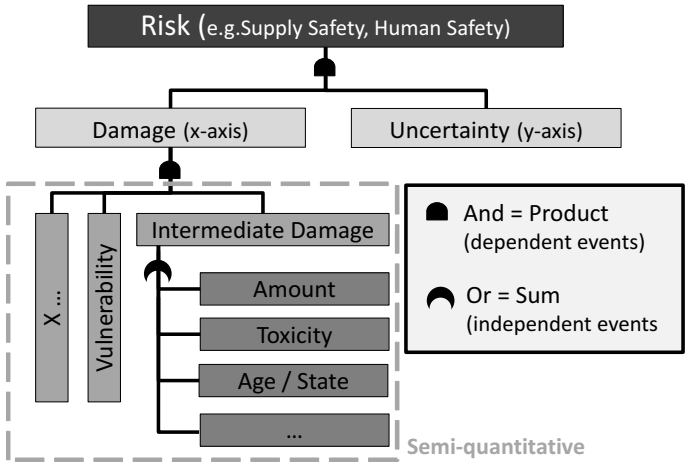


Figure 3.7.: Conceptual illustration of parameter combination in qualitative and semi-quantitative risk assessment, adopting a fault-tree flow diagram.

Risk Intensity by COST (Zwahlen, 2004)

The European COST Action program 620 (Zwahlen, 2004) proposed to develop risk intensity maps based on the spatially distributed vulnerability values. These risk intensity values RI depend on local-scale processes, neglecting the transport to the groundwater well. The COST risk intensity is thus a map showing the risk potential of groundwater contamination:

$$RI = \frac{1}{HI} \cdot \pi, \tag{3.4}$$

with HI being the hazard index, describing the harmfulness of each hazard and π being the vulnerability value (see Eq. 3.3) at the hazard location. The risk intensity index shown here provides smaller risk intensity numbers for higher risk levels. This is due to the fact, that risk sources with higher pollution potential receive a larger hazard index and more vulnerable locations a smaller value. In order to ensure non-ambiguous risk values, the reciprocal hazard index was used (Zwahlen, 2004).
This hazard index depends on the harmfulness of a risk source in general (H), the site-specific likelihood of groundwater contamination (R_f) and a ranking factor (Q_n) that expresses how large and how toxic the stored substances are:

$$HI = H \cdot Q_n \cdot R_f. \tag{3.5}$$

Hazard values H range between 1 and 100 and were proposed by an expert panel of the COST program. Within look-up tables it is possible to quickly assign a hazard point value to the risk source (e.g., nuclear waste site $= 100$, graveyard $= 25$). In order to distinguish between the harmfulness of individual hazards within one hazard category H, the weight

Q_n has been introduced and ranges from 0.8 to 1.2 to indicate the amount of toxic substances compared to the general average. The likelihood R_f represents the information basis about the hazardous site, which can be interpreted as the inherent uncertainty associated with the risk estimation due to lack of information.

Well Production Risk by Haakh et al. (2013)

Semi-quantitative risk assessment is widely accepted and especially favored by practitioners, which is shown by the more recent study of Haakh et al. (2013). The goal of this method is to estimate the risk of source water contamination, accounting for catchment-scale processes through modularizing the risk model according to the source-pathway-receptor concept (top level of Fig. 3.7). The risk R of drinking water contamination at the production well is given by Eq. (3.6).

$$R = S_{Br} \cdot E, \tag{3.6}$$

with S_{Br} being the damage at the well and E the probability of hazard failure. The damage at the well is estimated as

$$S_{Br} = S_O \cdot \pi \cdot t_c, \tag{3.7}$$

with S_O being the damage at the local scale (=spatial location of the hazard), π the vulnerability value of the aquifer according to Goldscheider (2005) and t_c the travel time a contaminant needs from the point of contamination to the well, tacitly assuming that longer travel times lead to more excessive contaminant dilution due to dispersion or to a larger degree of degradation. The damage at the local scale is estimated as the sum of hazards (various chemical compounds) at this site, accounting for toxicity Q and amount of mass m being available for accidental spills:

$$S_O = \sum_{i=1}^{n} m_i \cdot \omega_i + \sum_{j=1}^{n} Q_j \cdot \omega_j. \tag{3.8}$$

The parameter ω is a weighting factor, following the concept of DRASTIC. All factors are categorized between *very low* ($= 1$) and *very high* ($= 5$), in accordance to a 5×5-matrix (see Fig. 3.6). A conceptual flow diagram of this method is shown in Fig. 3.7, with $X = t_c$ being the travel time of a contaminant to the well t_c and S_O the intermediate damage at the local scale. The evaluation of risk levels between hazards is strictly limited to one hazard type sector (e.g., agriculture, industry, waste water, and so forth). There exist seven sectors, which are defined by DVGW (2006).

In the following, I will highlight some critical remarks of mine that each risk manager using semi-quantitative risk assessment methods should be aware of. The first one is to pay attention when using mathematical operations on ordinal scales that are only valid for calculations on the cardinal scale. This issue has been fully discussed in Section 3.5.2. Concerning the issues discussed below, not every semi-quantitative or qualitative risk assessment tool is affected by the proposed critical views. The following listing should only be seen as a check list to interact in depth with the chosen model:

- *Are all key factors relevant and appropriate for the risk assessment purpose?*
  The model choice to estimate risk and thus fulfilling pre-defined risk objectives is crucial. Not only the choice of factors, but also the validity of individual factors to predict the actual risk objective has to be checked. As an example, the travel time t_c in Eq. (3.7) assumes less impact (smaller impact value on the consequence axis, see Fig. 3.6) due to dilution or degradation with longer travel times. Travel time is an inappropriate factor (or at least an incomplete description) as two contaminants with different degradation potential would receive the same point value, although they would arrive at the well with different concentration levels. Also, this model implies that peak concentrations are relevant. Stronger plume dispersion after longer travel times may actually lead to longer time windows where the well is subject to concentration values larger than a critical threshold. Thus, one has to pose the question how the chosen key factors actually influence the risk objective.

- *Is zero-risk achievable?*
  Fig. 3.7 shows that sub-factors are aggregated by summation to receive an intermediate damage. This is the case in DRASTIC (see Eq. 3.3) and for the damage estimation at local scale (see Eq. 3.8). Let us assume, in the following that the mass at one site is very high, but the toxicity is negligible. Nevertheless, according to summation there still exists a damage potential at local scale. In this situation, it would help to condition factors on each other, e.g., by building a product (see Eq. 3.5).

- *Are hazards impacting all at the same time?*
  Based on probability theory, summation of probabilities indicates the independent failure of events (at least approximately, for very low probabilities), leading to a higher likelihood of impact. For example, fault tree analysis uses „OR-Gates“ to indicate these independent hazard failures. Summation of impact, however assumes that all events happen at the same time, or more precisely occur at the same time at the receptor or aggregation level. The same is true about the summation of factors in qualitative risk assessment. The summed mass of n-compounds in Eq. (3.8) tacitly assumes that all mass is released at the same time. Of course, there are good arguments for summation as, relatively to other sites, the larger mass storage causes a larger risk potential and needs a higher awareness than at other sites given a constant summed toxicity factor. The risk of overestimating the joint impact is nevertheless given in many risk assessment studies. Jamin et al. (2012) summed up groundwater quality indices across various contamination sites. US EPA (2007) proposes to sum up hazard indices across contaminants, assuming the perfect simultaneous occurrence of all contaminants at the receptor. When using summation, risk is misinterpreted and overestimated, as arrival times of individual compounds at the well will most likely differ in practice. Furthermore, the toxicity effect at the receptor is also assumed to be cumulative and occurring at the same time. There are two critical views. First, human health-effects due to toxic impact might not behave linearly, such that summation of toxic health effects is acceptable. US EPA (2007) shows that the exposure-dose response relationships is non-linear, especially in case of chemical mixtures. Second, hazards that are more

toxic fulfill in general higher safety requirements, such that the probability of failure is reduced. This leads to less frequent and thus possibly non-overlapping arrival times at the well. A detailed discussion on failure rates and their effect on cumulative concentration breakthrough curves is discussed in Chapter 7.

- *Are points across different factors comparable?*
  Points are assigned to the factors shown in Fig. 3.7, indicating the severity or impact of that factor. The points across factors are summed up, leading to the overall risk. DRASTIC evaluates their factors mainly on a 10-point-scale (Aller et al., 1987). Are 10-points assigned to a gravel soil layer as severe as a $0-2\%$ slope in topography? Haakh et al. (2013) evaluate the travel time on a 5-point-scale and COST uses a 100-point-scale for their hazard weight (Zwahlen, 2004). The hazard weight H for surface fuel storage tanks in COST equals $H_{fuel} = 50$ and for municipal sanitary landfills $H_{landfill} = 50$. Thus, the cumulative hazard severity of the fuel tank and the landfill equals the impact of nuclear waste sites $H_{nuclear} = 100$. Even more critical, four cemeteries $H_{cemetery} = 25$ in the catchment are as dangerous as the nuclear waste site. Therefore, aggregation of ordinal factors leads to risk levels that are misleading, due to the unknown interval scale.

- *Are points within one factor comparable?*
  Due to the above reason, Haakh et al. (2013) propose to compare only those risk sources, that are deemed to be within one hazard sector. These hazard sectors are defined according to DVGW (2006) into seven areas (e.g., agriculture, industry). The authors of that study wish to better reflect the unknown interval scale by evaluating hazards only within one hazard category. Nevertheless, the problem is persistent even within one factor. Taking the example of DRASTIC (Aller et al., 1987), assigning 10-points to the factor „depth to the water table“, if it falls between $0m - 1.5m$ and only half of the points (5), if it is between $9.1m - 15.2m$. This interval is non-linear and thus leads to the problem of translation. Doessing Overheu et al. (2013) propose for the mass discharge a log-scale instead of a linear or arbitrary interval width to assign points from 1 to 10 to the contaminant volume. Thus, the choice of interval distances is subjective and therefore also the severity points are subjective. Due to these reasons, summation and multiplication should only be used on cardinal scales.

3.5.4. Limitation of Qualitative Risk Assessment - A Summary

Qualitative (semi-quantitative) risk assessment methods as introduced above need to be critically reviewed. In the following, I would like to only shortly summarize the findings of the discussion in the previous two subsections. In most cases, all conclusions apply to both risk matrices and semi-quantitative risk approaches.
Risk matrices ...

- are subjective and thus ignore the value of transparency in decision making, as risk categories are stakeholder-dependent.

- show the hyperbolic character of the risk product (Kaplan and Garrick, 1981).

- can lead to symmetric or asymmetric risk categories.
- work on ordinal scales, such that mathematical operations defined for cardinal scales are prohibitive.
- fail to account for cumulative risk aspects, as summation assumes perfect simultaneous impact of all hazards at the receptor.
- are sufficiently resolved by three risk categories. With increasing risk categories, they lead to a lower probability in a meaningful hazard ranking.
- enable consecutive risk ranking only for positive inclined hazards.
- can only distinguish between hazards located in separate risk categories, and are unable to rank hazard within one risk category due to the use of an ordinal scale.

Semi-quantitative and qualitative risk assessment approaches need to be critically viewed, as...

- semi-quantitative approaches are formerly the same as qualitative risk approach, only resolving the risk situations in more than the factors „probability“ and „consequences“. Thus, semi-quantitative risk assessment codes background knowledge to several sub-factors. Still, they inherit most disadvantages of risk matrices.
- non-zero risk levels are estimated (depending on the risk model), although there exist an actual zero-risk situation, such as non-toxic compounds in Eq. (3.8).
- some of the chosen risk factors, e.g., travel time, age, may unintentionally reflect a different target than the pre-defined risk objective.
- the point or ranking interval within one risk factor is non-linear.
- the same severity level (assigned points) of two different risk factors might lead to a different magnitude of damage at the receptor.

3.6. Quantitative Risk Assessment

Quantitative risk assessment models estimate risk on a cardinal scale. There are two main quantitative risk modeling communities (see Fig. 3.2). First the deterministic risk modelers that perform impact rather than risk assessment of contaminated sites, accounting in most cases for human-health and environmental risk (e.g., McKnight et al., 2012; Troldborg et al., 2008). Second, the probabilistic risk assessment community that accounts for uncertainty aspects in risk assessment, such as resolving heterogeneity by geostatistical hydraulic conductivity fields (e.g., de Barros et al., 2013, 2009; Rubin, 2003) or using fault-tree analysis methods (e.g., Bolster et al., 2009; de Barros et al., 2011; Rodak and Silliman, 2012).
Deterministic models are calibrated software tools that fit a parameterized model to observed data the best possible way, neglecting the fact of parsimony. These risk models deliver a single risk estimate. Nevertheless, often, there exist several parameter set-ups that lead to the similarly good results. Probabilistic models account for this effect by assigning

probability density functions to these parameters (Öberg and Bergbäck, 2005), solving the model in the most simplest form several times with randomly chosen parameter sets. The results produce a probability density function, providing a measure of confidence in the prediction.
In the field of drinking water supply multiple quantitative risk models exist to ensure water safety. This section will introduce quantitative risk assessment studies in more depth, separated into deterministic (Section 3.6.1) and probabilistic risk assessment models (Section 3.6.2). Furthermore, quantitative risk assessment approaches have several limitations. These challenges include the aspect of dispersion (Section 3.6.3), leading to diluted well concentration levels, the necessity in accounting for uncertainty (Section 3.6.4) and risk aggregation (Section 3.6.5). A comprehensive discussion on all limiting factors in quantitative risk assessment would go beyond the scope of this short review.

3.6.1. Deterministic Risk Assessment Approaches

Deterministic risk assessment approaches mostly originate from the physics-based community, with their roots within the field of risk assessment of contaminated sites (Section 3.3.2). There are several types of physically-based risk models to ensure water safety. Here, I distinguish between *backward* and *forward* risk assessment models (more details, see Section 3.3.1). The following is a review of the most relevant studies in this field, serving to identify their short-comings.

Wellhead Protection Delineation - A Backward Risk Concept

A classical approach for a well safety and thus risk control is to delineate time-related wellhead protection areas by calculating hypothetical travel-time zones in a deterministic fashion, as suggested by the US EPA (1993). These capture zones most commonly rely on purely advective steady-state transport considerations, e.g., based on forward or backward particle tracking (e.g., Moutsopoulos et al., 2008; Pollock, 1988), neglecting dispersive transport effects (see discussion in Section 3.6.3).
Frind et al. (2006) transferred the concept of qualitative vulnerability mapping into a quantitative, mass discharge-based model. With the help of four intrinsic well vulnerability criteria, these authors assess the severity of impact at the well (=receptor) via a macrodispersive Eulerian approach. Their approach is purely vulnerability-based, neglecting uncertainty aspects. Furthermore, these authors present intrinsic vulnerability conditions, neglecting land use activities and thus hazard properties at the upstream end. In this context, intrinsic means that the vulnerability criteria depend only on the hydrogeological characteristics of the aquifer rather than on contaminant-specific transport properties. The well vulnerability criteria consider to some extent the dilution of potential spill events due to dispersive mechanisms. They also deliver additional information for well catchment managers and stakeholders beyond the classical information on travel time-based wellhead protection zone delineation, such as mean breakthrough or peak arrival time, peak concentration levels or well down-time.

Forward Risk Models

An overview on forward mass discharge-based risk approaches is given in Tab. 3.1, complementing the extensive literature review by Troldborg et al. (2008). In the following, each study is evaluated in more detail.

Troldborg et al. (2008) outline the importance to focus on catchment-scale processes instead of local-scale plume concentrations when assessing the impact on water supply wells from multiple sources. These authors show the cumulative aspect of mass discharge-based aggregation at the well across spatially distributed contaminant sites, resolving their risk model into a leaching (local scale) and an advective catchment-scale transport model via particle tracking (MODPATH, Pollock, 1994). In order to prioritize between hazards, the relative ratio of mass discharge of each hazard in comparison to the total mass discharge has been proposed. This enables stakeholders in the light of constrained remediation budgets to prioritize individual contaminant sites according to their relative impact at the well, focusing on an individual contaminant type.

Verreydt et al. (2012) published a mass-flux-based method, that *measures* the mass discharge from known or unknown contamination sites at pre-defined control planes, which are located upstream of a plane of compliance. The upstream distance is chosen in such a way that it is far enough upstream to initiate remediation efforts for well protection, if the maximum allowed mass discharge is exceeded. For each contaminant type, a specific control plane for monitoring mass discharge needs to be installed, making this approach unfeasible for well catchments with highly heterogeneous land-use distributions. Hazards that may occur between the control plane and the well are not considered, and the uncertainty of missing a contaminant discharge in an insufficiently dense sampling network at the control plane is not considered. Nevertheless, it is the only study that enables to assess the impact from non-point and point hazards (see Tab. 3.1).

Tait et al. (2004) introduced a multi-component mass flux-based approach to optimally position new boreholes within an urban catchment. The approach consists of three models: (i) The catchment zone probability model, which identifies the probabilistic advection-based capture zone of new wells by calculating the advective travel time of particles to the possible wells over a set of Monte Carlo simulations, (ii) the land-use model, which identifies all industrial sites from past and present within a pre-defined probabilistic borehole catchment and (iii) the probabilistic risk model, which solves an analytical one-dimensional equation for the solute transport from each industrial site located in the catchment to the possible boreholes, using the probabilistic arrival times from the catchment zone probability model. This approach is only applicable for continuous point sources and underestimates the actual catchment area due to purely advective flow calculations. This limitation brings several disadvantages. First, one may fail to account for sites that contribute to well contamination by dispersion mechanisms. Secondly, it is impossible to treat non-point threats (e.g., nitrate in agricultural use). Thirdly, accidents that are of short duration but occur more often (e.g., sewage overflow) cannot be treated. Despite of this, the implemented transport models overestimate concentration and thus risk levels. Tab. 3.1 shows that the authors neglect the effect of dispersion using a particle tracking code for transport calculations. Dispersion lowers the impact at the well and thus result in smaller risk levels (e.g., Cirpka et al., 1999).

Jamin et al. (2012) introduced a mass-discharge-based model to calculate a cumulative

groundwater quality index across hazard types and spatially-distributed risk sources. They follow a modular approach with three segments, following the source (a geospatial hazard database) - pathway (flow and transport in the aquifer) - receptor (well) concept. To evaluate the combined impact at the well, they calculate the cumulative groundwater index by transferring concentrations into the SEQ-ESO ('Systme d'Evaluation de la Qualit des Eaux Souterraines', e.g., groundwater quality evaluation system) ranking system. The proposed risk concept assumes that all hazards occur perfectly simultaneously. This overestimates the risk because the actual probability of all events occurring at the same time is rather low.

Shortcomings of Current Forward and Backward Risk Models

Most studies calculate the level of severity by deterministic transport modeling, neglecting the fact of uncertainty, e.g., in hazard characteristics (e.g., *hazard characterization - failure severity*) or hazard failure (*hazard characterization - failure probability*). In addition, dispersion is omitted or treated macroscopically, although dispersion on smaller scales has a significant impact on contaminant dilution and weakens the meaning of purely advective catchment outlines or impact (Section 3.6.3). All of the risk models reviewed here, as listed in **Tab. 3.1**, fail to aggregate mass discharges over a history of multiple spill events at a single hazard location (*aggregation - temporal*), nor do they perform a quantitative risk accumulation across hazard types (*aggregation - contaminant types*). Jamin et al. (2012) achieved the latter only with qualitative methods. Thus, these approaches are vulnerability instead of risk analysis tools after the definition used in this thesis.

3.6.2. Probabilistic Risk Assessment Approaches

Pollard et al. (2008) claim that environmental policies need to be evidence-based and risk-informed, so that prioritized, high quality and implementable decisions can be made with the confidence of stakeholders. Bakker (2003) states that good governance requires an open, transparent decision making process. These requirements are only achievable, if uncertainty quantification is part of the risk analysis. This confronts each risk manager, engineer and scientist with the task to quantify model uncertainty, taking the next step from deterministic to stochastic modeling, in order to provide a robust and risk-informed decision analysis framework for stakeholders. In contrast to deterministic risk models, probabilistic risk assessment models quantify epistemic uncertainty. They account for the fact, that the level of severity at the target is highly variable due to model-related uncertainty.
Throughout my thesis, I will use the term probabilistic risk assessment to infer risk assessment based on parameter uncertainty quantification within a Bayesian framework. Probabilistic risk assessment is an integral part of risk assessment, as risk covers epistemic and aleatory uncertainty aspects. Please see a short explanation on aleatory and epistemic uncertainty in Section 3.6.4. The following provides a review of *structured*, graph-based and *unstructured*, physically-based methods for probabilistic risk assessment. Again this serves to identify their shortcomings.

		STORM (this study)	Tait et al. (2004)	Troldborg et al. (2008)	Verreydt et al. (2012)	Jamin et al. (2012)
Hazard characterization (source)						
	Failure probability	x	o	o	o	o
	Failure severity	x	o	(o)	o	o
	Trigger events	(x)	o	o	o	o
	Point source	x	x	x	x	x
	Non-point source	x	o	o	x	o
Transport & Fate (pathway)						
	Soil Passage	(x)	x	x	o	o
	NAPL-phase	x	o	x	o	o
	Retardation	x	x	x	x	x
	Degradation order	1^{st}	1^{st}	1^{st}	1^{st}	1^{st}
	Dispersion	x	o	(o)	o	x
Aggregation (receptor)						
	Spatial (several locations)	x	x	x	o	x
	Temporal (history of events)	x	o	o	o	o
	Contaminant types	x	o	o	o	(o)
Risk type (objective)						
	human health	x	x	x	x	x
	environment	x	x	x	x	x
	Economical	x	o	o	o	o

Table 3.1.: Overview on mass-discharge-based regional risk assessment approaches, focusing on well safety.

Structured Risk Quantification

Structured risk quantification is defined by non-transport-based risk models, such as fault trees (e.g., Bedford and Cooke, 2001; Bolster et al., 2009), fuzzy models (e.g., Lee et al., 1992; Sadiq et al., 2007), regression-type models (Tesfamichael et al., 2005) and Bayesian networks (e.g., Fienen et al., 2012; Thomsen et al., 2012). The structural and modular approach of these methodologies is fully in line with the source-pathway-receptor concept. The modular character allows using exchangeable models for each module and flexibility in the level of detail per module (de Barros et al., 2011). Even physically-based (and hence also mass discharge-based) models can be used within individual modules (e.g., Nilsen and Aven, 2003; Rodak and Silliman, 2012).
The most frequent used approach is **fault tree analysis**. They are able to account for different hazard locations and contaminant types across the entire catchment. Examples can be found in Bolster et al. (2009); Lindhe et al. (2009); Rodak and Silliman (2012). However, fault trees are strictly Boolean. This means that they can only propagate yes/no information of failure events, and fail to provide information on the possible severity of impact at the well. Also, they have to assume complete independence of all events. For example, the inter-dependencies of system behaviors enforced by the flow and transport physics in the aquifer that is common to the entire catchment cannot be accounted for.
Bayesian networks (e.g. Perl, 2009) have the same structural and modular flexibility as fault tree approaches and are a promising next step for hierarchical modularized risk assessment (e.g. Thomsen et al., 2012). Bayesian Networks are based on graph theory, similar to fault tree analysis, where the relations between two objects in a graph are mathematically defined. Bayesian networks are directed acyclic graphs that allow to describe the conditional dependence structure between variables in the graph, and they can handle more than just Boolean variables. Only recently, Bayesian networks have been discovered for environmental problems (e.g., Fienen et al., 2012; Thomsen et al., 2012). However, the drawbacks are similar to fault-tree methods, with the exception that Bayesian networks allow inter-dependencies.

Unstructured Risk Quantification (physically-based risk modeling)

Unstructured risk quantification is, according to Tartakovsky (2013), related to physically-based transport models without being embedded into a modularized logic structure, but consider uncertainty aspects. The community around probabilistic risk assessment focuses on parameter uncertainty due to lack of knowledge in input variables, e.g., hydraulic conductivity. Varljen and Shafer (1991) were the first to use random space functions for the hydraulic conductivity in order to delineate well capture zones probabilistically, performing conditional Monte Carlo simulations. Other early work in this field was done, for example, by Franzetti and Guadagnini (1996) and by van Leeuwen et al. (1998). Many more studies followed, such as Jacobson et al. (2005) using analytical solutions, Stauffer et al. (2004) using semi-analytical solutions and Feyen et al. (2001); Moutsopoulos et al. (2008); Vassolo et al. (1998) using numerical approaches to delineate well capture zones while considering uncertainty. An overview as well as a comparative analysis on probabilistic methods for capture zone delineation is given by Stauffer et al. (2005) as a conclusion from the European

Commission funded project W-SAHaRA „stochastic analysis of well head protection and risk assessment“ ($2000 - 2003$). The uncertainty in delineating wellhead protection areas is, of course, expressed by probability density functions. Cole and Silliman (2000) expressed the uncertainty in capture zone location by „percentile capture contours“, whereas Guadagnini and Franzetti (1999) introduced the concept of „probabilistic isochrones“.
The community mainly investigates the uncertainty that stems from the description of subsurface properties. By casting geostatistical models into a Bayesian framework these models are able to condition the model with additional data based on hydraulic conductivity, hydraulic head or mass discharges (e.g., Bakr and Butler, 2004; Feyen et al., 2003). Overall, there exist only few research groups that employ geostatistical models within a Bayesian updating framework to condition model results and thus making drinking water risk assessment and management more robust and accurate (e.g., de Barros et al., 2012, 2013; Nowak et al., 2010).

3.6.3. Important Aspects - Why Dispersion Matters

Andersson and Destouni (2001) and Cushman et al. (2001) argue that risk assessment of contaminated sites tends to overestimate risk as contaminant transport models cannot accurately predict concentration levels due to unresolved natural attenuation and transport processes in the subsurface. The overestimation of concentration levels is acceptable in the light of conservative decision making. Nevertheless, conservative decisions may lead to unbearable costs in order to guarantee water supply safety. This also somehow contradicts the principle idea of risk assessment (RAGS Manual) to mitigate only those hazards within limited investigation budget that pose the largest risk. Due to these reasons, it is necessary to consider the aspect of natural attenuation and dispersion in risk assessment and management. For now, I only assume the effect of dispersion. I will show at a later point, how to incorporate natural attenuation to the presented risk models.
The better a model can capture the details of transport behavior, the more valuable is the impact for decision making that can be anticipated. For example, the arrival time and peak concentration of contaminants strongly depend on the structure of the granular soil matrix. The contaminant dilutes and spreads due to Brownian motion and due to the flow through the heterogeneous soil matrix (e.g., Scheidegger, 1954). The following aspects influence the scale-dependent spreading of contaminants:

(a) The velocity profile between a cross-section of two grains,

(b) The different pore throat sizes,

(c) The variability of grain size and geometry within Darcy scale, leading to tortuous flow paths,

(d) Spatially varying hydraulic conductivities between several Darcy REVs after upscaling to the continuum scale and

(e) Large-scale geologic features, such as geologic windows, lenses or geologic layering.

Dispersion of solute mass is composed of four main processes. First, the center of contaminant mass follows the tortuous path (e.g., due to geologic features) of the macroscopic mean flow. Second, the contaminant plume is spreaded by smaller-scale processes (meso-scale dispersion), stretching the plume and thus leading to cigar-shaped plumes with finger-like features. Third, hydromechanical dispersion mainly acts as a boosted version of diffusion. Fourth, the diffusion of contaminant mass to groundwater is enhanced due to the larger surface area of plumes induced by spreading at the meso-scale. This process is called dilution. There have been several studies that investigated the interaction and impact of spreading, mixing and dilution to concentration levels (e.g., Cirpka et al., 1999; Kitanidis, 1994; Nowak, 2005). For a more detailed discussion please go to the relevant literature.
Most well safety studies, such as wellhead delineation (e.g., Moutsopoulos et al., 2008; Pollock, 1988) or forward mass-discharge-based risk models (e.g., Tait et al., 2004) often use advective transport calculations, neglecting the effects of dispersion and dilution. These studies assume a plug displacement of contaminant mass, depending only on the macroscopic groundwater flow velocity. This has several implications within risk assessment:

1. Dispersion reduces the peak concentration and thus leads to overestimation of concentration levels at locations downstream of the contaminated site when not considered (e.g., Cirpka et al., 1999; Kitanidis, 1994).

2. Dispersion allows particles to cross streamlines and so reduces the meaning of bounding streamlines in a manner relevant for wellhead delineation or site characterization to identify relevant hazards. Thus, contaminators just outside of the investigation area or conventional advection-based protection zones may pose a risk, although deemed to be safe.

3. Dispersion leads to an arrival time distribution, providing information on first, peak and bulk arrival.

Therefore, to consider dispersion delivers important information for improved risk assessment and management. Nevertheless, depending on a list of relevant scaling metrics, the advection-only assumption ranges between acceptable to crude (de Barros et al., 2012, 2013). The most relevant scales that determine whether dispersion is relevant are the width of a possible spill, the duration of contaminant impact, the distance to the well, and characteristic length scale of heterogeneity.
The degree of local scale dispersion (see items above (a) to (c), de Barros et al., 2012; Rubin, 2003) is characterized by the Péclet number. The Péclet number is a dimensionless number to characterize the ratio between the characteristic time scale for advective and diffusive flux for the problem-relevant length scale, such as the characteristic length scale of heterogeneities. de Barros et al. (2012) show that advective assumptions may be valid even for small Péclet numbers, as contaminants close to the end point may not be significantly influenced by local scale dispersion due to short travel times (distances) to the well. Also, strong pumping may lead to large Péclet numbers, justifying the advection-dominated description of transport processes.
Despite the theoretic justification for advection-dominated transport models and thus risk assessment, many water suppliers measure tailed contaminant breakthroughs, indicating the influence of dispersion to the contaminant. Due to these reasons, Frind et al. (2006)

were among the first to introduce a (parameterized Fickian) macro-dispersive transport approach. This approach reflects the ensemble-averaged or large-scale transport behavior that results from transport in heterogeneous aquifers. It upscales the microscopic dispersion processes in combination with the uncertainty caused by heterogeneities onto the modeling scale (macro-scale). Heterogeneity below this scale is averaged out to an equivalent homogeneous component. The chosen dispersion coefficient, expressed by longitudinal and transversal dispersivity, is scale-dependent and increases with scale. The macrodispersion tensor describes the irregular displacement of the center of contaminant mass based on an upscaled homogeneous flow assumption that neglects that heterogeneity below this scale. Therefore, the diffusion-like character of macrodispersion may not be misunderstood as actual diffusion and dilution as described above (e.g., Kitanidis, 1988). It accounts for the overall sum of uncertainty in plume centroid positions, in plume outlines and in the dissipation rates of peak concentration (e.g., Dagan, 1984; Gelhar and Axness, 1983). Thus, under non-ergodic transport conditions, the upscaled Fickian macrodispersion approach fails to capture peak concentrations (Andricevic and Cvetkovic, 1998) and natural attenuation effects. In the other extreme, models that represent the existing heterogeneity only in parts or not at all underestimate dispersion and thus the impact of concentration levels at the well (e.g., Nowak, 2005), thus failing to accurately reflect the risk of well contamination.
In order to resolve heterogeneity, at least up to the grid scale, random time-independent hydraulic conductivity fields can be introduced (e.g., Dentz et al., 2000) and dispersion is parameterized only on the local scale. With the help of Monte Carlo simulation of heterogeneous conductivity fields in combination with only hydromechanical dispersion, it is possible to separate between the effects of uncertain plume position and dilution of peak concentrations (e.g., Kitanidis, 1994). Thus, the transport of contaminants is calculated with these random fields several times, in order to receive a set of potential distributions of contaminant concentration plumes (see Fig. 3.8). Fig. 3.8 (right) shows the ensemble average and individual concentration distribution across space for the obtained ensemble of contaminant plumes, here schematically shown for $n_r = 10$ realizations.
Each realization represents one possible real solution with local-scale dispersion coefficient. The ensemble average breakthrough curve represents the macro-dispersive solution, showing concentration levels where there are none in comparison to possibly true individual solutions. This is due to spatial integration over irregular plume outlines (spreading) and ensemble averaging over uncertain plume positions. The only exceptions are at the limits of large plumes at late travel times (Dagan, 1984; Dentz et al., 2000). Hence, macrodispersion approaches cannot separate the dilution of peak concentrations from the uncertainty in plume geometry or location. Macrodispersion averages also across peak concentrations at the target position, delivering smaller overall concentration levels and thus underestimating risk. Furthermore, positioning sampling locations according to the macrodispersive concentration profile can possibly miss the actual plume pathway. These are three good reasons to consider heterogeneity as random fields within probabilistic risk assessment instead of using deterministic upscaled model results. Therefore, this local-scale dispersive transport approach is pursued within this thesis. More reasons, such as the consequences related to risk assessment, are presented for an illustrative example in Section 8.3.

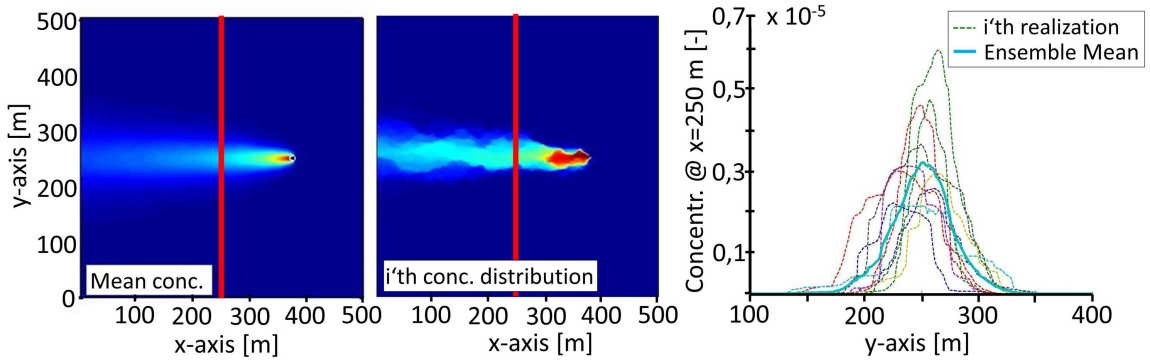


Figure 3.8.: Illustrative example to show the difference between macrodispersion and local-scale dispersion. Plume based on macrodispersive transport calculations (left). Plume based on local-scale dispersion (middle). Schematic plume distribution across $x = 250\,m$, showing $n_r = 10$ local-scale based plume distributions and the ensemble average (= macrodispersion solution).

3.6.4. Important Aspects - Why Uncertainty Matters

Aven and Renn (2009) state that „uncertainty analysis constitutes an integral aspect of the risk analysis". It is therefore indispensable to cast the question of well safety into a probabilistic framework to admit and quantify uncertainty. Unfortunately, each model used for risk analysis is a mere representation of the real world, taking model assumptions and simplistic mathematical approximations of observed and unobserved behavior of the physical subsystem or the risk source into account. These assumptions affect, among others, initial and boundary conditions, failure rate, discretization schemes and magnitude of mass release (e.g., Hoeting et al., 2006; Neuman, 2003; Park et al., 2010).
Therefore, computational models and thus the resulting risk estimates are subject to three main error sources: model conceptualization, parameterization and numerical implementation. These errors can be expressed by probability density functions, providing a prediction estimate (ensemble mean) and information on how accurate the prediction is (variance). In order to explore the uncertainty in conceptualization (e.g., Sousa et al., 2012), which model provides the best estimate, one could use methods, such as Bayesian model averaging (e.g., Hoeting et al., 1999). Numerical modeling error is getting smaller with increasing available computer power that allows finer discretization. Nevertheless, discretization (e.g., Finite-Difference-Methods) as performed in computational fluid dynamics or iterative solvers (e.g., Newton-Raphson) inhibit error sources. In this thesis, I only consider parameter uncertainty, which stems from three sources: measurement error, spatial heterogeneity and temporal fluctuation. An overview on parameter uncertainty is provided by many authors, such as Refsgaard et al. (2007); Tartakovsky (2007).
Often these uncertainties are distinguished into epistemic and aleatory uncertainty (e.g., Nilsen and Aven, 2003). Epistemic uncertainty is reducible and describes the state of imperfect knowledge. Therefore, epistemic uncertainty is interchangeably used for parameter uncertainty in the field of geostatistics. Aleatory uncertainty stems from the fact of natu-

ral variability, a process in nature that is random, such as climate or population variability. Aleatory uncertainty is non-reducible and is represented by probabilistic expressions (e.g., the choice of normal or Poisson distribution), which are part of the model (Nilsen and Aven, 2003).
Thus, considering both definitions, the stochastic hydrology community quantifies and reduces epistemic uncertainty. It is important to notice that Walker et al. (2003) refers to stochastic uncertainty as aleatory uncertainty. This is contrary to the definition before, where stochastic uncertainty is referred to parameter and epistemic uncertainty. An in depth discussion would be beyond the scope of this thesis. For a distinction between aleatory (stochastic) and epistemic uncertainty, I would like to refer to Aven and Renn (2009).
In addition, the lack of knowledge about the soil structure and scarcity of data, such as pointwise information (e.g., bore core) across the catchment, makes risk assessment a complex task. These uncertainties in material properties affect physical subsurface processes such as dilution and spreading of contaminant plumes to a large extent (Rubin, 2003). Therefore, data assimilation or conditioning tools are inevitable to reduce uncertainty as far as possible and so provide a robust and reliable decision basis. In order to improve model predictions, several studies in the past adopted the idea of incorporating additional data to reduce uncertainty (e.g., de Barros and Rubin, 2008; Leube et al., 2012; Maxwell et al., 1999), using inverse methods (Tarantola, 2004). There exist several conditioning methods for uncertainty reduction, such as the Bayesian (formal) GLUE approach (e.g., Feyen et al., 2003), Ensemble Kalman Filters for parameter estimation (e.g., Nowak et al., 2010; Schöniger et al., 2012), Markov-Chain-Monte Carlo methods (e.g., Zanini and Kitanidis, 2009), the Quasi-linear geostatistical approach (e.g., Kitanidis, 1995) and upgrades (e.g., Nowak and Cirpka, 2004) and many other methods (e.g., Alcolea et al., 2006; Hendricks Franssen et al., 2009).
Maxwell et al. (1999) introduces three questions related to risk management in an uncertain framework, that are in line with this thesis : (1) *„Where will the contaminant plume be located at a later time?“* (2) *„When will the contaminant plume arrive at some downgradient location?“* and (3) *„What is the risk to individuals utilizing the contaminated aquifer for household water?“* Deterministic models provide answers, but cannot predict the decision uncertainty. Only by admitting parameter uncertainty, the three questions can be answered robustly in an uncertain environment. From an even wider point of view, not only parameter uncertainty quantification is important, but also further uncertain aspects need to be considered (Aven, 2010). For example, the health-related impact (type of illness) by a chemical is unknown or the exposure pathways and duration of individual and related properties such as age, body weight, consumption attitude, and so forth. de Barros and Rubin (2008) even argues that untackled uncertainty in toxicity transfer functions may be outweighing the uncertainty in heterogeneity of the subsurface. Here, in this thesis I avoid the problem of specifying levels of human-health risk by measuring risk mostly by the characteristics of breakthrough curves in the drinking water well. If health-risk considerations are relevant to the involved stakeholders, the information obtained by a breakthrough curve can be transformed to human-health risk in post-processing.

3.6.5. Important Aspects - Why Aggregation Matters

Cumulative impacts refer to the total harm to human-health, environmental systems, businesses or any target object that results from different stressors over time. A stressor is a risk source that can negatively impact the receptor that is defined through the risk objective. Aggregation is the technical step to assess the cumulative impact or risk.
The European Commission (2003) requires to assess the cumulative environmental impacts to a target objective. The water framework directive evaluates the environmental status of a water body in the light of multiple stressors (e.g., McKnight et al., 2012). The US EPA promulgated in 2007 a guideline with *„concepts, methods and data sources for cumulative health risk assessment of multiple chemicals, exposures and effects "* (US EPA, 2007). Assessing human-health risk, the US EPA proposes already in US EPA (1989) to consider all potential exposure routes, such as inhalation, dermal contact or direct ingestion. Not only the sum of different exposure pathways (RAGS Manual, US EPA, 1989) but also different chemical compounds (US EPA, 2007) lead to cumulative health risk. The consideration of cumulative aspects will ultimately lead to better and more accurate predictions of risk levels and thus decision making.
In qualitative risk assessment, aggregation of impacts is commonly done by summation as shown in the Section 3.5.2. Jamin et al. (2012) follows the same route, calculating quantitative concentration levels and converting these into a groundwater quality index in order to aggregate across contaminant types through summation. This summation of ordinal values is critical. The proposed method for aggregation by the US EPA also relies on summation (US EPA, 2007). The cumulative hazard index is summed up via exposure pathways and the total hazard quotient is the sum across adverse health effects due to different chemicals. Further aggregation-based quantitative risk approaches are fault-tree based (e.g., Lindhe et al., 2009; Rodak and Silliman, 2012).
Following literature, summation of risk levels is widely used and accepted for cumulative risk assessment. Summation is valid as long as the system is linear. Summation for non-linear systems, such as threshold-based risk analysis, fails to provide accurate risk estimates and thus provide a poor decision basis.
Troldborg et al. (2008) show that, for spatially distributed contaminated sites aggregation on mass-discharge level is necessary at the receptor level to assess the cumulative impact of individual hazards in risk assessment. Tait et al. (2004) emphasized that contaminated sites, which are spatially separated and fail at different times, may lead to mass arrival at the downstream receptor during a same time period. This is shown within the example of Fig. 3.9 and will be discussed in-depth in the following.
Fig. 3.9 shows three breakthrough curves of chemicals with same chemical properties, resulting from three independent hazards distributed across the catchment. Two out of these three hazards actually lead to well exposure times, $t_{exp,2} = 123\,days$ and $t_{exp,3} = 114\,days$, when considering them individually. Thus, one hazard pose no risk to the drinking water well, as their impact on the receptor is below a given threshold level (here red line, Fig. 3.9). The summation of both risk estimates (exposure time) equals $t_{exp,sum} = 237\,days$. The summed up breakthrough curve (black line) is the sum of the three mass fluxes, leading to an accumulated exposure time of $t_{exp,con} = 353\,days$. This disparity is expressed

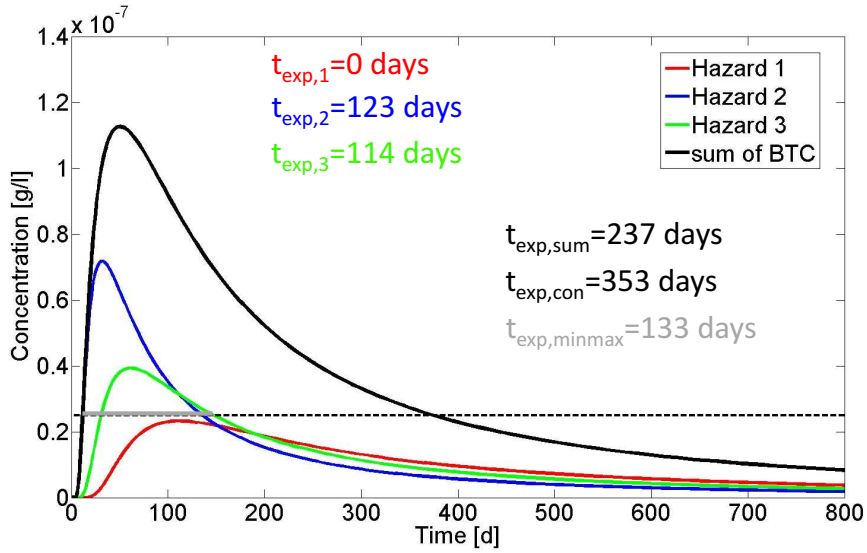


Figure 3.9.: Breakthrough curves of three individual hazards and the mass-discharge aggregated breakthrough curve (black), showing the risk measure: exposure time t_{exp}.

by Eq. (3.9):

$$\sum_{i=1}^{n} R_i \neq R(\sum_{i=1}^{n} x_i), \tag{3.9}$$

where R represents some risk or impact measure, here t_{exp}, and x some sub-process, that influences the risk or impact level, here aggregation of mass discharge or concentration levels $c_i(t)$.

There exists another possibility to look at the aggregation problem. Thinking of risk in terms that a system operates or not ($0/1$ in fault tree analysis), risk managers only focus on the time, where a system is out of operation (0-mode). The duration of failure (0-mode) under the assumption that failures happen jointly (worst case scenario) leads in this example of Fig. 3.9 to an exposure time of $t_{exp,minmax} = 133\,days$.

As simple as it sounds, the only correct way of aggregation is based on mass discharge aggregation, leading in this example to well down-time of $t_{exp,con} = 353\,days$. Adding the individual risk levels, however, is incorrect. The summed-up exposure time of two independent hazards leads to well down-times that underestimate risk. This underestimation is risk measure-dependent. In case of considering peak concentration, the summed-up maximum concentration level $c_{peak,sum} = 1.35 \cdot 10^{-}7\,g/l$ overestimates peak concentration of the total breakthrough curve $c_{peak,con} = 1.13 \cdot 10^{-}7\,g/l$.

Summation of risk levels always assumes that the impact is independent of each other and, here, in the case of well down-time, occur in a consecutive order. Simply aggregating the individual concentration levels would assume perfect concurrency. Often, this hard assumption of independence is justified to consider a worst case situation. For non-threshold-based risk measures, this worst-case assumption may be valid, if considering all hazards and not only the critical ones. Nevertheless, threshold-based risk levels such as well exposure time, time to react, or maximum concentration ratios fail to project the worst case situation.

Certain risk sources, such as combined sewage overflow, pipe leakage, fertilizing, and so on, fail frequently and thus impose a threat to the drinking water source more than once. The arrival of contaminant mass at the well from these type of hazards may temporally overlap quite often, adding up the contaminant mass and thus leading to higher concentrations levels in the well water. A detailed discussion is provided in Section 7.5.
Fault-tree models take the aspect of failure frequency into account, by assigning a constant failure rate (exponential lifetime functions) to each component. Each Component (node) is treated as being stochastic independent of each other. Due to the Markov assumption fault trees estimate a future state based on the present situation, such that there exist no time axis that allow aggregation of impacts over a history.
Aggregation is not only necessary across space, time and frequency, but also across hazard type. Often, the impacts are transferred according to a utility function in one common unit, such as costs (e.g., Lindhe et al., 2011) or the hazard quotient (US EPA, 1989). These impact factors are often aggregated by summation (see hazard quotient), assuming that each factor is independent of each other. In fact, summation of impact values within a non-linear system will misjudge the actual risk level (see discussion before).
In conclusion, cumulative risk assessment in the field of drinking water protection needs to account for mass discharge aggregation across space, time, frequency and type. Only the combination of space, time and frequency aspects in aggregation lead to the true impact at the receptor. Changing regulatory or critical levels leads to a non-linear impact response at the well. Therefore, risk management schemes that rely on simple summation-based risk models will drastically misjudge the actual risk situation.

3.7. Findings on Risk Assessment and Management

- Risk terminology is disunited even within the field of environmental engineering. Nevertheless, all risk assessment and management frameworks follow similar concepts and include steps such as definition of risk objectives, hazard identification, risk estimation, evaluation and treatment.
- Semi-quantitative and qualitative risk assessment models are formally the same and work on ordinal scales. All mathematical operations on ordinal scales are prohibitive and lead to faulty risk interpretation. The use of risk matrices and qualitative risk models, in general, for decision making and hazard prioritization is limited, such that a consecutive ordering of hazards is close to impossible.
- Despite of this, qualitative risk assessment is a helpful and valuable tool, especially in the sense of a screening tool and taking the first steps to approach risk assessment and management to better understand the system.
- Quantitative risk assessment methods avoid some of the pitfalls of qualitative approaches and seem to be a promising next step in more accurate risk management (consecutive ranking). Nevertheless, the increased accuracy goes along with increased costs in model set up, calibration, expert knowledge, and so forth.

- Structured risk estimation methods, such as fault tree-based ones, fail to adequately assess the level of severity, missing mass aggregation across modules. However, the modular character has its advantages.
- Physically-based risk assessment models allow to consider dispersion and degradation effects in the transport impact calculations. The actual risk situation is misjudged, when risk or impact levels are summed up, considering advection-only assumptions and neglecting uncertainty.
- Physically-based risk assessment models account for cumulative effects of past, present and future stressors, while considering hazard type, failure frequency and the spatial and temporal distribution of these hazards and follow the source-pathway-receptor model. This is achieved by aggregation of mass-discharges on the level of the receptor.
- Risk-informed decisions need to be based on probabilistic and physically-based risk models, that account among others for the variability in risk sources (space, time and type) across the entire catchment, dispersion effects, and the uncertainty in hydrogeology and geostatistical model selection.
- Robust and risk-informed decision making is available, when epistemic (parameter) uncertainty is quantified.

4. VIP - A Risk Quantification Approach[1]

VIP stands for *V*ulnerability *I*so*P*ercentiles. It is a vulnerability concept that provides a framework for probabilistic, quantitative and physically-based risk assessment and management, based on so-called vulnerability isopercentile maps (VIP maps). It aims at supporting wellhead delineation with information on probabilistic compliance levels of maximum allowable vulnerability. VIP maps display these information by showing isolines of non-compliance levels.
The aim of this chapter is to introduce the methodological approach to obtain vulnerability isopercentile maps, featuring the conceptual and non-technical level. Mathematical and technical issues will be addressed in Chapter 4. The possible risk management options in wellhead delineation offered by VIP will be introduced in Chapter 6. The VIP concept is based on nine steps that are introduced in Section 4.1. The benefit of setting well vulnerability criteria into a risk context and calculating vulnerability isopercentiles is given in Section 4.2. These nine steps of the VIP approach are demonstrated for two different novel method combinations. The modeling tools used in the two VIP frameworks and their differences are presented in Section 4.3.

4.1. The Vulnerability Concept in Nine Steps

Fig. 4.1 shows the step-by-step procedure of the VIP concept. In the following, these nine steps are qualitatively described, highlighting the benefits of the risk quantification and management framework VIP.

Step 1: Model Set up

Model-based delineation of capture zones is available by simulation tools such as MODFLOW (Harbaugh et al., 2000). Such models are required to predict travel times to the well and other characteristics that describe how vulnerable the well is to contaminant release at each point in the catchment. Setting up such a model requires, among other things, an assumption on the spatial structure of heterogeneous aquifer parameters (e.g., hydraulic conductivity, porosity), a corresponding parameterization and a flow and transport description. The spatial structure can be parametrized, e.g., via zone-wise values, as values of pilot points or as more general geostatistical models. These models can easily be exchanged

[1]Parts of this chapter have been published in Enzenhoefer et al. (2012) and (Enzenhoefer et al., 2013).

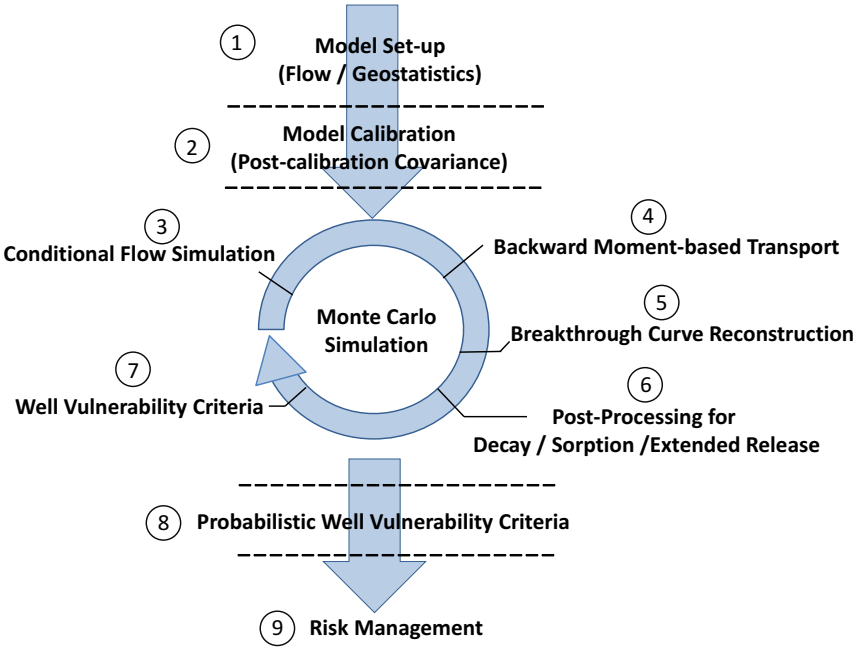


Figure 4.1.: VIP concept to determine probabilistic well vulnerability criteria.

depending on software accessibility and financial budgets.

Geostatistical Model

In order to account for spatial variability and parameter uncertainty, the hydraulic conductivity K can be treated as a random space function (e.g., Delhomme, 1979). K is the most sensitive parameter to assess wellhead location (e.g., Feyen et al., 2003). Further parameters that may be assumed to be uncertain include recharge q_r (e.g., Hendricks Franssen and Kinzelbach, 2004) and porosity n_e. The latter is often assumed to have a smaller influence due to its narrow range in aquifers, but may in fact be relevant for the assessment of time-of-first-arrival (e.g., Riva et al., 2006). Due to the Monte Carlo approach, any other kind of additional uncertainties would be easy to implement, such as uncertain boundary conditions (e.g., Kitanidis, 1995), and so forth.
In case boundary conditions or conceptual aspects of the catchment model are uncertain, it is also possible to randomize them and, if necessary, also employ Bayesian model averaging (e.g., Hoeting et al., 1999). A good reading on the analysis of different conceptual models of the aquifer heterogeneity structure is Riva et al. (2006). These authors show how such assumptions affect the uncertainty for delineating time-related wellhead protection zones.

Groundwater Flow Model

It is most common to use steady-state flow fields and neglect the effect of dispersion, while calculating wellhead protection zones (e.g., Moutsopoulos et al., 2008). The importance of dispersion has been shown in Section 3.6 and is considered in all risk models developed in this thesis. For flow and transport, I assume steady-state flow fields. This is despite the fact that flow fields may actually be subject to seasonal, diurnal or other fluctuations, such as pumping rates, recharge levels or background flow direction. Also, the concept of spatially fixed bounding streamlines for well catchments breaks down under transient conditions (Festger and Walter, 2002). Transient flow conditions may influence the extent and shape of the capture zone delineation. Furthermore, well catchments under transient flow conditions are subject to additional parameter uncertainty related to time/frequency analysis. According to Reilly and Pollock (1996), seasonal variations or small daily fluctuations are of minor importance. Nevertheless, this simplification may not be valid in some situations, such that future research is needed.
The governing equations for flow and transport are provided in Section 5.1.1. The geostatistical model to account for uncertainty in hydraulic conductivity is discussed in Section 5.1.3.

Step 2: Model Calibration

The flow model is typically calibrated to available flow and transport data. Due to epistemic uncertainty, model and measurement error or other simplifying assumptions, it is often impossible to obtain one unique parameter set in light of low data availability to represent reality and thus the observed data. The goal of model calibration and inverse modeling is to find one or more parameter sets that represent observation data as close as possible.
The classical optimization approach formulates an objective function, based on the residuals between model and data. This objective function is minimized. Optimization algorithm are available such as gradient-based (e.g., Steepest-descent, Levenberg-Marquardt) or stochastic-based algorithms (e.g., simulated annealing, bootstrapping). All over-determined and linear problems are well-posed, such that gradient-based methods are applicable. Mildly non-linear parameter dependence can be linearized, e.g., through Taylor series expansion. PEST (Doherty and Hunt, 2010) is a non-linear parameter estimator for problems that are mildly non-linear. PEST is capable to regularize such that under-determined and some classes of ill-posed problems can be made well-posed, allowing to obtain a single best parameter set.
Geostatistical methods generate a large number of uncertain parameters. Then, the calibration problem is heavily under-determined, such that the statistical approach to inversion, has to be used, and non-linear and global optimization techniques are in demand. Stochastic calibration, such as Bootstrapping, allows to establish many possible parameter sets to represent reality in an under-determined situation.
Depending on the parametrization of the spatial structure, adequate calibration schemes are available. Using pilot-points (e.g., Alcolea et al., 2006) or zone-wise description of the hydraulic conductivity fields lead to a smaller amount of uncertain model parameters. It is possible to receive the post-calibration covariance matrix by any deterministic calibration

model, such that it can be used for conditional flow simulations (step 3). A closer discussion on linear and non-linear calibration and well-posed problems is given in Chapter 5.

Step 3: Conditional Simulation

Calibration and conditioning are closely related. Calibration traditionally refers to deterministic parameter estimation of a single parameter set that represents the observed data best (step 2), but the post-calibration parameter uncertainty still needs to be quantified and accounted for. Conditioning is classically used in stochastic hydrogeology, where the geostatistical model based on a random space function is updated with information that was not available when the geostatistical model was formulated (Rubin, 2003). Due to epistemic uncertainty, model parameters are non-unique and lead to uncertain model predictions. This directly leads to the use of Monte Carlo simulations. In order to reduce uncertainty, one can couple Monte Carlo simulations with any kind of conditioning schemes, such as Ensemble Kalman Filters (e.g., Nowak et al., 2010; Schöniger et al., 2012), Bayesian GLUE (e.g., Feyen et al., 2003), Markov-Chain-Monte Carlo methods (e.g., Zanini and Kitanidis, 2009), the quasi-linear method of geostatistical inversion (Kitanidis, 1995), the successive linear estimator (e.g., Vargas-Guzmán and Yeh, 2002) or Null-space Monte Carlo methods (e.g., Jaime Gómez-Hernánez et al., 1997).
Because there are many possibilities, each with specific advantages and disadvantages (see, e.g., Hendricks Franssen et al., 2009), the main idea of the VIP concept is independence on the actual choice of conditioning or calibration methods. In any case, the output of steps 2 and 3 is a set of many parameter sets that are consistent with the data, and statistically represent the uncertainty after calibration.

Step 4: Moment-based Backward Transport Calculations

The Monte Carlo simulations used in step 3 lead to high computational costs. To reduce the computational effort, model reduction techniques are in demand. A promising model reduction approach is the combination of temporal moments (e.g., Harvey and Gorelick, 1995) and reverse flow formulation (Neupauer and Wilson, 2001).
The most efficient way to delineate wellhead protection zones is to solve the flow and the advective-dispersive transport problem reversely (Neupauer and Wilson, 2001). The reason is that the reverse transport approach delivers the required information about solute transport from all possible spill locations to the well in a single transport simulation. Instead of releasing a solute tracer at each location $\mathbf{x}_i$ within the domain and then solving many separate transport simulations within the same (forward) flow field, it reverses the direction of flow and injects a virtual tracer into the groundwater well, from which the tracer is transported reversely into the aquifer. The reverse modeling approach is formally based on the adjoint-state solution of solute transport and is conceptually similar to backward particle tracking (e.g., Uffink, 1989).
The classical or reverse advection-dispersion equation can be solved by Eulerian schemes or by particle tracking random walk (Lagranian) for transport calculations. The VIP approach

is independent on the choice of Lagranian or Eulerian transport description. In this thesis, I will present two VIP applications, utilizing the Eulerian (Chapter 8) and the Lagranian (Chapter 9) transport framework.
The computational demand within a particle-tracking random walk (PTRW) approach and the accuracy of model predictions are directly dependent on the number of particles being used. Like-wise, Eulerian transport description asks for finely-resolved simulations in order to avoid high levels of numerical dispersion that may bias the risk estimates, and small time steps are required to guarantee an accurate resolution of the temporal behavior. As a counter measure to keep computational costs low, moment-based transport simulations in combination with robust breakthrough curve reconstruction techniques are adopted to simulate transport in both VIP models. A mathematical description will follow in Chapter 5. For more information on time-related moments and their information content, see Leube et al. (2012).

Step 5: Breakthrough Curve Reconstruction

Although step 4 considers only temporal moments of breakthrough curves for the sake of computational speed up, the vulnerability analysis of the VIP concept requires the full time behavior of breakthrough curves. Analytical and numerical approaches, such as kernel density estimators or polynomial function fitting, are available to reconstruct contaminant breakthrough curves based on time-related moments. Harvey and Gorelick (1995) proposed to use the method of Maximum Entropy in log-time to reconstruct breakthrough curves from temporal moments. Among all reconstruction options, the maximum entropy principle assumes as little as possible and is therefore the least subjective way of reconstruction (e.g., Jaynes, 1957).
Folks and Chhikara (1978) proposed to fit a so-called inverse Gaussian distribution (IGD) for a wide range of fields. Hathhorn and Charbeneau (1994) proposed to fit the inverse Gaussian distribution as parametric breakthrough curve model to small numbers of particles from particle tracking random walk simulations, and Hathhorn (1996) showed the direct correspondence between the inverse Gaussian distribution and an analytical solution for advective-dispersive transport in semi-infinite domains. It corresponds to the analytical solution for tracer breakthrough of the advection dispersion equation (ADE) in three-dimensional uniform flow fields, with instantaneous point-like injection. Mathematical details follow in Section 5.5. The mathematical background of reconstructing the full breakthrough curve information by maximum entropy in log-time is provided in Appendix A.1. Also, it is possible to post-process the data to account for retardation and degradation.
Please note, that steps 4 and 5 depend on the available flow and transport software. Again, the mentality of VIP is to be independent of software choices.

Step 6: Post-processing for Fate and Extended Release

Frind et al. (2006) followed a conservative approach based on so-called intrinsic well vulnerability criteria. In this context, intrinsic stands for the contaminant transport that is

independent of contaminant properties, i.e., without decay or sorption. Specific vulnerability additionally accounts for contaminant-specific properties, i.e. degradation and retardation by sorption. Andersson and Destouni (2001) showed the importance to consider sorption within risk analysis. Within a simple post-processing procedure, the methodology proposed here is capable to include linear sorption and first-order degradation. In the same post-processing, the VIP concept can be applied to extended release conditions. All post-processing steps can be applied in any arbitrary combination, depending on the given situation of assumed wellhead contamination. A mathematical description is provided in Section 5.6, and Section 9.2.5 illustrates the difference between contaminant-specific and intrinsic concentration breakthrough curves at the well in a test case.

Step 7: Well Vulnerability Criteria

The intrinsic well vulnerability criteria by Frind et al. (2006) deliver information on peak arrival time, peak concentration, time of contaminant exposure at the well and the time of first-arrival. These four criteria can be computed from the reconstructed and post-processed breakthrough curves that result from steps 5 and 6 for each location across the domain. They serve as fundamental basis for the vulnerability isopercentile concept and the risk quantification approach, STORM (see Chapter 7). The well vulnerability criteria depend on regulatory concentration-based threshold levels and on the concentration breakthrough curves at the well. For more details to well vulnerability criteria in a risk context, see Section 4.2. Appendix A.3 presents a semi-analytical solution to directly solve for two of the four well vulnerability criteria.

Step 8: Probabilistic Well Vulnerability Criteria

The information from step 7 on well vulnerability criteria is deterministic. Therefore, these four criteria are cast into a probabilistic framework to assess for epistemic uncertainty. This is achieved by evaluating the pixel-wise statistics of well vulnerability criteria across the conditional realizations from step 3. The resulting probabilistic well vulnerability criteria deliver indispensable information as required by US EPA (2001) on the cumulative distribution of the well vulnerability criteria across the whole modeling domain. These cumulative distribution functions can be evaluated in light of non-compliance to pre-defined critical threshold values. With this information, water managers are able to plot vulnerability isopercentile lines (VIP) that directly suggest outlines of wellhead protection areas with consistent reliability levels. A detailed explanation follows in Section 4.2.

Step 9: Risk Management

The VIP concept provides probabilistic information that supports stakeholders to develop and improve complete risk management schemes as recommended by water safety plans. The corresponding VIP maps provide access to stakeholder-specific performance indicators

or risk measures such as (1) *the possible impact and range of damage due to hazardous events in the catchment (e.g., likelihood and duration of well down-time)*, (2) *the prioritization of contaminant sources at different locations (e.g., via expected peak concentrations)* and (3) *the economic value of possible mitigation measures (e.g., cost effectiveness of re-delineating the wellhead protection area)*.
To illustrate the usefulness of the vulnerability concept, several studies related to data worth, prioritization and risk-aware delineation are presented throughout the thesis. Furthermore, the VIP concept improves the discussion basis with all stakeholders involved in the risk assessment process.
The resulting probabilistic VIPs will allow water managers to

(a) quantify the current safety level of existing delineations,

(b) newly delineate protection areas in an area-neutral fashion to achieve higher consistent safety levels,

(c) control the costs for larger risk-aware delineations and other risk mitigation measures,

(d) render the decision making process more robust against uncertainties, and

(e) improve transparency towards consumers and stakeholders.

A detailed discussion related to risk management within the VIP framework is given in Chapter 6.

4.2. Vulnerability used in a Risk Context (step 7 & 8)

The central point of the VIP concept are the four well vulnerability criteria by Frind et al. (2006) (step 7), set into a probabilistic context (step 8). These two steps are the essential conceptual steps of the VIP framework and need much more explanation than it has been provided in the nine-step overview of Section 4.1. Therefore, both steps are discussed in more detail in the following.

4.2.1. Step 7: Well Vulnerability Criteria (WVC)

The four intrinsic and deterministic well vulnerability criteria are (see Fig. 4.2):

1. The time t_{peak} between a spill event and arrival at the well.
2. The level of peak concentration c_{peak} relative to the spill concentration c_{spill}.
3. The time t_{react} to breach a given threshold concentration c_{crit} (e.g., a drinking-water standard).
4. The time of exposure t_{exp} during which the threshold concentration is exceeded.

The first criterion t_{peak} represents the most common time-related capture delineation scheme. For example, German guidelines (e.g., DVGW, 2006) state, that the critical

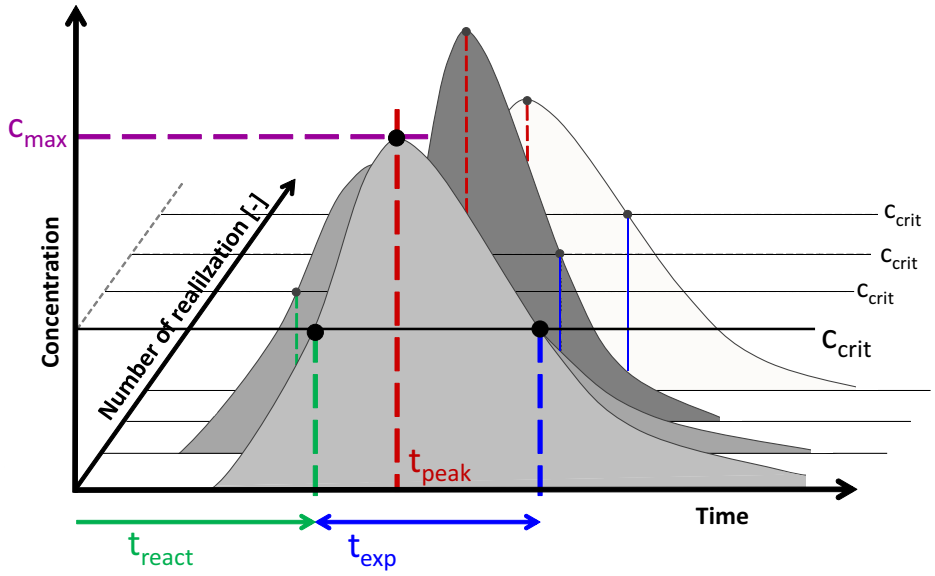


Figure 4.2.: Well vulnerability criteria after Frind et al. (2006), cast into a probabilistic framework.

travel time to ensure microbiological safety of drinking water is $\tau_{crit} = 50d$. By original definition, the capture zone is delineated according to the bulk arrival of concentration, often denoted as t_{50}. Advective-based or advection-dominated transport models commonly use the measure of bulk arrival time, as peak and bulk arrival times are identical for symmetric (i.e., almost Gaussian) breakthrough curves and also at the advection-dominated limit. The VIP concept uses the arrival time of peak concentration instead of bulk arrival in order to delineate wellhead protection zones, as the often observed tailing of breakthrough curves typically leads to earlier peak arrival t_{peak} than bulk arrival t_{50} in heterogeneous media (see Section 3.6). Therefore, t_{peak} is the more conservative and relevant criterion to be considered in risk analysis. The effect of using t_{peak} instead of t_{50} is shown and discussed further in Section 8.3.

The second criterion, peak concentration c_{peak}, accounts for dilution of peak concentrations by pore-scale dispersion, sub-grid dispersion, heterogeneity and direct dilution within the pumping well. As discussed in Section 3.6, assessing this criterion excludes all upscaled (e.g., parameterized Fickian) macrodispersive approaches, because they fail to reflect actual levels and arrival times of peak concentrations (see discussion on bulk t_{50} and peak arrival time t_{peak}, Section 8.3). The peak concentration c_{peak} forms the basis for human-health risk assessment related to acute doses (see Fig. 7.2 and Tab. C.1).

The third criterion t_{react} tells water managers the time available to react after a spill event before critical contaminant levels are being exceeded at the well. This is the most important information to design early-alert sensor or monitoring systems, and to plan emergency measures, even for worst-case most early arrival of low concentrations on any desirable confidence level. Knowing the probability distribution of t_{react} delivers

the information basis for operational risk assessment for setting up mitigation measures.

The fourth criterion, the time t_{exp} above a certain critical concentration level, is equivalent to well down-time. It can serve as a measure for the damage in economic risk analysis, that a catchment manager or the water supply company has to cope with. The time out of operation can then easily be expressed monetarily, and the expected financial loss can be compared to the costs of alternative risk treatment methods within risk-informed management decisions. If a spill remains undetected, exposure time t_{exp} is also an important impact factor for chronic health risk. Furthermore, well down-time indicates the consequences related to customer being without drinking water (e.g., Lindhe et al., 2009), due to failure within the supply system. Therefore, exposure time t_{exp} forms the basis for economic and operational risk management, and assessment of chronic health-effects (see Fig. 7.2).

Thus, well vulnerability criteria play a critical role in risk management from the economic, environmental or human safety point of view. Furthermore, they can be used as *Environmental Performance Metrics* as introduced by de Barros et al. (2012). In summary, well vulnerability criteria deliver indispensable risk-based information on:

1. the possible detriment, such as the maximum extent of acute health effects via peak concentrations (WVC 2), the possible extent of chronic health effects by long-term exceedance of threshold levels (WVC 4), or the magnitude of economic loss due to well down-time (via the duration of excessive concentration levels in the well, WVC 4),

2. the prioritization of contaminant sources at different locations within the catchment, e.g., due to shorter available reaction time (via the time until threshold arrival, WVC 3), or due to larger potential of adverse health effects or economic loss (see item above),

3. the economic value of suggested risk mitigation measures, e.g., by cost efficiency analysis or by conducting a cost benefit analysis with the customers willingness to pay (e.g., Lindhe et al., 2011), and

4. the suggested areal outlines for wellhead protection zones that are directly linked with the transport physics to the well (WVC 1) that deliver management options as described in detail in Section 6.2.

Please note, that the third and fourth well vulnerability criterion require a pre-defined threshold concentration level, c_{crit}, on which the resulting well vulnerability values depend on. This concentration-based threshold level is not necessarily coinciding with the critical performance levels, introduced for the probabilistic vulnerability concept (see next). In order to avoid misunderstanding, the critical performance levels are denoted by greek letters or by WVC_{crit} and the concentration threshold level as c_{crit}. The well vulnerability criteria only depend on the shape of the concentration breakthrough curve. Thus, also monitored breakthrough curves can be interpreted in the light of the well vulnerability criteria with the limitation that monitoring-based VIP cannot be used for risk prevention.

Possibilities to Combine the Four Criteria

Together, the third (time to react) and the fourth vulnerability criterion (well down-time) provide the necessary information for financial optimization of risk treatment alternatives, while criteria one and two yield the essential information for toxicity assessment in human-health risk assessment and for compliance with classical time-related delineation laws. An example for a contaminant-specific aggregation of vulnerability characteristics to augment the second and fourth well vulnerability criteria is

$$\int_0^\infty (c(t) - c_{crit})^\psi \, \mathrm{d}t, \tag{4.1}$$

with ψ being a dimensionless weighting factor that water managers are able to non-linearly weigh the severity of impact (response), e.g., within a non-linear dose-response relationship. Eq. (4.1) can also be interpreted as a linear non-carcinogenic human-health risk model (due to the threshold-based concentration) with $\psi = 1$ and, when the averaged concentration over time is multiplied with some toxicologically-based parameter (e.g., the lethal dose LD_{50} or slope factor in cancer risk assessment) that depends on a dose-response relationship.
As a simplistic approach for combining all four intrinsic well vulnerability criteria, the overall maximum extent of the delineated areas for a certain isopercentile for all four criteria is available. This would yield the highest reliability in providing safe and clean drinking water. This is similar to Sousa et al. (2012), who dealt with capture zone delineation uncertainty by taking the maximum extent from all model scenarios. More sophisticated ways to combine the four intrinsic vulnerability criteria to one simple rule of delineation could be defined on a case-to-case basis, reflecting the individual preferences of the water manager. Although a „unified“ risk measure is possibly available, I do not recommend this type of aggregation, but rather propose a multi-objective optimization to find along a Pareto front the best solution, that is most suitable for the different interests of stakeholders (see Section 11.4). Risk assessment and management is based on a singular risk objective (see Section 3.2) to which the well vulnerability criteria provide quantitative risk measures. Considering a unified approach blurs the actual risk picture. This is demonstrated later in this thesis (see Section 10.3), where I compare the prioritization of hazards given two different risk objectives. The discussion on stakeholder conflicts and a unit risk measure is continued in Section 7.1.

4.2.2. Step 8: Vulnerability Isopercentiles (VIP)

In order to account for epistemic uncertainty, which evolves through the lack of knowledge about the system to be modeled, the four deterministic well vulnerability criteria by Frind et al. (2006) are set into a probabilistic framework (see Fig. 4.2). As discussed in Section 3.6.4, it is irresponsible to perform risk analysis and management without considering epistemic uncertainty. Only by admitting and considering the effects of model and parameter uncertainty, the required probability distributions for contaminant peak concentrations, travel times, threshold arrival times and well down-time can be simultaneously obtained

for each location across the domain. These probability density functions enhance the information available for risk managers. They provide information on expected damage and the damage related to a pre-defined confidence interval. This is in line with the US EPA (2001) guidelines, calculating human-health risk from individual contaminants and comparing the probability distributions to contaminant specific critical values.
The easiest way to set a deterministic model into a probabilistic framework is by Monte Carlo simulation, varying the input parameters for each realization to account for epistemic uncertainty. The four intrinsic well vulnerability criteria are evaluated for each Monte Carlo realization $j = 1, \ldots, n_r$ at all points $\boldsymbol{x}_i$ in the domain Ω. As a result, the full probability distribution of the corresponding well vulnerability criteria k for each location $\boldsymbol{x}_i$ is available. If additional information is available in form of direct and/or indirect measurements, the model can be calibrated on the available data $\boldsymbol{d}_0$ (step 3), leading to the posterior or conditioned probability density distribution for the k-th criterion:

$$P_k\left(WVC_k \geq WVC_{k,crit} | \mathbf{x} = \mathbf{x}_i, \mathbf{d}_0\right) \approx \frac{1}{n_r} \sum_{r=1}^{n_r} I_{k,r}(\mathbf{x}_i), \tag{4.2}$$

where $I_{k,r}(\boldsymbol{x}_i)$ is an indicator function that assumes a value of unity in realization r, if the k-th criterion is larger than the value $WVC_{k,crit}$ at location $\boldsymbol{x}_i$ within the domain Ω, and zero otherwise. When choosing a certain critical value $WVC_{k,crit}$, Eq. (4.2) yields catchment-wide probability maps $P_k\left(\mathbf{x}\right)$ of exceeding (or falling below) a critical level $WVC_{k,crit}$ for wellhead protection. The concept of critical values is in line with the concept of performance criteria introduced by James and Freeze (1993) to judge the risk of system failure. When plotting these probability maps as iso-probability contours, one obtains so-called vulnerability isopercentiles (VIPs). This follows the idea of percentile capture contours (e.g., Cole and Silliman, 2000) and probabilistic isochrones (e.g., Guadagnini and Franzetti, 1999), which express the uncertainty in capture delineation of wellhead protection zones.
Compared to the non-probabilistic well vulnerability criteria, the VIPs deliver additional probabilistic information, such as:

1. The *probability* of peak arrival from all potential spill locations to be faster than a required minimum time (e.g., 50 days). The probability distribution of t_{peak} delivers the information necessary to assess the risk of not meeting the legal regulation about time-related delineation. This allows to rationally choose larger catchment outlines for safety reasons.

2. The *probability* of peak concentrations in the well to be larger than some maximum allowed level (e.g., an MCL). The statistics of c_{peak} form the basis for human-health risk assessment, providing information on acute doses (see Fig. 7.2 and Tab. C.1), and allows to judge the compliance with legal threshold concentrations.

3. The *probability* of the time window available to react (=reaction time) after a spill event until a critical concentration level is exceeded in the well (e.g., drinking water standard) to be smaller than a minimum critical value (e.g., 10 days) required to take adequate counter measures. This is of importance to asses the reliability of monitoring systems and thus allows to judge the compliance of the operational processes.

4. The *probability* that the well has to be shut down or is exposed to a non-compliance contamination level for more than a given critical duration. The statistics of exposure time t_{exp} form the basis for economic and operational risk management, and assessment of chronic health-effects (see Fig. 7.2).

The threshold level c_{crit} for the third and fourth vulnerability criteria and the critical values $WVC_{k,crit}$ for calculating the VIPs used in Eq. (4.2) are specific to each investigated situation and supply company. Often, hard data are given by regulatory agencies (e.g., drinking water standards), depending on the properties of the contaminants under investigation (bacteriological, chemical, and so forth). In all other cases, the technical situation or the risk acceptance level of the water supply company will determine the threshold values. Also, the risk measures in the second risk quantification framework (STORM, see Chapter 7) will depend on the a-priori defined critical threshold value c_{crit}.
Like the well vulnerability concept, the presented VIP concept is independent on the modeling framework. In this thesis, I will show how VIP maps are calculated for two different modeling problems in the upcoming Section 4.3.

4.3. Introduction of Two Numerical Toolkits for VIP

This Section links the VIP concept from Sections 4.1 and 4.2 with the numerical techniques in Chapter 4, the management framework from Section 6, and the test and demonstration cases in Chapter 8 and Chapter 9. A this point of the thesis, I introduce two possible numerical toolkits for the VIP framework with their set of tools used to assess probabilistic well vulnerability criteria. This exploits the freedom and independence of the VIP framework from technical choices in its steps 1 to 5. First, I discuss the reason why two VIP frameworks are presented (Section 4.3.1). Second, the approach for each VIP framework and their methodological differences within the nine steps are illustrated (Section 4.3.2).

4.3.1. Why Two VIP Modeling Frameworks?

The VIP concept does not depend on a specific hydrogeological setting nor on the processes occurring within a catchment. Therefore, I demonstrate the usefulness of VIP maps to a synthetic (see Chapter 8) and to a real catchment (see Chapter 9). The synthetic test case is a simple rectangular sandbox model, whereas the real catchment features a karstic hydrogeological setting. A second major advantage of the VIP concept is the independence on methodological tools, such that risk managers can use their own software tools to perform probabilistic risk assessment by the VIP concept. Here, I use two novel model combinations to calculate the vulnerability-based risk measures (see Fig. 4.3).

Why VIP in a Synthetic Catchment?

The synthetic test case provides a simple domain to investigate well vulnerability criteria (Frind et al., 2006) within a stochastic framework in more detail. The used numerical framework and test case demonstrates the usefulness of separating between actual dilution and uncertainty in plume spreading and location. This separation delivers indispensable risk-based information, which are not available by deterministic (macro-dispersive) approaches.

Why VIP in a Real Catchment?

Practitioners still refrain from applying new techniques related to probabilistic risk assessment for mainly three reasons.

1. They fear the possibly cost-intensive additional areal demand of probabilistic safety margins,
2. probabilistic approaches are allegedly complex, not readily available, consume huge computing resources and require many complex assumptions (e.g., Renard, 2007), and
3. uncertainty bounds are fuzzy, whereas final decisions are binary (e.g., Pappenberger and Beven, 2006).

The primary goal to illustrate the VIP concept to a real catchment is to show that these reservations are unjustified and that probabilistic information is available with readily accessible software tools.

4.3.2. Two Approaches Applied to Calculate VIP Maps

Two different modeling tools for exposure risk quantification are shown in Fig. 4.3. The combination of methods proposed here for the two VIP frameworks are merely two sets of possible options, chosen according to the ease of use and the ready availability of tools. The individual methods applied to each VIP framework, demonstrated in Fig. 4.3, are mathematically explained in Chapter 5.

Approach for VIP applied to the synthetic catchment

The step-wise approach for this VIP framework uses state-of-the-art tools mostly found in academia, as explained next.

Step 1: The groundwater flow model is described by *Eulerian flow formulation* and solved by a Finite Element code developed by Nowak (2005).
The aquifer heterogeneity is resolved on and above the relatively fine grid scale in each realization with the help of Monte Carlo simulations, using *log-hydraulic conductivity fields* that follow the concept of random space functions (e.g., Delhomme, 1979).

	VIP in General	VIP I (synthetic test case)	VIP II (real test case)
Step 1	Reverse steady-state Flow & transport specification	Eulerian Approach	Lagrangian Approach
Step 1	Geostatistical fields	Random Space Function	Zonation-based
Step 2	Calibration	Bayesian stochastic a-priori distributions	Linear (PEST resp. by-hand)
Step 3	Conditioning realization	Fast-kriging-like technique Formal Bayesian GLUE	Post-calibration covariance matrix
Step 4	Advective-dispersive transport calculation	Temporal Moments	Moments by Particle Tracking Random Walk
Step 5	Breakthrough curve reconstruction	Maximum Entropy in log-time	Inverse Gaussian distribution
Step 6 & 7	Vulnerability assessment	intrinsic	specific

Figure 4.3.: Showing two different method combinations to achieve VIP maps.

Only the unresolved variability on the sub-grid scale remains parameterized via local (grid-size) dispersivities (e.g., Rubin et al., 1999). This is due to the fact that vulnerability criteria are sensitive to the difference between uncertainty in plume location and actual dilution (Kitanidis, 1994), which prohibits to use solution techniques on the macroscale. I also allow the geostatistical model to be uncertain within the framework of Bayesian geostatistics (e.g., Kitanidis, 1986), by using uncertain mean, trend, covariance parameters and shape (e.g., Feyen et al., 2003; Nowak et al., 2010, see Section 5.1.4). I do so, because uncertain covariances add substantially to the uncertainty of transport (e.g., Riva and Willmann, 2009, see Section 5.3.1).

Step 2: The geostatistical modeling approach directly leads to large numbers of uncertain parameters and thus to an under-determined calibration set up. Therefore, classical calibration tools, such as PEST (Doherty and Hunt, 2010) will lead to no solution. Stochastic calibration techniques are in demand (see step 3).

Step 3: Here, I condition the randomized geostatistical model to direct and indirect data, as proposed by Feyen et al. (2003) with the formal Bayesian *G*eneralized *L*ikelihood *U*ncertainty *E*stimator (GLUE). This weighting scheme quickly leads to excessive computation times, such that I recommend to process direct point-scale measurements of parameters with fast Kriging-like conditional simulation techniques (e.g., Fritz et al., 2009) in combination with conditional sampling of covariance parameters (e.g., Pardo-Igúzquiza, 1999) in order to save computer time. The uncertain mean and trend coefficients are directly included in the kriging procedure (e.g., Fritz et al., 2009; Kitanidis, 1986). As a next step, indirect (head, tracer) data is included to the weighting scheme

in order to condition the geostatistical model. For both steps, processing direct and indirect measurements, I utilize rejection sampling in order to avoid weighted realizations as they would appear in Bootstrap filters. A description of the conditioning scheme is provided in Section 5.3.1.

Step 4: With the conditioned geostatistical models, the contaminant transport for each location across the domain can be calculated. Here, I follow the idea of backward transport approach by Neupauer and Wilson (2001) to reverse the flow field. Secondly, I solve directly the steady-state governing equations for temporal moments of transport (e.g., Harvey and Gorelick, 1995), delivering the information on the transport behavior in a highly condensed form (see Section 5.4.2).

Step 5: In order to exploit the full time-dependent concentration history at any point across the domain, the concentration breakthrough curves have to be reconstructed according to the temporal characteristics. Harvey and Gorelick (1995) recommend using the maximum entropy method in log-time to find the Lagrangian coefficients by nonlinear optimization, such that the temporal moments are met by the parameterized exponential function. In order to further reduce the computational demand in the algorithms for breakthrough curve reconstruction, a situation-adapted Gauss-Hermite integration has been implemented. A mathematical description of the breakthrough curve reconstruction is given in Section 5.5.1 and Chapter A.

step 6: With the available information on the full breakthrough curve, it is now possible to evaluate the four well vulnerability measures in each realization.

Step 7: The resulting breakthrough curves are not modified in a post-processing, assuming *intrinsic* conditions. Intrinsic conditions denote the maximum impact to be expected as no natural attenuation is available that could lower the four vulnerability measures.

Step 8: The cumulative vulnerability criteria distributions are evaluated in light of noncompliance, such that vulnerability isopercentile lines can be mapped across the catchment.

Step 9: The available probabilistic information are used to demonstrate the effect of conditioning, by adding additional data to reduce epistemic uncertainty. It is analyzed, how sensitive the uncertainty reduction is to quality, quantity and accuracy of data measurements. The difference between macrodispersive solutions and the probabilistic information is obtained and discussed in the context of risk management (see Chapter 6). Furthermore, the vulnerability results are used within robust decision analysis, supporting stakeholders to find the cost-optimal risk mitigation scenario.

This VIP framework is demonstrated along a 2D synthetic test case with a single pumping well in a porous aquifer. The results are presented in Chapter 8.

Approach for VIP applied to the real catchment

The second risk model provides a toolbox to achieve probabilistic vulnerability maps with readily available software as it can be found on the market for practitioners. The step-by-step approach is explained next.

Step 1: The steady-state *flow field* is calculated by MODFLOW (Harbaugh et al., 2000). MODFLOW is widely accepted, and thus dramatically reduces the developing time for consultants. The parameterization of the flow and transport model is based on a zonation approach for hydraulic conductivity. The heterogeneous aquifer parameters to be parameterized are hydraulic conductivity and recharge. Due to the low number of uncertain parameters, classical optimization-based calibration techniques can be employed.

Step 2: The MODFLOW model has been calibrated by a local consultant kup (engineering company Prof. Kobus and Partner, Lang and Justiz, 2009), in an trial-and-error procedure. Here, I demonstrate the approach as model calibration would have been done with PEST (Doherty and Hunt, 2010). PEST can return a post-calibration covariance matrix (the inverse of the Hessian matrix used during optimization) to quantify parameter uncertainty (e.g., Fienen et al., 2010) via the additional utility PREDUNC. A detailed explanation how to obtain a post-calibration covariance matrix is given in Section 5.2.

Step 3: The uncertain parameters to model conditional flow fields are randomly generated, based on the post-calibration covariance matrix, and a Cholesky decomposition scheme (see Section 5.3.2).

Step 4: Contaminant transport is calculated by particle tracking random walk (PTRW, e.g., Uffink, 1989) within a reverse flow field. Here, the widely used transport code MODPATH (Pollock, 1994) was extended with a random walk scheme and applied to the reverse transport problem. The PTRW approach is further explained in Section 5.4.3.

Step 5: The accuracy of PTRW is proportional to the square root of the number of particles in each cell (e.g., Kinzelbach, 1988), which easily leads to millions of particles to be tracked (e.g., Hassan and Mohamed, 2003). As a counter measure, the IGD proposed by Folks and Chhikara (1978) is introduced to approximate the full information of particle arrival with extremely low numbers of particles (see Section 5.5.2).

Step 6: The reconstructed breakthrough curves are evaluated for each location in the domain according to the four well vulnerability criteria.

Step 7: Before the vulnerability criteria are evaluated, the reconstructed breakthrough curves are post-processed in order to translate intrinsic (contaminant-independent) to *contaminant-specific* WVC (e.g., Kourakos et al., 2012). In this thesis, I present the effect of retardation on the contaminant breakthrough curve (see Section 5.6).

Step 8: Again, the probabilistic information are obtained via the Monte Carlo approach, here only using a computationally small number of realizations. VIP lines are mapped, showing areas of non-compliance with pre-defined critical threshold levels.

Step 9: The probabilistic information is used for risk-aware delineation of wellhead protection zones. In this work, I present three pragmatic management options (see Section 6.2).

This VIP framework is demonstrated along a real test case that is actually used in practice for the Burgberg catchment (Zweckverband Landeswasserversorgung). The Burgberg

catchment is located in western Germany and features a karstic setting. The results are presented in Chapter 9.

4.4. Findings of VIP

The **technical advantages** are:

- The VIP concept is modular, based on nine steps and conceptually straightforward.
- The VIP approach is formulated independently from the hydrogeological settings and geologic conditions, such as alluvial or karst aquifers.
- The tools used within each step of the VIP framework are independent of each other, such that any new combination of models would lead to similar results.
- It only needs minimal code development, and it is fully compatible with commercial, closed-source simulation software.
- The VIP concept can be applied to arbitrarily complex problems that include ...
  - any kind of model uncertainty, uncertainty in boundary conditions, geostatistical assumptions, non-stationarity, and all other sources of uncertainty that might be important to consider in probabilistic risk assessment.
  - well- and ill-posed uncertain parameter estimation problems.

The VIP concept in a **management** context:

- The VIP concept is a catchment-based risk approach, following a backward risk approach (e.g., Cushman et al., 2001).
- The four well vulnerability criteria deliver information for human-health, ecological, technical or economical risk assessment.
- The probabilistic information on the vulnerability criteria allows water stakeholder to evaluate the cumulative probability distribution for each vulnerability criterion at any location across the catchment in light of critical threshold values that may possibly cause damage.
- It allows prioritization of hazards and possible risk mitigation scenarios.
- By incorporating additional data, the VIP concept allows to render the decision making process more robust against uncertainty.

The **limitations** of the VIP concept:

- ... does not allow cumulative impact evaluation, because the possible interaction between hazards is not (yet) accounted for.
- ... neglects the probability of hazard failure, because it is a vulnerability.

One key goals of the STORM concept will be to overcome these two limitations.

5. VIP - A Mathematical Description[1]

This chapter takes a closer look at the physical, statistical and mathematical formulations and numerical choices necessary on a technical level, required to fill the VIP concept with numerical life. The focus is on steps 1 to 6.

5.1. Step 1: Model Set-up

The governing flow equations are introduced, that are used for both the synthetic and the real test case in Chapters 8 and 9 (Section 5.1.1). Bayes theorem (Section 5.1.2) and the geostatistical models are discussed next (Section 5.1.3). The real test case in Chapter 9 is zonation-based, whereas subsurface heterogeneity for the synthetic test case in Chapter 8 is resolved by a random space function. In Section 5.1.4, the Bayesian geostatistical approach as used in this thesis is expained.

5.1.1. Governing Equations

The **groundwater flow equation** at steady state is

$$-\nabla \cdot (K \nabla \phi) = q_s \quad \text{in } \Omega, \tag{5.1}$$

with locally isotropic hydraulic conductivity $K(x)$, hydraulic head ϕ, the source and sink term q_s (including wells) and the domain Ω. Boundary conditions for Eq. (5.1) are:

$$\begin{aligned} -(K \nabla \phi) \cdot \mathbf{n} &= \hat{q} \quad \text{on } \Gamma_1, \\ \phi &= \hat{\phi} \quad \text{on } \Gamma_2. \end{aligned} \tag{5.2}$$

Here $\hat{q}$, and $\hat{\phi}$ are prescribed fluxes and heads on the defined boundaries $\Gamma = \Gamma_1 \cup \Gamma_2$ of the domain Ω, respectively, and $\mathbf{n}$ is the normal vector pointing outward of the domain. In order to investigate the contaminant-independent transport properties of the system and obtain intrinsic vulnerability criteria, the **advective-dispersive transport equation** (Eulerian description) for a conservative tracer is considered:

$$\frac{\partial c}{\partial t} + \nabla \cdot (\boldsymbol{v} c - \boldsymbol{D} \nabla c) = 0 \quad \text{in } \Omega. \tag{5.3}$$

[1] Parts of this chapter have been published in Enzenhoefer et al. (2012) and (Enzenhoefer et al., 2013).

Here c is concentration, t is time, velocity $\boldsymbol{v} = \boldsymbol{q}/n_e$, $\boldsymbol{q} = -K\nabla\phi$ is the Darcy velocity, and n_e is the effective porosity. The local-scale hydromechanic dispersion tensor $\boldsymbol{D}$ is usually parameterized according to Scheidegger (1954):

$$\boldsymbol{D} = (\alpha_L - \alpha_T)\frac{\boldsymbol{v}\boldsymbol{v}^t}{|\boldsymbol{v}|} + (\alpha_T|\boldsymbol{v}| + D_m)\,\boldsymbol{I}, \tag{5.4}$$

with $\boldsymbol{I}$ being the identity matrix, α_L the longitudinal and α_T the transversal dispersivity. D_m is the molecular diffusion coefficient. Boundary conditions regarding Eq. (5.3) are

$$\begin{aligned} -\boldsymbol{n}\cdot\boldsymbol{v}c + \boldsymbol{n}\cdot(\boldsymbol{D}\,\nabla c) &= 0 \quad \text{on } \Gamma \setminus \Gamma_2 \\ c &= \hat{c} \quad \text{on } \Gamma_2 \\ -\boldsymbol{n}\cdot\boldsymbol{v}c + \boldsymbol{n}\cdot(\boldsymbol{D}\,\nabla c) &= -\boldsymbol{n}\cdot\boldsymbol{v}\hat{c}_{spill}\delta(t_0) \quad \text{on } \Gamma_{\boldsymbol{x}_0}, \end{aligned} \tag{5.5}$$

with $\hat{c}$ being the prescribed concentrations, here $\hat{c} = 0$, on Γ_2. $\hat{c}_{spill}\delta(t_0)$ is an instantaneous contaminant release at time t_0 with concentration $\hat{c}_{spill}$, here localized to a small element inside the domain Ω at the location $\boldsymbol{x}_0$, enclosed by the internal boundary Γ_{x_0}. For generality, I use a dimensionless normalized spill concentration of unity. No-flux conditions on all boundaries except $\Gamma \setminus \Gamma_2 \cup \Gamma_{x_0}$ have been assumed here for simplicity of notation in the synthetic test case (see Chapter 8). The boundary conditions for the MODFLOW model (real test case in the Burgberg catchment) are presented in Section 9.1.1.

5.1.2. Bayes' Theorem

Bayes' theorem is named after Thomas Bayes ($1701 - 1761$), who first proposed to update belief with available knowledge. This is particularity interesting for updating a model with available knowledge, such as measurement data as demonstrated in this thesis. In a Bayesian sense, probability is defined by a degree of belief. According to Christakos and Olea (1992), stochastic hydrogeology is based on the subjectivist view that underlies Bayes theorem. The theorem reads:

$$P(A|B) = \frac{P(B|A)\cdot P(A)}{P(B)}, \tag{5.6}$$

where $P(A|B)$ is the probability of A given the event, knowledge or extra information of B, $P(A)$ being the a-priori belief on the probability of A, often expressed by a subjective probability distribution, $P(B|A)$ is the likelihood of B given A, and $P(B)$ is the probability of B, the evidence or knowledge that has not been used before. $P(B)$ is also referred to as the marginal likelihood or the model evidence. The probability given an event is often referred to as conditional probability, such as $P(A|B)$ and $P(B|A)$.
Bayesian updating is the process of combining newly available knowledge that has not been available before to previous knowledge, leading to a more informed posterior belief. It is applied in this thesis to generate conditional simulations (step 3). This is achieved by generating unconditional realizations based on a prior assumption about geostatistics, followed by calculating the likelihood of each realization to comply with observed data. The variance

of the posterior belief is not per se smaller than before. A reduction of variance can only be guaranteed in linear inference problems. However, the conditional expected value of the prediction result is mostly closer to the real value. The methodology of Bayesian updating is often criticized by frequentists due to the subjective information content related to prior assumptions. Therefore, Bayesianists or subjectivists often try to choose weak a-priori distributions that are completely overwritten by the data within the posterior, such that new information in form of measured data drives the model to the correct prediction results.

5.1.3. Geostatistical Model Set-up

The impact at the well (i.e., the breakthrough curve) is subject to flow and transport processes. Knowing hydraulic conductivity or the exact location of geological windows at each point in space across the domain is impossible. The measurement information that is available are point information that need physical interpretation and extrapolation in space. Therefore, the description of heterogeneity across space requires a mathematical framework, such as geostatistical tools, to represent and interpolate or extrapolate the knowledge gained by point-wise observations made in the catchment.
Geostatistics provide statistical methods and tools to estimate the spatial pattern for each point within the domain, given relatively sparse and point-like data. Deterministic interpolation techniques, such as spline or inverse distance weighted methods, are almost prohibitive to use for interpolation of hydraulic conductivity fields, as the required amount of data and density is hard to achieve, especially thinking about the costs for detailed hydrogeological surveys. Thus, geostatistical interpolation is embraced. In geostatistics, the spatial pattern is assumed to follow some sort of coordinated spatial pattern, which can be described via mathematical functions. This „interpolation“ is based on an autocorrelation function that characterizes the spatial correlation of the pattern. Parameter values that are spatially closer are more correlated than parameter values that are farther away.
Sampled data is interpreted as a result of such a natural, random process. This stochastic description of hydraulic conductivity is known as random space functions (RSF) (Delhomme, 1979). The RSF model characterizes the pattern of spatial variability of a random space function $Z(x)$. Each point $\boldsymbol{x}$ in the domain has a random value $z(\boldsymbol{x})$, following the random variable $Z(\boldsymbol{x})$. The random space function is typically characterized through statistical moments such as correlation values. In order to infer the spatial structure from observed data, two statistical assumptions are necessary:

a) Stationarity (statistical homogeneity) assumes that the values of all random variables $z(x)$ originate from a single random space function $Z(x)$ over the entire domain Ω.

b) Ergodicity assumes that the ensemble statistics of Z (e.g., mean, covariance) coincide with the spatial averages that can be found within a single realization (Christakos and Olea, 1992).

The covariance function is defined as:

$$\boldsymbol{C}_Z(\boldsymbol{h}) = \mathrm{E}\left[(Z(\boldsymbol{x}) - m(\boldsymbol{x}))(Z(\boldsymbol{x}+\boldsymbol{h}) - m(\boldsymbol{x}+\boldsymbol{h}))\right], \tag{5.7}$$

with $\boldsymbol{h}$ being the separation vector between two locations (e.g., Rubin, 2003). The semi-variogram γ_z is strongly related to the covariance matrix $\boldsymbol{C}_Z$ by

$$\gamma_Z(\boldsymbol{h}) = \sigma_Z^2 - \boldsymbol{C}_Z(\boldsymbol{h}), \tag{5.8}$$

with σ_Z^2 being the variance of the random variable Z across the domain. The (semi-)variogram is a measure of variability, whereas the covariance function is a measure of similarity. The shape of the covariance function is obtained by analyzing measurement data and then fitting a covariance model (or theoretical variogram) to the data. Several covariance models exist, such as the exponential, Gaussian, authorized, power law, spherical model and so on. In the remainder of this thesis, I assume a Matérn model because it has a shape parameter that can be used for allowing uncertainty in the correlation model and because it embraces several other models on special cases (see Section 9.1).

Kriging

A well-known approach to estimate spatial variability is Kriging. Kriging uses geostatistical interpolation to inter- and extrapolate data into space. It provides a best estimate of the random space function $Z(x)$ at each location $\boldsymbol{x}$, and also provides the estimation variance as a measure for uncertainty around the best estimate. First, Kriging uses an empirical variogram or covariance function that depends only on the separation distance h, and fits a theoretical variogram or covariance function to the data in order to characterize the spatial structure of the hydraulic conductivity field. Kriging relies on the stationarity assumption introduced above. To be more specific, it typically assumes:

1. Stationarity of the first moment, i.e., the expected value of $Z(x)$ is constant in the domain Ω
$$\mathrm{E}\left[Z(\boldsymbol{x})\right] = \mu. \tag{5.9}$$

2. Stationarity of the second moment, i.e., the spatial correlation only depends on the separation distance h.
$$E\left[\left(Z(\boldsymbol{x}) - Z(\boldsymbol{x}+\boldsymbol{h})\right)^2\right] = 2\gamma_Z(\boldsymbol{h}). \tag{5.10}$$

The semi-variogram function is similar to Eq. (5.7) and is given by:

$$\gamma_Z = \frac{1}{2n(\boldsymbol{h})} \cdot \sum_{i=1}^{n(\boldsymbol{h})} \left(Z(\boldsymbol{x}) - Z(\boldsymbol{x}+\boldsymbol{h})\right)^2. \tag{5.11}$$

With the linear Kriging estimator it is now possible to infer the unobserved locations $\boldsymbol{x}_0$ from the observed data values:

$$\hat{z}(x_0) = \sum_{i}^{n(\boldsymbol{h})} \omega_i z(x_i), \tag{5.12}$$

with $\boldsymbol{\omega}$ being a weighting vector. The weighting vector is based on the proximity of the sampled locations to the unobserved point and on the spatial correlation structure of the sampled values. The weights can be obtained by solving the following system:

$$\boldsymbol{C} \cdot \boldsymbol{\omega} = \gamma_0, \tag{5.13}$$

with $\boldsymbol{C}$ being the matrix of all variogram / covariance values (calculated by the theoretical model) between all known locations (measurement location) in the domain Ω that is $\boldsymbol{C}_{ij} = \boldsymbol{C}_Z(x_i - x_j)$. γ_0 is the semi-variogram vector, containing all variogram values between the observed and the unobserved locations, leading to the following three properties.

a) The weights have to sum up to one: $\sum_{i=1}^{n_h} \omega_i = 1$ (ordinary and universal Kriging)

b) The estimation error ϵ of Eq. (5.12) is unbiased: $E\left[\hat{Z}(x) - Z(x)\right] = 0$

c) The variance of the estimation error $E\left[\left(\hat{Z}(x) - Z(x)\right)^2\right]$ is minimal.

Depending on the properties of the random space function and the various possible degrees of stationarity assumed, different methods for calculating the weights of Eq. (5.12) can be derived. For example, well-known Kriging types are simple Kriging (mean is stationary and known), ordinary Kriging (mean is stationary but unknown) and universal Kriging (mean can be non-stationary, e.g., exhibit a trend). A good reading on Kriging methods used in the field of groundwater modeling is De Marsily (1986).

5.1.4. Bayesian Geostatistical Model Formulation

For the synthetic test case, I assumed a geostatistical model that resolves heterogeneity via a random space function, now denoted as $\boldsymbol{s}$ instead of Z. Let $\boldsymbol{s}$ be a discretized $n_s \times 1$ random space vector $\boldsymbol{s} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon_s}$ where the mean vector $E[\boldsymbol{s}] = \boldsymbol{X}\boldsymbol{\beta}$ represents the trend model, and $\boldsymbol{\epsilon}_s$ denotes zero-mean fluctuations. The distribution of $\boldsymbol{s}$ follows $\boldsymbol{s} \sim N(\boldsymbol{X}\boldsymbol{\beta}, \boldsymbol{C_{ss}})$, i.e., $\boldsymbol{s}$ is multi-Gaussian with covariance matrix $\boldsymbol{C_{ss}}$ (see Eq. 5.7). Thanks to the flexibility of the Monte Carlo simulations and Monte Carlo-based conditioning scheme (e.g., GLUE), arbitrarily complex non-multi-Gaussian models could be employed as well. $\boldsymbol{X}$ is a $n_s \times p$ matrix with p deterministic trend functions, and $\boldsymbol{\beta}$ is the corresponding $p \times 1$ vector of trend coefficients. For a spatially constant mean of $\boldsymbol{s}$, $\boldsymbol{X}$ is a $n_s \times 1$ vector with unit entries, and $\boldsymbol{\beta}$ is the actual mean value. In this specific case, the trend coefficients are uncertain and will follow a normal distribution $\boldsymbol{\beta} \sim N(\boldsymbol{\beta}^*, \boldsymbol{C}_{\beta\beta})$ with the expected value vector $\boldsymbol{\beta}^*$ and the $p \times p$ covariance matrix $\boldsymbol{C}_{\beta\beta}$ (e.g., Kitanidis, 1986). The distribution of the fluctuations $\boldsymbol{\epsilon}_s$ are defined by the vector of structural parameters $\boldsymbol{\theta}$, containing, for example, variance and scale parameters of the covariance function. Subsequently, $\boldsymbol{\epsilon}_s$ has a covariance matrix $\boldsymbol{C_{ss}} = \boldsymbol{C}(\boldsymbol{\theta})$.

5.2. Step 2: Model Calibration

The model for the real test case has been calibrated by hand by the consultant „kup“ (Chapter 10). Here, I demonstrate classical deterministic model calibration, as if the calibration

would have been done with PEST, as PEST is widely used in deterministic model calibration. If catchment-specific data are available from past or current site characterization campaigns, it is desirable to calibrate or condition (see Step 3) the model to a given data set arranged in the $m \times 1$ vector $\boldsymbol{d}_o$. The data set may comprise direct or indirect data, such as conductivity data from grain size analysis or permeameter tests, drawdown data from hydraulic tomography, well testing or past well production data or tracer data.
In general, $\boldsymbol{d}_0$ is related to the parameter vector $\boldsymbol{p}$ by some model: $\boldsymbol{d} = \boldsymbol{f}(\boldsymbol{p})$. Here, $\boldsymbol{f}$ is a model that relates observable variables (e.g., conductivity measurements, head observations, well concentrations) to the parameters $\boldsymbol{p}$. The goal of calibration is to choose the model parameters such that simulated data are as close as possible to the observed ones.
There are two approaches that can be used to calibrate linear or weakly non-linear models. The first calibration technique is „*manual trial-and-error*“, where a parameter set in a forward model is changed and the results are compared to the measurement data. The parameters are adjusted by hand in an iterative process to receive the smallest residual between observation and simulation. The second calibration technique is „*automated parameter estimation*“, where the residuals are minimized by a suitable optimization algorithm. Automated calibration techniques are superior especially in the context of incorporating additional knowledge to the calibration process via regularization terms that are similar to Bayes theorem (see Section 5.1.2). A well-known automated (quasi-)-linear parameter estimator in the field of hydrogeology is PEST (Doherty and Hunt, 2010). PEST is capable of solving all well-posed calibration problems. A *well-posed problem* fulfills the following requirements:

(1) the formulated problem is *overdetermined*,

(2) the parameters are sensitive to observation data and sufficiently independent of each other (uniqueness),

(3) small data changes do not significantly change the calibration results, thus the solution is robust with respect to noise in the data (stability).

Next, I briefly explain how to set up the objective function for linear parameter calibration and then I extend formulation to weakly non-linear problems and introduce regularization. A linear forward model is defined by the following equation:

$$\boldsymbol{X} \cdot \boldsymbol{p} = \boldsymbol{d}_{sim}, \tag{5.14}$$

with $\boldsymbol{X}$ representing the model, containing $m \times n$ constant model elements, and $\boldsymbol{p}$ is the n-parameter vector which has to be determined. The data vector $\boldsymbol{d}_{sim}$ contains the m model predictions that should coincide as closely as possible to the observed data $\boldsymbol{d}_0$. The distance between model prediction and observation is a measure for the goodness of fit. Two approaches, the least-square and the maximum likelihood approach, are commonly used. Both approaches use an *objective function* Φ. The objective function of the least-square approach reads:

$$\Phi = \boldsymbol{\epsilon}^T \boldsymbol{W} \boldsymbol{\epsilon} = (\boldsymbol{d}_{sim}(\boldsymbol{p}) - \boldsymbol{d}_0)^T \boldsymbol{W} (\boldsymbol{d}_s im(\boldsymbol{p}) - \boldsymbol{d}_0), \tag{5.15}$$

where $\boldsymbol{\epsilon}$ denotes the error vector, expressing the deviation of simulated measurements to the observed data and $\boldsymbol{W}$ is the weighting vector. The parameters are adjusted such that Φ

decreases to a minimum. For linear models, the optimal parameter vector has an analytical solution given by

$$P_{opt} = \left(\boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X}\right)^{-1} \boldsymbol{X} \boldsymbol{W} \boldsymbol{d}_0. \tag{5.16}$$

The maximum-likelihood approach tries to find a parameter set $\boldsymbol{p}$ with the highest probability $P(\boldsymbol{\epsilon})$ to describe the observed data set. Here, the observed data is assumed to be subject to statistical noise, leading to an assumed statistical distribution for the residuals ϵ. The most frequently used error model is that ϵ is normally distributed with zero mean, some physically plausible level of variance, and independent. Under this assumption, the objective function for the maximum likelihood approach reads:

$$L = P(\boldsymbol{\epsilon}|\boldsymbol{p}) = \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi}\sigma_i} \cdot \exp\left[-\frac{\epsilon_i^2}{2\sigma_i^2}\right], \tag{5.17}$$

with σ^2 being the variance of the measurement error. Maximizing L under this assumption and minimizing Φ, both yield identical optimal parameter sets and the same objective function as above (Eq. 5.15) is obtained.
The calibration problem can be formulated as an optimization problem to find the parameter set $\boldsymbol{p}$, providing the best estimate $\boldsymbol{d}_{sim}$. The parameter vector $\boldsymbol{p}$ minimizes Eq. (5.15) by

$$\boldsymbol{p} = \left(\boldsymbol{X}(\boldsymbol{p})^T \boldsymbol{W} \boldsymbol{X}(\boldsymbol{p})\right)^{-1} \boldsymbol{X}(\boldsymbol{p})^T \boldsymbol{W} \boldsymbol{d}_0. \tag{5.18}$$

If the model is non-linear, it can be written as

$$\boldsymbol{d}_{sim} = \boldsymbol{X}((p)) \cdot \boldsymbol{p} = \boldsymbol{d}_{sim,0} \cdot \boldsymbol{p}. \tag{5.19}$$

After linearization by Taylor expansion around $\boldsymbol{p} = \boldsymbol{p}_0$ with the sensitivity matrix $\boldsymbol{J}_0 = \boldsymbol{X}(\boldsymbol{p}_0)$, the model is approximated by

$$\boldsymbol{d}_{sim} = \boldsymbol{X}\left(\boldsymbol{p}_0\right) \boldsymbol{p}_0 + \boldsymbol{J}_0 \left(\boldsymbol{p} - \boldsymbol{p}_0\right). \tag{5.20}$$

This leads to the classical Gauss-Newton iteration solution:

$$\boldsymbol{p}_{k+1} = \boldsymbol{p}_k + \left(\boldsymbol{J}_k^T \boldsymbol{W} \boldsymbol{J}\right)^{-1} \boldsymbol{J}_k^T \boldsymbol{W} \left(\boldsymbol{d}_0 - \boldsymbol{d}_{sim,k}\right) \tag{5.21}$$

with iteration index k. PEST is based on this algorithm. There exist several optimization techniques, such as gradient-based approaches (e.g., Gauss-Newton-method, Levenberg-Marquardt) or global optimization approaches (e.g., simulated annealing). Gradient-based approaches are well suited for well-posed problems. Most calibration methods, including those that follow classical optimization ideas (e.g., PEST or Pilot Points), can be set up to work with a-priori estimates of parameter values and with a $n \times n$ pre-calibration covariance matrix $\boldsymbol{C_{pp}}$ between the parameters to be calibrated. $\boldsymbol{C_{pp}}$ can be used to implement expert knowledge and thus restrict parameter uncertainty even prior to calibration, by following Bayes theorem.
From the viewpoint of classical optimization, the inverse of the prior covariance matrix is a Thikonov regularization matrix $\boldsymbol{\Gamma} = \boldsymbol{C_{pp}}^{-1}$, which penalizes deviations from $\boldsymbol{p}_0$. The regularization results in an objective function $\Phi(\boldsymbol{p})$ with one additional term:

$$\Phi(\boldsymbol{p}) = \frac{1}{2} \left(\boldsymbol{d}(\boldsymbol{p}) - \boldsymbol{d}_0\right)^T \boldsymbol{W} \left(\boldsymbol{d}(\boldsymbol{p}) - \boldsymbol{d}_0\right) + \lambda \left(\boldsymbol{p} - \boldsymbol{p}_0\right)^T \boldsymbol{\Gamma} \left(\boldsymbol{p} - \boldsymbol{p_0}\right). \tag{5.22}$$

Under the statistical point of view, $\boldsymbol{W}$ is the inverse of an error covariance matrix $\boldsymbol{R}$, which characterizes the magnitude of measurement (and model) error for the calibration data used, expressed as variance. λ is a regularization parameter, here $\lambda = 1/2$ as the natural (statistically motivated) value, if setting $\boldsymbol{\Gamma} = \boldsymbol{C_{pp}}^{-1}$ and $\boldsymbol{W} = \boldsymbol{R}^{-1}$. In absence of regularization or prior knowledge, the second term in Eq. (5.22) simply vanishes. Regularization can also be used without Bayesian thoughts, simply in order to overcome problems with ill-posedness. In the statistically motivated context, following the concept of successive linearization, the calibrated parameter set $\hat{\boldsymbol{p}}$ is found via the Gauss-Newton iteration:

$$\hat{\boldsymbol{p}}_{k+1} = \hat{\boldsymbol{p}}_k + \boldsymbol{C_{pp}}\boldsymbol{J}_k^T \left(\boldsymbol{J}_k \boldsymbol{C_{pp}} \boldsymbol{J}_k^T + \boldsymbol{R}\right)^{-1} \left(\boldsymbol{d}_0 - \boldsymbol{J}_k \hat{\boldsymbol{p}}_k\right). \tag{5.23}$$

5.3. Step 3: Conditional Simulation

Conditional realizations are statistically equally likely parameter sets, and are used to quantify the remaining uncertainty after calibration or Bayesian updating. Subsequently, within a Monte Carlo mentality, one can simulate flow in the catchment for each set $\boldsymbol{p}_r$, $r = 1, \ldots, N$ with any flow model, such as MODFLOW. The geostatistical model for the synthetic test case generates realizations conditioned on direct and indirect data with the help of formal Bayesian GLUE (Section 5.3.1). The zonation model of the real test site has been calibrated by trial and error. The conditional realizations are generated based on a post-calibration covariance matrix (Section 5.3.2).

5.3.1. Bayesian Conditioning

In the context of Bayesian updating, the model for the data is written as $\boldsymbol{d} = f(\boldsymbol{p}) + \boldsymbol{\epsilon_s}$. In the following, the notation for model parameter will be submitted from $\boldsymbol{p}$ to $\boldsymbol{s}$, referring to the geostatistical model formulation from Section 5.1.4. The $m \times 1$ measurement error vector $\boldsymbol{\epsilon}_r$ follows an error model, here, with the distribution of $\boldsymbol{\epsilon}_r \sim N(\boldsymbol{0}, \boldsymbol{R})$, i.e., with zero mean and $m \times m$ error covariance matrix $\boldsymbol{R}$ that characterizes the magnitude of measurement error. If deemed appropriate, $\boldsymbol{\epsilon}_r$ may also include model error. Then, for known $\boldsymbol{s}$, the measurements have the distribution $\boldsymbol{d}|\boldsymbol{s} \sim N(\boldsymbol{f}(\boldsymbol{s}), \boldsymbol{R})$. According to Bayes theorem (see Section 5.1.2), the distribution of $\boldsymbol{s}$ conditioned on a given data set $\boldsymbol{d}_0$ and known $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ is:

$$p(\boldsymbol{s}|\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{d}_0) = \frac{p\left(\boldsymbol{d}_0|\boldsymbol{s}\right)\, p\left(\boldsymbol{s}|\boldsymbol{\beta}, \boldsymbol{\theta}\right)}{p\left(\boldsymbol{d}_0\right)}. \tag{5.24}$$

The Bayesian distribution (marked by a tilde) for uncertain $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ is obtained by marginalization (e.g., Kitanidis, 1986):

$$\tilde{p}\left(\boldsymbol{s}|\boldsymbol{d}_0\right) = \int_{\beta}\int_{\theta} p(\boldsymbol{s}|\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{d}_0)\, p\left(\boldsymbol{\beta}, \boldsymbol{\theta}|\boldsymbol{d}_0\right) d\boldsymbol{\theta} d\boldsymbol{\beta}. \tag{5.25}$$

In this procedure, the entire joint distribution of $\boldsymbol{s}, \boldsymbol{\beta}$ and $\boldsymbol{\theta}$ is jointly conditioned on $\boldsymbol{d}_0$ (e.g., Pardo-Igúzquiza, 1999; Woodbury and Ulrych, 2000). Using the Bayesian GLUE approach

(e.g., Feyen et al., 2003), conditioning the probabilistic WVC is achieved by

$$\widetilde{P}(WVC \geq WVC_{crit} \,|\, \boldsymbol{x}_i,\, \boldsymbol{d}_0) \approx \frac{1}{n_r} \cdot \sum_{j=1}^{n_r} w_j \cdot I_j(\boldsymbol{x}), \tag{5.26}$$

with $I(\boldsymbol{x})$ being an indicator function that assumes a value of unity, if the impact value of the respective criterion exceeds the critical value WVC_{crit} in realization j at location x_i, and zero otherwise. The weight $w_j = \frac{L_j}{\sum_{j=1}^{n} L_j}$ of realization j represents the likelihood L given $\boldsymbol{d}_0$:

$$L\left(\boldsymbol{s}_j, \boldsymbol{\theta}_j, \boldsymbol{\beta}_j | \boldsymbol{d}_0\right)_j = \left(\frac{1}{2\pi \cdot ||\boldsymbol{R}||}\right)^{m/2} \exp\left[-\frac{1}{2}\left(\boldsymbol{d}_0 - \boldsymbol{d}_{sim}(\boldsymbol{s}_j)\right)^T \boldsymbol{R}^{-1} \left(\boldsymbol{d}_0 - \boldsymbol{d}_{sim}(\boldsymbol{s}_j)\right)\right], \tag{5.27}$$

where $\boldsymbol{d}_{sim}(\boldsymbol{s}_j) = \boldsymbol{f}(\boldsymbol{s})$ is the corresponding simulated data set of realization j.
For large data sets, only a few realizations may be assigned with significant weights. This problem can be alleviated, when first processing direct point-scale measurements of parameters with Kriging-like techniques (Fritz et al., 2009, in combination with conditional sampling). The uncertain mean and trend coefficients may be directly included in the Kriging procedure (e.g., Fritz et al., 2009; Kitanidis, 1986). Then, only indirect (head, tracer) data remain for processing with the weighting scheme. When replacing Eq. (5.26) by a rejection sampling scheme, the weights ω_j remain constant for all conditional realizations to direct and indirect data. The rejection sampling reads:

$$L_r = \frac{L_j}{L_{max}} = \begin{cases} \omega_j = 1, \text{if } L_r \geq z(\text{accepted } \boldsymbol{s}_j) \\ \omega_j = 0, \text{if } L_r < z(\text{rejected } \boldsymbol{s}_j), \end{cases} \tag{5.28}$$

where z is a random number drawn from the uniform distribution between 0 and 1.

5.3.2. Conditional Simulation based on Post-Calibration Covariance

Classical calibration tools such as PEST and any Gauss-Newton algorithm provide a post-calibration covariance matrix $\boldsymbol{C_{pp|d}}$ to represent the remaining uncertainty in the calibrated parameters around $\hat{\boldsymbol{p}}$. Analytically, it can be written as the inverse of the Hessian matrix $\boldsymbol{H}$ of second derivatives of the weighted least squares objective function, and has the form:

$$\boldsymbol{C_{pp|d}} = \left(\boldsymbol{J}^T \boldsymbol{R}^{-1} \boldsymbol{J} + \boldsymbol{C_{pp}}^{-1}\right) = \boldsymbol{C_{pp}} - \boldsymbol{C_{pp}} \boldsymbol{J}^T \left(\boldsymbol{J} \boldsymbol{C_{pp}} \boldsymbol{J}^T + \boldsymbol{R}\right)^{-1} \boldsymbol{J} \boldsymbol{C_{pp}}, \tag{5.29}$$

where $\boldsymbol{J}$ is the Jacobian matrix evaluated at $\boldsymbol{p} = \hat{\boldsymbol{p}}$. With the post-calibration covariance matrix $\boldsymbol{C_{pp|d}}$ and the calibrated parameter set $\hat{\boldsymbol{p}}$, it is possible to generate many alternative parameter realizations $\boldsymbol{p}_r$, $r = 1, \ldots, N$ using Monte Carlo simulations. Each parameter set $\boldsymbol{p}_r$ is physically plausible for the model and statistically legitimate in the light of the available calibration data and prior assumptions. A detailed explanation is given by Fienen et al. (2010). The randomized model parameters $\boldsymbol{p}_r$ are the sum of the calibrated values and a random component that corresponds to $\boldsymbol{C_{pp|d}}$:

$$\boldsymbol{p}_r = \hat{\boldsymbol{p}} + \text{chol}\left(\boldsymbol{C_{pp|d}}\right)^T \boldsymbol{\epsilon}_r, \tag{5.30}$$

with chol being the Cholesky decomposition and $\boldsymbol{\epsilon}_r$ being a vector of independent, standard normal distributed random numbers.

5.4. Step 4: Advective-Dispersive Transport Formulation

For both test catchments, the reverse flow formulation to calculate transport is adopted (Section 5.4.1). The transport for the Eulerian flow description is directly calculated via temporal moments (Section 5.4.2). Transport for the real test site is calculated with MODPATH that has been extended with a random walk scheme, and temporal moments can be evaluated even from low particle numbers (Section 5.4.3).

5.4.1. Reverse Flow and Transport Simulation

Monte Carlo-based techniques for transport-based wellhead delineation directly lead to high computational demands, unless using smart concepts. To decrease computational costs, the reverse formulation of transport can be applied. It allows to map well vulnerabilities and capture zones onto the entire domain with a single model run. The key idea is to reverse the direction of flow in the entire computational domain and inject a unit-mass, instantaneous tracer pulse into the drinking water well (e.g., Frind et al., 2006; Kunstmann and Kinzelbach, 2000; Neupauer and Wilson, 2001). This avoids the necessity to perform many transport simulations with contaminants injected at each possible spill location within the entire catchment. The transport is subsequently solved reversely, using

$$\frac{\partial c}{\partial t} = \nabla \cdot (\boldsymbol{v}c + \boldsymbol{D}\nabla c) \text{ in } \Omega, \tag{5.31}$$

and the boundary conditions change to

$$\begin{aligned} -\boldsymbol{n} \cdot (\boldsymbol{D}\,\nabla c) &= 0 \quad \text{on } \Gamma \setminus \Gamma_2 \\ c &= 0 \quad \text{on } \Gamma_2 \\ \boldsymbol{n} \cdot \boldsymbol{v}c + \boldsymbol{n} \cdot (\boldsymbol{D}\,\nabla c) &= \boldsymbol{n} \cdot \boldsymbol{v}\hat{c}_{spill}\delta(t_0) \quad \text{on } \Gamma_{well}, \end{aligned} \tag{5.32}$$

where Γ_{well} is an internal boundary that encloses the well, and $\boldsymbol{n}\cdot\boldsymbol{v}$ is the velocity perpendicular to Γ_{well}. More explanation for the backward transport approach is given, e.g., by Uffink (1989), Cornaton (2003) and by Neupauer and Wilson (2001).

5.4.2. Temporal Moment Approach

The time-consuming computations to receive time-dependent breakthrough curves can be reduced by directly solving for the temporal moments of contaminant transport (Harvey and Gorelick, 1995). Temporal moments are capable of representing the characteristics of a concentration breakthrough curve. Their physical meaning is further explained by Cirpka and Kitanidis (2000a) and has recently been discussed by Leube et al. (2012).
The k-th temporal moment m_k of a breakthrough curve $c(\boldsymbol{x}, t)$ at location $\boldsymbol{x}$ is defined as:

$$m_k(\mathbf{x}) = \int_0^\infty t^k \cdot c(\boldsymbol{x},\, t)\, \mathrm{d}t. \tag{5.33}$$

The zeroth moment m_0 represents the accumulated mass over time that passes by a location $\boldsymbol{x}$. The normalized first temporal moment m_1/m_0 represents the mean arrival time of a solute at location $\boldsymbol{x}$. The normalized second central temporal moment m_{2c}/m_0 can be interpreted as local dilution. The physical meaning of several lower-order temporal moments is discussed in more detail by Cirpka and Kitanidis (2000b). Higher order temporal moments describe characteristics such as skewness, peakedness, and more complex characteristics of the temporal breakthrough curve's time behavior that are also known from statistics (e.g., Wackerly et al., 2002). These characteristics are illustrated in Fig. 5.1.
Moment generating equations can be derived from Eq. (5.3) or Eq. (5.31) and their respective boundary conditions by multiplying the equations with t^k and then integrating over time as in Eq. (5.33) (Cirpka and Kitanidis, 2000b). For Eq. (5.31), partial integration leads to:

$$\begin{aligned} \nabla \cdot (\boldsymbol{v} m_0 + \boldsymbol{D}\,\nabla m_0) &= 0 \quad \text{for } k = 0 \\ \nabla \cdot (\boldsymbol{v} m_k + \boldsymbol{D}\,\nabla m_k) &= k \cdot m_{k-1} \quad \forall\, k > 0, \end{aligned} \tag{5.34}$$

with the boundary conditions:

$$\begin{aligned} -\boldsymbol{n} \cdot (\boldsymbol{D}\,\nabla m_k) &= 0 \qquad \text{on } \Gamma \setminus \Gamma_2 \\ m_k &= 0 \qquad \text{on } \Gamma_2 \\ \boldsymbol{n} \cdot \boldsymbol{v} m_k + \boldsymbol{n} \cdot (\boldsymbol{D}\,\nabla m_k) &= \boldsymbol{n} \cdot \boldsymbol{v} \hat{m}_{k,well} \quad \text{on } \Gamma_{well},\ \forall\, k \geq 0. \end{aligned} \tag{5.35}$$

Here, $\hat{m}_k$ is the k^{th}-raw temporal moment of $\hat{c}$ on the boundaries Γ_2 and Γ_{well}. Because the contaminant release at the well on boundary Γ_{well} is instantaneous at time $t_0 = 0$ and with unit spill concentration, $\hat{m}_{k,well}$ is one for $k = 0$ and zero for all $k \geq 1$. On the boundary Γ_2, $\hat{m}_k$ is zero for all k, because the second line in Eq. (5.32) is zero due to the reverse approach, regardless of the corresponding $\hat{c}$ in Eq. (5.5). Eq. (5.34) is formally identical to a steady state partial differential transport equation, which eliminates the need of numerical time integration and directly yields temporal characteristics at very low computational costs.

5.4.3. Backward Random Walk Particle Tracking

The computational costs of finely-resolved transport simulations with Eulerian methods are almost prohibitive, especially in large catchment simulations and in three-dimensional cases. This is even more so, when insisting on acceptably low levels of numerical dispersion, in order to not bias the risk estimates. The particle tracking random walk (PTRW) method (e.g., Prickett et al., 1981) is free of numerical dispersion and artificial oscillation (e.g., Salamon et al., 2006). PTRW is simple and robust in application. For these reasons, the use of PTRW methods is recommended. PTRW has already been used to compute advective-dispersive capture zones around groundwater wells (e.g., Uffink, 1989), yet without Monte Carlo simulation of uncertain conductivity fields.
The PTRW approximates the solution to Eq. (5.31) by (e.g., Kinzelbach, 1988):

$$\boldsymbol{X}_p(t + \Delta t) = \boldsymbol{X}_p(t) + (\boldsymbol{v}(\boldsymbol{X}_p, t) + \nabla \cdot \boldsymbol{D}(\boldsymbol{X}_p, t))\,\Delta t + \boldsymbol{B}(\boldsymbol{X}_p, t) \cdot \boldsymbol{\xi}(t)\sqrt{\Delta t}, \tag{5.36}$$

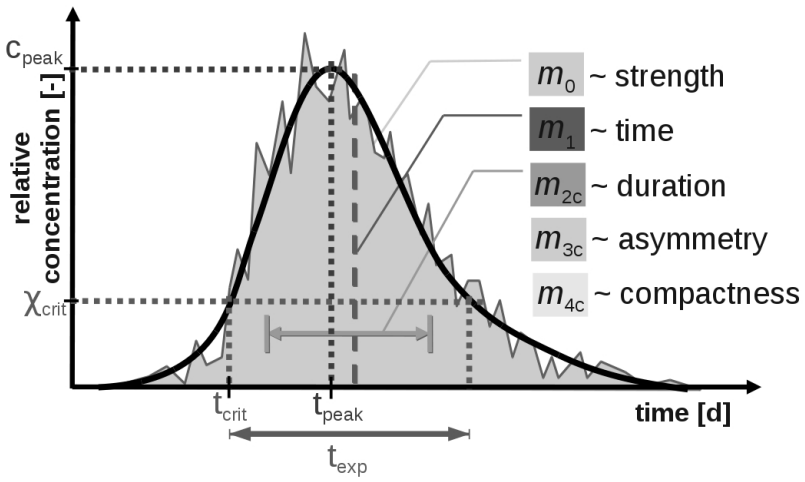


Figure 5.1.: Illustrative sketch, showing the four intrinsic well vulnerability criteria and temporal moments characterizing the concentration breakthrough curve $c(t)$.

where $\boldsymbol{X}_p(t)$ is the particle position at time t and $\triangle t$ is the current time step. $\boldsymbol{B}$ is the displacement matrix that has to fulfill $B \cdot B^T = 2D$. $\boldsymbol{\xi}(t)$ is a vector with three normal distributed uncorrelated random variables with zero mean and unit variance. The displacement matrix $\boldsymbol{B}$ after Salamon et al. (2006) equals in 3D:

$$\boldsymbol{B} = \begin{pmatrix} \frac{v_x}{|v|}\sqrt{2(\alpha_L|v|+D_m)} & -\frac{v_x v_z}{|v|\sqrt{v_x^2+v_y^2}}\sqrt{2(\alpha_T^V|v|+D_m)} & -\frac{v_y}{\sqrt{v_x^2+v_y^2}}\sqrt{2\left(\frac{\alpha_T^H(v_x^2+v_y^2)+\alpha_T^V v_z^2}{|v|}+D_m\right)} \\ \frac{v_y}{|v|}\sqrt{2(\alpha_L|v|+D_m)} & -\frac{v_y v_z}{|v|\sqrt{v_x^2+v_y^2}}\sqrt{2(\alpha_T^V|v|+D_m)} & \frac{v_x}{\sqrt{v_x^2+v_y^2}}\sqrt{2\left(\frac{\alpha_T^H(v_x^2+v_y^2)+\alpha_T^V v_z^2}{|v|}+D_m\right)} \\ \frac{v_z}{|v|}\sqrt{2(\alpha_L|v|+D_m)} & \frac{\sqrt{v_x^2+v_y^2}}{|v|}\sqrt{2(\alpha_T^V|v|+D_m)} & 0 \end{pmatrix}, \quad (5.37)$$

where α_T^V and α_T^H are the vertical and horizontal transverse dispersivities, respectively, and Eq. (5.4) can be obtained by $1/2B \cdot B^T$ when setting $\alpha_T^V = \alpha_T^H = \alpha_T$. The desired concentrations are evaluated by computing the particle density in each grid cell $k = 1, \ldots, n_c$ and in each time step t_ℓ, $\ell = 1, \ldots, n_t$.

$$c_k(t_\ell) = \frac{1}{n_{e,k} V_k} \sum_{i=1}^{n_p} m_i \cdot I_{ik\ell}. \quad (5.38)$$

Here, $n_{e,k}$ is the porosity in cell number k, V_k the volume of the k-th grid cell, n_p is the total number of particles used in the simulation, m_i the mass of the i-th particle and $I_{ik\ell}$ is an indicator that assumes a value of unity, if the i-th particle is in the k-th grid cell at time step t_ℓ, and zero otherwise.

One key feature of PTRW that the accuracy of PTRW is proportional to the square root of the number of particles in each cell (e.g., Kinzelbach, 1988), which easily leads to millions of particles to be tracked (e.g., Hassan and Mohamed, 2003). As a counter measure, a strongly reduced particle number is used, and Eq. (5.38) is replaced by a more robust breakthrough curve reconstruction method instead, based only on few lower order temporal moments.

5.5. Step 5: Moment-based Breakthrough Curve Reconstruction

The contaminant breakthrough curves are reconstructed from transport-based moments via the maximum entropy method (Section 5.5.1). The particle-based moments are cast into an analytical reconstruction scheme (inverse Gaussian distribution, Section 5.5.2).

5.5.1. Maximum Entropy Method

Using the Maximum Entropy method (e.g., Jaynes, 2003) in log-time (e.g., Harvey and Gorelick, 1995) to recover the full concentration profile yields for a breakthrough curve:

$$c(t) = \frac{1}{t} \exp\left(-\sum_{\ell=0}^{n_\ell} \lambda_\ell \cdot \ln t^\ell\right), \tag{5.39}$$

where $\lambda_\ell = [\lambda_0, \ldots, \lambda_{n_\ell}]$ are Lagrangian parameters which are obtained by ensuring that the temporal moments of Eq. (5.39) are identical to those obtained from the transport simulation.

$$m_k = \int_{-\infty}^{+\infty} t^k \cdot \frac{1}{t} \exp\left[-\sum_{\ell=0}^{n_\ell} \lambda_\ell \cdot \ln t^\ell\right] \mathrm{d}t. \tag{5.40}$$

Here, n_ℓ is the highest order of moments considered and must be an even number for the integral to exist. The non-linear optimization problem in Eq. (5.40) can be solved by the standard Newton method (Mohammad-Djafari, 2001) or any other suitable optimization algorithm. In this thesis I evaluate the integral in Eq. (5.40) by Gauss-Hermite integration (e.g., Abramowitz and Stegun, 1964) after transforming $s = \ln t$. This numerical integration scheme is computationally less demanding and, with only 40 Gauss-Integration-Points, I received a speed-up of more than ≈ 20 in comparison to Riemann or trapezoidal integration. More information on setting up the Gauss-Hermite integration for non-linear breakthrough curve reconstruction in log-time and assessing the third and fourth WVC numerically is provided in the Appendix.

5.5.2. Inverse Gaussian Distribution

For simpler breakthrough curve reconstruction, the first two temporal moments obtained from the transport code are inserted into Eq. (5.41). The inverse Gaussian distribution is given by

$$c(t) \approx \mathrm{IGD}\,(t;\, \mu,\, \eta, m_0) = m_0 \cdot \sqrt{\frac{\eta}{2\pi t^3}} \exp\left[-\frac{\eta\,(t-\mu)^2}{2\mu^2}\right], \quad t > 0, \tag{5.41}$$

with parameters μ being the first temporal moment and $\eta = \mu^3/\sigma^2$ being a shape parameter depending on variance and mean of t (viewed as a random variable of particle arrival time). The parameter m_0 re-scales the entire distribution to the total observed mass m_0, accumulated over time within the breakthrough curve in a given cell. The inverse Gaussian with

$\eta/\mu \to \infty$ shows asymptotic convergence to normality (Folks and Chhikara, 1978).
If desired and computationally feasible, larger particle numbers for direct interpretation of $c(t)$ or more elaborate breakthrough curve reconstruction techniques may be applied, such as kernel density estimates (Fernàndez-Garcia and Sanchez-Vila, 2011) or maximum entropy reconstruction (Harvey and Gorelick, 1995). In this thesis, the PTRW code has been linked to the inverse Gaussian distribution, such that the first two temporal moments based on low particle numbers are sufficient. The accuracy of using low particle numbers in combination with the inverse Gaussian distribution is presented for the real test case in Section 9.2.6.

5.6. Step 6: Post-Processing for Contaminant Fate

For linear contaminant-specific sorption processes, the velocity $\boldsymbol{v}$ and dispersion tensor $\boldsymbol{D}$ in Eq. (5.31) are modified as follows

$$\widetilde{\boldsymbol{v}} = \frac{\boldsymbol{v}}{R_d} \quad \text{and} \quad \widetilde{\boldsymbol{D}} = \frac{\boldsymbol{D}}{R_d}, \tag{5.42}$$

with R_d being the retardation factor. This simply means to linearly scale the time in Eq. (5.41) with R_d, and to re-normalize the resulting breakthrough curve with $1/R_d$:

$$c_{R_d}(t) = \frac{1}{R_d} \cdot c\left(t'\right), \tag{5.43}$$

with $t' = \frac{t}{R_d}$. To account for first-order degradation processes with decay rate ξ_d, the concentration obtained by Eq. (5.41) or Eq. (5.39) is modified as

$$c_D(t) = c(t) \cdot \exp\left[-\xi_d t\right]. \tag{5.44}$$

The breakthrough curves from any extended release history can be calculated by convoluting the solution of the instantaneous injection problem with the release strength of the extended release of single spill events:

$$c_e(t) = \int_0^t c(t-\tau)\dot{m}(\tau)\mathrm{d}\tau, \tag{5.45}$$

where $c_e(t)$ is the breakthrough curve for extended release, which is characterized by the time-dependent injected mass flux $\dot{m}(\tau)$, and τ is the time coordinate for the release history. The latter requires steady-state flow conditions. The extension to transient flow is provided by Srinivasan et al. (2012).

5.7. Steps 7 to 9: Evaluating VIP and Decision Support

A detailed explanation of vulnerability values in a probabilistic context has been given in Section 4.2 (steps 7 and 8). The evaluation of VIP is demonstrated for two test cases in Chapter 8 and Chapter 9. An introduction to use vulnerability values in a risk management context (step 9) will be given in Chapter 6.

6. VIP - A Risk Management Approach[1]

Risk managers are asked to evaluate the calculated risk against their pre-defined safety margins (compliance level). In case risk is unacceptable, risk mitigation opportunities are in demand. Thus, stakeholders are confronted with the task to find the optimal mitigation option, under the given situation, to treat risk the best possible way. This risk treatment may address risk reduction based on different scales, such that the consequences of each alternative cannot be assessed equally and compared. Another puzzle in decision making is the environmental complexity penetrated by uncertainty. Therefore, rational decision support frameworks are required to aid stakeholders in evaluating and comparing risk mitigation alternatives in the light of uncertainty and across impact dimensions.
The goal of this chapter is to introduce the fundamental concept of decision analysis (Section 6.1) to improve robust decision making in a complex environmental setting. Furthermore, a specific decision analysis framework for risk-aware delineation based on the VIP concept is presented in Section 6.2. Here, I define the risk as the probability of (non-)compliance with pre-defined acceptable vulnerability levels. Section 6.3 defines two uncertainty measures and two economic risk measures that can be used to assess the performance of data acquisition and risk reduction plans. The chapter closes with a summary on the main findings of risk management in a VIP context (Section 6.4).

6.1. Introduction to Decision Analysis

In case risk is unacceptable, risk managers are pre-dominantly confronted with two questions:

i. how effective a proposed measure for risk reduction below an acceptance level, recognizing that zero-risk is unattainable (Pollard et al., 2008) and

ii. what costs are related to available risk reduction options.

In order to mitigate risk effectively (i), three general decision options are available for risk managers:

(1) *Reduce uncertainty*, e.g, through collecting additional information in sampling campaigns or by reducing the probability of hazard failure occurrence.

(2) *Reduce impact*, e.g., install risk mitigation approaches that reduce mass release, decrease receptor vulnerability or other options to reduce the damage at the receptor.

(3) *Reduce uncertainty and impact*.

[1] Parts of this Chapter have been published in (Enzenhoefer et al., 2013) and Enzenhoefer et al. (2011)

Uncertainty Reduction

Risk measures, such as peak arrival time, depend on the geostatistical description of the hydraulic conductivity field, site characterization (amount and quality of data) and uncertainty reduction. Collecting data yields additional information to reduce uncertainty in predicting risk levels or well vulnerability criteria. Yet, sampling and investigation campaigns are often restricted by limited budgets, or by other constraints such as inaccessibility of sampling locations. The amount of data required to reduce uncertainty to an acceptable level has to be determined rationally and in a goal-oriented manner. This leads to the question how to optimize experimental designs (e.g., Pukelsheim, 2006), i.e., where to sample, which types under which circumstance and in which time frame to reduce uncertainty the most. Often, these optimal design studies are evaluated in the context of least costs (see discussion below on decision framework). Optimal design has been transferred to geostatistics (e.g., Christakos and Olea, 1992; Müller, 2007) and related inverse problems (Herrera and Pinder, 2005; Nowak et al., 2010). Furthermore, optimal design is applied to many environmental engineering problems, such as cost-minimal remediation of contaminated sites (e.g., Feyen and Gorelick, 2005) or the best set-up to monitor contaminant plumes (e.g., Meyer et al., 1994). Maxwell et al. (1999) showed that increased sampling numbers on hydraulic conductivity reduces uncertainty in the geostatistical model. This had a positive effect on the prediction goal (here: human-health risk), reducing the prediction variance and thus uncertainty. Bakr and Butler (2004) investigated the required amount of data to delineate a wellhead protection area up to a certain accuracy level. They also showed that the gain in information content is limited with increasing amount of data (see also Feyen et al., 2001). However, the information gain for a given amount of data can be maximized by optimizing data type and sampling location (e.g., Nowak et al., 2010).

Impact Reduction

Following the source-pathway-receptor concept, there are three options available to reduce the level of impact severity at the well:

(i) *Source-level*, e.g., decreasing the amount of mass being released to the aquifer in case of an accident.

(ii) *Pathway-level*, e.g., installing degradation tools to reduce concentration levels, such as reactive barriers (Freeze and McWhorter, 1997).

(iii) *Receptor-level*, e.g., enhancing the resilience of the receptor, which might be complemented by post-treatment facilities (Lee et al., 1992; Lindhe et al., 2011).

All these studies rely on scenario-based risk mitigation strategies and follow the suggestion by Aven and Kørte (2003) to perform decision and formal risk analysis in order to aid risk-informed decisions, followed by an informal judgment and review process.

Decision Framework

Finding the best risk mitigation option can be interpreted as balancing the effectiveness of risk mitigation options (question i.) and the related costs (question ii.). To achieve this, authors have used cost-benefit (e.g., Lindhe et al., 2011; Massmann and Freeze, 1987) and cost efficiency analysis (e.g., Lindhe et al., 2011). These are methods that rationally support decision making. Overall, there exist two distinct decision-making frameworks in engineering literature: optimization and decision analysis (e.g., Freeze and Gorelick, 1999). Tartakovsky (2013) reviews both frameworks, showing similarities of the two approaches. The first decision framework, *optimization*, provides the optimal decision or risk mitigation opportunity from all possible (continuous) alternatives by maximizing or minimizing an objective function (e.g., Bakr and Butler, 2004; Nowak et al., 2010). The second framework, *decision analysis* provides the best option out of a pre-defined set of scenario alternatives (e.g., Andersson and Destouni, 2001; Freeze and McWhorter, 1997; Massmann and Freeze, 1987). According to Massmann and Freeze (1987), „decision analysis is the branch of systems analysis that allows for the determination of the best alternative from a set of alternative courses of action". Thus, decision analysis is a special case of optimization, according to Aven and Kørte (2003). Nevertheless, these authors warn that optimization studies often neglect factors that influence the decision making. Some of these factors cannot be adequately captured in strict mathematical expressions, i.e. via utility values (Shrader-Frechette, 1991). A good reading on the differences between decision analysis and optimization is Freeze et al. (1990) and Tartakovsky (2013). This decision analysis framework is applied to the VIP framework to assess the reliability of drinking water supply safety by delineating more accurate wellhead protection zones as proposed by Feyen et al. (2003) (see Chapter 8).

Economic-based Decision Analysis

In economic-based decision analysis frameworks, risk managers are able to compare the costs ($\boldsymbol{C}$) and economic benefit ($\boldsymbol{B}$) between different risk reduction options (see Fig. 6.1). Lindhe et al. (2011) were the first to introduce an economic-based decision framework (cost efficiency ratio and cost benefit analysis) in drinking water supply systems to choose between several scenario-based decision options that reduce risk, measured in customer minutes lost. Tesfamichael et al. (2005) investigated the trade-off between economic benefits and health-risk for using atrazine in the agricultural sector. Andersson and Destouni (2001) introduced an abatement cost analysis to reduce concentration levels at a compliance boundary below a pre-defined threshold level.

In Fig. 6.1, the marginal economic costs Y (i.e., the marginal costs with respect to X) are constant with additional sampling X, which is denoted by the marginal cost curve u_1. With increasing sampling size or better sampling, the uncertainty in areal delineation is reduced. The amount of additional areal reduction Y decreases with increasing X. Thus, the additional information (benefit) retrieved from additional samples decreases as shown in Fig. 6.1 for the benefit curve u_2. Any scenario where the marginal benefit outweighs the marginal cost function is a good choice (e.g., James and Gorelick, 1994). Comparing several scenarios with each other will deliver the best scenario that is closest to the intersection of the marginal

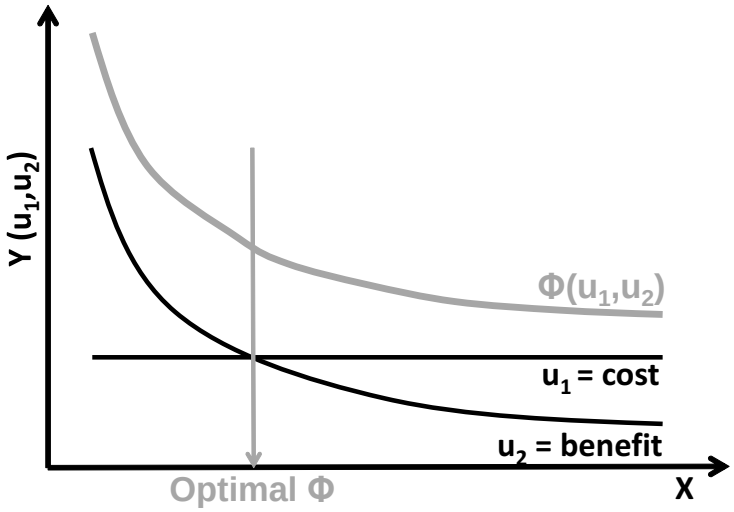


Figure 6.1.: Optimal risk concept and tradeoff between marginal cost of utility u_1 and u_2. Here, u_1 equals incremental sampling costs C_{samp} and u_2 equals the information gain measured as additional areal reduction (A_{err}) per sampling increment.

cost and marginal benefit (risk reduction) curves (u_1, u_2). In an optimization framework this leads to the following objective function Φ:

$$\Phi = |\boldsymbol{B}(X) - \boldsymbol{C}(X)|\,, \tag{6.1}$$

with $\boldsymbol{B}(X)$ being the marginal benefit and $\boldsymbol{C}(X)$ the marginal costs associated to each scenario X. The scenario yielding a value closest to zero of Eq. (6.1) delivers the best decision choice (see Fig. 6.1). Considering Eq. (6.1) shows, how closely decision analysis and optimization is connected.

The marginal costs Y of X can be calculated by any economic valuation technique, such as damage replacement costs, the hedonic price method, the bids received by consultants and so forth. The first two are indirect pricing methods, whereas the latter is a direct pricing technique incorporating stakeholders. According to Freeze and Gorelick (1999), costs in general are related to capital C_{cap}, operational C_{op} and costs C_f that are associated to sampling failures:

$$C_i = \int \frac{1}{(1+g)^t}\left[C_{cap}(t) + C_{op}(t) + C_f(t)\right]\mathrm{d}t \tag{6.2}$$

with g being the discount rate and C_i the total costs associated to the $i-th$ risk treatment alternative or scenario. A good reading to get a first overview on economic valuation is provided by Turner et al. (2004). There are valuation methods that help to transfer benefits $Y(X)$, which are not measured on a monetary scale, to the cost scale, such as market analysis methods. Often, the utility concept is used to relate benefits to costs by willingness-to-pay studies (e.g., Fishburn, 1970).

Within the VIP concept, economic value is determined by the reduction of areal demand for delineation (see Section 6.3.1). This area depends on the resolved degree of epistemic uncertainty and the chosen level of non-compliance, among others. According to Fig. 6.1,

a data acquisition campaign should be conducted to an increasing degree X, as long as the benefits u_2 (additional areal reduction) outweigh the cost curve u_1 for sampling (e.g., James and Gorelick, 1994). This type of comparison between information gain due to sampling and economic costs of sampling is known as data worth analysis (e.g., Freeze et al., 1990; James and Gorelick, 1994). The concept of data worth is illustrated within a synthetic case study (see Chapter 8). The worth of additional data to decrease uncertainty in terms of wrongly delineated area is also presented in Section 8.5. In this section, I will provide answers related to the following three questions:

1. How many measurements are needed?
2. What type of measurements are best?
3. How well does the sampling quality have to be?

6.2. Risk-aware Delineation of Wellhead Protection Zones

A classical approach to manage the risk of well exposure is to ban all hazards from a travel-time based area, which is known as delineating a wellhead protection zone. Questions by water stakeholders and managers such as (1) *'What is the current overall safety level of an existing delineated wellhead protection area?'* or *'Can we provide the same reliability level with a smaller protection area at lower costs than the existing one?'* (2) *'How much can the system reliability be improved without additional areal costs?'* and (3) *'What reliability level can be achieved at only little additional areal costs?'* (4) *'How much can the system reliability be improved with additional and more accurate data?'* will now be answered by providing four corresponding management options based on the VIP maps.
The four suggested management strategies are conceptually shown in Fig. 6.2. These strategies will help water stakeholders to support robust and transparent decisions under uncertainty, while keeping the additional costs relatively small or even leading to smaller delineated areas. Costs in this context are so-called areal costs associated with the area of land delineated as wellhead protection zone, as land has a monetary value.
The required size of delineated wellhead protection zones under uncertainty is easily larger than the actual size of the truly required wellhead protection area A_{true} (see Fig. 6.3). The delineated area depends on the degree of uncertainty (areal demand) and on the desired safety level $\beta = 1 - \alpha$ of the stakeholders. α is denoted as the risk acceptance level, i.e., the accepted probability that the system may fail. The choice of β or α is often limited by financial constraints of the water stakeholder. A no risk ($\alpha = 0$) situation is unachievable (e.g., Pollard et al., 2008). There always exists the chance of something unexpected happening to the groundwater, endangering safe water supply. Furthermore, delineation outline of $\beta = 100\%$ would lead to excessively large wellhead protection zones that are in conflict with the stakeholders interest in the delineated land. In particular, the agricultural sector is a strong stakeholder blocking many delineation processes in Germany[2].

[2]Personal communication with Mr. Müller, Ministry for Environment, Agriculture, Food, Wine and Forest of the State of Rheinland-Pfalz

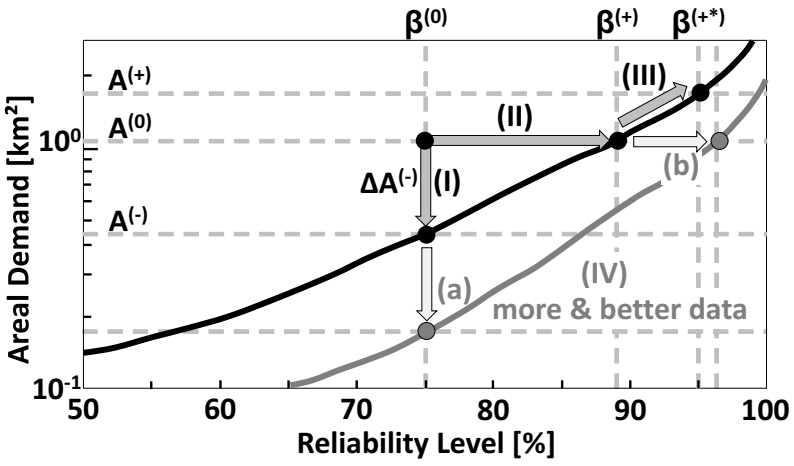


Figure 6.2.: Risk management with the VIP framework, showing increased areal delineation costs with increasing reliability level.

If a delineation process has not been considered yet, water managers are free in their decision to choose an acceptable probability level of non-compliance with the well vulnerability criterion (i.e. an (non-)exceedance probability level, depending on the respective criterion). The desired reliability level β^* defines a consistent outline of the wellhead protection zone, corresponding to the probability isolines of each VIP.

(I) Current Delineation Reliability Level and Smaller Consistent Area

As basis for further analysis, the current safety level of an existing wellhead protection area $A^{(0)}$ is assessed first. This is of special interest for water supply companies with no sophisticated post-treatment facilities, as they usually fully rely on perfect groundwater quality, or for supply wells, which have to be abandoned from the first day of contamination. The areal demand for source water protection increases with the desired reliability level β^*, as illustrated in Fig. 6.2. The area A_{β^*} needed to maintain this reliability level β^* is the integral over the entire domain Ω:

$$A_{\beta^*} = \int_{\Omega} I\left(P_1 \leq \beta^*\right) \mathrm{d}\Omega, \tag{6.3}$$

where P_1 is the exceedance probability for a given well vulnerability criteria, here peak arrival ($k = 1$), and a certain critical level WVC_{crit} according to Eq. (4.2).
Now, the maximum consistent safety level $\beta^{(0)}$ of an existing protection zone $A^{(0)}$ is defined. In case of an existing protection zone it is the β-value of the largest isopercentile, which is still fully enclosed by the wellhead area $A^{(0)}$.
In many practical situations, parts of the current protection zone $A^{(0)}$ will be located outside the VIP contour line $\beta^{(0)}$. Therefore, already a smaller protection zone $A^{(-)}$ with the same consistent reliability level $\beta^{(0)}$ as the current wellhead protection can be defined. This can be achieved by simply cropping the wellhead protection area exactly along the $\beta^{(0)}$ VIP line, freeing up an excess area $\Delta A^{(-)} = A^{(0)} - A^{(-)}$ (see Fig. 6.2).

(II) Area - Neutral New Delineation

As a possible next step, the area $A^{(0)}$ of an existing protection zone can be re-allocated within an optimized outline. To this end, the excess area $\Delta A^{(-)}$ from step (I) is re-invested towards an increased safety level $\beta^{(+)}$. The new outline will have the same area $A^{(0)}$ but with the largest possible consistent reliability level $\beta^{(+)}$ that can be achieved with $A^{(0)}$. The new outline is found by choosing a consistent safety level $\beta^{(+)}$, such that the new area $A^{(+)}$ is just as large as the old area $A^{(0)}$, but the new outline follows exactly a VIP line β^* (see Fig. 6.2):

$$\beta^{(+)} : A^{(+)} \equiv A^{(0)}. \tag{6.4}$$

(III) Reliability Increase at Minimum Costs

Reallocation or cropping of wellhead protection zones is not always possible, e.g., due to political interests, or due to existing long-term rights and concerning land use. Also, the general public, regulators and customers might not accept the fact that some parts of the existing protection zone are removed in the above management option, as there may or may have been other plausible or even possibly subjective (based on the public risk perception) reasons behind the existing delineation. Please remember, wellhead delineation is enforced by the regional municipalities in Germany (see Section 3.4). Even in such a complex local situation, the probabilistic reliability information contained in the VIPs may be used to improve the safety of water supply, without competing against the locally established need for the current protected area. The key argument is the model-based and probabilistic character of VIPs, and the plain, yet objective, quantitative and transparent language spoken by the contour maps.
Therefore, the optimized area $A^{(+)}$ from step (II) is complemented by minimal additional area in order to increase the reliability level $\beta^{(+)}$ to some larger level $\beta^{(++)}$. The area $\triangle A^{(+)}$ needed for this improvement is exactly the area between the VIPs with the two safety levels $\beta^{(+)}$ and $\beta^{(++)}$. Thus, the cost function for the increase of reliability $\triangle\beta = \beta^{(++)} - \beta^{(+)}$ is directly given by $\Delta A^{(+)} = A\left(\beta^{(++)}\right) - A\left(\beta^{(+)}\right) = A^{(+)} - A^{(0)}$ in Fig. 6.2.
An intermediate alternative is to keep $A^{(0)}$ without cropping and to complement it with pieces $\triangle A^{(*)}$ that are missing in order to reach any desired safety level $\beta^{(*)} > \beta^{(0)}$. The corresponding cost function is more complex and cannot be plotted in a generalized manner (as a function of isopercentiles and their enclosed area only), as it depends on the actual geometry of $A^{(0)}$. However, it would better respect the specific local situation, and leave the other plausible and possibly subjective arguments untouched.

(IV) Increased Reliability Level with Additional and More Accurate Data

The area sacrificed to uncertainty can be reduced, when additional, more accurate and the most beneficial type of data is incorporated to the delineation model. This inevitably leads to the following two questions:

(a) What is the best catchment investigation strategy to achieve a pre-defined reliability level at lowest costs?

(b) What is the highest achievable reliability level at no additional costs?

Fig. 6.2 illustrate these two questions schematically. With additional data, the reliability of the wellhead delineation (x-axis) is expected to increase up to a certain limit at no additional or only small pre-defined costs (Fig. 6.2, IV(b)). Likewise, a given reliability level can be achieved with less areal demand, when uncertainty in areal delineation is reduced (Fig. 6.2, IV(b)). Each sampling design would result in an additional risk-aware delineation curve (gray curve) and thus add to the increased information gain. The benefit of using additional and better data decreases with more sophisticated sampling techniques and more data (distance between lines). Thus, an uncertainty measure is defined that can follow the marginal benefit curve given in Fig. 6.1 in order to know how beneficial additional or better data is.
In this study, this is done for sampling type, quality of measurement and sample size and illustrated for a synthetic test case (Section 8.5). In order to find an optimal sampling design case in light of monetary costs, decision analysis is employed to identify the best scenario for question (a) or (b) among pre-defined sampling designs. Given the fact that stakeholders pre-define an acceptable safety level (compare Fig. 3.2, step: risk objective), it is now possible to find the cost-optimal situation, where the marginal costs for sampling outweigh the benefit in reducing areal demand as far as possible (see IV(a), Fig. 6.2). The same is true for fixed costs, where additional data pay for themselves in order to improve system reliability β (see IV(b), Fig. 6.2).
The concept of risk-aware delineation of wellhead protection zones (I-III) is demonstrated for a real test site in Chapter 9. The uncertainty reduction due to additional and better data is demonstrated along a synthetic test case in Chapter 8.

6.3. Uncertainty and Economic Risk Measures within the VIP Framework

In order to measure the effect of uncertainty reduction in the current problem setting for wellhead delineation, I introduce two area-based utility functions that measure the amount of excessive delineated area sacrificed to uncertainty (Section 6.3.1). The first uncertainty measure is based on the delineated *areal error*. The second measure estimates the area that is enclosed between two vulnerability isopercentile lines. Thus, the transfer into economic analysis is straightforward (see Section 6.1), as areal demand can be priced. As direct economic measures for risk, I will consider customer minutes lost (Section 6.3.3) and damage replacement costs (Section 6.3.4).

6.3.1. Uncertainty Measure: Areal Delineation Error

The benefit of using additional or better data in Fig. 6.1 is expressed by epistemic uncertainty reduction ΔU, leading to VIP lines that are closer to reality. The error of delineation, A_{err}, is

the excessive delineated area A^+ plus the undelineated area A^- (see Eq. 6.5) compared to the truly required (but unknown) protection area A_{true}. A^+ is part of the current delineated area A_{curr} for the sake of confidence (see Fig. 6.3). A^- is the area that is not delineated although it is part of the actually required area, A_{true}. The light gray and white area in Fig. 6.3 indicates the area, that is truly required A_{true} and the dark gray and white area indicates the currently delineated area A_{curr}. The two partial errors can be combined to the overall error A_{err} by:

$$A_{err} = \omega_1 \cdot A^- + \omega_2 \cdot A^+. \tag{6.5}$$

with ω being the weights assigned to the two different types of area delineation error. Inter-

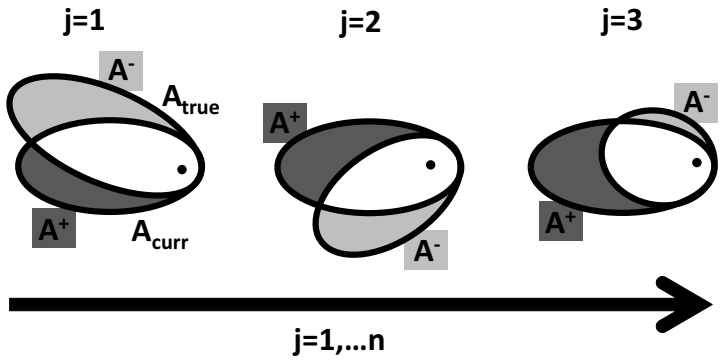


Figure 6.3.: Schematic concept of delineation error due to epistemic uncertainty.

preting the areal error in the light of human-health risk, the non-delineated error area A^- is more severe (because contamination can enter the well unnoticed) than area A^+ that is protected but located outside the actual wellhead protection zone. For simplicity, and avoiding discussion on how to value health in monetary terms, I keep these weight equal $\omega_1 = \omega_2$ for the remainder of this thesis, and omit them from notation.
Within a Monte Carlo context, the expected areal delineation error U as a measure for uncertainty can be calculated as

$$U = \frac{1}{n_r} \sum_{j=1}^{n_r} \omega_j \cdot A_{j,err}, \tag{6.6}$$

with n_r Monte Carlo simulations and ω_j being the weight of the $j-th$ realization as it may arise from Bayesian updating. Here, I will assume equal weights for each model realization as each ensemble-based realization is equally likely in this thesis by rejection sampling (see Section 5.3.1). The ensemble-averaged outline A_{50} is denoted as the actual outline A_{curr}. The smaller the delineation error, the more accurate the delineation is and thus the smaller the measured uncertainty (see Eq. (6.6) and Fig. 6.2, IV(a)). This uncertainty measure is used in the remainder of this sampling decision analysis framework and in assessing the data worth.

6.3.2. Uncertainty Measure: Area-Enclosed by the VIP Lines

Another uncertainty indicator measures the areal demand in delineation between the $10^{th}-$ and $90^{th}-$percentile contours of the well vulnerability criteria, normalized by the area within the $50^{th}-$percentile contour:

$$U = \frac{A_{10} - A_{90}}{A_{50}}, \tag{6.7}$$

where U is some measure for the fraction of the area a planner has to sacrifice due to uncertainty. This definition of U is related to the idea of interpercentile range, normalized by a median (Enzenhoefer et al., 2012).

6.3.3. Customer Minutes Lost as Economic Risk Measure

Customer minutes lost (CML) is a performance indicator for water managers to measure the lack of water supply due to quality failure (e.g., a contaminant concentration within the well above drinking water standard) or quantity failure (e.g., due to water pressure loss) in minutes per year. The measure, customer minutes lost, is widely used for drinking water supply systems (e.g., Blokker et al., 2005) and reads within the VIP framework after Lindhe et al. (2009):

$$CML(\boldsymbol{x}_0) = t_{exp}(\boldsymbol{x}_0) \cdot P_e(\boldsymbol{x}_0) \cdot n_{r,co_{aff}}, \tag{6.8}$$

with P_e being the entry probability of a spill event occurring at location $\boldsymbol{x}_0$. The exposure time t_{exp} indicates the damage intensity in the sense of minutes or days, when the well is out of operation due to quality failure. $n_{r,co_{aff}}$ is the relation between affected consumers ($n_{co_{aff}}$) and all possible consumers ($n_{co_{all}}$) within the supply system, which also denotes a consequence attribute. The number of consumers being affected, $n_{co_{aff}}$, depends on the current water demand within the supply system. In the unlikely case that pumping is constant over time, but traditionally assumed for wellhead protection delineation, the consumer ratio $n_{r,co_{aff}}$ is:

$$n_{r,co_{aff}} = \frac{d_{min}}{d}, \tag{6.9}$$

with d_{min} being the pre-defined amount of water demand per consumer standardized by the actual demand d of an individual consumer for each model realization.
As the additional information through data sampling reduces uncertainty, VIP lines move closer to reality. Thus, the risk measure, CML, should become more realistic with improved sampling schemes. An illustration of customer minutes lost in the light of uncertainty reduction is provided in Section 8.5.

6.3.4. Damage Replacement Costs as Economic Risk Measure

According to Turner et al. (2004), there are several economic valuation techniques available to transfer different mitigation options to a common cost scale. Here, I adopt the method of replacement costs. It calculates the avoidable damage costs related to each mitigation

alternative. Here, the damage replacement cost method is a function of well down-time, t_{exp}. Due to the spatial information provided by the VIP maps, it is possible to calculate the expected damage D for contaminant release at each location $\boldsymbol{x}_i$ across the domain:

$$D(\boldsymbol{x}_i) = t_{exp,i} \cdot Q_p \cdot \gamma, \tag{6.10}$$

with Q_p being the well pumping rate and γ representing an a-priori defined water prizing function. This water prizing function can be constant, such as the given water market price or a more sophisticated cost function. A demonstration of using damage replacement costs in a decision analysis framework for risk mitigation strategies is given for a synthetic test case in Section 8.6.

6.4. Findings of VIP-based Risk Management

- Decision analysis frameworks transparently reflect uncertainty with associated costs, such that optimal, objective and transparent decision support is available among a set of pre-defined scenarios.
- The reliability of wellhead delineation can be improved with no or only minimal costs, when considering VIP-based re-delineation.
- More and better data, such as data type and quality of sampling, reduce parameter uncertainty and thus help to increase reliability levels at no additional costs. Alternatively, costs can be reduced at a constant safety level.
- The first well vulnerability criterion, t_{peak} is used to assess the areal demand in wellhead delineation and thus indirectly used to measure uncertainty.
- The fourth well vulnerability criterion, t_{exp} is suitable for economic risk analysis, and allows to calculate damage replacement costs and customer minutes lost.

7. STORM - A Risk Quantification Framework

This chapter introduces the forward risk assessment and management framework, called STORM. STORM is an acronym for **ST**akelholder-**O**bjective **R**isk **M**easures. The name indicates the flexibility of the risk framework to calculate risk measures that are of specific interest for various stakeholders. Two key points that motivated STORM are introduced in Section 7.1.
As already presented in Chapter 2, Fig. 2.1, STORM is modularized into ten modules to quantify and manage risk from source to receptor. The regional risk assessment is penetrated by uncertainty, such as frequency of hazard failure, type and quantity of resulting contamination loads, subsurface heterogeneity, and so forth. Therefore, it is necessary to quantify uncertainty and support stakeholders to take robust and meaningful decisions in an uncertain problem setting based on rational and transparent modeling results.
The prerequisite for applying STORM is a hazard database obtained by hazard identification (Section 7.2). The ten modules that make up the STORM framework are set into a structured framework, where each module represents a sub-model within STORM (see Fig. 7.1). These ten sub-models can be classified into four main modeling steps, including an (1) *event model*, (2) *immission model*, (3) *transport and impact model* and (4) *risk evaluation model* (see Fig. 7.1).
In more detail, STORM considers the impact and the mean time to failure of individual hazards (modules 1-2: *event model*, Section 7.3), the retarding process through the unsaturated zone (modules 3-4: *immission model*, Section 7.4), contaminant transport to the drinking water well via a physically-based flow and transport model (modules 5-8: *transport and impact model*, Section 7.5) and cumulative risk estimation with prioritization of the most severe risk sources (modules 9-10: *risk evaluation model*, Section 7.6).

7.1. Introduction to STORM

This section will summarize and extend the motivation and key aspects that lead to the development of STORM. Water utility managers and stakeholders are required by regulations, e.g., the water safety plan (Davison et al., 2005), to perform risk analysis, prioritize hazards and thus control risk within the catchment. Risk management within the catchment is often performed in a forward mode, taking the complex hazard properties into account and assessing their impact to the receptor. Two key challenges arise.
First, the hazards within the catchment have properties that differ depending on the hazard type (e.g., chemical properties, spatial and temporal release pattern). For example, hazards may include natural risk phenomena or human activities, such as industrial, agricultural,

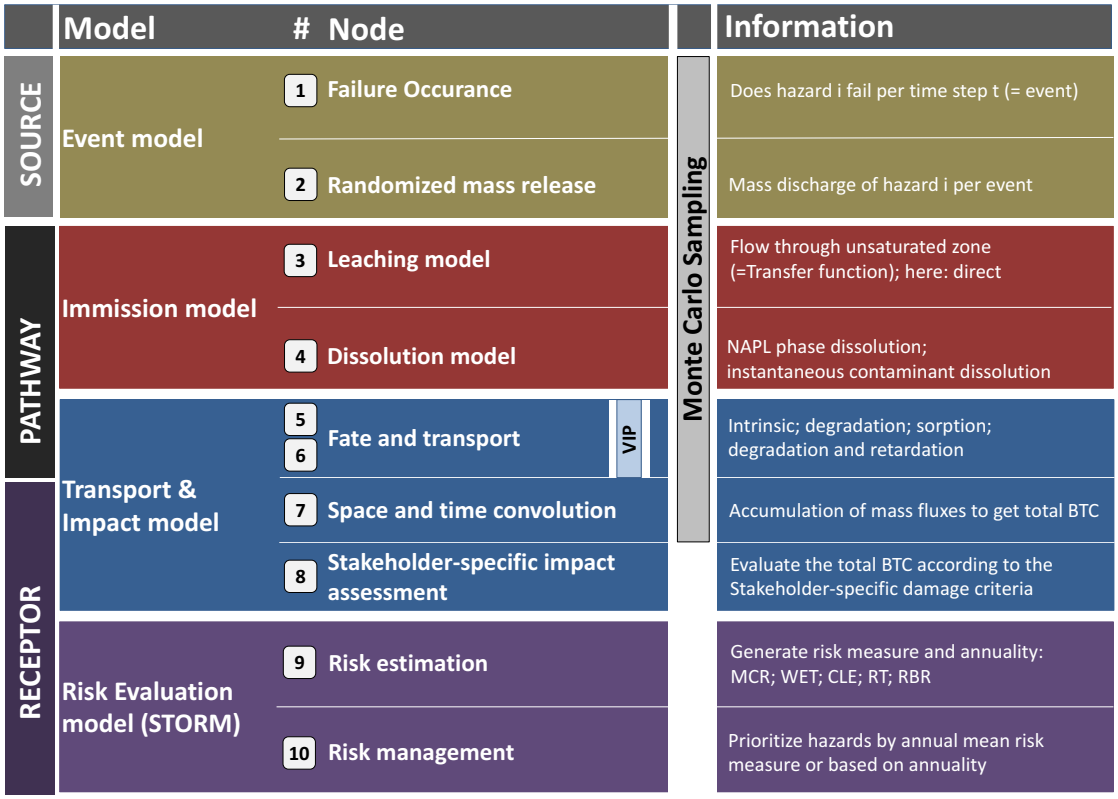


Figure 7.1.: Stakeholder-specific mass discharge-based risk assessment concept (STORM), showing ten individual modules (=nodes) separated into four major model segments, following the source-pathway-receptor model.

transport-related hazards and geogenic hazards (see Fig. 7.4). These risk sources are spatially distributed across the catchment, fail at different times, lead to different temporal impact distributions at the receptor and are of different contaminant type. Still, in the sense of cumulative risk assessment, their individual impacts have to be considered to measure overall risk. Therefore, smart aggregation techniques beyond simple risk summation are in demand. More details are presented in Section 7.1.1.
Second, different risk management objectives with different underlying risk definitions may lead to contradicting management actions. For example, MacGillivray et al. (2006) state that the overarching public health goal may be in conflict with goals regarding economically efficient water supply. The German guideline DVGW (2009) introduces a risk-based and process-oriented management for technical, economical and health-related hazards. Water managers and stakeholders have different interests in groundwater and thus follow activities that may be conflicting within the well catchment. The hazard ranking, risk perception and derived risk management strategies depend on the stakeholder-objective view for a specific situation, such that human-health risk assessment needs different risk measures than assessing the technical performance. More details are presented in Section 7.1.2.
Due to these two above mentioned reasons, *cumulative impact assessment* and *multi-objective stakeholder views*, a flexible risk quantification and management framework is in demand.

STORM aims to prioritize hazards within a cumulative risk setting in ten modules and investigates the factors that influence this ranking. These factors may include different risk perception, different critical cut-off levels or different risk objectives. Chapter 10 demonstrates the influence of these three factors in an application.

7.1.1. Mass-Discharge-based Risk Aggregation

Most aspects in risk aggregation, such as summation or worst-case scenarios, have been discussed in Section 3.6. As a summary of that section, aggregation has to be based on the mass discharge at the receptor. It accounts for hazards that are distributed across the whole catchment and for simultaneous contaminant well arrival, although hazards have failed independently of each other at different times. These two aspects have been already investigated in literature (see Tait et al., 2004; Troldborg et al., 2008).
Aggregation across contaminant types, such as chlorinated solvents and BTEX compounds, is less profound and only available by introducing the concept of utility theory (e.g., Fishburn, 1970). The basic principle is to unify the consequence unit, such as expressing the severity scale in terms of costs, disability-adjusted life years (DALY), customer minutes lost (e.g., Lindhe et al., 2009), cancer risk (e.g., Freeze and McWhorter, 1997), hazard quotient (e.g., US EPA, 2007) or other risk estimates. Here, I introduce four mass-discharge-based well vulnerability criteria as utility values that can serve as *intermediate risk* estimates. Intermediate risk estimates allow the quantification of transport-based risk measures that comply with stakeholder objectives (see next Section 7.1.2). These risk estimates host aggregation information only on mass-discharge level (module 8). Summation of these vulnerability-based risk levels (utility values) across hazard types is prohibitive in a non-linear risk situation. Thus, threshold-based risk values can only be aggregated by statistical aggregation of the underlying mass fluxes (module 9).
Tait et al. (2004) state that not the time of failure, but the temporal arrival of contaminant at the well is important for risk assessment. In fact, this is true as long as hazards only fail once. In all other cases, the spatial and temporal arrival at the well and the arrival due to periodic failure have to be considered. Therefore, the novelty in STORM is aggregation of recurring hazard failures over time, such as combined sewage overflow due to heavy rainfall events, regular fertilizing over a well catchment lifetime (e.g., Kourakos et al., 2012) and fecal transport (e.g., Page et al., 2012) from deer feeding places or cow pastures. The concept of recurring events is commonly accepted in flood risk and reliability engineering. In flood risk management, the severity is plotted versus the recurring interval. Thus, risk managers are able to install mitigation measures or emergency plans based on their risk perception and accept system failure once in a prescribed (large) recurrence period. Risk managers using tools of reliability engineering, such as fault-tree analysis, calculate the expected number of failure events for a given time period. Lindhe et al. (2009) estimated the expected failure in minutes within a one hundred year time period. Here, I'll introduce a mean annual impact measure, where the impact measure is defined by a corresponding stakeholder objective (Section 7.1.2).
Section 7.5 introduces module 8, convoluting mass discharges of hazards at the receptor level (space, time and frequency aggregation). Module 9 (annuality and annual mean risk

measure) is presented in Section 7.6, showing statistical mass-discharge-based aggregation across contaminant types. Module 10 (see Section 7.6) introduces prioritization of individual hazards within a cumulative risk situation.

7.1.2. Vulnerability-based Risk Objectives

Stakeholders that participate in risk assessment for drinking water catchments have different opinions and interests, such that also their objectives differ. Öberg and Bergbäck (2005) state that there exist two risk analysis interests in environmental engineering, ecological (e.g., McKnight et al., 2012) and human-health risk (e.g., Freeze and McWhorter, 1997). As previously discussed, technical and economic risk analysis is as relevant as the former two. These different stakeholder views may possibly lead to different rankings of risk sources within the catchment and depend on the impact unit used to estimate risk.

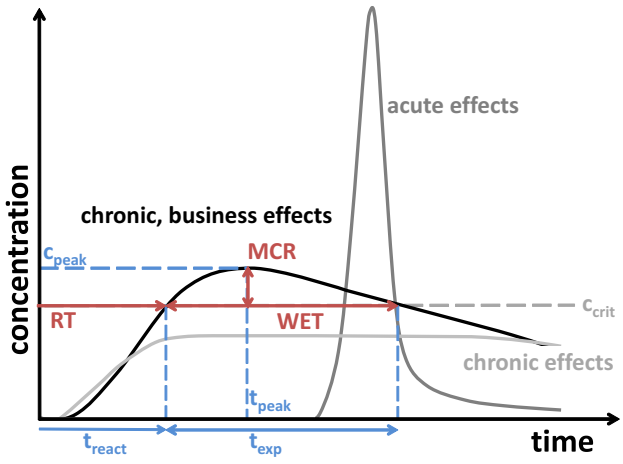


Figure 7.2.: Adverse effects triggered by type of contaminant breakthrough in the well water, leading to stakeholder-objective-risk measures.

A similar situation occurs, if one stakeholder follows more than one suggested risk objective. This is often tacitly assumed, when asking for a unified vulnerability-based risk measure or when different risk measures are compared to each other (e.g., qualitative with quantitative ones). The resulting dilemma becomes very obvious, even when prioritizing only individual hazards without any aggregation. For example, human-health related questions may focus on concentration levels, assessing the acceptable or chronic daily intake rate (see Eq. 3.1) as proposed by the US EPA (1989). Hazards that lead to a high and extended concentration dose in the supplied water are prioritized. Nevertheless, in an early-alert system, these hazards may be less prioritized due to longer reaction times, t_{react}. In this case, hazards that threaten the supply safety with shorter arrival time, even at lower peak concentration levels, are prioritized (see Fig. 7.2). Therefore, I claim that hazard prioritization is only meaningful, if it consistently considers only one pre-defined risk objective (see Section 3.2). The „unifying“ idea is proposed for further research, by casting the different stakeholder views into a

multi-objective optimization framework.
To allow a flexible analysis, to address the needs of many stakeholders, and to allow for multi-objective approaches in later research, it is necessary to provide a framework that is flexible to adapt to the risk objectives demanded by different stakeholders. A similar idea has been followed by de Barros et al. (2012). They introduced environmental performance metrics, accounting for many objectives separately in one model. Here, the mass-discharge-based well vulnerability criteria provide the necessary information to derive all transport-related risk measures (see Fig. 7.2) and account for the multiple stakeholder objectives. For example, from the intermediate risk level of well down-time (WET), the performance of risk measure „customer minutes lost“ (CML, e.g., Lindhe et al., 2009) can be derived. Cancer risk (e.g., Rodak and Silliman, 2012), human-health-related hazard quotients (HQ, US EPA, 1989) or DALY can be derived from the cumulative contaminant load for a certain time period (Contaminant Load Exposure, CLE). Toxic units (TU, e.g., McKnight et al., 2012) can be derived from the maximum concentration load (MCR) and so forth.
As mentioned before, another advantage of using well vulnerability criteria in the STORM context is that they allow aggregation of risk sources across different contaminant types (see Section 7.5). Contrary, the information value of a chosen STORM (e.g., well exposure time, maximum concentration ratio) strongly depends on the hazard types (e.g., long-term or pulse source) that may pose a risk to water supply in the catchment. Tab. C.1 provides a brief overview on STORM and which metrics are suitable to assess the risk of certain hazard types. More details will be provided in Section 7.5.

7.2. Hazard Identification

The STORM approach assumes as a prerequisite, that sufficient information about the hazards within the catchment is available, as a result of the hazard identification process (see Fig. 3.2). Drinking water is prone to human activities and natural phenomena within the catchment area. There are agricultural activities (e.g., fertilizing), communal infrastructures (e.g., sewage treatment plants), industrial activities (e.g., varnishing) and many other risk sources threatening safe drinking water supply (see Fig. 7.4). Therefore, as a first step in risk analysis, water stakeholders are asked to identify past and present risk sources (e.g., DVGW, 2009) that may have or will alter the groundwater quality.
The available information for each risk source h is documented within GIS-systems or other file types, thus constituting a hazard database. The stored information is discussed, item by item, in the following subsections. STORM uses an interconnected database to store information on contaminant properties and the hazardous site itself (see Fig. 7.3). Furthermore, information on the *spatial*, *temporal* and *chemical hazard type* is stored to support and control the individual sub-elements of STORM (modules 1 - 10). All aspects related to source geometry in the subsurface are neglected in the current research, but could easily be included due to the modular approach. For a detailed discussion on how source geometry influences risk levels, please see Thomsen et al. (2012).

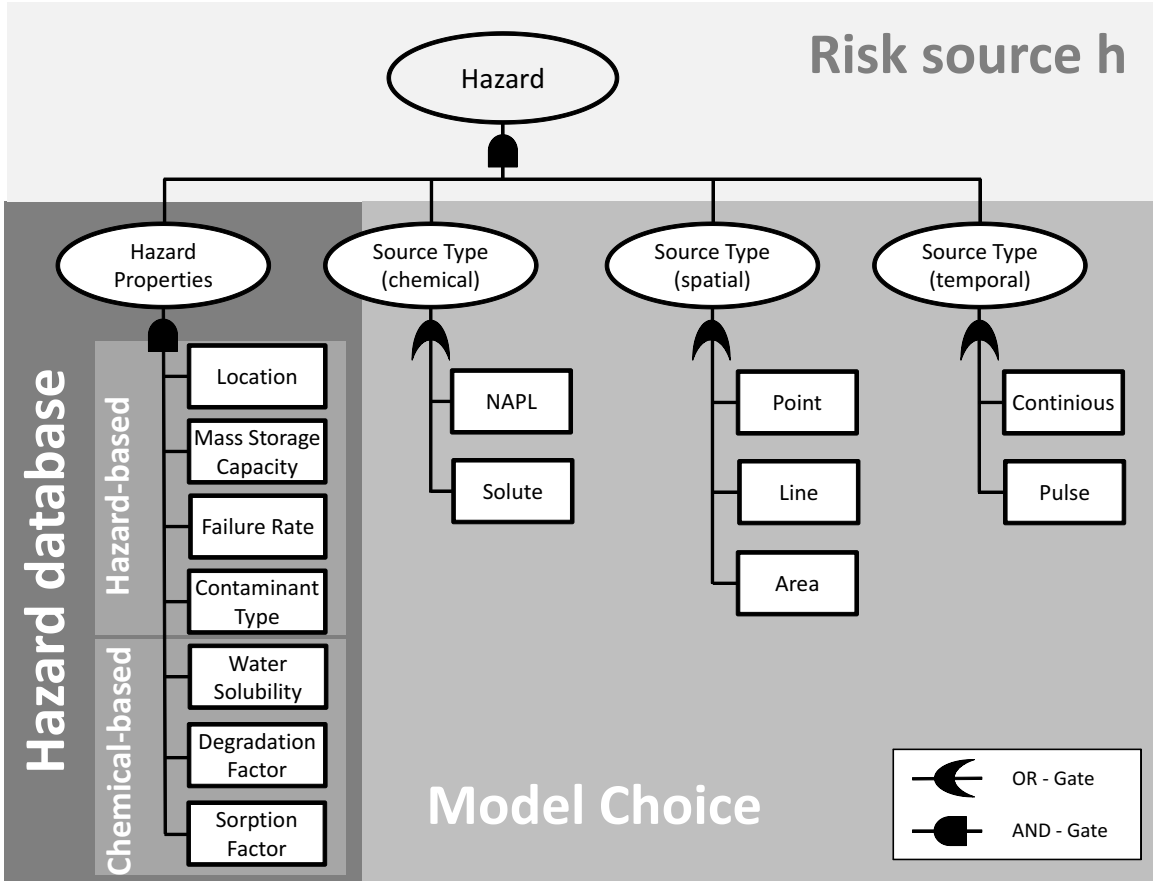


Figure 7.3.: Necessary hazard information to assess hazard-specific mass flux estimates at the supply well, by first retrieving hazard property information from a hazard database and secondly choosing the individual physically-based models within the logical risk assessment model.

7.2.1. Hazard Database

The hazard database contains all information collected for each risk source h. The hazard database stores information on all hazard locations, such as the amount of mass m_0 being handled, the location $\boldsymbol{x}$ a risk source is mapped to, the failure frequency λ for a given time unit, and which contaminant type j is stored at the risk location. A risk source may contain several contaminants j, which are treated independently as individual risk sources h in the hazard database. Within a second database, the contaminant-specific transport properties such as water solubility, retardation and degradation factors of each contaminant type j are stored. Thus, both databases are linked to each other. Fig. 7.4 shows several risk sources that are categorized according to spatial (column) and temporal (row) hazard type. The additional information on *spatial*, *temporal* and *contaminant hazard type* triggers the model choice within the individual modules of STORM.

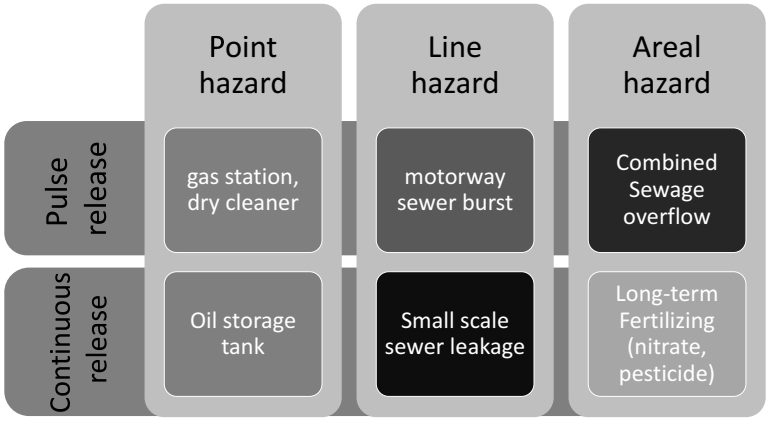


Figure 7.4.: Examples of risk sources, categorized according to spatial and temporal release pattern.

7.2.2. Spatial Hazard Types

Kourakos et al. (2012) showed that non-point sources need a different treatment than point sources in monitoring, concentration level assessment and implementation of regulatory frameworks. Indeed, transport time statistics and magnitude of impact at the well depend on the location and areal distribution of the contaminant release. Therefore, the risk assessment approach has to be flexible to adapt to the spatial properties of hazards. Risk assessment of contaminated sites often uses point sources, neglecting the aspect of line or areal sources (e.g., Tait et al., 2004; Troldborg et al., 2008). Punctual risk sources, such as underground storage tanks, gas stations, sewage treatment plants, etc. are mapped in STORM to one location and to one discretization cell of numerical transport models (see Fig. 7.4). *Line hazards*, such as pipe networks, roads, etc., may fail at independent locations at different times, acting as individual point sources after failure. However, they may also emit contaminant loads along an entire line segment that spans several grid cells. *Areal risk sources*, such

as farm fields, cemeteries, etc., release harmful mass over a larger area that often extends over many discretization cells of the numerical transport model. This inevitably leads to different arrival times and impact intensity behavior at the well for a single hazard, even without considering dispersion effects.
Hazards that follow a line source are discretized into $l = 1, \ldots, L$ single points. The initial mass $m_{0,h,j}$ released from a risk source at each location $\boldsymbol{x}_l$ is assumed to be constant for line hazards. The failure rate per location l decreases uniformly to:

$$\lambda_{h,j}^{l} = \frac{\lambda_{h,j}}{L}, \tag{7.1}$$

Nevertheless, it is possible to assign higher failure rates $\lambda_{h,j}^{l}$ to individual locations l and vice versa, such as representing dangerous road sections, corroded pipe material and so forth. This procedure allows failure of one single risk source at different locations at the same time. This might be true, e.g., for sewage canals that leak at more than one position at the same time.
Areal risk sources that release contaminant mass on a larger area than individual discretization cells are treated the same way as line hazards (e.g., Srinivasan et al., 2012). However, the mass $m_{0,h,j}$ being released due to an event is now distributed across all L locations, leading to a lower mass release potential $m_{0,h,j}^{l}$. The likelihood of failure for each location stays constant ($\lambda_{h,j}^{l} = \lambda_{h,j}$). After pre-processing the risk source according to the spatial hazard type (line/area), the consecutive modules 1 to 10 of STORM can be performed as if all hazards were point hazards. For prioritization of hazards (module 10), the individual point sources of a line and area hazard are convoluted to assess the total impact by the spatially hazard source on the overall risk.

7.2.3. Temporal Hazard Types

Hazards not only vary in their spatial extension, but also in their release duration (see Fig. 7.4). This is most obvious by comparing a rather short event, such as a sewage overflow due to heavy rainfall, and a more extended event, such as leachates from municipal waste landfill sites. The temporal distribution of mass release influences the level of risk and the strategy to encounter risk by available management options. Pulse events may have a higher impact at the receptor, but are limited in duration, whereas continuous releases are often less severe in their resulting concentration levels, but more severe in their duration of impact. The most tragic combination is long-term contamination at unacceptably high concentration levels. According to Fischhoff (1990), these long term pollutions may lead to closure of the facility, depending on the risk perception of water stakeholders, the capability of post-water quality treatment and the economic feasibility to deal with this long-term contamination.
Pulse hazards, such as animal fecal releases, truck accidents, etc., are of short duration, enabling stakeholders to manage risk on a shorter time scales, e.g., compensating well downtime by prior water storage. This is different with chlorinated solvents and other contaminants with slow dissolution, high retardation rates and constant sources. Such contaminants can possibly act as sources for long-term contamination (e.g., Chambon et al., 2011). Thus, risk management options may unfold their effectiveness only after years.

Contaminant types of continuous hazards, such as leakage of municipal waste sites at far distance to the well, may lead at some point to long-term contamination with low concentration levels. In general, these background concentrations are often negligible in risk assessment as being too small to directly cause harm. Nevertheless, it is important to notice that background concentration reduces the distance to a critical concentration level c_{crit}. Thus, the dilution potential of the pumped drinking water to cope with contaminants of the same type is lowered, and the risk of concentration levels induced by the hazards within the same chemical class exceeding a given threshold level increases.

7.2.4. Chemical Hazard Types

There are many different types of contaminants, ranging from microorganisms to chlorinated solvents. Tait et al. (2004) and Verreydt et al. (2012) assume contaminants as solutes. Troldborg et al. (2008) account for non-aqueous phase liquids (NAPL) that form either pools on top of the water table (LNAPL) or below the groundwater table (DNAPL). These pools lead to retarded and extended concentration levels at the well, as only a small amount of NAPL per time is dissolved into the by-passing groundwater flow. Therefore, transport-based risk estimates are sensitive to the chemical characteristics of hazards that dictate its miscibility with water.
The hazard database stores information on chemical properties of individual contaminant types (chemical database). Furthermore, it allows for each hazard to set the sub-model choice for dissolution into groundwater.

7.3. Event Model

The event model is modularized into two sub-models - the failure model and the mass release model. Both models account for parameter uncertainty.

Module 1: Failure Model

The failure model characterizes the probability that a risk source may be subject to a failure event (see Fig. 7.1, module 1). A failure event F is the situation that a risk source is not properly functioning at a given point in time. The failure event may possibly cause a negative effect at the receptor, depending on the amount of mass being released, the attenuation and degradation potential of the aquifer and the amount of clean water being pumped parallel to the contamination load.
The event occurrence follows a stochastic process, which is used to simulate the total number of events and the time of occurrence. The stochastic process is expressed by a time-dependent failure rate $\lambda(t)$. Failure rate is not a probability of failure occurrence and thus can be larger than one. It is denoted as

$$\lambda(t) = \frac{p_f(t)}{\beta(t)} = \frac{p_f(t)}{1 - P_f(t)} = \frac{P_f(t) - P_f(t + \Delta t)}{\Delta t \cdot P_f(t)}, \tag{7.2}$$

with $p_f(t)$ being the time-to-failure distribution (i.e., failure density function) and $\beta(t)$ being the reliability or survival function. The reliability function describes the probability of no failure before time t and is defined as $\beta(t) = 1 - P_f(t)$, with $P_f(t)$ being the cumulative failure distribution (Bedford and Cooke, 2001). The failure occurrence and the time between two events depend on the lifetime function $P_{f,h}(t)$ of the risk source h. The failure rates over lifetime follow a bathtub curve with three distinct failure periods assigned to a component (hazard), called „infant mortality“, „useful life“ and „wear-out-time“ (e.g., Wilkins, 2002). There are several lifetime functions available, such as exponential, beta and Weibull distributions. The last two mentioned functions are capable to capture failures by a lifetime-dependent failure rate $\lambda(t)$.
For simplicity, I only consider the useful life period, neglecting the fact of high failure rates at the beginning and at the end of a hazard lifetime. This is commonly accepted in reliability engineering (e.g., Bedford and Cooke, 2001; Lindhe et al., 2009). In fact, risk sources that are constantly maintained to state-of-the-art technologies operate over a long time in their useful life period. As the failure density is nearly constant over this life period, one can assume a constant failure rate λ, such that the stochastic process follows a Poisson process. Therefore, the time intervals between events are assumed to be independent of each other and the probability distribution $P_f(t)$ for the failure-free time until the next event follows an exponential distribution:

$$P_f(t) = \int_0^t \lambda e^{-\lambda\tau} d\tau = 1 - e^{-\lambda t} \tag{7.3}$$

The advantage of Eq. (7.3) is the simple use of constant failure rates to describe the probability of system failure, with a smaller data demand than in the more complex models.

Module 2: Mass Release Model

The amount of contamination being released to the subsurface after an event at any hazard has occurred is often uncertain (see Fig. 7.3, module 2). Thus, the mass being released can be described by a probability distribution function $P(m_{0,h,j}(\boldsymbol{x}))$. For simplicity, I assume a uniform probability distribution of mass being released to the unsaturated zone, specified by a lower and upper bound.

7.4. Immission Model

The immission model characterizes the flow through the unsaturated zone. The immission model consists of two sub-models. One characterizes the flow through the vadose zone, the other one describes the dissolution process of the contaminant mass to the aquifer.

Module 3: Leaching Model

Contaminant mass is typically released onto the surface or within the unsaturated zone up to the dissolution into groundwater. Then, it propagates through the unsaturated zone to

the aquifer (see Fig. 7.1, module 3). The contaminant transport through the vadose zone depends on the complex interaction between different factors, such as hydrogeological properties (soil type, porosity, layer thickness, etc.) and contaminant types (retardation and degradation factor, density, etc.). Several models exist to approximate the travel time through the unsaturated zone, ranging from simple transfer functions (e.g., Tait et al., 2004) to complex physically-based models with non-linear functions (e.g., Pollock et al., 2002). Sousa et al. (2012) investigated within their study the ratio between travel time through unsaturated zone and total travel time ($t_{tot} = t_{unsat} + t_{sat}$). They concluded that considering unsaturated transport time is only reasonable, when there is a large time lag ratio and when enough data is available on unsaturated soil parameters. For simplicity, I neglect the transport segment through the unsaturated zone, but future extensions are straightforward. However, the unsaturated zone often has a large degradation potential.

Module 4: Dissolution Model

Contaminants reaching the aquifer are either instantaneously dissolved within groundwater or pool up as NAPLs (see Fig. 7.1, module 4). In case of non-aqueous phase liquids, the groundwater discharge in the aquifer dissolves the entrapped NAPL phase, depending on the water solubility of the contaminant and the cross-sectional area of the NAPL source zone. Here, I assume the Darcy flux $\mathbf{q}$ in flow direction to be perpendicular to the cross-sectional NAPL area A_h, leading to the mass discharge $\dot{m}_{h,j}$:

$$\dot{m}_{h,j}(\boldsymbol{x}_h) = \int_{A_h} \boldsymbol{v}_f(\boldsymbol{x}_h) \cdot c_{sat,j} \mathrm{d}A, \tag{7.4}$$

with $c_{sat,j}$ being the saturated concentration, depending on the water solubility of contaminant j. The duration $t_{NAPL,h,j}$ of mass release of this particular event is given by

$$t_{NAPL,h,j}(\boldsymbol{x}_h) = \frac{m_{0,h,j}(\boldsymbol{x}_h)}{\dot{m}_{h,j}(\boldsymbol{x}_h)}. \tag{7.5}$$

7.5. Fate and Transport Model

The fate and transport model describes the contaminant transport from the point of release in the aquifer to the receptor. The model consists of three sub-models. The first sub-model follows the advective-dispersive flow and transport formulation. The second sub-model accounts for degradation and attenuation. The third sub-model accounts for the aggregation of all mass discharges of one contaminant type j.

Module 5: Transport Model

The transport model calculates the impact (=amount of mass discharge) to the drinking water well, depending on the mass m_0 that has been released to the aquifer after an event F

has occurred (see Fig. 7.1, module 1-4).
Practically, all methods solving the advection dispersion equation can be used to get the individual mass discharges at the well. Nevertheless, methods that calculate mass discharge-based impact of multiple hazards within the drinking water catchment can easily lead to high computational demands, especially when multiple transport simulations with contaminants injected at each possible spill location $\boldsymbol{x}_h$ within the entire catchment have to be performed. To decrease computational costs, smart simulation concepts are needed. Here, I will use the tool set as already introduced for the VIP framework (see Section 4.3). In specific, the application of STORM in Chapter 9 will use MODFLOW (e.g., Harbaugh et al., 2000) and PEST (e.g., Doherty and Hunt, 2010). The mathematical description for modeling and simulating flow and transport in a Bayesian geostatistical framework has been given in Chapter 5.
Following the same approach as Einarson and Mackay (2001), background concentration values $c_{h,j}$ of contaminant j of hazard h in the well can be estimated for steady-state flow fields with a constant mass release (continuous plume) as

$$c_{h,j} = \frac{\dot{m}_{0,h,j}}{Q_P}, \tag{7.6}$$

with Q_P being the pumping rate, and $\dot{m}_{0,h,j}$ being the mass release rate for hazard h and contaminant j.
The effect of longitudinal dispersion is neglected in case of background concentration as transport in the aquifer is assumed to be at steady-state. Transverse dispersion only plays a role close the bounding streamlines. The background concentration, here simultaneously used for continuous sources, can be added to the overall breakthrough curve $c_{Tot,j}$ of contaminant j.

Module 6: Fate Model

It is possible to account for linear sorption and first-order degradation processes in a post-processing to module 5. A full mathematical description has been provided in Section 5.6.

Module 7: Space and Time Convolution

The drinking water well acts as an integral receptor for all contaminants that have been released on the catchment level at different locations and time scales (see Fig. 7.5). All hazards $h = 1, \ldots, n_h$ are mapped to individual location $\boldsymbol{x}_h$ in the catchment. Mass release events (see Fig. 7.5, stronger coloring) are mapped over time, showing multiple failure events over simulation time (see Hazard 2 in Fig. 7.5). The uncertainty in risk source failure and thus mass release is resolved via Monte Carlo simulation of possible failure and release histories in the catchment.
Fig. 7.6 illustrates temporal, spatial and frequency aspects within a single random release history, leading to the cumulative concentration history $c_{Tot,j}$ in the well of the three hazards (green line). In the given example (taken from Chapter 10), these three hazards $h = 7, \ldots, 9$

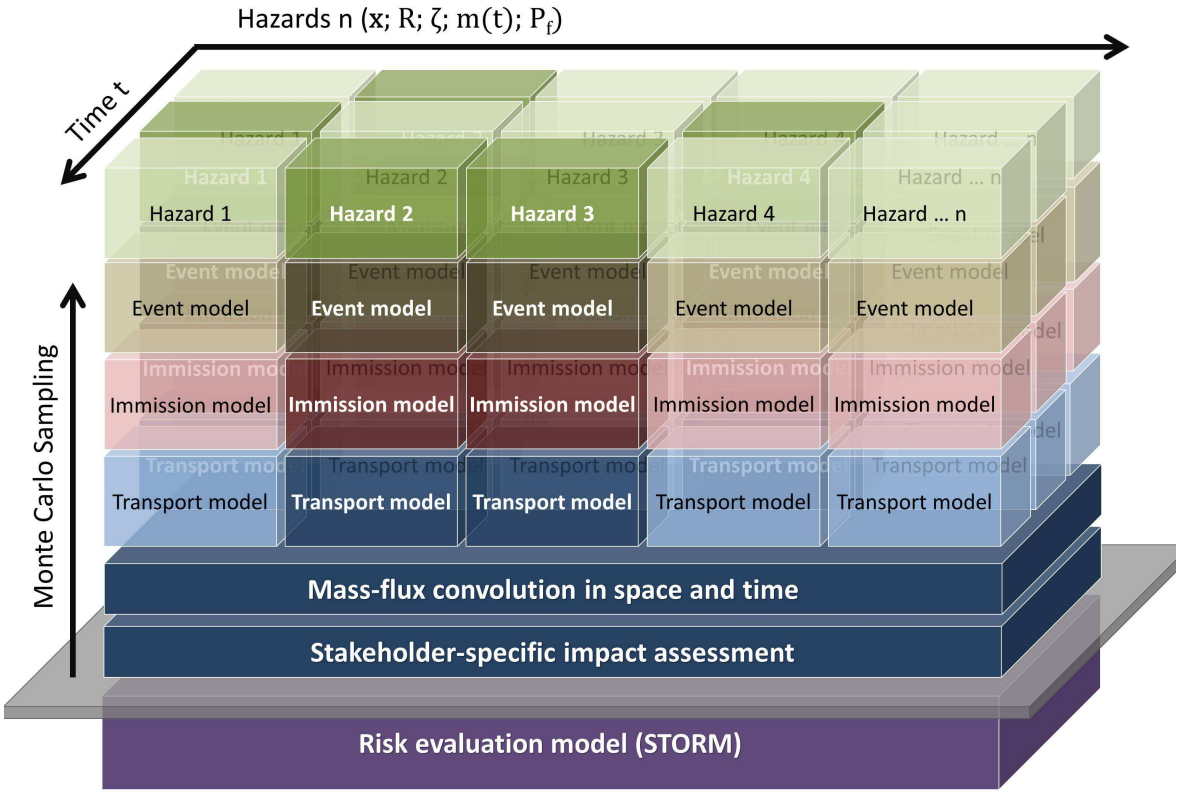


Figure 7.5.: Conceptual visualization of mass-flux aggregation (module 6) due to event occurrences (module 1) of spatially-distributed hazards within the whole catchment and temporal offsets, showing potential hazardous events by highlighted areas (stronger coloring).

are of the same contaminant type $j = 3$. Hazard 7 is an areal hazard source type, discretized in $l = 1, \ldots, 9$ point sources.
None of the individual hazards within the presented time period would exceed a critical concentration level with $c_{crit,1} > 3.5 \cdot 10^{-6}$. However, there are two important aspects for aggregation. First, risk managers often deem the situation to be safe, when individual hazards are treated independently of each other. In fact, only considering the breakthrough curves of hazard 6, 7 and 8 in a cumulative setting leads to an overall failure. Second, considering the frequency of hazard failure also adds to the cumulative contaminant impact ($t_{sim} > 1400\,d$). Aggregation of single events may not lead to critical threshold exceedance ($t_{sim} < 1400\,d$). Only after a total of ≈ 72 failures (hazard 7), the critical level $c_{crit,2}$ is exceeded. A one-time failure of hazard 7 does not lead to contaminant well exposure with a critical value of $c_{crit,2} = 2 \cdot 10^{-6}$. Please note that the initial buildup of the time-related aggregation in an initially clean aquifer leads to a statistically not representative time period within the simulation. This start-up will have to be discarded from statistical analysis.
The decisions and risk measures derived from the cumulative breakthrough curve strongly

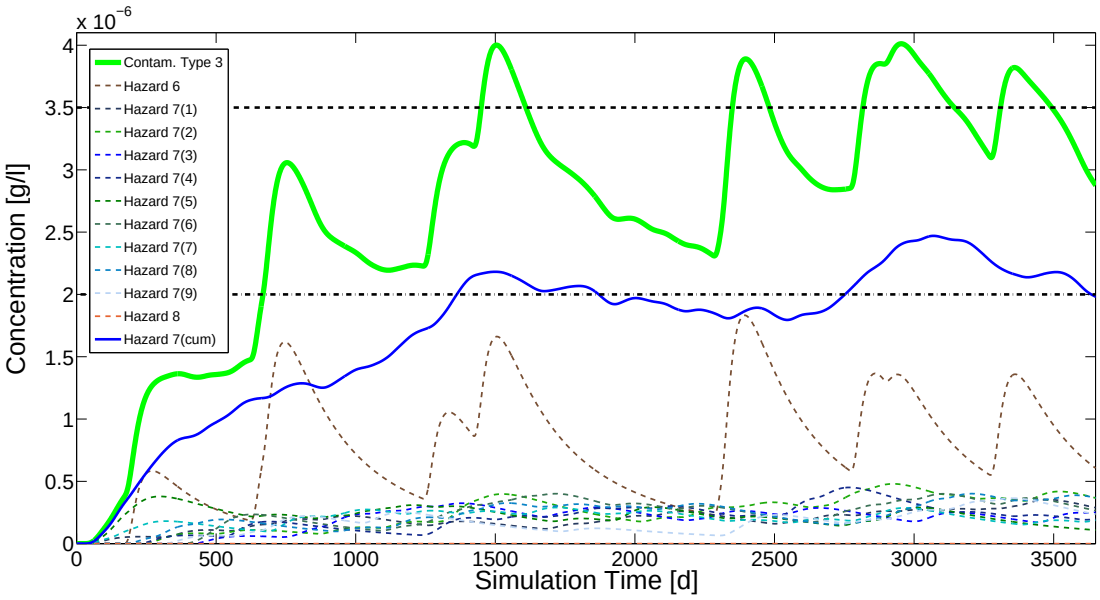


Figure 7.6.: Example to add mass-discharge for aggregating threshold-based values, due to a non-linear behavior in adding threshold-dependent damage values (module 6).

depend on regulatory levels or critical levels, which are denoted as c_{crit}. Please note, any changes in critical levels c_{crit} lead to a non-linear response in risk (see Fig. 7.6). For example, $c_{crit,1}$ only leads to a relevant impact, if the total concentration breakthrough curve is considered. Choosing $c_{crit,2} = 2.0 \cdot 10^{-6}$ at a lower level already leads to relevant well contamination by hazard 7 alone, although it would have been no risk before. The total breakthrough curve (green) is continuously above the threshold level, leading to long-term well down-times. This might lead to unacceptable risk levels and such to well closure in the most unfavorable case (e.g., Fischhoff, 1990).

Module 8: Stakeholder-objective Risk Measures

The STORM concept is based on four well vulnerability criteria (see Section 4.2). Due to multiple interests and available risk measures (e.g., DALY, CML, Micro morts, etc.), I introduced in Section 7.1.2 the motivation to use well vulnerability criteria as intermediate risk levels. These intermediate risk levels are called stakeholder-objective risk measures ξ. Here, I define the stakeholder-objective risk measures listed in Tab. C.1 in more detail:

Maximum concentration ratio, MCR: The maximum concentration ratio, MCR, is the ratio between the maximum concentration c_{peak} to be expected at the well and the contaminant-specific critical concentration level c_{crit}. The ratio is defined as

$$MCR_{h,j} = \frac{c_{peak,h,j}}{c_{crit,j}}. \tag{7.7}$$

Maximum concentration ratio MCR is a dilution measure, showing two dilution aspects. First, the degree of reduced contaminant concentration due to plume dispersion

within the catchment is considered (e.g., Cirpka and Kitanidis, 2000b) . Second, the dilution of concentration due to pumped clean well water is included to the assessment (Einarson and Mackay, 2001). Third, c_{crit} can be adopted accordingly, if water treatment facilities are available. Due to the contaminant-specific normalizing factor $c_{crit,j}$, hazards can be prioritized across contaminant types j.
Maximum concentration ratio forms the basis for all health-related and concentration-based risk measures, such as toxic units in ecological risk assessment (e.g., McKnight et al., 2012). Limitations occur for long-term contaminant release types, as long-term contaminant uptake leads to chronic rather than acute health effects.

Contaminant Load Exposure, CLE: Concentration profiles are continuous functions, which can be integrated over time in order to get the total mass of contaminant load that a consumer is exposed to over a given time period. This is under the assumption that contamination stays undetected. Contaminant load exposure CLE is most suitable for chronic health-related risk measures. The total load takes short and long-term contamination effects into account. It is defined as the load above a given threshold level, e.g., drinking water standard or RfD (US EPA, 2012).

$$CLE_{h,j} = \int_0^\infty c_{h,j} \left(c_{h,j}(t) \geq c_{crit,j}\right) \mathrm{d}t, \tag{7.8}$$

where $c_{crit} = 0$ can be used to assess the total load, such as necessary in carcinogenic health-risk assessment (e.g., US EPA, 1989).

Well Exposure Time, WET: Well exposure time, WET, is the time that concentration levels in the well are above a given threshold level c_{crit} and is identical to the fourth well vulnerability criterion t_{exp}. It can be used to assess exposure duration, i.e. for cancer risk assessment (e.g., US EPA, 1989) or the risk of well down-time due to qualitative water unavailability. The duration of well down-time enables stakeholders to evaluate the risk economically, as the total unavailable volume of drinking water for exposure duration can be priced. The measure is further discussed in Chapter 10.

Response Time, RT: The response time, RT, returns the available time between hazard failure and the point in time, where the concentration exceeds a given threshold level in the production water. It is identical to the third well vulnerability criteria t_{react}. The response time can be used to install a smart network of monitoring wells, with travel times long enough for detection and installing mitigation measures or installing control planes (e.g., Verreydt et al., 2012).
Prioritization of hazards by response time depends on the gradient of the following two aspects in a cumulative risk setting. First, the hazardous concentration level $c_{h,j}$ being present in the well that day of exceedance. And second, the temporal duration since the hazard has failed until the day of exceedance.

Required Blending Ratio, RBR: Risk of supply failure can be reduced as long as the water distributor has alternative water sources (e.g., river, lake, sea, ...) or technical storage capacities that can be used to dilute critical concentration in the well to non-critical values. The compensation capacities are accounted for when working with the required

blending ratio:

$$RBR_j = \frac{\dot{m}_{tot,j}}{Q_{Tot}} \cdot \frac{1}{c_{crit,j}}, \tag{7.9}$$

with Q_{Tot} being the total amount of raw water and $\dot{m}_{tot,j}$ being the total mass discharge of contaminant j in the well, where in all other water sources that contribute to Q_{Tot} the contaminant j is absent. In reality, this might not be true, such that $\dot{m}_{tot,j}$ is the aggregated mass discharge of contaminant j over all water sources.
A similar concept is available with well down-time, as the time of water being unavailable can be compensated by the time water is available through storage. The compensation time is dependent on the water demand and may vary over time. In both cases, the water constructor can pre-define a maximum possible blending ratio, such that risk is defined as the probability of exceeding the pre-defined maximum blending ratio given a fixed water demand.

These stakeholder-specific risk measures deliver additional and valuable information for catchment managers, such that the following questions can be answered:
„What is the risk ...“

1. *„... of water treatment failure due to high concentration loading? (MCR)“*
2. *„... to encounter adverse health effects due to acute and chronic concentration loading? (MCR, CLE)“*
3. *„... of closing a well due to qualitative issues? (WET)“*
4. *„... of suffering quantitative supply failure due to well down-times? (WET)“*
5. *„... of the population to be exposed to contamination accumulated over time? (CLE, MCR, WET)“*
6. *„... of detection failure due to the installed groundwater monitoring network? (RT)“*
7. *„... that maximum blending capacities fail to guarantee concentration levels below drinking water standards? (RBR)“*

7.6. Risk Evaluation

The risk evaluation model consists of two modules. First, it quantifies the risk (module 9). Second, it serves to find risk mitigation alternatives that reduce the risk to an acceptable level, if required in case of non-compliance with the pre-defined objective (module 10).

Module 9: Risk Estimation

Risk is calculated as a combination of uncertainty and level of severity at the receptor. In case hazards are treated independently of each other and one assumes the vulnerability

(impact) mentality, the risk R equals the hazard-wise expected damage, measured through the severity measure ξ (e.g., t_{peak}, WET, MBR):

$$R(\xi|F) = E(\xi|F) = \int_0^\infty \xi \cdot p(\xi|F)\mathrm{d}\xi, \tag{7.10}$$

with $p(\xi|F)$ being the probability density distribution of the stakeholder-objective severity measure ξ in case of hazard failure F. $p(\xi|F)$ stems from the fact of parameter uncertainty. $R(\xi|F)$ expresses the mean damage given a failure F, i.e., a vulnerability. The risk perception of Eq. (7.10) is the most wide-spread view in environmental risk assessment. Considering the time time between failure in risk assessment, Eq. (7.10) changes for WET to:

$$R_a(t_{exp}) = E(\xi(t)) = \frac{1}{T}\int_0^T \frac{1}{n_{MC}} \sum_{k=1}^{n_{MC}} I_0\left(c_{k,j}(t) \geq c_{crit}\right) \mathrm{d}t, \tag{7.11}$$

where R_a is the risk (defined as expected damage) measured through the impact metric ξ (here: well exposure time to concentrations $c_{k,j}$ larger than c_{crit}), $k = 1, \ldots n_{MC}$ denotes the Monte Carlo realization k and I is an indicator function that returns unity, if its argument is true, and zero else. The cumulative concentration breakthrough curve $c_{k,j}$ for each Monte Carlo realization k and contaminant j in the well includes model and parameter uncertainty as well as the stochastic history of failure events. Depending on the desired level of aggregation, one can either use all breakthrough curves from all contaminant types or perform individual analysis for each contaminant type j. With appropriately chosen time units $R_a(\xi)$ is the mean annual risk measure, which denotes the expected impact per year.
A more detailed illustration of the annual risk R_a, is based on the idea from flood risk management. Impacts are mapped against a recurrence interval, often denoted as annuality. According to this procedure, we follow a similar idea, estimating the impact strength of annual well-down time:

$$R_J(t_{exp}) = \frac{1}{T}\sum_{l=1}^{n_l} t_{exp,l}^2 \cdot I\left(t_{exp,l} \geq t_{exp,crit}\right), \tag{7.12}$$

with n_l being the number of events and $t_{exp,crit}$ a critical well exposure time to be expected within a given time period. For convenience we speak of annuality curves and impact impacts per return period, although in a strict mathematical sense R_J is the strength (variance) of the mean annual impact. The annual mean risk measure R_a is related to Eq. (7.12) by

$$R_a(\xi) = \int_0^\infty \xi_{crit}\mathrm{d}J_a. \tag{7.13}$$

The annuality hosts valuable information on temporal recurring severity levels, such that water managers are able to install emergency plans for a certain recurrence period. In many situations the interpretation of annuality is more intuitive in communicating risk to stakeholders. Nevertheless, there are examples, such as customer minutes lost, where risk is expressed in consequences per year (e.g., Lindhe et al., 2009).

Module 10: Risk Management based on STORM

In case, risk levels (module 9) are unacceptable, risk mitigation alternatives are in demand. The most severe hazards within the catchment are determined by prioritization. When considering hazards independently and without aggregation over time, prioritization can directly be performed by individual event-wise risk estimates (e.g., Eq. 7.10). However, when acknowledging the need for spatial and temporal aggregation, prioritization of individual hazards becomes a complex task.
Troldborg et al. (2008) suggested to take the mass ratio between single concentration and total breakthrough curve. I follow the approach by Troldborg et al. (2008). Adapted to the current context, this means to estimate the ratio between the total aggregated risk with and without an individual hazard. The risk ratio R_r reads:

$$R_{r,h} = \frac{R_0 - R_{c,h}}{R_0}, \tag{7.14}$$

with R_0 being the cumulative risk considering all hazards, and $R_{c,h}$ being the cumulative risk without hazard h. The hazard with the largest risk ratio is ranked to be most severe. Chapter 10 will demonstrate this approach within an illustrative example.

7.7. Findings of STORM

- Hazard prioritization is only meaningful, when one pre-defined risk objective is consistently considered within the risk analysis.
- STORM improves risk management
  - It delivers system and hazard information, where mitigation measures are most beneficial, depending on the expected annual impact or the severity of a $10-$ or $100 - year$ event.
  - Risk can be evaluated in light of the stakeholder attitude to risk, e.g., accepted impact only statistically once in $100 years$.
  - Stakeholders can establish emergency plans or improve the system, such that it is capable to compensate the expected impact for a given return period.
- Small concentration levels lower the distance to critical concentration levels and therefore have to be considered in cumulative risk assessment.
- There exist more than one way to measure risk, such as consequences per year (=damage times probability) or the damage to be expected within a certain recurrence period.
- Prioritization in a cumulative setting is based on fraction by which any individual hazard changes the cumulative risk estimate.

8. Application of VIP - Synthetic Test Case

This chapter demonstrates the vulnerability-based probabilistic risk assessment and management concept from Chapters 4 to 6 on a synthetic test case. The value of mapping vulnerability isopercentile lines is discussed as well as the advantage of using local-scale versus macrodispersive transport calculations. Among other risk management options, I will introduce a framework to judge the worth of data related to drinking water safety as proposed in Chapter 6.
Section 8.1 introduces the synthetic test case. The resulting intrinsic vulnerability isopercentile maps are presented in Section 8.2. Section 8.3 illustrates the necessity of using probabilistic instead of deterministic risk approaches. The benefit of data acquisition for more accurate wellhead delineation and sound decision analysis is shown in Section 8.5. The application of decision analysis within an economic context is presented in Section 8.6. The chapter closes with a brief summary in Section 8.7. When appropriate, references to the nine steps of the VIP approach as defined in Chapter 4 are provided.

8.1. Synthetic Test Case

The VIP concept is demonstrated on two variants of a synthetic test case (Section 8.1.1 to 8.1.3 and Section 8.1.4). The flow and transport properties are kept identical for both versions with the exception of the mean recharge rate and standard deviation. The main difference between the two test cases is the model domain size, which had to be enlarged for demonstrating the benefit of additional data.

8.1.1. Model Set-up (Step 1)

For simplicity, the VIP concept is illustrated on a rectangular 2D example with domain size $300\,m \times 300\,m$ (see Fig. 8.1). This example only serves for illustrative purposes, as the method is independent of dimensionality, complex geometries and boundary conditions. A hydraulic background gradient from west to east with $\nabla\phi = 0.005$ is assumed with appropriate fixed head conditions on both Dirichlet boundaries (west and east). Within the reverse approach (Section 5.4.1), a Dirac-pulse with concentration $\hat{c}_{spill}$ is introduced at a single well at $x = 225\,m$ and $y = 150\,m$ with a pumping rate of $Q_P = 1 \times 10^{-4}\,m^3 s^{-1}$. All other boundaries (north and south) are no-flux boundaries. The aquifer is assumed to be leaky and confined with an uncertain normally distributed and spatially constant recharge rate of $q_{rg} = 120\,mm\,a^{-1}$ and standard deviation $\sigma_{rg} = 10\,mm\,a^{-1}$.

The discretization of the domain equals $dx = dy = 1\,m$ with assumed subgrid-scale dispersivities of $\alpha_L = 2.5\,m$ and $\alpha_T = 0.25\,m$ (including the dispersion due to depth averaging, e.g., Dentz et al., 2000). The total number of nodes to solve are $n_{tot} = 90.601$. As covariance model for log-transmissivity $Y = \ln(T)$, the Matérn correlation function is chosen (e.g., Handcock and Stein, 1993) because it has an additional shape parameter κ. Treating κ as uncertain resembles Bayesian model averaging over a continuous spectrum of covariance shapes (e.g., Nowak et al., 2010). The parameters of the structural model are $\boldsymbol{\theta} = \left(\mu, \sigma^2, \kappa, \lambda_x, \lambda_y\right)$, where μ is the mean value of log-transmissivity $Y = ln(T)$ (with T in units of $m^2 s^{-1}$), σ^2 is the variance of Y, κ is the shape parameter, and λ_x, λ_y are the length scales. First, in absence of site-specific data, I assume the uncertain structural parameters to follow uniform prior distributions with lower and upper bounds, $\mu = [-7.5 \;\; -5.5]$, $\sigma^2 = [1 \;\; 3]$, $\kappa = [0.5 \;\; 5]$, $\lambda_x = [10 \;\; 25]\; m$ and $\lambda_y = [5 \;\; 15]\; m$. Nevertheless, due to the Monte Carlo framework, it would be easy to implement any other prior distribution assumption.
Unconditional transmissivity fields are generated and flow and transport simulations are performed with the same numerical implementation as in Nowak et al. (2008). The simulations were run on a computer cluster with 36 cores at 3.0GHz and a total of 300Gb RAM. The computational time for $n = 5000$ unconditional realizations is $28\,h$ and, in the conditional case, $222\,h$. In real applications, lower numbers of realizations (e.g., $n = 500$) could also be used, but should be confirmed by convergence analysis (e.g., using bootstrap or jackknife methods, Shao and Tu, 1996).

8.1.2. Model Calibration and Conditioning (Steps 2-3)

In order to reduce epistemic uncertainty, the VIP framework is conditioned to direct and indirect measurements. Here, I generate a random realization, which is denoted as the „synthetic truth“. From this realization, five artificial measurements of head ϕ_0 and ten measurements of log-transmissivity Y_0 (see Fig. 8.1) are drawn. These measurements are perturbed with random measurement error that has standard deviation of $\sigma_Y = 1$ and $\sigma_\phi = 0.25\,m$ for measurements of Y and ϕ, respectively. The structural parameters used to generate the synthetic „true“ random field are $\mu_o = -6.83$, $\sigma_o^2 = 1.91$, $\kappa_o = 0.49$, $\lambda_{x,o} = 9.11\,m$ and $\lambda_{y,o} = 5.17\,m$. Fig. 8.1 shows exemplary the synthetic „true“ log-transmissivity field and the corresponding streamlines with the given initial and boundary conditions. For the conditional simulation, I used the Bayesian Generalized Likelihood Uncertainty Estimator in combination with a fast Kriging-like conditioning scheme for direct measurements and rejection sampling (Section 5.3).
The actual outlines for the critical values (see Tab. 8.1) that apply in the synthetic „real“ realization are shown in Fig. 8.2. The ten artificial log-transmissivity and five head measurements are obtained from this synthetic „true“ random realization.
The plus sign in Fig. 8.2, Fig. 8.3 and Fig. 8.4 marks the location S for which a contaminant spill (e.g., virologically or microbial loaded water) within the aquifer is assumed, yielding well vulnerability quantities and probabilities of exceeding a critical level $WVC_{k,crit}$ (see Eq. 4.2). The location S will serve as a possible spill location for comparison in Section 8.2.
In the synthetic „truth“, location S has peak arrival time $t_{peak,S,obs} = 80\,d$, dilution of peak concentration by the factor of $c_{peak,S,obs} = 1.99 \times 10^{-8}$, time to react $t_{crit,S,obs} = \infty\,d$ (as the

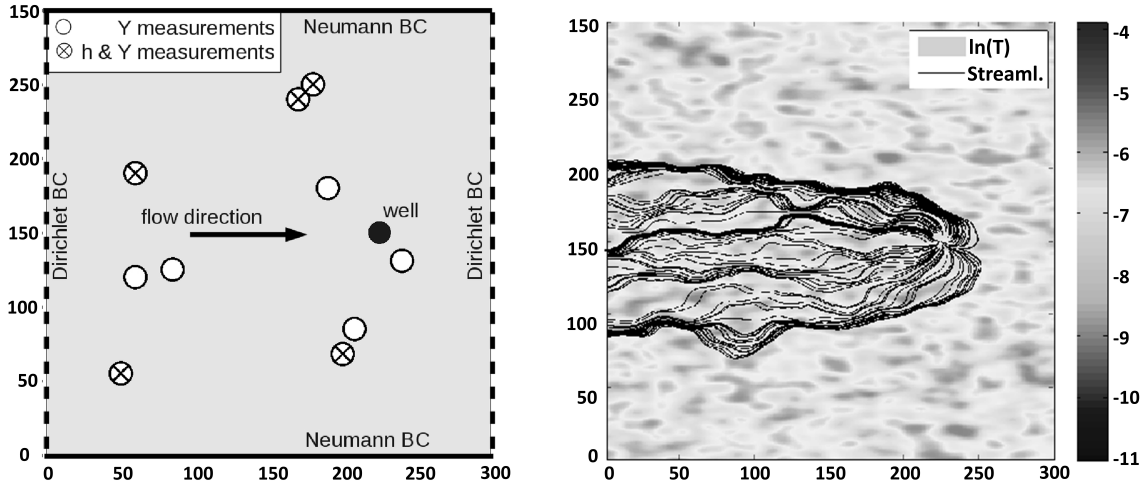


Figure 8.1.: Illustrative example, showing location of measurements (left). Log-transmissivity field and the corresponding streamlines of the well capture zone for synthetic „truth“ with structural parameter set ($\mu_0 = -6.83, \sigma_0^2 = 1.91, \lambda_{x,0} = 9.11\,m$ and $\lambda_{y,0} = 5.17\,m$) (right).

resulting concentration in the well never rises above the critical threshold value c_{crit}) and exposure time $t_{exp,S,obs} = 0\,d$ (as there is no exposure due to concentration levels below c_{crit}).

8.1.3. Vulnerability Set-up (Steps 4-8)

The breakthrough curves for each location in the domain are reconstructed from temporal moments (Section 5.4.2) by non-linear maximum entropy-based parameter estimation in log-time as given in Eq. (5.40). The resulting well vulnerability measures are obtained by comparing the contaminant impact to pre-defined critical values. The critical values are chosen according to Tab. 8.1. The first well vulnerability criterion represents the German

k	WVC	Level	Dim.
1	τ_{crit}	50	$[d]$
2	ζ_{crit}	1×10^{-7}	[-]
3	τ_{react}	21	$[d]$
4	τ_{exp}	2	$[d]$

Table 8.1.: Critical threshold values for the synthetic test case.

wellhead protection area with $WVC_{1,crit} = \tau_{crit} = 50\,d$ (DVGW, 2006), but here evaluated for the arrival time of peak concentration instead of bulk arrival time. The second vulnerability criterion shows the area within which a contaminant is being diluted by less than a factor of $WVC_{2,crit} = \zeta_{crit} = 1 \times 10^{-7}$. The third criterion shows the probabilistic extent of the capture zone, in which a critical reaction time $WVC_{3,crit} = \tau_{react} = 21\,d$ is exceeded,

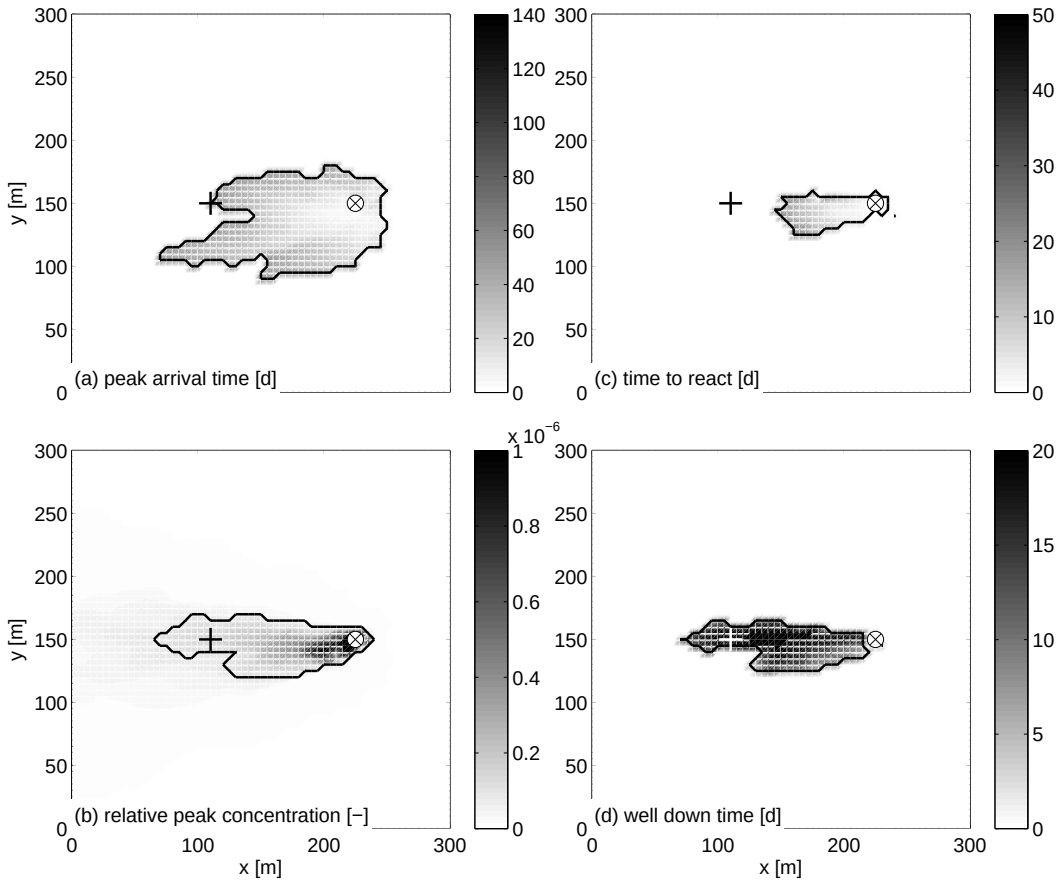


Figure 8.2.: Outlines and criteria values of the „real“ realization.

thus indicating the confidence in the reaction time for a water manager until the contamination breaches the given threshold level $c_{crit} = 1 \times 10^{-7}$ and mitigation measures have to be installed. The fourth criterion indicates the potential area for spills where the well is exposed to contamination above the threshold for more than $WVC_{4,crit} = \tau_{exp} = 2\,d$. This critical exceedance level τ_{exp} represents the maximum time, for which one can compensate the contaminated pumped raw water by water stored in reservoirs for the exposure duration.
Due to the Monte Carlo framework, the posterior empirical probability distributions of the four intrinsic vulnerability criteria are available for each location across the domain. Thus, vulnerability-based isopercentile lines of non-compliance (Eq. 4.2) can be mapped and used for decision analysis (step 9).

8.1.4. Vulnerability-based Decision Analysis Set-up (Step 9)

Additional data that is obtained from the synthetic true test case will be used to investigate the benefit of data according to Section 6.3. In Section 8.5, I will show the impact of data type, quality, and quantity for pre-defined sampling designs (see Fig. 8.7) on the delineation

accuracy and then investigate the influence on the economic risk measure, customer minutes lost (Eq. 6.8). The effect of uncertainty on delineation accuracy will be measured by the areal delineation error A_{err} as defined by Eq. (6.5).

8.2. VIP Maps

Fig. 8.3 displays the four intrinsic well vulnerability criteria with isopercentiles of $[0.1,\ 0.5,\ 0.9]$, based on the set-up from Section 8.1.1 with unconditional realizations. The 10^{th} isopercentile line, for example, encloses the area A_{10} outside of which the respective vulnerability criterion is exceeded with less than 10% probability. The smaller the percentile, the larger the enclosed area, and the safer is the corresponding wellhead delineation. Fig. 8.4 shows the corresponding results for the conditional Monte Carlo simulations using the synthetic data set, obtained from the synthetic truth shown in Fig. 8.2. The color shading in Fig. 8.2, Fig. 8.3 and Fig. 8.4 show the actual values or ensemble average of the underlying exposure risk quantities (t_{peak}, c_{peak}, t_{react} and t_{exp}). The ensemble average is a pixel-wise statistical value over n_r Monte Carlo simulations. The **choice of critical level** c_{crit} determines the second to fourth well vulnerability criteria substantially and thus the shape and size of the corresponding vulnerability maps. If the critical peak level ζ_{crit} for the second vulnerability criterion (see (b) in Fig. 8.3, Fig. 8.2 and Fig. 8.4) equals the threshold level c_{crit} for the third and fourth vulnerability criterion (see (c) and (d) in Fig. 8.3, Fig. 8.2 and Fig. 8.4), the isopercentiles of the reaction time (see third criterion and (c) in Fig. 8.3, Fig. 8.2 and Fig. 8.4) can be at most as wide as the isopercentiles of peak concentration. For large critical values of reaction time τ_{react} and $\zeta_{crit} = c_{crit}$, the isopercentiles of the third well vulnerability criterion become equal to the isopercentiles of the second criterion, because the third criterion will degenerate to the information whether *any* reaction is necessary at all, regardless of the time to react. The same effect occurs for the fourth vulnerability criterion for small values of the critical exposure levels τ_{exp}, because non-zero exposure times appear wherever the critical threshold level c_{crit} is breached.
According to Aven (2011), it is important to determine for any risk assessment study both the exposure level (severity) of the hazardous contamination and the likelihood of its occurrence at the receptor. Both of these risk-based information types are contained within the VIP maps, showing for each location the existing intrinsic well vulnerability of the drinking water well and its exceedance probability.
In the following, the results for the four VIP maps will be discussed, based on the example of a spill at location S.

The first VIP (peak arrival time) is estimated (ensemble mean) for the unconditional case with $t_{peak,S,uncond} = 61\,d$ (see Fig. 8.3 (a)) and for the conditioned example $t_{peak,A,cond} = 68\,d$ (see Fig. 8.4 (a)). In a conventional approach, the stakeholder would assume that there will be no exposure risk for the drinking water well by the spill event at S in both cases, as microbial safety is defined in Germany by transport times larger than $\tau_{crit} = 50\,d$. Taking the new probabilistic information into account, the vulnerability maps for peak arrival show exceedance probabilities $\widetilde{P}(t_{peak,S,uncond} > \tau_{crit}) = 55.2\,\%$

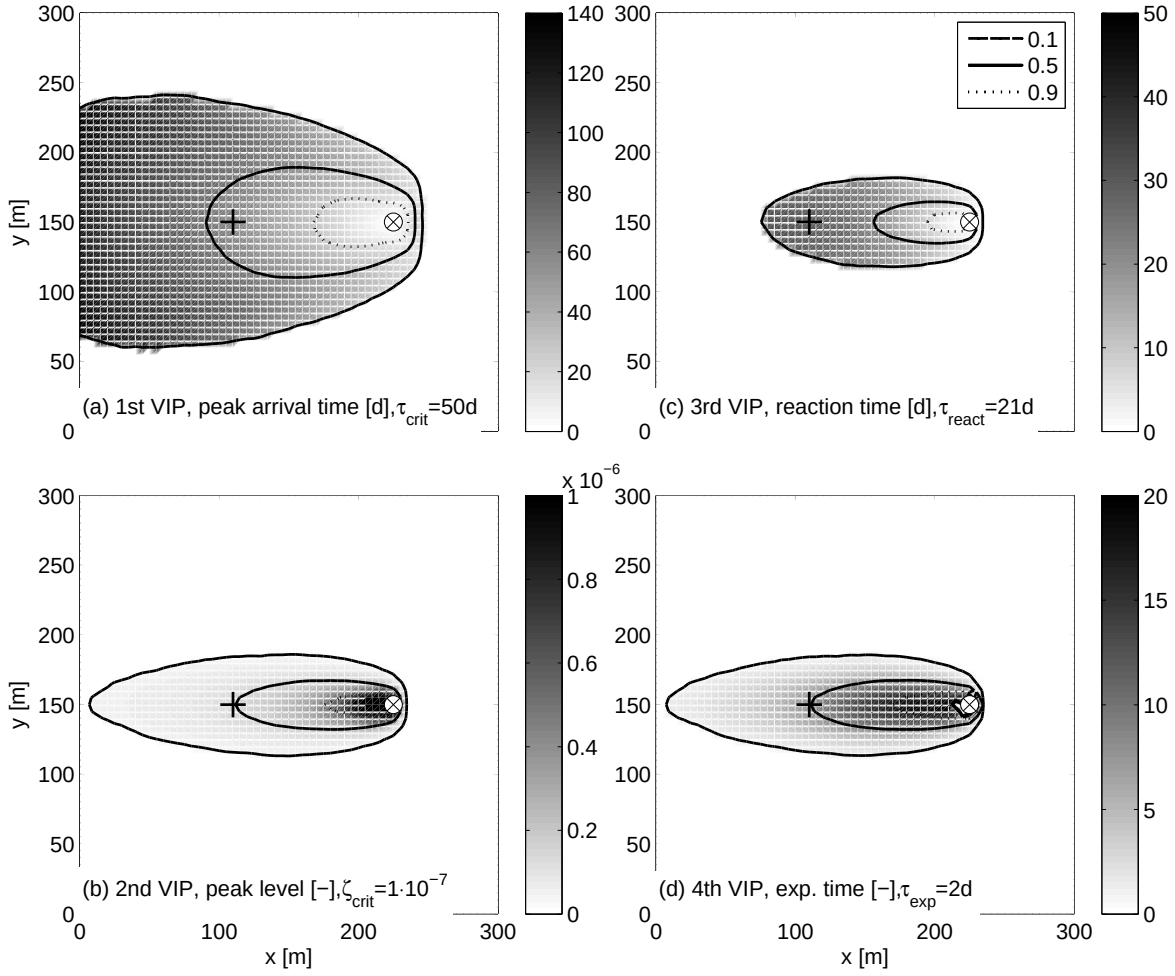


Figure 8.3.: Probabilistic isopercentiles $[0.1,\ 0.5,\ 0.9]$ for the four intrinsic well vulnerability criteria (a)-(d) from $n = 5000$ unconditioned simulations. Grey-scale maps show the ensemble mean of the respective well vulnerability criteria (a)-(d).

and $\widetilde{P}(t_{peak,S,cond} > \tau_{crit}) = 39.8\,\%$. This is indeed a substantial risk, and would be missed within conventional deterministic approaches.

The second VIP (peak concentration) directly shows the relevance of distinguishing between dilution and the uncertainty of plume location. The expected maximum concentration at the well is, on average, diluted by the factor of $c_{peak,S,uncond} = 1.21\times10^{-7}$ and $c_{peak,S,cond} = 1.12 \times 10^{-7}$, if contamination occurred at location S (see Fig. 8.5). Single realizations can yield higher and lower dilution factors. The ensemble average breakthrough curve is presented as the solid thick line. A detailed discussion on this topic is given in Section 8.3. The exceedance probabilities are $\widetilde{P}(c_{peak,S,uncond} > \zeta_{crit}) = 43.2\,\%$ and $\widetilde{P}(c_{peak,S,cond} > \zeta_{crit})$=$46.8\,\%$.

The third VIP (reaction time) delivers information about the time available to react before the well has to be shut down. In this example, the average values for location S are $t_{react,S,uncond} = 30\,d$ and $t_{react,S,cond} = 42\,d$, which is substantially smaller than the

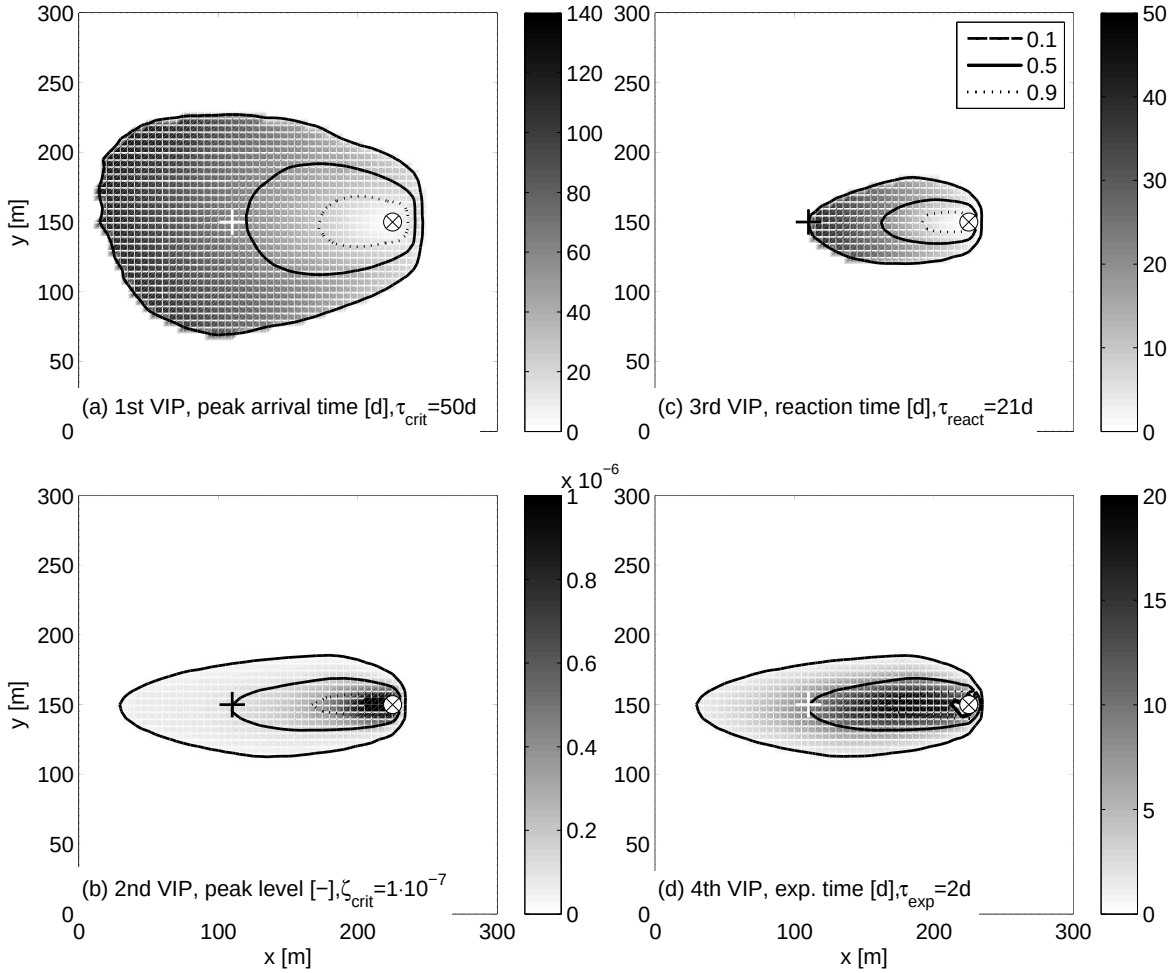


Figure 8.4.: Probabilistic isopercentiles $[0.1,\ 0.5,\ 0.9]$ for the four intrinsic well vulnerability criteria (a)-(d) from $n = 5000$ conditioned simulations. Gray-scale maps show the ensemble mean of the respective well vulnerability criteria (a)-(d).

numbers for peak arrival time. All realizations, which do not breach the threshold level $c_{crit} = 1 \times 10^{-7}$, are not considered within the ensemble averaging as no reaction time is required at all. The 10^{th} percentile of available reaction time at location S is as low as $12\,d$ (uncond) and $17\,d$ (cond), indicating how fast early alert systems and emergency management decisions would have to be, such that well safety can still be guaranteed in adverse cases.

The fourth VIP (well down-time) gives the necessary information about the well down-time to be expected after contamination by a spill event within the catchment. The average exposure time related to a spill at location S is $t_{exp,S,uncond} = 7\,d$ and $t_{exp,S,cond} = 8\,d$, showing exceedance probability of $\widetilde{P}(t_{exp,S,uncond} > \tau_{exp}) = 43.8\,\%$ and $\widetilde{P}(t_{exp,S,cond} > \tau_{exp}) = 46.1\,\%$. This indicates the time frame and the associated uncertainty that the well will be out of operation.

8.3. The Effect of Macrodispersion in Risk Management

Macrodispersive approaches or models without separation of dilution and uncertainty in position cannot map isopercentile lines of non-compliance, but only one capture contour line (e.g., Frind et al., 2006). In that case, uncertainty in size and position is lumped together within an implicitly averaged transport equation, blurring the overall picture and destroying a large part of the probabilistic information.
In Section 3.6, I briefly introduced the shortcoming of macrodispersive transport models, related to represent the uncertainty of plume location along a cross-sectional depth-averaged control plane. The ensemble-averaged vulnerability criteria are, per definition, equivalent to solutions based on upscaled macrodispersion approaches (see background gray-scale of Fig. 8.2, Fig. 8.3 and Fig. 8.4). The spatial distribution and features of macrodispersive well vulnerability results are discussed by Frind et al. (2006). Here, I compare the consequences that arise from using a macrodispersive risk approach for risk-informed decisions.
Deterministic approaches with upscaled macrodispersive parameterization mix uncertainty in location of the contaminant plume with actual dilution, leading to delineation outlines based on the average of all possible geostatistical hydraulic conductivity fields. This approach cannot transparently reflect the uncertainty in plume location and actual dilution. Here, I perform a pixel-wise evaluation within a Monte Carlo framework, addressing epistemic uncertainty. The relevance of uncertainty quantification has been demonstrated for peak arrival time (see Section 8.2). Although S is assumed to be safe, the probabilistic peak arrival distribution provides valuable information on exceedance probability. Thus, following a deterministic approach would have led to underestimation of the actual risk.

8.4. Considering Peak or Bulk Arrival Time

Fig. 8.5 (left) shows contaminant breakthrough curves, when parameter uncertainty within the risk quantification framework is transparently revealed. The ensemble mean breakthrough curve is the actual solution provided, in case a deterministic macrodispersive transport model would have calculated the impact of a contaminant spill at location S to the well. The statistical information on peak concentration levels at the well in case of a deterministic model run would not be accessible. Much worse, the average over strongly peaked distributions with peaks at different peak arrival times leads to a much smaller peak level of the macrodispersive (implicitly ensemble averaged) breakthrough curves (see Fig. 8.5). This illustrates best, why macrodispersive approaches are not adequate for probabilistic risk assessment, if transport is non-ergodic (e.g., Hassan et al., 2002).
The typically positively skewed breakthrough curves of transport in heterogeneous formations yield earlier arrival time for peak concentrations t_{peak} than for the arrival of bulk mass t_{50} at the well (Haggerty and Gorelick, 1995). Fig. 8.5 (right) illustrates this with a scatter plot between t_{peak} and t_{50}. In this example, $t_{peak,S,uncond}$ ($t_{peak,S,cond}$) is on average $15\,\%$ ($11\,\%$) smaller than t_{50}, leading to $7\,\%$ ($6\,\%$) larger catchment delineation on average, as shown in Fig. 8.6. The size difference depends on the degree of heterogeneity, and will be more drastic for high variability cases or fractured media. In risk analysis, the underestimation of

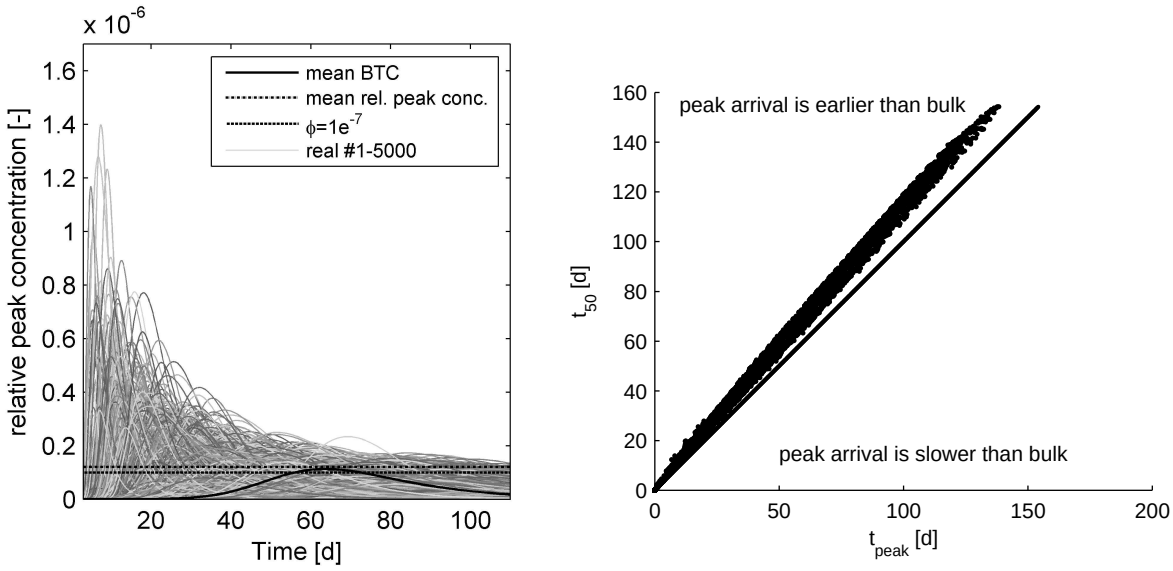


Figure 8.5.: Breakthrough curves (BTC) of all realizations and the average breakthrough curve (bold) of $n = 5000$ unconditional realizations at the drinking water well, if a hazardous spill occurred at location S (left). Scatter plot of mean peak arrival t_{peak} versus mean bulk arrival time t_{50} at location S (right).

protection zones when using bulk arrival time can have crucial liability issues in risk-based decisions. The advantage of the first vulnerability criterion over the traditional bulk-related assessment is that it takes the more conservative peak time instead of the bulk arrival time into account. Therefore, considering peak arrival time in risk analysis is the more relevant measure of contaminant arrival.

8.5. The Effect of Conditioning

The effect of reducing epistemic uncertainty U (see Eq. 6.7) by the conditional realization with 15 measurements for all four vulnerability criteria in the illustrative example of Section 8.1.2 is provided in Table 8.2. Here, U is a measure for the area a planner has to sacrifice due to uncertainty according to Eq. (6.7).

As shown in Tab. 8.2, conditioning to data reduces uncertainty, thus increasing the robustness of any decision that has to be made regarding water supply safety. Clearly, it is worth to spend money on catchment investigation to receive more confident and smaller probabilistic wellhead protection delineation outlines. The quantity and quality of data or the type of data to reduce uncertainty to an acceptable level can be assessed in a rational manner by decision analysis. Here, the utility value to measure uncertainty is areal error A_{err} as defined in Eq. (6.5).

The results in Section 8.2 have shown that the unconditional VIP line A_{10} for peak arrival time extends beyond the eastern model boundary, such that the model set up is not suitable for assessing data worth by areal error A_{err} as defined in Eq. (6.5). Thus, I extend the above

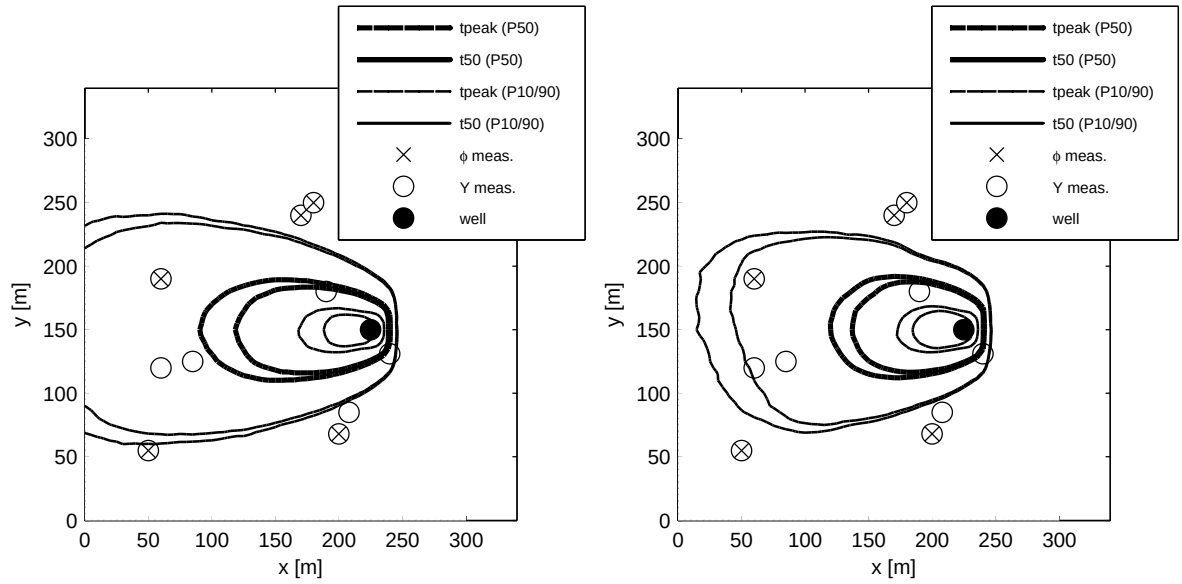


Figure 8.6.: Showing the size of the time-related wellhead protection zones, depending on t_{peak} and t_{50} for different isopercentile levels in the unconditional (left) and conditional (right) case.

VIP	critical value	unconditional uncertainty U_{uc}	conditional uncertainty U_c
t_{peak}	$\tau_{crit} = 50d$	43.1%	25.2%
c_{peak}	$\zeta_{crit} = 1 \times 10^{-7}[-]$	14.6%	10.4%
t_{crit}	$\tau_{crit} = 50d$	14.6%	10.4%
t_{exp}	$\tau_{exp} = 2d$	14.5%	10.3%

Table 8.2.: Showing the fractional area $[\%]$ of delineated catchments according to the four VIP maps that is sacrificed to uncertainty for the conditioned and the unconditioned case.

model domain to a model domain outline of $500\,m \times 300\,m$ with 150.801 total number of nodes to be solved (see Fig. 8.7). The mean annual and spatial constant recharge rate is increased to $q_{rg} = 280\,mm/a$ with standard deviation of $\sigma = 40\,mm/a$. The well position changes to $x_{well} = 375\,m$, $y_{well} = 150\,m$. All other settings, such as initial, boundary conditions, model discretization and structural parameter assumptions, are kept constant to the previous model set up (see Section 8.1.1).

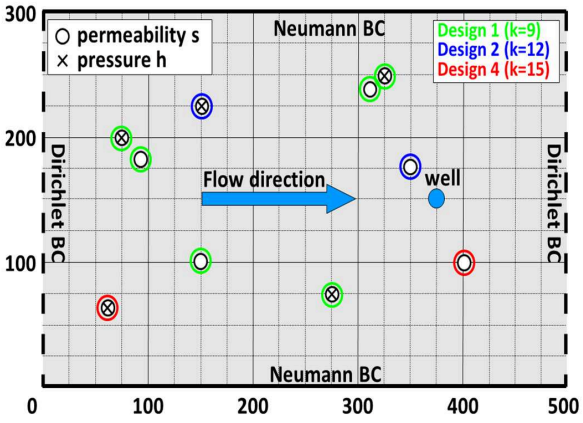


Figure 8.7.: Illustrative synthetic example, showing location of measurements for assessing data worth.

8.5.1. Uncertainty Reduction

Scenarios of Catchment Investigation

Four scenarios are compared with each other, ranging from the unconditional model simulation (scenario 0) to model predictions being conditioned to 9 (scenario 1), 12 (scenario 2) and 15 (scenario 3) direct and indirect measurements (see Fig 8.7). In comparison to the previous model set-up, only a total of $n_r = 1000$ conditional transmissivity fields are generated, which is sufficiently large to demonstrate the effect of model conditioning. The synthetic „truth“ is generated in the same spirit as explained before, using the same structural parameter set up $\boldsymbol{\theta}$ as used in Section 8.1.1. From this „true“ realization, the corresponding observed log-transmissivity and head measurements for scenario 1 to 3 (see Fig. 8.7) are drawn. The random measurement error is assumed to follow a zero-mean normal distribution with standard deviation σ_Y for direct measurements of hydraulic conductivity and σ_H for indirect measurements of hydraulic head, as shown in Fig.8.8. The matrix shown in Fig. 8.8 is used to assess the impact of measurement error (e.g., low vs. high technology measurement devices) on the resulting accuracy of delineation outlines.

The Overall Monitoring Effect

Fig. 8.9 shows that, with increasing confidence level β, the areal error and thus the delineation area increases. The increase in areal demand is non-linear, following roughly a Weibull distribution. The reference case in Fig. 8.9 is chosen for design case 2 with 12 sampling locations (8 direct and 4 head measurements), denoted as D2-RMM. D2 stands for design case 2 and RMM for the medium hydraulic conductivity (2nd letter) and medium head (3rd letter) measurement error σ_R (see Fig. 8.8). There are two ways to read Fig. 8.9 that corresponds to two risk reduction options. First, one can see how more or better data reduce the areal error to deliver a more robust decision basis. Second, higher reliability levels

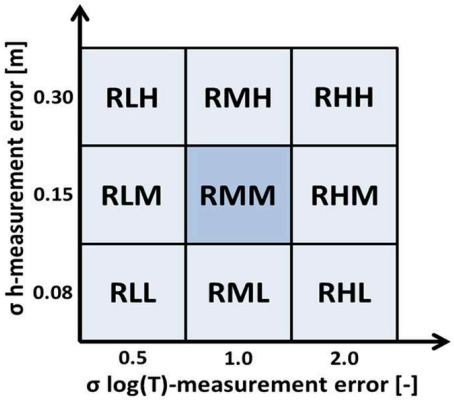


Figure 8.8.: Sampling scenario matrix, showing the measurement quality with pre-defined standard deviation.

are available at the same areal error level, when improving the available data basis. For example, let us assume that water supply safety is currently (see D2:RMM) at a reliability level of $85\,\%$. The areal error can be reduced (1st option), when the sampling accuracy of head (D2:RML) or hydraulic conductivity (D2:RLM) measurements is doubled or if additional data is added to the conditioning process (D4:RMM). The largest uncertainty reduction to be achieved within design case 2 is to double the sampling accuracy of hydraulic conductivity (D2:RLM). During all these changes, the reliability level was kept constant. For the second option, sampling campaign 2 (D2:RLM) leads to a reliability level of $\beta \approx 90\,\%$, when keeping the expected areal error constant at $A_{err} \approx 10^4\,m^2$. The decision between these two options depends on the risk perception of stakeholders. Stakeholders that are risk averse use the least uncertain decision and prefer an increase in reliability level (option 2). Risk sympathetic stakeholders prefer decisions that lead to greater benefits, here the delineation outline with a smaller area and smaller areal error (option 1).

More or Better Data?

The next question to be pursued is whether the accuracy in measurements leads to more information gain than incorporating more data? Fig. 8.10 shows the information gain in reducing uncertainty (A_{err}) for a given reliability level $\beta = 90\,\%$ by changing the measurement error or by adding more samples at pre-defined locations.

Fig. 8.10 shows that most excess area is delineated for the model run D1:RHH with the least accurate measurement devices and the design with only 9 measurements instead of 12 and 15 sampling points. Contrary, the least areal error demand is given by the sampling campaign D4:RLL, with 15 sampling locations and low measurement error in hydraulic head and log-transmissivity. Increasing the quantity of sampling from D1 to D2 (three additional samples) shows a significant uncertainty reduction. There is hardly any information gain, when considering another three additional measurements. This significant decrease follows the benefit curve of Fig. 6.1, where with increasing sampling size the utility gain decreases.

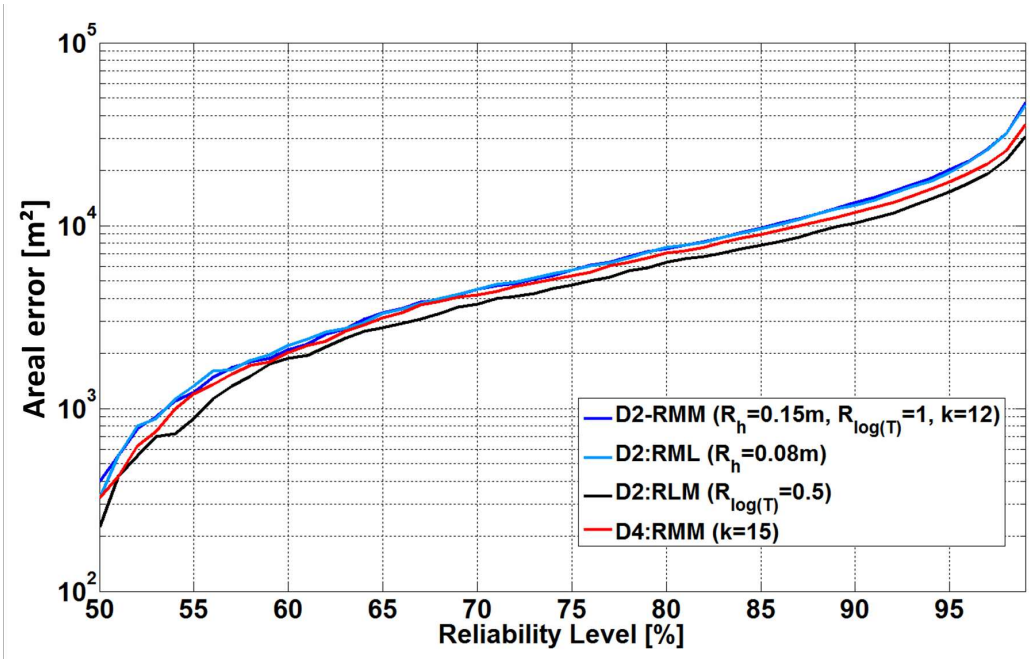


Figure 8.9.: Relationship between reliability level and area sacrificed due to uncertainty, showing four sampling scenarios based on measurement accuracy and type at constant sampling size.

In comparison, increasing the measurement accuracy always leads to information gain for the proposed sampling campaigns. The largest information gain is achieved by increasing the measurement accuracy in both sampling types from RHH to RMM.
The presented results depend on the pre-defined sampling campaign and the assumed synthetic truth that generates the data for conditioning. However, with these limitations, sampling campaign D4:RLL (high measurement accuracy in both sampling types and maximum sampling number) provides the best sampling scenario for uncertainty reduction.
Overall, Fig. 8.10 shows that, with increasing sampling number and accuracy, the areal error demand reduces. Furthermore, it provides information that uncertainty reduction reacts non-linear to the chosen sampling parameters. Therefore, epistemic uncertainty reduction is catchment- and scenario-specific but leads to more robust and confident decisions.

Which Data Should Be More Accurate

Fig. 8.10 shows that measurement accuracy plays a significant role in epistemic uncertainty reduction. A comparison of measurement accuracy in hydraulic head and log-transmissivity data is shown in Fig. 8.11 for a pre-defined reliability level $\beta = 90\%$. Doubling or reducing the measurement accuracy in hydraulic head has less impact on the areal delineation error. The accuracy of log-transmissivity influences the areal error and decreases the uncertainty demand with more accurate sampling. Therefore, the accuracy in direct measurements is more important than to indirect measurement data in this scenario set-up.

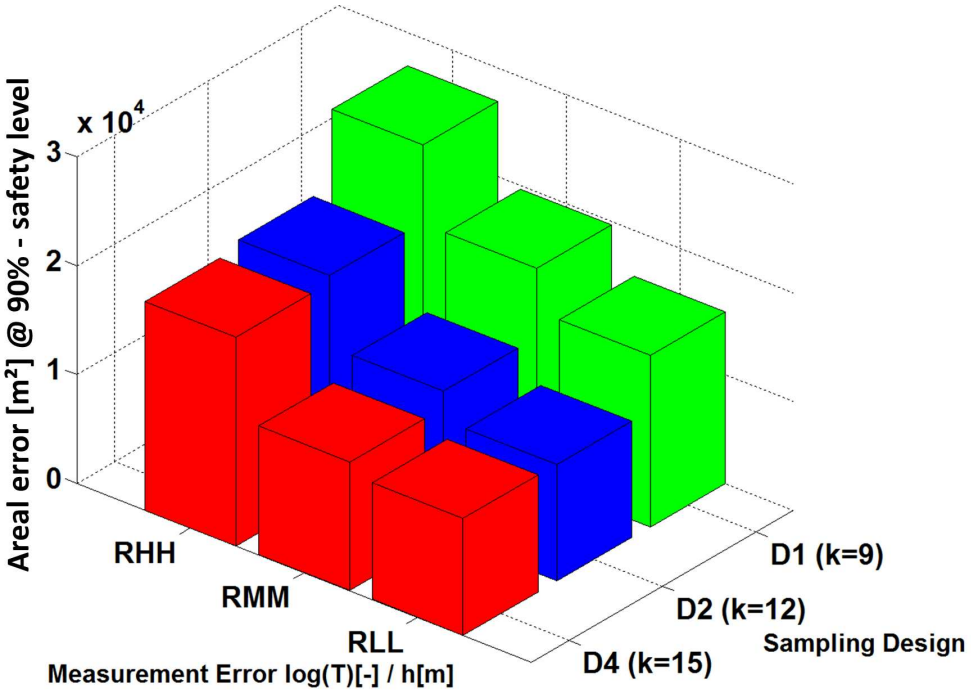


Figure 8.10.: Comparing measurement quantity and quality by showing utility value areal error A_{err} for three design cases ($D1, D2, D4$) and three measurement qualities from low (RHH) to high measurement accuracy in hydraulic transmissivity $log(T)$ and piezometric head h (RLL).

8.5.2. Improved Risk Estimation

In Section 8.5.1, the effect of uncertainty reduction is shown by reducing the wrongly delineated area due to uncertainty. This effect has been measured with the utility value A_{err} (Eq. 6.6). de Barros et al. (2012) argue that uncertainty reduction depends on the chosen prediction goal or the environmental performance metric. Here, I investigate the effect of increasing sampling size on the prediction goal of customer minutes lost (CML). Customer minutes lost is a risk measure used in drinking water supply to assess the time of water being unavailable within 100 years.
$n_{CML} = 50,000$ Monte Carlo simulations are performed to calculate the water supply reliability, measured by customer minutes lost (see Eq. 6.8). The exposure time t_{exp} at location S follows the cumulative density distribution of exposure time (fourth vulnerability criterion). The probability P_f of Eq. (6.8) for a major aquifer contamination to occur at location S, is assumed to follow a β-distribution (see Lindhe et al., 2009) with failure occurrence of at least once every 20 years and in a most likely event of once every 5 years. The demand behavior d of a consumer follows a positively skewed β-distribution with given pre-defined lower ($d_{min} = 80\,ld$) and upper bounds ($d_{max} = 290\,ld$).
In contrast to the previous three sampling designs (9 (D1), 12 (D2) and 15 (D3) measurements), I only use the sampling design D1 and D3. Lindhe et al. (2009) defined for the Gothenburg system an acceptable level for customer minutes lost of $R_{crit} = 144\,CML$,

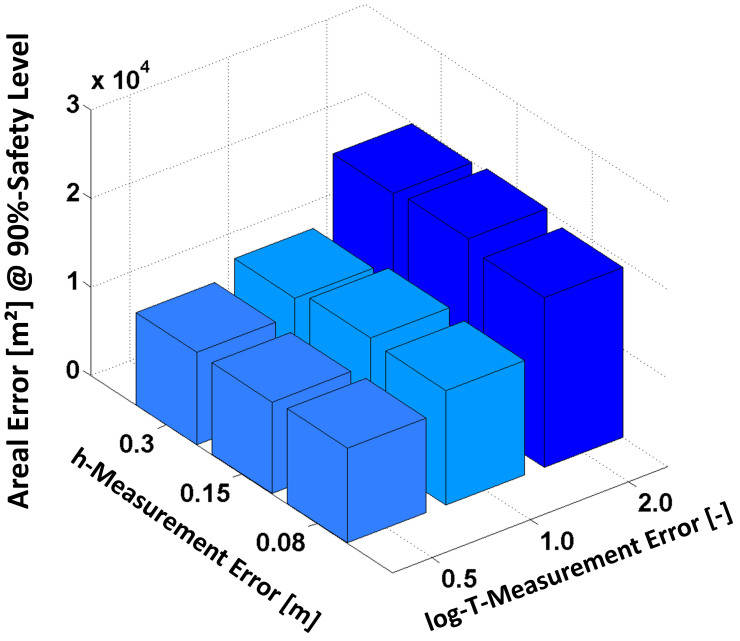


Figure 8.11.: Assessing for design $D2$ ($k = 12$) the utility for doubling or halving the quality in head or conductivity measurements.

which indicates that a consumer has to suffer $10\,days$ of water shortage within $100\,years$ after a single spill event. In the synthetic case study, a spill event at location S drastically exceeds this acceptable risk level R_{crit}, as aquifer contamination would last for several days or even years. Fig. 8.12 shows the cumulative CML distribution for location S. The mean customer minutes lost for scenario 0 is $\bar{R}_{scen0} = 760\,CML$, which indicates a mean supply failure of approximately $53\,days$ within $100\,years$.

The results show, that with increasing reliability level, the customer minutes lost value increases in all three scenarios. This is due to the fact that the more severe events, causing longer supply failure, are more seldom. Taking additional data into account will decrease the possible errors in the risk prediction. Thus, at first sight, one is surprised that the customer minutes lost increases ($R_{75,scen0} = 1188\,CML$ and $R_{75,scen2} = 2029\,CML$) with increasing sampling size. Please note that, as conditioning reduces uncertainty, it moves VIP lines and thus the estimates of well down-time closer to reality. This indicates that the calculated exposure time ($t_{exp,S,scen0} = 11\,d$) of the unconditional model (scenario 0) is underestimated in comparison to the more realistic exposure time ($t_{exp,S,scen2} = 16\,d$). Therefore, uncertainty reduction helps to estimate the expected impact more accurate. While the probability distribution of impact will surely become more information with more data, one cannot predict prior to sampling in which direction it will shift. Here, it shifts towards larger impact values, thus providing a more confident but larger risk estimate.

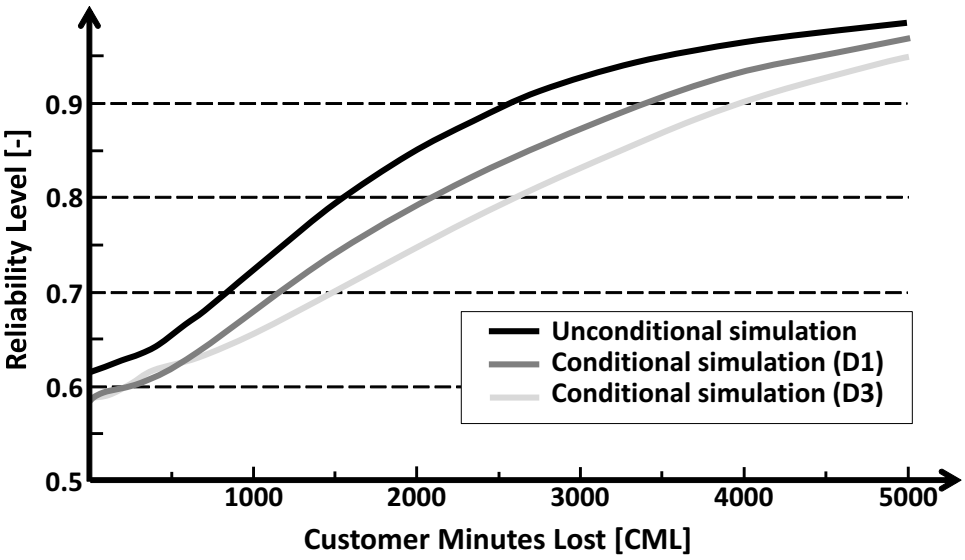


Figure 8.12.: Cumulative probability distribution of customer minutes lost for the three given scenarios evaluated at location S.

8.6. Application of Economic Decision Analysis Concepts

As introduced in Chapter 6, two economic decision analysis concepts are illustrated. The first concept follows the delineation concept based on peak arrival time (VIP 1) and the economic benefit of using additional data to reduce epistemic uncertainty (more accurate delineation) compared to sampling costs. The second approach demonstrates a prioritization scheme of risk mitigation alternatives based on economic valuation techniques. The study is based on the model domain introduced in Section 8.5.1.

8.6.1. The Economic Value of Additional Data

The previous results (see Section 8.5) show insights on the value of additional data in comparison to more accurate measurements. Here, I want to demonstrate the worth of additional data expressed through their economic value as described in Section 6.1. The economic value of information gained by additional data has to outweigh the costs related to data acquisition. In this study, the utility value to measure uncertainty is areal error, A_{err} as defined in Eq. (6.5). It is a translation of uncertainty effects into a monetary context by assuming an economic value for each parcel land within the catchment. Here, I assume a unit monetary value per square meter. Thus, it allows to compare sampling designs according to their costs and benefits regarding system safety as illustrated in Fig 6.1.
For simplicity, I assume for Eq. (6.2) that total costs equal the installation costs of monitoring wells $C_i = C_{samp}$. Monitoring costs (C_{op}) and restoration costs (C_f) are neglected. Thus, the costs $Y = u_1 = C_{samp}$ increase with an increasing sampling network X (Fig. 6.1). The sampling costs (see Eq. 6.2) are assumed to be 5000€ per data acquisition (well). The

measurement error to assess the sampling quality is chosen according to the medium values ('M') of Fig. 8.8. In the scenario chosen here, the costs associated to area being wrongly delineated due to uncertainty equal $1 €/m^2$.

Fig. 8.13 shows five safety levels β of delineation and their behavior in the light of sampling. The total costs are calculated according to Eq. (6.2). It is the sum of monetized benefits and sampling costs. The points located on the sampling-design axis represent the three conditional sampling scenarios with 9, 12 and 15 measurements. For a given reliability level, such as $\beta = 90\,\%$, the total costs decrease until 9 measurements, despite the fact that the sampling size increases. This is due to the information gain obtained by additional data acquisition to reduce delineation uncertainty. Thus, the benefit in saving areal costs is larger than the costs for data acquisition. Data sampling is performed as long as the total cost curve for any given reliability level decreases. The return point (minimum of the objective function, see Eq. 6.1) denotes the cost-optimal solution given a pre-defined safety level β. Here, taking $\beta = 90\%$, the cost-optimal solution is sampling campaign 1 with 9 sampling locations (see Fig. 8.7). Again, stakeholders could choose to accept the total amount of costs (here delineated area), such that a higher safety level $\beta = 95\%$ is available at no additional costs in comparison to the initial situation or a fixed investigation budget.

Fig. 8.13 strongly depends on the sampling costs C_{samp}. Increasing sampling costs (e.g., $10,000 €/well$) reduces the information value of additional sampling much faster. Therefore, the pay back to invest into data is depleted earlier. In this specific case, the minimum for $\beta = 90\,\%$ is reached with 8 sampling locations and costs increase dramatically beyond. Contrary, cheaper available measurements (e.g., $2,500 €/well$) will lead to larger cost-optimal sampling campaigns. Here, the optimum would be incorporating 12 measurements (D2) for the reliability level of $\beta = 90\,\%$.

In summary, the efficiency and worth of data depends on pre-defined costs of sampling and of wrongly delineated area, the pre-defined sampling campaigns (scenario or design cases) and thus on the sampling locations. Measurement locations are often fixed, so that such a scenario-based decision analysis is acceptable. Nevertheless, the decision analysis framework would benefit from casting the problem into a formal optimization framework (e.g., Nowak et al., 2010).

8.6.2. Economically-based Prioritization of Risk Mitigation Alternatives

The following is an example to demonstrate the use of VIP-based information for economically-based prioritization of risk mitigation alternatives. Fir simplicity, the example is based on two mitigation alternatives.

Fig. 8.14 shows the two risk treatment alternatives. The first scenario follows the sampling campaign as discussed before (uncertainty reduction due to quantity of data, Section 8.5) with $\beta = 90\,\%$. The total costs for sampling increase with each additional sampling. Here, exposure time increases with additional sampling as collected data indicate a smaller wellhead protection outline. The second scenario is related to insurance premiums. The insurance will cover the damage after an event has happened. Here, I assume a constant payment of $150 €/a$ over a $20\,year$ period. Each mitigation alternative is monetary valued and plotted in Fig. 8.14.

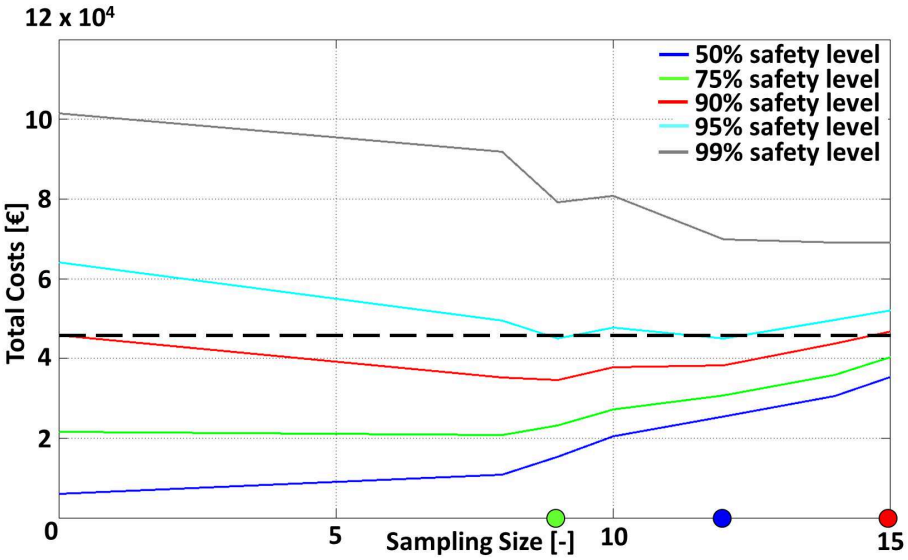


Figure 8.13.: Showing for three given design cases (colored circles) the safety level depending on total costs for drilling ($5000\,€/well$) and areal error ($1\,€/m^2$) versus sampling size.

These two simplistic risk mitigation scenarios are set into an economically-based decision framework and compared to the economic valuation method, damage replacement costs, as defined in Eq. (6.10). In this specific example, the damage replacement costs measure the unavailability of drinking water, which is expressed by the vulnerability criterion of well down-time t_{exp}. The costs related to water production are assumed to be $\gamma = 0.45\,€/m^3$. For illustrative purposes, the pumping rate is assumed to be $Q_p = 100\,l/s$ and not $Q_p = 1 \cdot 10^{-4}\,m^3/d$ as described in Section 8.1. With increasing exposure time, the consequences increases (see Fig. 8.14, red line).

The results of Fig. 8.14 show that, for an expected exposure duration of $t_{exp}(\boldsymbol{x}_S) \leq 18\,d$, the investment in monitoring is less expensive than to pay the damage caused by well contamination. For exposure durations of $t_{exp}(\boldsymbol{x_S}) > 17\,d$, insurance is the economically most cost-effective option. The pay back of the premium starts with exposure durations of $t_{exp}(\boldsymbol{x}_S) \geq 8\,d$. For exposure durations of $t_{exp}(\boldsymbol{x_S}) < 17\,d$, the sampling scenario is less expensive than the premium, such that monitoring should be the preferred choice up to total costs of $C(t_{exp} = 7\,d) = 27,216\,€$ as larger costs are better to handle with an insurance policy. This is due to the constant cost assumption and would need more sophisticated insurance cost functions that account for discount and inflation rates, and so forth.

Of course, sampling does not mitigate actual damage but it provides valuable information on VIP lines that are closer to reality. Here, I translated the information gain into a monetary context by saving costs related to wrongly delineated area. The sampling scenario is a risk mitigation opportunity that reduces the uncertainty component (*probability*) within the risk definition. Other scenarios may reduce the contaminant load (*impact*) to the well and actual costs are related to these actions. The premium scenario follows an *end-of-pipe strategy*, accepting the current risk situation, assuming that insurance money can compensate

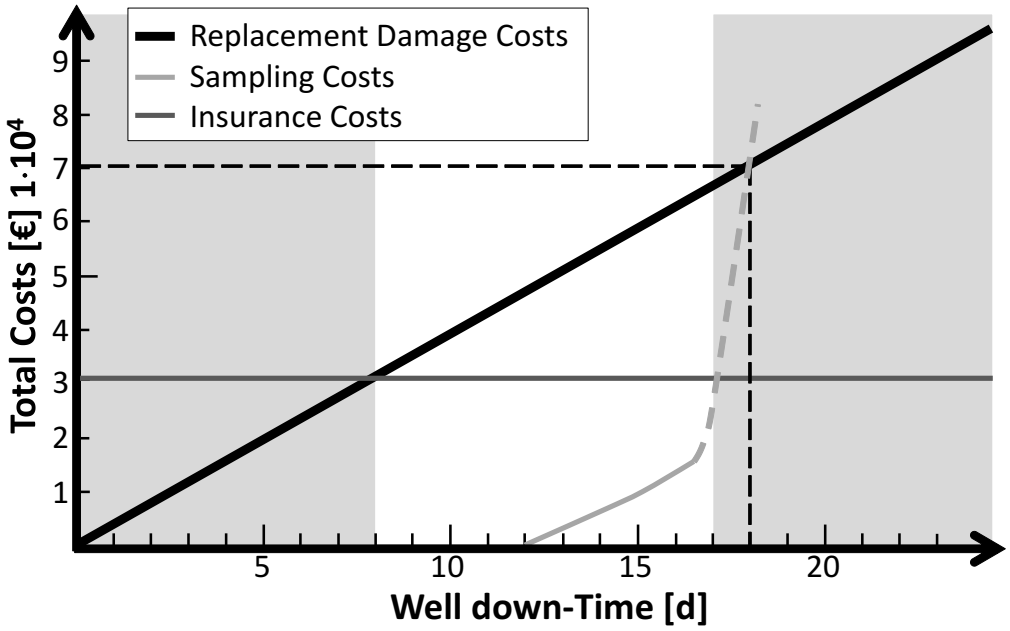


Figure 8.14.: Rational risk-based decision support, comparing here two alternatives (insurance premium (green), data sampling (blue)) to damage replacement costs (red). The dashed-blue line is a projection of costs for further sampling.

for all damages. Relating these three types of risk mitigation strategies to each other in an economic optimization context will need future research.

8.7. Findings of VIP Applied to a Synthetic Test Case

In this chapter, the vulnerability concept has been applied to a synthetic example with a 2D semi-confined heterogeneous aquifer and a single pumping well. Via finely-resolved Monte Carlo simulations, maps of vulnerability isopercentiles (VIP maps) were calculated. The computational costs are kept moderate by combining the reverse formulation of advective-dispersive transport and the concept of temporal moments.
This probabilistic information provides additional information and enables risk managers to map non-compliance lines for pre-defined critical levels. In the specific example, it turned out that deterministic approaches assume a spill location to be no risk due to larger travel times than the critical value. Nevertheless, the VIP lines show that in more than $\approx 40\,\%$ (conditioned model result) or $\approx 55\,\%$ (unconditioned model result) of all cases the contamination is a risk to the well source. Thus, in about half of all cases, deterministic results would have underestimated the risk of that contaminant source.

- The finely resolved geostatistical and probabilistic approach allows to separate between dilution of concentration and uncertainty in plume position.

- Despite the (small) loss of information due to the temporal moment approach, the gain in computational efficiency and the resulting accessibility to probabilistic information is more valuable in my opinion.
- It considers peak instead of bulk arrival time, thus rendering larger and safer wellhead-protection zones.
- The probability distribution of each vulnerability criteria at a location S provide information on all possible solutions.
- The design of risk mitigation options (e.g., wellhead protection zone delineation) can be performed, depending on the desired reliability level.
- The VIP framework is able to update model predictions with additional and better data, such that the prediction is closer to reality and more robust. The technique allows to compare the benefits of using more or less data, and/or data with smaller or higher quality.
- The degree of uncertainty reduction responded non-linear to a linear change in quantity and quality of data. Quite intuitively, the sampling campaign with the best measurement accuracy and with most sampling locations reduced the uncertainty most.
- The amount of sampling has been evaluated in an economic context, providing the cost-optimal choice (1) to increase the safety level at no additional costs (here, from $90\,\%$ to $95\,\%$) or (2) to improve the robustness in decision-making as VIP lines move closer to reality (here, 9 measurements for reliability level $\beta = 90\,\%$). In addition to this, customer minutes lost increased with reduced uncertainty, as the sampling data used in the example pointed to larger necessary wellhead protection zones with increasing well down-time.
- Economic valuation techniques can be applied to compare several risk mitigation strategies, such as additional catchment investigation, insurance premiums, and so forth. The chapter concluded with setting up a rational risk-based decision framework to compare different risk mitigation alternatives with each other in an economic context, evaluating mitigation alternatives that either reduces probability or impact.

9. Application of VIP - Burgberg Test Case[1]

This chapter demonstrates the application of the VIP concept to a real case study. The key point is to demonstrate that risk-aware wellhead protection delineation to guarantee safe water supply is ready for practical application by water stakeholders and consultants. The Burgberg catchment operated by the „Zweckverband Landeswasserversorgung“ will serve as demonstrator. The delineation of the Burgberg catchment has recently been reconsidered. Thus, comparing the actual decision to a robust and risk-based decision framework is of special interest. Also, a second VIP tool set will be compiled to show that the VIP framework can be operated with readily available end-user software. The case study features a karstic aquifer, represented by a model, which has been set up by an engineering company. A methodological description is provided in Section 4.3. With the presented tool set, the current reliability level is assessed, and the risk mitigation options are presented to support stakeholders in taking future risk-informed decisions. The mitigation options are compared based on the utility value of additional delineated area.
Section 9.1 introduces the model set up for Burgberg catchment. Section 9.2 is analogous to Section 8.2, showing the vulnerability maps for the real test case. Section 9.3 provides a risk-aware delineation management concept and Section 9.4 concludes this chapter with a brief summary.

9.1. Burgberg Case Study

The Burgberg catchment, operated by the Zweckverband Landeswasserversorgung (LW), is located on the eastern Swabian Alb in the south-west of Germany, and is part of one of the largest drinking water catchment in Germany. The total area is $A_{II} = 1.1\,km^2$ for the inner wellhead protection zone defined by solute travel times of $\tau_{crit} = 50\,days$ or less (zone II) and $A_{III} = 458.12\,km^2$ for the extended protection zone (zone III) according to the applicable guidelines (e.g., DVGW, 2006). The well catchment is karstic and includes an Upper Jurassic aquifer that is partly covered by Quaternary gravel channels with varying thickness of up to $10\,m$. The karst aquifer mainly consists of two aquifer layers. The upper aquifer consists of strongly karstified reef limestone. The lower part of the Upper Jurassic is less porous and consists of upper and lower limestone units. The Lacunosa marl (Lower Jurassic) follows the lower limestone and marks the aquitard level (e.g., Schloz et al., 2007). A well field extracts groundwater for drinking water use from three deep karst wells, screened in the upper aquifer layer.

[1] Parts of this chapter has been published in Enzenhoefer et al. (2013).

9.1.1. Numerical Model Set-up (step 1)

Lang and Justiz (2009) developed a numerical model for fully saturated, single-porosity, steady-state and semi-confined groundwater flow in the Burgberg catchment based on MODFLOW (e.g., Harbaugh et al., 2000). The hydrogeological model consists of three layers, representing (1) the Quaternary gravel channels, (2) the highly conductive Upper Jurassic aquifer and (3) the lower part of the Jurassic karst aquifer with lower conductance (limestone). Due to hydrogeological surveys (e.g., tracer and pumping tests) and known geological settings, the hydrogeological model consists of one zone for the first layer, 12 zones for the second layer and 9 zones for the third layer. The zonation and grid density increases around the well (see Fig. 9.1). The third layer is shown in the Appendix (Fig. B.1). The discretization within the vicinity of the well is $\Delta x_w = \Delta y_w = 12.5\,m$ and extends to $\Delta x_d = \Delta y_d = 100\,m$ in the remote parts of the catchment, resulting in $n_c = 75,797$ actively used model cells.
Three rivers and two karst springs are incorporated as leakage boundaries. The north and south boundaries are modeled as no-flow boundary conditions. They follow the geometry of streamlines that have been calculated from a larger catchment model (e.g., Lang and Justiz, 2009). The boundary in the west is again a no-flow boundary, where the catchment touches the great European water divide. The great European water divide marks the border between the catchments of the river Rhine and the Danube river, which end in the North Sea and the Black Sea, respectively. The east boundary is implemented as a Dirichlet boundary condition with $H = 445m$, representing a gallery of springs. The pumping rate of the well field is on average $Q_p = 300\,l/s$. The groundwater recharge rates within the domain are calculated on the basis of data given by the federal environmental protection agency of Baden-Württemberg (LUBW) from 1961 to 1990. The mean recharge rate for this time period is $q_r = 10.5\,l/(s\,km^2)$ ranging from zero in the east to $q_{r,max} = 50\,l/(s\,km^2)$ in the west, equipped with a global multiplier for calibration.

The current model is a single-porosity model, and could be upgraded to a dual-porosity (e.g., Barrenblatt et al., 1960), multi-rate mass transfer (e.g., Haggerty and Gorelick, 1995) or continuous time random walk (e.g., Berkowitz and Scher, 1995) model for better representation of transport in karst. However, the single-porosity model is the actual used model in practice for that catchment. Such upgraded models would require much more calibration data that are not available. Also, changes in the model would not affect the methodology that is in the focus of this thesis.

9.1.2. Model Calibration and Conditioning (step 2-3)

The model was calibrated carefully by trial and error at steady-state, using 27 hydraulic head measurements taken on two reference days with two different pumping rates of the pumping wells, pump test information, sieve curves from bore core analysis and tracer tests (e.g., Lang and Justiz, 2009). Parameters for calibration included a total of 12 hydraulic conductivity values and the multiplier for recharge. The involved objective function in Eq. (5.22) is the same as used by PEST. The calibrated hydraulic conductivity for the upper layer equals $K_1 = 4 \cdot 10^{-3}\,m/s$, whereas the hydraulic conductivity values for the zones within the sec-

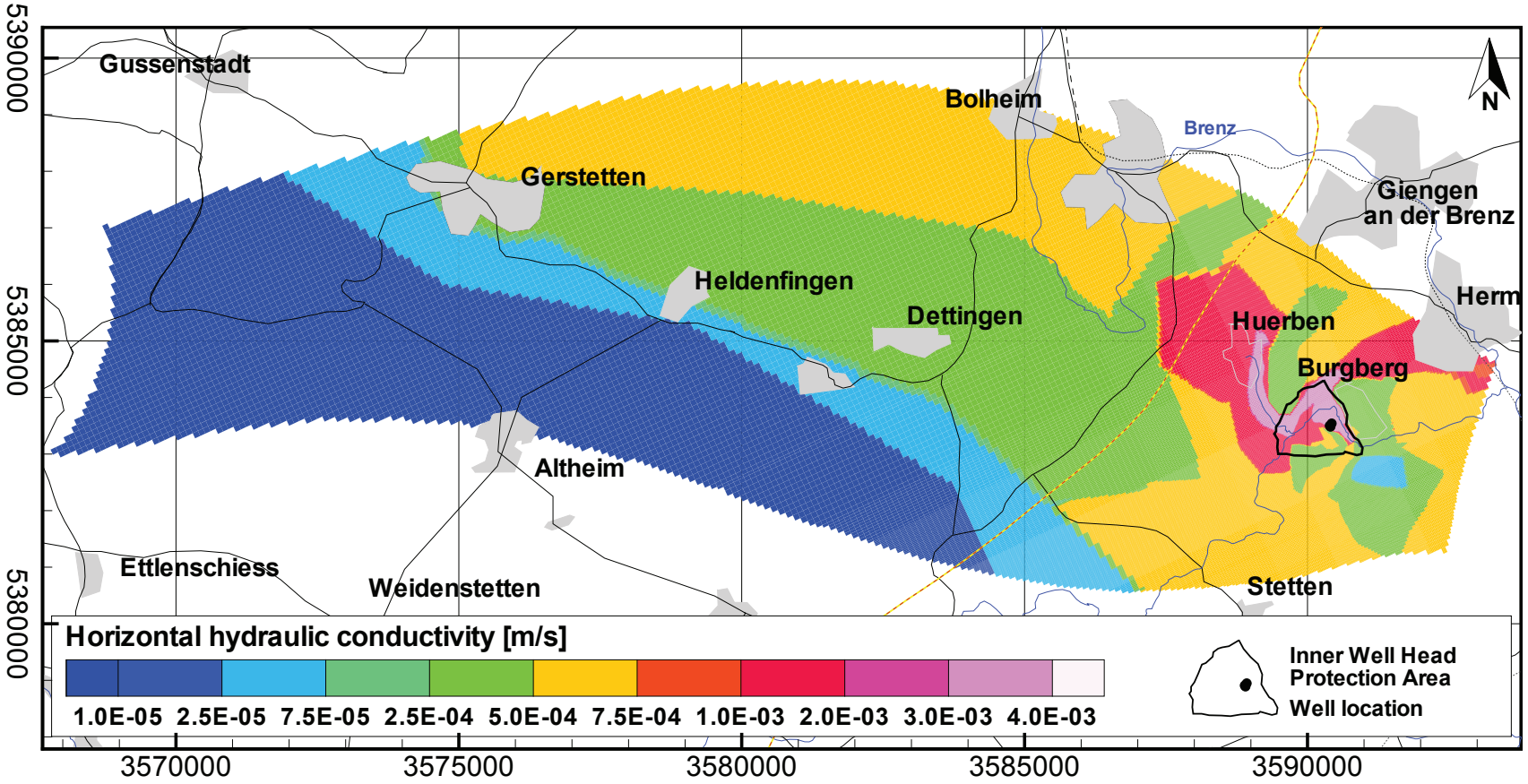


Figure 9.1.: Zonation for the upper Burgberg aquifer, showing layers 2 of the MODFLOW model by Lang and Justiz (2009). The position of the wells are marked by the black dot, black lines show the road system, blue lines mark creeks and rivers and the gray areas show cities and villages.

ond layer range between $K_{2,min} = 1 \cdot 10^{-5}\, m/s$ and $K_{2,max} = 5 \cdot 10^{-3}\, m/s$. For the third layer, values fall between $K_{3,min} = 1 \cdot 10^{-6}\, m/s$ and $K_{3,max} = 5 \cdot 10^{-3}\, m/s$ (see Fig. 9.1). Lang and Justiz (2009) assumed that the porosity is $n_{e,1} = 10\%$ for the gravel material in the top layer and $n_{e,2,3} = 3.6\%$ for the two layers that represent the Upper Jurassic karst aquifer. The calibrated K values show an increasing trend towards the well. This is fully consistent with the theory of effective conductivity values in convergent flow fields through heterogeneous (here: karstic) media (e.g., Indelman et al., 1996).
For all 13 calibrated parameters, a linear sensitivity analysis is performed to obtain the Jacobian $\boldsymbol{J}$. For the head measurements, I assume a measurement error variance of $\sigma^2_\phi 2 \cdot 10^{-4}\, m/s$ (corresponding to $\approx \pm 3\, cm$ as the $95\,\%$ confidence interval) and error correlation of 0.5 to generate the error covariance matrix $\boldsymbol{R}$. The a-priori variance for the log-conductivity values in each zone is set to $\sigma^2_{\boldsymbol{pp},K} = 0.25$ (i.e. the $95\,\%$ confidence interval is one order of magnitude) under the assumption that conductivity is uncorrelated between the zones before calibration. For the groundwater recharge multiplier the variance is assumed to be $\sigma^2_{qr} = 0.156$. All these assumptions are inserted into Eq. (5.29) to obtain the post-calibration error covariance. The average individual post-calibration variance for all conductivity values (on the diagonal of the matrix) is $\sigma^2_{pp|\boldsymbol{d}} = 0.031$, and the average correlation (all non-diagonal elements) is $\rho_{\boldsymbol{pp}|\boldsymbol{d}} = -0.011$ (Nowak et al., 2010).
In total, $n_r = 1000$ conditional flow realizations were generated. For each realization, the hydraulic conductivity K and the recharge multiplier are randomly drawn according to Eq. (5.30). Because calibrating mostly to pressure data in deterministic calibration leaves all transport parameters uncalibrated, independent statistical assumptions are necessary for them. Their range is given in Table 9.1. Porosity in all three layers and the longitudinal and transversal horizontal dispersivities are kept uncertain, following for simplicity uniform distributions. The porosity range is assumed to take karstic features, such as faults and fractures, into account. The transverse dispersivity in the vertical direction is kept constant at $\alpha_{TV} = 0.01\, m$. The horizontal transverse dispersivity is assumed to be linked to the longitudinal one via $\alpha_{TH} = \alpha_L/20$.

9.1.3. Vulnerability Set-up (step 4-8)

The simulation end time for the reverse transport is set to $T_{Tot} = 30\, years$ and the total number of particles is $n_P = 500,000$. The accuracy of breakthrough curve reconstruction to obtain well vulnerability criteria from this relatively small particle number is compared to the results from a reference simulation with $n_P = 15,000,000$ particles in Section 9.2.6. The simulations were run on a personal computer with 2 Intel cores at 2.26 GHz and 4 GB RAM

			Layer I	Layer II	Layer III	pdf-type
porosity	n_e	[%]	$8-23$	$2-30$	$2-30$	uniform
dispersivity	α_L	[m]	$20-500$	$20-500$	$20-500$	uniform
	α_{TH}	[m]	$1-25$	$1-25$	$1-25$	$\alpha_{TH} = \alpha_L/20$

Table 9.1.: Uncertain transport parameters for the Burgberg test case.

resulting in a simulation time of $\approx 7 days$. This should be perceived as a reference simulation effort. Much smaller time may be achieved, when reducing the number of $n_r = 1000$ conditional realizations, or when reducing the total simulation time of $T_{Tot} = 30\, years$ or with more computational power, if available. With only $n_r = 100$ realizations (sufficient to approximate a $90\,\%$ probability reasonably well) and only $10\, years$ of total simulation time, the computational time reduces to $\approx 6\, hours$. As a last step, the four vulnerability criteria are evaluated in each realization.
By combining the vulnerability criteria from all conditional realizations, the empirical probability distribution for each well vulnerability criterion at each location in the domain is available. The underlying critical levels for non-compliance levels (see Eq. 4.2) have been chosen in close cooperation with the water supplier „Zweckverband Landeswasserversorgung“ as given in Table 9.2, and correspond at least to the maximum contaminant levels specified by the applying German regulation (DVGW, 2006), compare also US regulation (e.g., US EPA, 2009).

9.2. Probabilistic Well Vulnerability Results - VIP Applied to Burgberg

The Fig. 9.2 shows the isopercentiles $\alpha = [0.1,\ 0.5,\ 0.9]$ for the four probabilistic well vulnerability criteria, based on the critical values from Tab. 9.2. Quite intuitively one can observe, that the smaller the non-compliance probability α (i.e., the larger the corresponding reliability level β), the larger the required delineated area. The 50th-percentile line, for example, encloses the area A_{50} outside of which the respective vulnerability criterion is violated with less than $50\,\%$ probability (see Eq. 4.2). The area between the A_{50}- and the A_{10}-line may be interpreted as the safety margin to achieve a reliability of $90\,\%$. Thus the reliability of the supply system based on the area $A(\beta)$ can be plotted for each possible reliability level β against the (areal) costs, see Fig. 6.2.
The filled (red) dot within Fig. 9.2 marks the location S ($1.2\, km$ upstream of the well field with grid coordinates $\boldsymbol{x}_{grid} = [239;\ 167]$; or in easting and northing coordinates $\boldsymbol{x}_s = [3589240;\ 5383470]$ m), where a spill event for the sake of discussion is assumed, in order to compare results and illustrate the VIP-based management concept.

Peak arrival time (VIP 1)	$WVC_{1,crit} = \tau_{crit}$	50	[d]
Maximum Concentration (VIP 2)	$WVC_{2,crit} = \zeta_{crit}$	$1 \cdot 10^{-8}$	[-]
Reaction Time (VIP 3)	$WVC_{3,crit} = \tau_{react}$	10	[d]
Well exposure Time (VIP 4)	$WVC_{4,crit} = \tau_{exp}$	30	[d]

Table 9.2.: Critical levels of well vulnerability for the Burgberg test case.

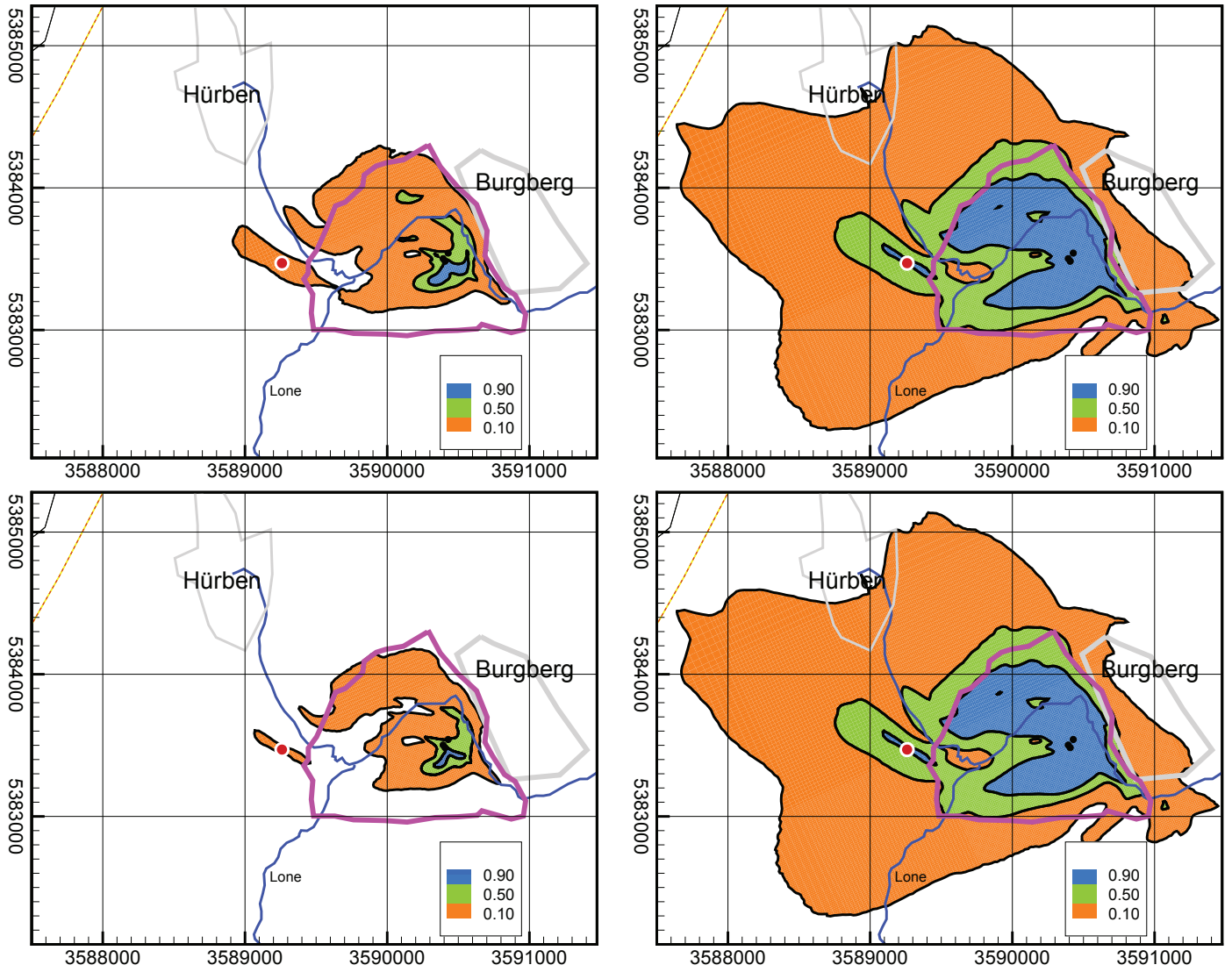


Figure 9.2.: Probabilistic vulnerability isopercentiles (VIP) for the four intrinsic well vulnerability criteria (VIP 1: top left, VIP 2: top right, VIP 3: bottom left, VIP 4:bottom right) from $n = 1000$ conditioned realizations.

9.2.1. VIP 1: Probability of Early Peak Arrival Time

The first VIP map (see Fig. 9.2, top left) represents the regulations for the German inner wellhead protection area, when using the critical travel time $WVC_{1,crit} = \tau_{crit} = 50\,d$ (see Tab. 9.2). Here, the peak arrival time is used instead of bulk arrival time, because this is the more conservative and more meaningful definition of arrival time in the advective-dispersive context (see Section 8.3). The thick solid line (purple) shows the existing well-head protection zone within the Burgberg catchment. At the spill site S (red dot) selected for illustration, the mean arrival time of peak concentration equals 142 days, with $23\,\%$ probability that peak arrival time is faster than $\tau_{crit} = 50\,d$. Apparently, the peak arrival times may easily fall below τ_{crit}, while the mean arrival time is much larger than $\tau_{crit} = 50$ days. This fact stems from the strong skewness of the arrival time distribution (see Fig. 9.3), which has often been found to be close to log-normal under uncertainty of hydraulic conductivity (e.g., Cvetkovic et al., 1992; Dagan and Nguyen, 1989; Nowak et al., 2008). The isopercentile lines of peak arrival time are shown for the non-compliance level $\alpha = [0.1, 0.5, 0.9]$, indicated by the three colors in Fig. 9.2. A more detailed discussion on the first VIP map, how peak arrival time is used for risk-aware delineation is provided in Section 9.3.

9.2.2. VIP 2: Probability of Insufficient Plume Dilution

The second VIP map (see Fig. 9.2, top right) shows the probability map that a point-like instantaneous contaminant spill of unit mass (e.g., $1\,kg$) causes a peak concentration in the well of more than $WVC_{2,crit} = \zeta_{crit} = 1 \cdot 10^{-8}\,\frac{kg}{m^2}$ (see Tab. 9.2). In case of a one-time spill event of a water-soluble compound at location S, I assume a total contaminant mass of $m_s = 10,000\,kg$ to be infiltrated to the aquifer. The expected peak concentration for this contaminant in the well is then calculated by multiplying m_s with the mean value of the second well vulnerability criterion. The expected concentration for a unit-mass spill in the test case at location S is $3.9 \times 10^{-8}\,[-]$ (see Fig. 9.2, top right), which leads to an expected peak concentration in the above contaminant example of $3.9 \times 10^{-1}\ mg/l$. Single realizations can be more or less diluted, thus leading to higher and lower possible peak concentrations as discussed in Fig. 8.5. The probability to not comply with the critical dilution factor ζ_{crit} at position S is 92 %. This indicates strongly that storing or handling such an amount of contaminants at location S should be avoided.

9.2.3. VIP 3: Probability of Too Little Available Response Time

The third VIP map shows the probabilities that an early-alert and emergency reaction system with a necessary reaction time of $WVC_{3,crit} = \tau_{react} = 10\,d$ will fail. Such an early-warning system could be a network of monitoring wells, sampled at least every τ_{react}- days, including the time for sample analysis and installation of mitigation measures to reduce or even avoid the impact on the drinking water supply. The reaction time is defined as the time available until a certain threshold concentration of $c_{crit} = 1 \cdot 10^{-7}\,mg/l$ is exceeded within the well for a unit mass release. The areal extensions of VIP contours for reaction time are smaller than

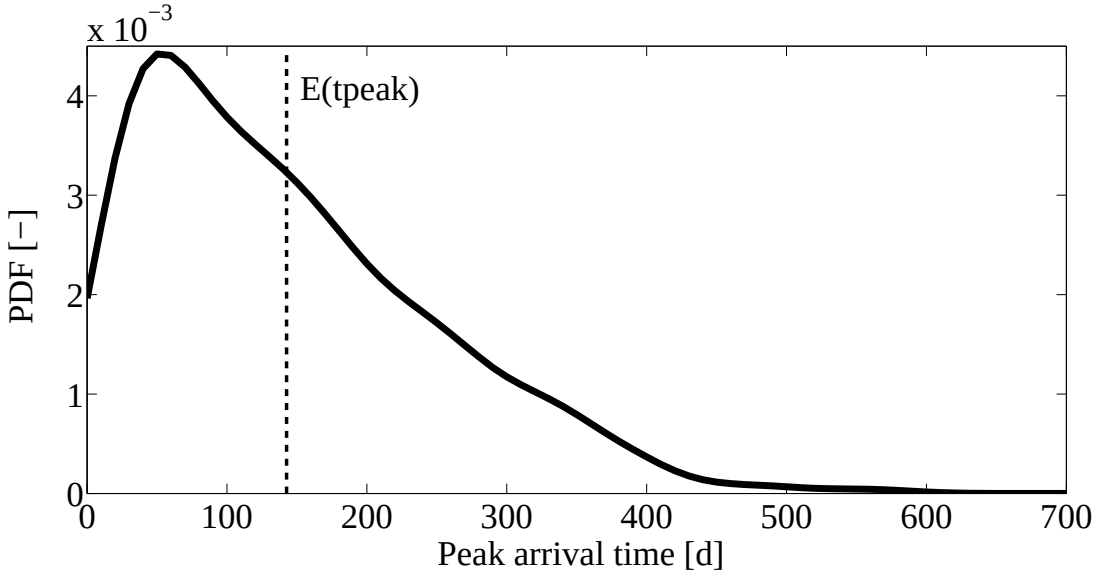


Figure 9.3.: Probability density distribution of peak arrival time for a spill event at location S.

those for the corresponding peak arrival time, because the first arrival of low concentrations c_{crit} is much earlier than peak arrival (compare Fig. 4.2 and Fig. 8.5). The mean reaction time for a unit spill at location S (in all realizations when c_{crit} is exceeded) is $66\,days$, which is by far larger than the critical value $\tau_{react} = 10\,days$. The probability that the reaction time is smaller (larger) than the given critical duration is $15\,\%$ $(85\,\%)$. The conditional probability of having a sufficient available reaction time in critical cases (where c_{crit} is exceeded) is not $85\,\%$ but only $22\,\%$ because realizations with $c > c_{crit}$ behave differently than those with $c \leq c_{crit}$.

9.2.4. VIP 4: Probability of Excessive Exposure Time

The fourth VIP map shows the probability of exposure times being larger than the given critical exposure time $WVC_{4,crit} = \tau_{exp} = 30\,d$ (see Tab. 8.1). Exposure time is defined as the duration a well is contaminated with solute concentration above a critical threshold concentration $c_{crit} = 1 \cdot 10^{-7}\,mg/l$, and thus unavailable for the water supply without treatment. The mean well down-time for location S equals 289 days. The probability of critical well contamination for more than $\tau_{exp} = 30\,d$ is $92\,\%$. $99\,\%$ of all realizations exceeding c_{crit} lead to well down-times of $\tau_{exp} \geq 30\,d$. In fact, contamination at almost any possible location within the catchment will lead to a well exposure duration larger than $\tau_{exp} = 30\,d$, for the considered concentration threshold level. Therefore, the map for VIP 4 and VIP 2 are practically identical (see Section 8.2). In order to avoid the risk of long well down-times, the protection area has to be enlarged up to the level, where either the risk acceptance level is fulfilled or the economic viability is over-strained.

9.2.5. The Impact of Considering Specific Vulnerability

Next, a brief example is presented on how the information in the VIP maps can be modified for contaminant-specific properties (here: linear sorption) and how this affects the VIP maps evaluated via Eq. (5.43). Fig. 9.4 shows the two ensemble-averaged breakthrough curves at location S, depending on the retardation factor R_d. The base case (dark gray line) is calculated without any attenuation potential besides the hydromechanical dispersion behavior of the aquifer. These intrinsic conditions are valid for conservative tracers or bacteriological and virological transport problems.
The other breakthrough curve (medium gray) shows a retarded concentration profile at the well with $R_d = 3.5$ due to linear sorption processes to the minerals. This value only serves to illustrate the flexibility of the vulnerability concept to account for fate and degradation processes due to the post-processing procedure. The retardation level is contaminant and site-specific and needs adaption to each catchment. The peak arrival time is shifted to later times, thus the isopercentile outlines within the first VIP map (peak arrival time) would become smaller than in the intrinsic case. Also, the intensity of the peak concentration in the well is reduced to much lower levels, causing smaller VIP contours for the second well vulnerability criterion. The corresponding change of the third VIP map (threshold arrival time) is more difficult to explain. On the one hand, the time of exceeding a given threshold level would be postponed to later times (compare first VIP map, Fig. 9.2, top left), thus leading to increased reaction times. On the other hand, all concentrations are attenuated more strongly (compare second VIP map, Fig. 9.2, top right), such that the critical concentration would be exceeded in fewer realizations. In this specific example, both effects would roughly cancel out, and the third VIP map would almost not change. The same would be true for the fourth VIP map (exposure duration), as exposure time shrinks with increasing attenuation, while it increases with the slower overall speed of transport. Tab. 9.3 summarizes the corresponding probability values of non-compliance obtained when assuming $R = 3.5$ for a spill at location S.

9.2.6. Accuracy of Breakthrough Curve Reconstruction from Small Particle Numbers

To assess the accuracy of breakthrough curve reconstruction from small particle numbers via the inverse Gaussian distribution (Eq. 5.41), two different solute breakthrough curves are demonstrated at the well (see Fig. 9.5). Again, a contamination event at location S serves

	VIP 1		VIP 2		VIP 3		VIP 4	
	$P_{1,S}$ [%]	$\bar{t}_1$ [d]	$P_{2,S}$ [%]	$\bar{c}_r$ [$-$]	$P_{3,S}$ [%]	$\bar{t}_3$ [d]	$P_{4,S}$ [%]	$\bar{t}_4$ [d]
intrinsic	23.2	142	91.8	$3.9e^{-8}$	22.1	66.4	8.3	289.1
with sorption	2.7	499	33.9	$1.1e^{-8}$	2.7	125.6	0.3	259.0

Table 9.3.: Ensemble-averaged well vulnerability criteria and the corresponding VIP value at location S for intrinsic transport conditions and with linear sorption ($R_d = 3.5$).

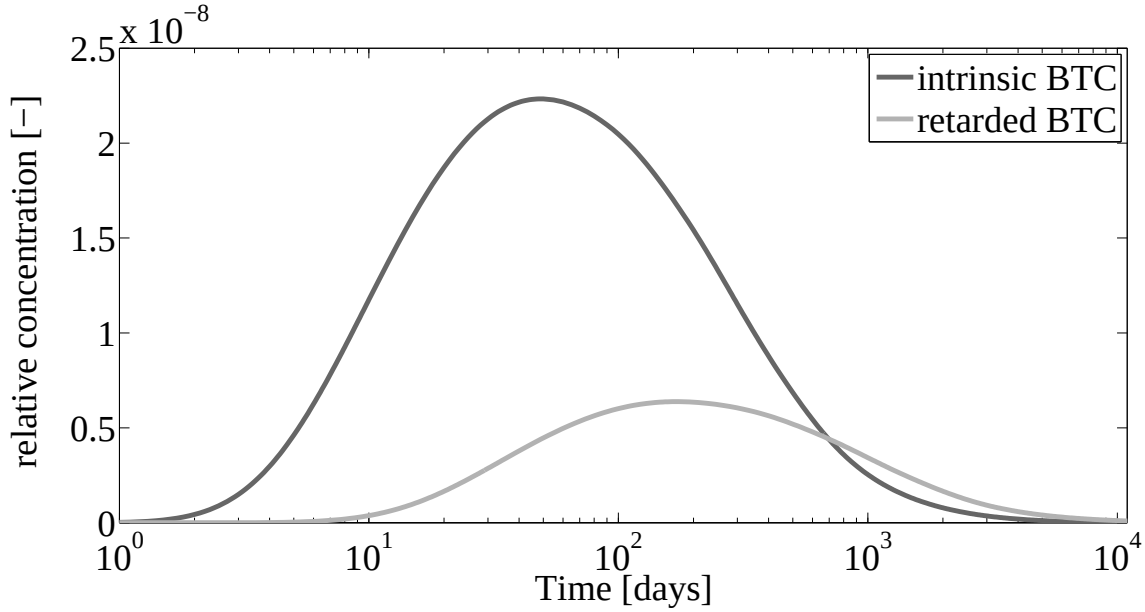


Figure 9.4.: Mean solute breakthrough curves from $n = 1000$ realizations within the well for a contamination spill at location S. Dark gray line indicates intrinsic conditions and medium gray line is the breakthrough curve including sorption effects with $R_d = 3.5$. Please note the logarithmic scale of the time axis.

for illustration. The solid black line is based on a reference simulation with $n_p = 15,000,000$ particles in a single realization, using the calibrated parameter set (e.g., Lang and Justiz, 2009). This number of particles follows a guideline of Kinzelbach (1988). However, all presented VIP calculations are based on calculations with a smaller number of particles $n_p = 500,000$ in the backward transport modeling case. The reconstructed breakthrough curve based on the inverse Gaussian distribution (IGD, Eq. 5.41) and the smaller particle number is shown as the blue line in Fig. 9.5. The values of the four well vulnerability criteria computed with these two methods differ by $\triangle t_{peak} = 2\,\%$, $\triangle c_{peak} = 26\,\%$, $\triangle t_{react} = 10\,\%$ and $\triangle t_{exp} = 7\,\%$. The high deviation in peak concentration $\triangle c_{peak}$ stems only from the high fluctuations in particle arrival even in the reference simulation, and would be reduced to below $10\,\%$, if applying a suitable noise filter to the reference BTC. These error levels fall far below the statistical fluctuation of the well vulnerability criteria between the realizations. Thus, the reduced model seems to be acceptable for the vulnerability concept, reflected against the fact that it yields a computational speed up by a factor of 30.

9.3. The Areal Re-delineation Risk Mitigation Concept

The VIPs discussed in Section 9.2 are domain-wide maps of exceedance probability, delivering the necessary statistical information for robust risk management decisions. Lines of equal probability, for example when choosing a maximum acceptable non-compliance probability α, are adequate outlines for risk-driven wellhead protection areas. Such areas offer

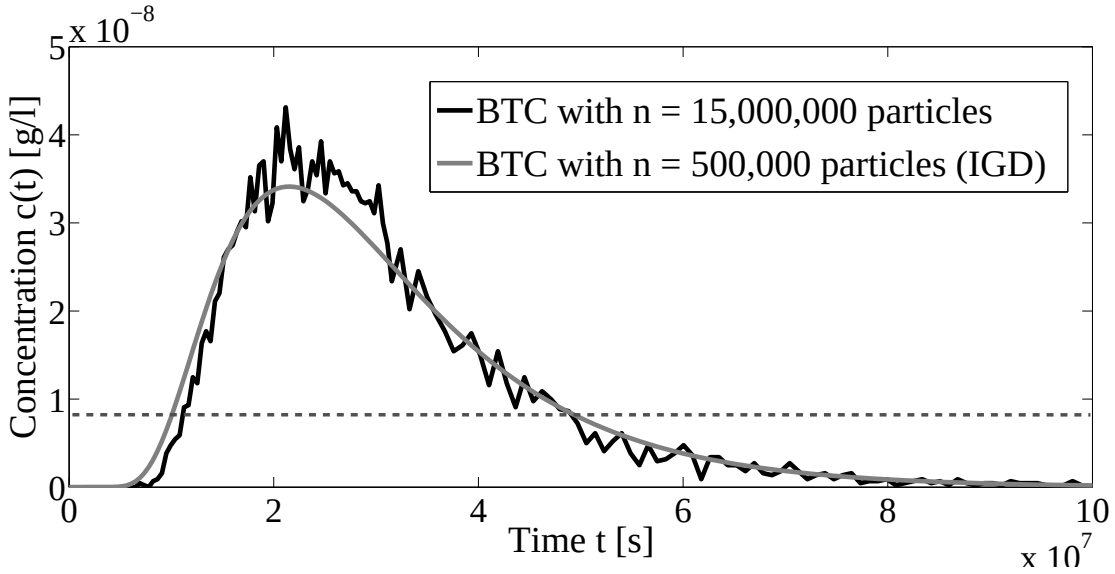


Figure 9.5.: Showing the accuracy of breakthrough curve reconstruction via the inverse Gaussian distribution for a single breakthrough curve. Black line is the reference simulation with $n_p = 15,000,000$ particles. Blue line: reconstruction based on $n_p = 500,000$ particles.

consistent reliability levels of $\beta = 1 - \alpha$ along their entire outline. In the following, the focus is only on the first probabilistic vulnerability criterion as the one being relevant for wellhead delineation to discuss the first three management options mentioned in Section 6.2.

9.3.1. Option I: Current Reliability Level and Smaller Consistent Area

Fig. 9.6 shows that the politically installed and currently used inner wellhead protection zone $A^{(0)}$ (purple line) for an arrival time of $\tau = 50\,d$ covers most of the area proposed by the 10-th isopercentile ($\alpha = 10\,\%$). However, there is excess protection area in the south and some unprotected area in the west. Expressed in numbers, the current delineation has a probability of non-compliance between $0\,\%$ in the south and $25\,\%$ in the northwest. Thus, the largest consistent reliability level of the current protection zone according to Eq. (6.3) is only $\beta^{(0)} = 75\,\%$ (see enclosed blue area, Fig. 9.6). The same consistent reliability level could already be achieved with a protection area enclosed by the 25-th isopercentile line, having only $A_{25} = 0.45\,km^2$ area instead of $A^{(0)} = 1.1\,km^2$ (see Fig. 6.2). The difference between these two areas is $\triangle A^{(-)} = 0.65\,km^2$. This freed up area is available for management option II.

9.3.2. Option II: Areal Reallocation

A much larger consistent reliability level can be achieved when re-allocating the excess area $\Delta A^{(-)} = 0.65\,km^2$ (see Eq. 6.4) according to a vulnerability-based isopercentile line that is

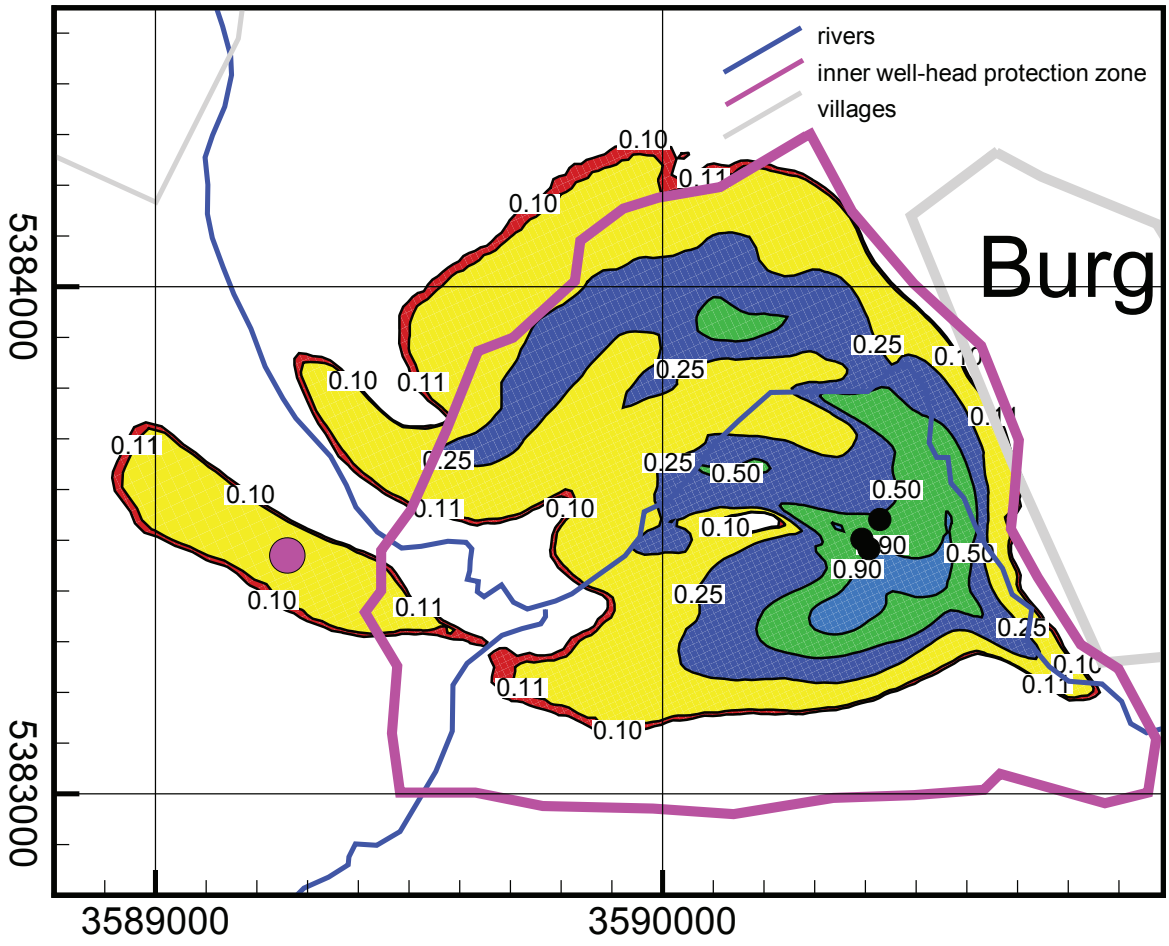


Figure 9.6.: Actual reliability level of current inner wellhead protection zone (blue) with $\tau_{crit} = 50$d at lower areal demand $A^{(-)}$and increased reliability levels at no (yellow, $A^{(0)}$) and only few additional costs (red, $A^{(+)}$).

found by $A_\alpha = A^{(0)}$. The dashed line (enclosing the yellow area) in Fig. 9.6 shows the resulting new protection zone, based on the 11-th percentile. It offers a reliability level increased from $75\,\%$ to $89\,\%$ at no additional areal costs.

9.3.3. Option III: Reliability Increase at Minimum Costs

Fig. 6.2 shows how the areal demand $A(\beta)$ of the overall protection zone depends on the desired reliability level β. Apparently, the areal demand increases more than exponentially with increasing reliability levels in the upper range. This curve roughly follows a Weibull distribution. Thus, very high reliability levels are only achievable with extraordinary costs, e.g., $\beta_1^{(++)} = 96\,\%$ with $A^{(+)} = 0.78\,km^2$, whereas a reliability level of $\beta_2^{(++)} = 90\,\%$ is comparatively easy to achieve with an additional area of $\triangle A^{(+)} = 0.08\,km^2$ (shaded red in Fig. 9.6).

9.3.4. Option III(b): Reliability Increase While Honoring the Current Delineation

Despite the larger delineated wellhead protection area $A^{(+)}$ for the chosen reliability level $\beta^{(++)}$, some areal patches in the south of the current wellhead protection area $A^{(0)}$ would become unprotected. This specific area was only delineated due to local political reasons, in order to protect springs close to the well gallery. Please remember, wellhead delineation is executed by state authorities (see Section 3.4). The second version of this option would take the existing delineation (purple line) and supplement it by all currently non-covered areas that fall within a desired reliability level. For example, the unification of the purple outline and the $11\,\%$ VIP line (yellow area) would achieve a consistent reliability level of $89\,\%$ while not giving up any of the currently delineated area. This would achieve the same consistent level as option II, yet at additional areal costs of $0.22\,km^2$.

9.4. Findings of VIP Applied to a Real Test Case

In this chapter, the VIP concept has been applied to a real test case and shows that accounting for uncertainty in risk management for wellhead protection is easily accessible with known software, such as MODFLOW, MODPATH and PEST. The real field site is located in a strongly karstified region on the eastern edge of the Swabian Alb, Germany.
In summary, this chapter demonstrated successfully that ...

- ... higher reliability levels of a given time-related wellhead protection zones demand a larger delineation area around the well.
- ... risk-aware delineation leads to smaller areal delineation demand at the same safety level, or that with the currently available delineated area a higher reliability level is achievable.

- ... politically and local circumstances within the delineation process can be transparently incorporated into the decision making process.
- ... contaminant-specific VIP maps are available through a fast post-processing procedure.
- ... saving computational simulation time is available, when using a low particle-based breakthrough curve reconstruction approach.

10. Application of STORM - The Burgberg Catchment

The goal of this chapter is to demonstrate the STORM concept on a realistic application, featuring a small set of synthetic hazards (Section 10.1). The application is based on the flow and transport calculations that have been presented for the Burgberg catchment (see Chapter 9). The risk quantification results (Section 10.2) are applied in a small risk management and prioritization study in order to show how STORM can support stakeholders to take risk-informed decisions and how different risk measures, risk objectives or risk estimation schemes lead to different risk management actions (Section 10.3). In this chapter, well exposure time (WET) is used as a demonstrator for choosing one stakeholder-objective risk measure. The chapter concludes with a brief summary (see Section 10.4).

10.1. Synthetic Hazard Test Case

Performing risk analysis for a real catchment is a challenging task. Up to date, it is not possible to find any information on hazard failure rates, nor do water managers know exactly how much mass enters the subsurface after a failure event has occurred. Water catchment managers are just starting to identify possible risk sources across the catchment, with only vague knowledge on actual mass storage at these hazardous sites. Due to these reasons, a synthetic hazard database is set up. It will serve the purpose to illustrate the benefit of using frequency analysis as a next step within risk quantification and management.
In the presented hazard test case, there are a total of three contaminant types $j = I$ to III, distributed over a number of hazards across the catchment. The hazards are defined such that each specific case of chemical, spatial or temporal hazard type (see Fig. 7.4, model choice) is present at least once. For illustration purposes, the number of hazards is kept small at nine hazards, in order to keep the case study simple and transparent. The nine hazards are listed in Tab. 10.1.
The assumed distribution of hazards within the Burgberg catchment is shown in Fig. 10.1. The three contaminant types may reflect any hazards such as chlorinated solvents, BTEX compounds or microorganisms. For convenience and not to be side-tracked on parameter accuracy for chemical properties, the three contaminant groups are kept generic, demonstrating the possibility to account for retardation, degradation or intrinsic effects. In the remainder of this thesis, I will therefore only refer to generic contaminant types. The current simplifications can be made without loss of generality, because the approach is not limited in hazard number or contaminant types. The available chemical property information are stored in the contaminant property list (see Tab. 10.2).

h	Location $\boldsymbol{x}_h$	Mass m_0 [kg]	Failure rate λ	CT	Haz. Type	Haz. Type	Haz. Type
	(grid coordinates)	[min / max]	$[F/a]$	j	(chemical)	(spatial)	(temporal)
1	(270 /169)	[240 / 360]	1/10	1	NAPL	Point	pulse
2	(242 /170)	[1200 / 1800]	1/15	1	Solute	Point	pulse
3	(243 /168)	[4000 / 6000]	1/25	2	Solute	Point	pulse
4	(273 /157)	[2000 / 3000]	1/15	2	Solute	Point	pulse
5	(264 : 270 /186)	[960 / 1440]	1/10	2	Solute	Line	pulse
6	(290 /212)	[10 / 40]	1	3	Solute	Point	pulse
7	(264 : 266 /200 : 202)	[25 / 50]	2	3	Solute	Area	pulse
8	(259 /156)	[60 / 140]	1/8	3	Solute	Point	pulse
9	(258 /193)	[2x / 3x]	1/100	1	Solute	Point	continuous

Table 10.1.: Hazard database with nine synthetic risk sources (CT = Contaminant Type).

CT	Examples	Critical level	Saturation conc.[1]	Retardation	Degradation
		$c_{crit}[g/l]$	$c_{sat}[mg/l]$	$R_d[-]$	$\xi_d[1/d]$
1	Chlorinated solvents	$5.0e^{-6}$	660	2.4	0
2	BTEX compounds	$5.0e^{-6}$	573	1.0	$3.7e^{-8}$
3	Microorganism, Metals	$5.0e^{-6}$	35	1.0	0

Table 10.2.: List of contaminant-specific properties.

Please note, that the critical concentration levels are set to $c_{crit} = 5e^{-6}\, g/l$ for the three contaminant types. These criteria are based on US EPA (2009) for chlorinated solvents and BTEX compounds and could be easily and independently of each other lowered to $c_{crit} = 1e^{-6}\, g/l$ or any other reasonable and situation-specific value. In general, retardation rates are site specific, depending on the organic content of the aquifer material. Here, I assume an artificial but still realistic retardation rate. The chosen degradation factor is small, using higher values is possible at any time. The water saturation concentration is based on the Integrated Risk Information System (IRIS) (accessible via US EPA, 2012). Although some of the data rely on regulatory standards, the test case is purely synthetic and designed with the sole purpose as an illustration of the risk quantification and management concept.

10.2. Risk Quantification - STORM Applied to Burgberg

The data from Section 10.1 are used to extend the VIP application study from Chapter 9 to stakeholder-objective risk values. This procedure includes the generation of randomized release histories from the hazards according to Section 7.3, the dissolution process of contaminants to the groundwater through the unsaturated zone (see Section 7.4), the transport and fate of contaminants to the well, including cumulative impact assessment (see Section 7.5), and using the post-processed stakeholder-objective risk measures (see module 8 in the same Section and Section 7.1.2) in risk management (see Section 7.6). The obtained results are his-

tograms of failure intensities, which can be analyzed and discussed in various ways. Here, it quantifies the mean annual risk of supply failure due to contaminant exposure above a critical threshold level. With the probabilistic information on temporal events, the annuality of well exposure time is available for each contaminant type and across contaminants.

10.2.1. Impact Maps

Fig. 10.1 shows intrinsic information on well exposure time and the minimum contaminant mass that already leads to a detectable impact at the well. The intrinsic impact neglects retardation and degradation and is valid for inert contaminants such as metals or microorganisms. Again, contaminant-specific maps that account for retardation and degradation are available (see Section 9.2.5). The maps display the distribution of generic risk sources (CT I =black, CT II =medium gray, CT III =light gray) across the Burgberg catchment, distinguishing point, line and areal source types. The pumping wells are marked by white dots.
Fig. 10.1 (top) shows the ensemble-averaged minimum impact, measured in well exposure time, due to a contamination at this specific location. In this context, minimum impact is defined as the severity of impact that has to be at least expected at the well, averaged over all realizations. The ensemble-averaged results coincide with the results obtained by a deterministic risk assessment (see discussion in Section 8.3). These deterministic (impact-based) models provide maps as shown in Fig. 10.1 and do not provide information on any cumulative probability distribution, such as knowing the $90\,\%$ confidence interval on impact prediction (e.g., US EPA, 2001). The difference between the presented maps in Fig. 10.1 and the VIP maps of the synthetic (Fig. 8.4) and Burgberg application (Fig. 9.2) is that here the *minimum impact* and necessary *minimum mass* are visualized and not the ensemble-averaged impact due to a unit mass. Therefore, these maps follow the classical backward risk assessment approach (e.g., Cushman et al., 2001).
Fig. 10.1 (top) shows that well exposure time increases with increasing distance to the well. This is due to transport-based dispersion, such as spreading of the contaminant plume. The dispersion effect leads to contaminant breakthrough curves at the well with increasing exposure time for more diluted curves, until dilution lowers the peak below the threshold level. The presented pane of the complete model domain only shows increasing exposure times. The minimum exposure time depends on the amount of mass being released to the aquifer (see Fig. 10.1, bottom). The bottom figure shows the minimum mass needed to impact the pumped raw water with a contaminant load that is just enough to cause a non-zero well exposure time, i.e., c_{crit} is exactly met by the peak concentration of the breakthrough curve in the well.
Based on these maps, risk managers can perform prioritization depending on the hazard location relative to the well. For example, hazards with lower mass release potential are acceptable for a specific location and pose no risk to the well (as measured by well exposure time), as dispersion lowers the peak concentration below the critical threshold level c_{crit}. Here, hazards 6 to 9 pose no risk, as their maximum mass m_0 to be released is below the minimum impact mass. In case that hazard 8 ($m_0 = [60\,/\,140]\,kg$) is located only a bit further south and east (see Fig. 10.1) the minimum mass to impact the well reduces to $m_0 < 100\,kg$.

Therefore, hazard impact and risk strongly depends on the location of hazards relative to the receptor. Furthermore, the assumption that an individual hazard does not impact the well is only true, if this risk source or contaminant type respectively can be treated independent of other risk sources.
Now I consider for prioritization the actual mass being released, using the mean mass m_0 (Tab. 10.1) in a deterministic fashion. Hazard 2 delivers the most severe impact with $t_{exp,2} = 1143\,d$, followed by hazard 3 $t_{exp,3} = 592\,d$ and hazard 4 $t_{exp,4} = 484\,d$. All other hazards do not exceed the critical threshold value, if considered individually. Thus, contaminant type $j = I$ contains the most severe hazard, whereas contaminant type II has two hazards that contribute to well failure, and hazards of contaminant type III do not impact the well at all.

Hazard prioritization based on a **deterministic** risk model (impact-based, B1):

$$h_2 > h_3 > h_4 > [h_1, h_5, h_6, h_7, h_8, h_9] = 0$$

10.2.2. Risk Estimation

In this thesis, I proposed the next step from deterministic to uncertainty-based risk assessment models, denoted in this thesis as probabilistic risk model. The hazard ranking, proposed in Section 10.2.1 is pragmatic and simple, based on individual hazard-wise analysis. Module 8 of STORM (see Section 7.5) explains the necessity of aggregation across space, time and frequency, such that hazards cannot be treated independently of each other at the well. If aggregation is neglected, risk sources may be deemed to be safe, although being dangerous, if assessed in combination with others (see hazard 7, Fig. 7.5).
All risk models that do not account for failure probability estimate risk given damage. The expected outcome of these risk assessment models that account for uncertainty aspects except for failure probability are similar to the deterministic „risk“ (=impact) approaches (see Section 8.3, Fig. 10.1). The advantage of risk models in comparison to vulnerability-based approaches is the availability of the probability density distributions of impacts (compare Fig. 10.2).
The current section follows this argumentation and performs an aggregated risk assessment. Compared to vulnerability-based approaches, not only risk given a failure but also the probability distribution of failure events will be considered.

Risk Estimation given Damage

The STORM framework is capable to estimate risk given damage (Eq. 7.10), assessing an intermediate risk level between deterministic risk given damage and risk. In the context of STORM, risk given damage considers the hazard-wise failure frequency as shown in the histogram (see Fig. 10.2). Thus, the risk given damage values are weighted by the hazard-wise

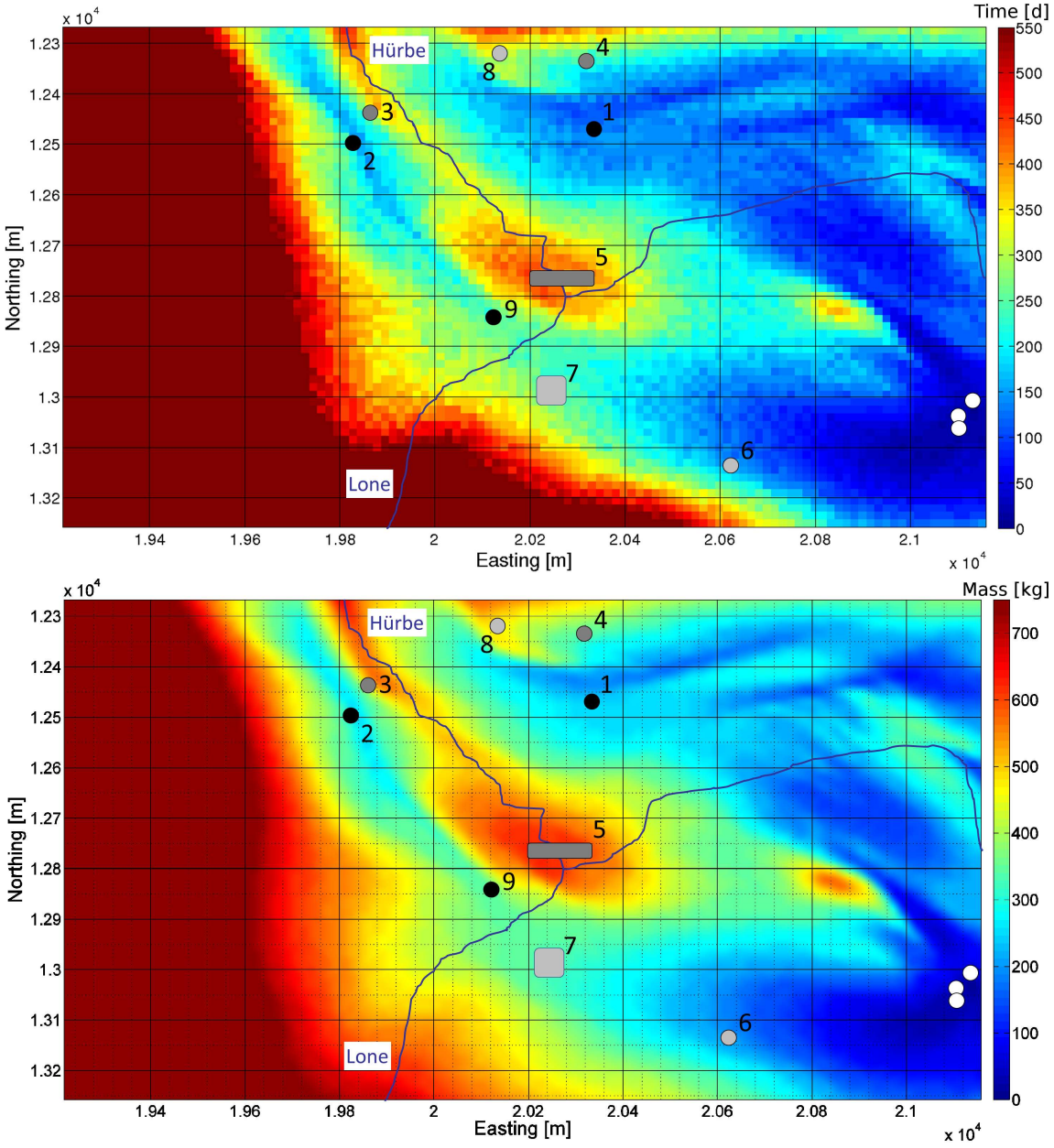


Figure 10.1.: Maps of minimum impact and mass to be expected at the well, showing the spatial distributed minimum well exposure time ($t_{exp}\,[d]$,top) and the required minimum mass ($m_{min}\,[kg]$, bottom) at intrinsic conditions ($R_d = 1, \xi_d = 0$).

failure rates. Still, Eq. (7.10) estimates the expected impact in case of a failure.
The histograms show the impact distribution for the three presented contaminant types (top three plots) over simulation time. The impact is a result from the contaminant-specific total breakthrough curve. Thus the risk given damage is the ensemble-average of all impacts per contaminant type and is also available for each individual hazard or the sum of all hazards (Fig. 10.2, bottom). The cumulative or total expected risk is calculated by Eq. (7.13). Please note, as soon as contaminant-specific aggregation occurs, the failure rate assigned to each hazard influences the prioritization in comparison to risk models that only account for parameter uncertainty. Fig. 10.2 (top three) shows the distribution of well exposure times based on contaminant-specific total concentration breakthrough curves. The bottom histogram shows the stacked plot of the three contaminant-specific distributions of well exposure time normalized by simulation time T_{Tot}.

The mean well down-time given a damage equals $R_{D,I}(WET|F) = 507\,d$ for the contaminant type I. The total breakthrough curve of contaminant type I consists of hazards 1, 2 and 9 and accounts for retardation processes. The expected well down-time for contaminant type II equals $R_{D,II}(WET|F) = 497\,d$ and $R_{D,III}(WET|F) = 195\,d$ for type III. This indicates that the contaminant type I poses the largest risk to safe water supply given a damage. The overall expected well down-time, taking all contaminants into account equals $R_{D,cum} = 393\,d$. Thus, summation of risk measures across contaminant type is critical as it assumes the occurrence of well exposure time per contaminant type in series (non-overlapping well exposure durations).

Ranking of contaminant types based on a **probabilistic** risk model (see Eq. 7.11):

$$R_{D,I} > R_{D,II} > R_{D,III}$$

Mean Annual Risk

The presented risk quantification approach extends the risk analysis by a frequency component of hazard failure. This allow providing information on how often and to which extent the receptor is exposed to contamination. According to Eq. (7.11) it is possible to estimate the mean annual well exposure time. Stakeholders are able to judge, if this expected annual exposure time is acceptable for them or not. This risk interpretation goes beyond the concept of risk given damage (e.g., VIP concept) as discussed in Chapter 7.

The expected annual risk (here: well exposure time) equals $R_{a,I} = 37\,d$, $R_{a,II} = 53\,d$ and $R_{a,III} = 19\,d$ for the three contaminant types. Thus, contaminant type II has a larger annual impact than contaminant I, which is contrary to the observation made before. The reason is due to the failure probability and the probability distribution of the mass. Hazards of contaminant type II fail less frequent, but with a higher mass load. In order to avoid a high impact at the well, which is on average $R_{D,II}(WET|F) = 497\,d$, catchment managers are asked to reduce the contaminant load of contaminant type II, such that hazards that contribute to the total breakthrough curve deserve a closer look. The total expected annual well down-time equals $R_{a,Tot} = 109\,d$, which is the sum of the three contaminant types, only

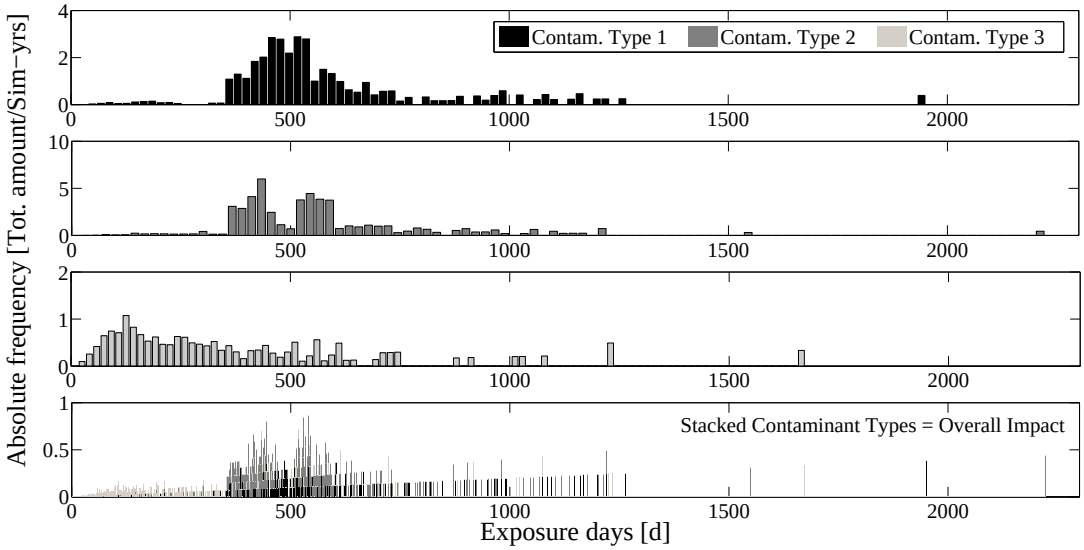


Figure 10.2.: Histogram of well exposure time (t_{exp}) by individual contaminant types I to III and in total across type (bottom).

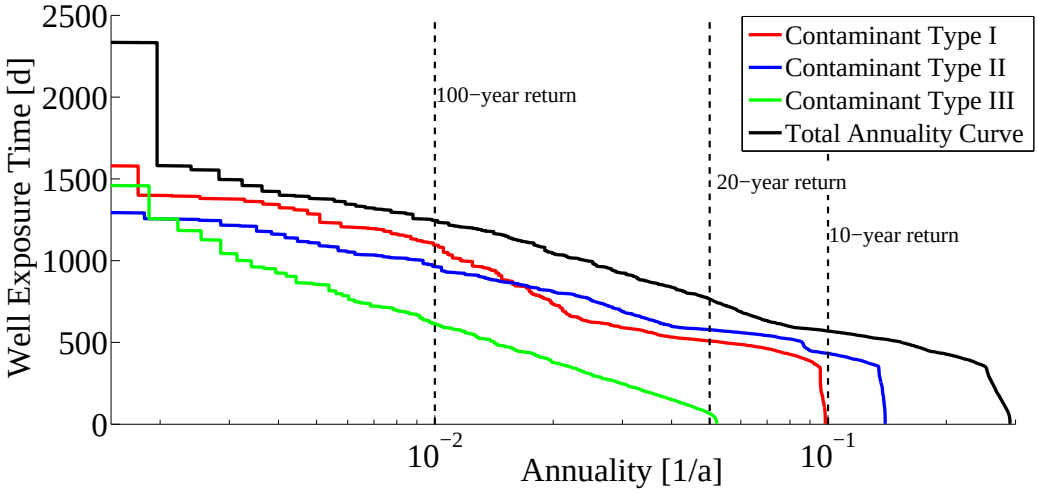


Figure 10.3.: Annuality J_a of well exposure times.

in the case, when well down-time due to any contaminant type occurs in a consecutive, non-overlapping, order. The estimation of the actual total expected well down-time needs further investigation and developing time.

Prioritization of contaminant types based on a **probabilistic** risk model (Annual mean risk):

$$R_{a,II} > R_{a,I} > R_{a,III}$$

Annuality-based Risk

Another way to illustrate the annual risk is by plotting the cumulative distribution function according to Eq. (7.12). Fig. 10.3 shows the total annuality curve in relation to the three contaminant-specific annuality curves. When annuality curves cross each other, the ranking order of hazards changes. Thus, prioritization of hazards depends on the return period. For this example, I compare the $100 - year$ and $20 - year$ return period for the contaminant-specific annuality curves. The well exposure time in $100\,years$ equals for the three contaminant types $R_{J,I}(T_a = 100) = 1096\,d$, $R_{J,II}(T_a = 100) = 962\,d$ and $R_{J,III}(T_a = 100) = 611\,d$. The well down-time for $T_a = 20\,years$ equals $R_{J,I}(T_a = 20) = 508\,d$, $R_{J,II}(T_a = 20) = 576\,d$, $R_{J,III}(T_a = 20) = 70\,d$. Contaminant type II is more responsible for smaller ($WET < 860\,days$) and more frequently occurring events than contaminant type I. As opposed to reducing impact by reducing stored contaminant mass of type II, risk management should focus on the contaminant type I to reduce risk, e.g., by improving the reliability of hazards that handle contaminant type I. It is very seldom that a well exposure time due to contaminant type III has to be expected. Nevertheless, within 100-years, the stakeholder has to expect a well down-time of $t_{exp,III} = 576\,d$. The steep incline at high frequency for each contaminant type is due to the sudden incline of well exposure time, if an event occurs (see Fig. 10.1, top). The total well exposure time for $T_a = 100\,years$ equals $R_{J,Tot}(T_a = 100) = 1245\,d$ and $R_{J,Tot}(T_a = 20) = 764\,d$. This clearly shows that the total risk is not the sum of individual hazards or contaminant types, respectively (Fig. 10.3). A detailed description on hazard prioritization based on two return periods is given in Section 10.3.

Prioritization of chemicals based on a 100-year **annuality**-based risk:

$$R_{J,I}(T_a = 100) > R_{J,II}(T_a = 100) > R_{J,III}(T_a = 100)$$

Prioritization of chemicals based on a 20-year **annuality**-based risk:

$$R_{J,II}(T_a = 20) > R_{J,I}(T_a = 20) > R_{J,III}(T_a = 20)$$

10.3. Risk Management - STORM Applied to Burgberg

Risk standards, such as DVGW (2009), challenge water stakeholders to prioritize risk sources in order to install effective risk mitigation measures that target the hazards of highest priority. This section addresses prioritization of individual hazards based on annuality (see Section 10.3.1). The current section also provides a comparison of three risk mitigation scenarios (see Section 10.3.2) and the influence of choosing a stakeholder-objective risk measure on the hazard ranking (Section 10.3.3). Section 10.3.4 goes a step further as three different hazard ranking lists are compared that are obtained from different risk estimation approaches, such as qualitative risk estimation with a 5×5 risk matrix.

10.3.1. Impact of Different Risk Attitudes on Prioritization

The prioritization of hazards depends on the chosen risk estimate (Section 10.2.2). In this section, I will discuss the impact task of risk attitude on hazard prioritization. Thus, risk managers are able to target their risk mitigation task more precisely, prioritizing hazards that account for frequent or less frequent damage. The prioritization is based on Eq. (7.14).

Fig. 10.4 shows the prioritization of hazards according to a short return period with $T_a = 10\,years$ and according to events that are occurring only once per $T_a = 100$ years. For both return periods, hazard 3 is most severe as it lowers the expected impact for the given return period by $\Delta R_{J,3}(T_a = 100) = 26\,\%$ and $\Delta R_{J,3}(T_a = 10) = 23\,\%$, if removed from the system. Hazard 2, which denotes the highest annual exposure risk, when hazards are analyzed independently of each other without accumulation, is ranked second for the $10 - year$ return period with $\Delta R_{J,2}(T_a = 10) = 17\,\%$. Thus, hazard 2 is relevant in risk management for more frequent events. The $100 - year$ return impact does not significantly change, when omitting hazard 2, with $\Delta R_{J,2}(T_a = 100) < 1\,\%$. For $T_a = 100$, hazards 3, 4 and 5 are more important than hazard 2. Overall, hazards of contaminant type II (hazards 3 to 5) and hazard 2 (contaminant type I) influence the more frequent events (see Fig. 10.4). Contaminant type II contributes most to seldom event periods. Likewise, contaminant type III contains relevant hazards that influence events with longer return period. Thus, although hazard 6 to 8 do not exceed the threshold individually nor in perfect concurrence (see discussion in Section 10.3.4, B2), they still pose a risk of well down-time for seldom events ($T_a > 25\,years$). This is due to the fact, that these hazards have a low individual impact but exceed the critical threshold value, when multiple events occur due to high failure rates that previous contamination has not ebbed away in the production well. The result of this effect is shown in Fig. 10.3.

Fig.10.4 shows clearly that hazard ranking changes, depending on the chosen return period (here based on $T_a = 10\,years$ and $T_a = 100\,years$). Fig. 10.4 provides an overview on contaminant types and hazards that are responsible for certain recurrence intervals. Thus, thanks to the additional frequency information provided by STORM, risk managers are able to identify hazards that are relevant for higher or for more frequent impacts.

In addition to this, prioritization depends on the chosen risk estimate. Here, the chosen risk estimates were the annual mean impact R_a (Section 10.3.4) and impact based on a given return period R_J (this section). Furthermore, the presented results are catchment-specific and

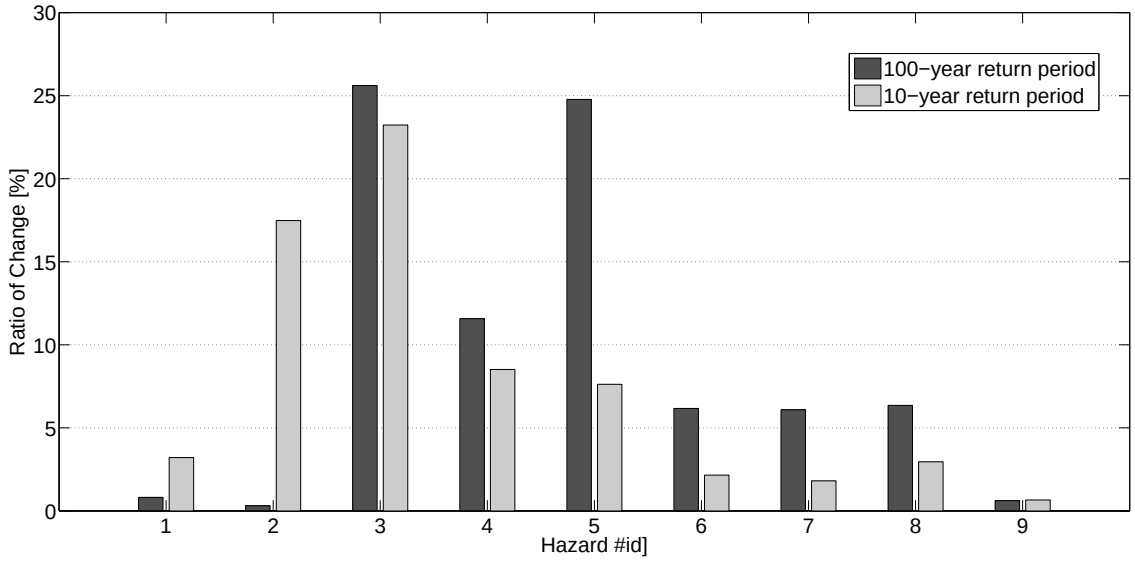


Figure 10.4.: Prioritization of individual hazards based on the $10-$ and $100-year$ return period of WET. The y-axis shows the change of risk according to Eq. (7.14)

depend on the location of the risk source (see Fig. 10.1).

Hazard prioritization based on **risk attitude** (D1, D2):

$$T_a = 10: \; h_3 > h_2 > h_4 > h_5 > h_1 > h_8 > h_6 > h_7 > h_9$$
$$T_a = 100: h_3 > h_5 > h_4 > h_8 > h_6 > h_7 > h_1 > h_9 > h_2$$

10.3.2. Choice of Risk Mitigation Strategy

Risk mitigation has several dimensions. First, it is possible to reduce the failure rate for the most severe hazard or a combination of hazards (HP). Second, stakeholders can introduce risk mitigation measures that reduce the contaminant impact at the well (HM). Third, both aspects could be combined (HB).
The effects of the first three mentioned risk mitigation options are illustratively shown for the top hazard 2 (ranked according to expected annual exposure time and impact assessment) and for top hazard 3 (ranked according to the $10-$ and $100-year$ return period). In the following, it is assumed that risk mitigation options are available that can be used to halve the impact, failure probability or both simultaneously. Fig. 10.5 demonstrates the results of this simplistic scenario-based illustration.
As expected, the risk option to reduce both aspects simultaneously delivers the highest risk reduction for hazard 2 and hazard 3. The most effective reduction in annual expected well exposure time is halving mass and probability for hazard 2 ($\Delta R_{a,Tot}(WET|HB) = 35\,\%$).

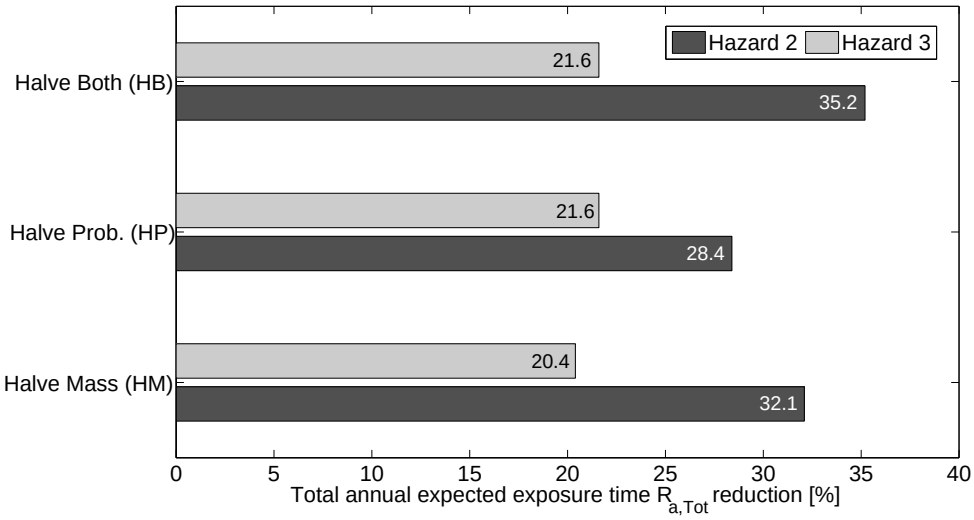


Figure 10.5.: Change of total annual expected risk $R_{a,Tot}(WET)$ [%] by applying three risk mitigation options (halved mass, halved failure probability, both halved) on hazard 2 and 3.

Installing risk mitigation options that tackle the contaminant load and thus the impact at the well would still lead in this case to a reduction of the expected annual risk by $\Delta R_{a,Tot}(WET|HM) = 32.1\,\%$. Both options clearly outweigh other options like better maintenance in order to decrease the failure rate, which would only yield $\Delta R_{a,Tot}(WET|HP) = 28.4\,\%$. Nevertheless, the most efficient way to mitigate risk is having mass. The combination of both options does not significantly reduce risk levels.
It is not always the case that mass-related risk mitigation options are superior to option that reduce probability (e.g., failure rate, uncertainty). For hazard 3, halving probability yields $\Delta R_{a,Tot}(WET|HP) = 21.6\,\%$ and is slightly more effective than reducing the contaminant load to the well with $\Delta R_{a,Tot}(WET|HM) = 20.4\,\%$. In fact, halving probability reduces the total annual risk likewise, when combining both mitigation options for hazard 3. Thus, mitigation options related to probability should be suggested for hazard 3.
The study shows (Fig. 10.5) that risk mitigation is highly hazard-specific and strongly depends on the local situation (e.g., catchment, hazards). The question to find the best risk reduction strategy could also be answered through optimization techniques. Such risk-optimized decision analysis within STORM may deserve more research in future studies.

10.3.3. Impact of Different Risk Measures on Prioritization

The determination of the risk objective is the first step in risk analysis. Unfortunately, stakeholder views may change due to several reasons, such that flexible tools are in demand to evaluate any possible pre-defined risk objective. The consequences to base decisions on unqualified or unreflected risk analysis and prioritization schemes may be unbearable. The impact of choosing different risk measures will be demonstrated now. Hazard prioritization is

based on well exposure time (WET) and maximum concentration ratio (MCR). „Maximum concentration ratio“ is relevant in human-health or ecological risk assessment, whereas the „well exposure time“ is appropriate to assess the technical or economical failure of a system. Fig. 10.6 shows the impact distribution of the maximum concentration ratio over the simulation time period T. The respective results for well exposure time have been presented in Fig. 10.2.

The overall risk given a damage is at $R_{D,Tot}(MCR) = 2.04$. Thus, peak concentration levels are on average twice the critical threshold value c_{crit}. The most severe contaminant type is II with values being four times c_{crit}, $R_{D,II}(MCR) = 3.78$. Within hazard ranking hazard 3 is most severe $R_{D,3}(MCR) = 9.47$, followed by hazard 4 with $R_{D,4}(MCR) = 3.46$ (see Fig. 10.6). Tab. 10.1 shows that these two hazards are the ones with the highest mass release potential.

The total annual risk equals $R_{a,Tot}(MCR) = 0.71$. Peak concentrations are on average below the critical threshold value in the cause of a year. In fact, the most relevant hazard that dominates peak concentrations and total risk is hazard 7 $R_{a,7}(MCR) = 0.98$, followed by hazard 3 $R_{a,3}(MCR) = 0.38$. This is due to the high failure rates (see Tab. 10.1). Furthermore, all peak concentrations, also those below C_{crit}, are considered within this analysis. Thus, concentration levels of hazard 7 are close to the drinking water limit c_{crit} most of the time.

Considering well exposure time, the hazard prioritization changes (see Fig. 10.2). The ranking order changes to hazard 2, 3 and 4, respectively ($R_D(t_{exp})$). Furthermore, well exposure time classifies contaminant type I as the most severe contaminant type according to Eq. (7.10) instead of contaminant type II as indicated by the maximum concentration ratio. In addition to this, hazards 7 to 9 are ranked last for annual mean well exposure time, whereas hazard 7 is the top hazard for $R_{a,7}(MCR)$. Hazard 7 is a line hazard and fails twice a year at low mass release rates. Although not exceeding c_{crit} it still poses a continuous source of risk that possibly needs treatment.

With this simple illustration, it is most obvious that risk mitigation options differ, depending on the chosen risk-objective. Of course, there are similarities, such as identifying hazards 3 and 4 as a common major risk source. However, there are also risk sources that are not identified as top risks in one or the other case (here, hazard 2 or hazard 7). For maximum concentration-based assessment, hazards 3, 4 and 7 are most relevant, whereas for well exposure time hazards 2, 3 and 5 attract the attention of the water manager. Therefore, decisions and risk mitigation options will differ. Also, the target contaminant group for the specific example catchment changes from contaminant type I (WET) to II (MCR).

Prioritization of hazards based on **stakeholder-objective** risk measures (MCR):

$$R_D(MCR) : h_3 > h_4 > h_2 > h_5 > h_1 > h_7 > h_6 > h_8 > h_9$$

$$R_a(MCR) : h_7 > h_3 > h_4 > h_6 > h_2 > h_5 > h_1 > h_8 > h_9$$

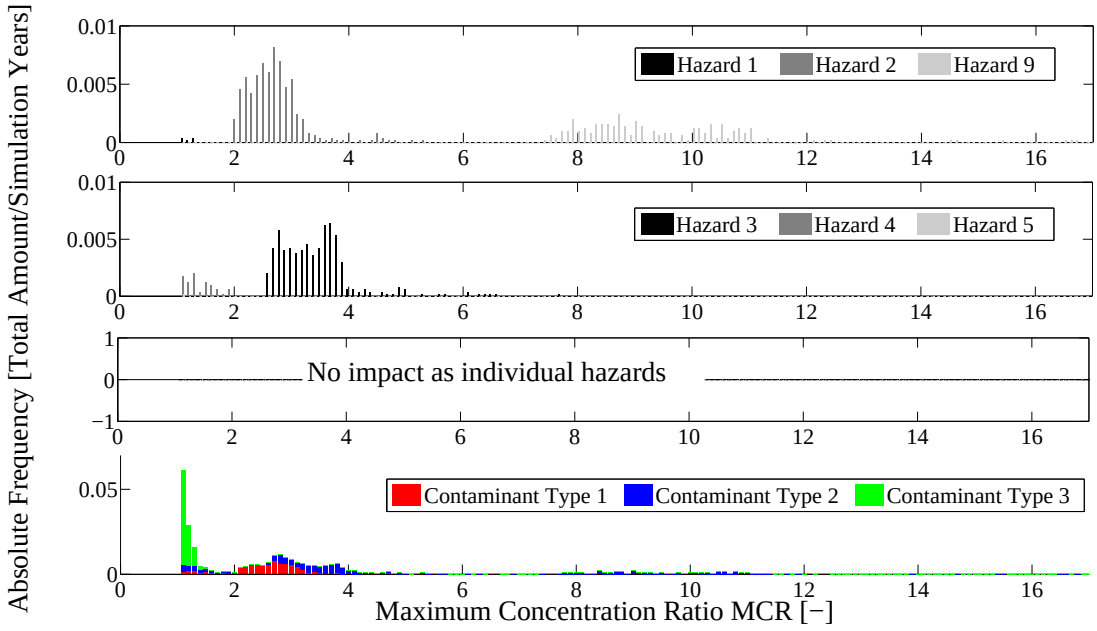


Figure 10.6.: Histogram of maximum concentration ratio (MCR) by individual contaminant types I to III and in total across type (bottom).

10.3.4. Impact of Different Risk Estimation Schemes on Prioritization

In literature, many risk analysis methods exist. In Chapter 3, the differences between qualitative and quantitative risk estimation and the importance of aggregation has been discussed. In the previous section, prioritization has been performed with one risk quantification approach, here, the quantitative STORM approach. Finally, I demonstrated the effect of using two different stakeholder-objective risk measures during prioritization. Now, I will compare three hazard rankings that are based on three different risk estimation approaches. The three approaches are qualitative risk assessment, quantitative impact assessment (e.g., Tait et al., 2004; Troldborg et al., 2008) and physically-based probabilistic risk assessment.

A: Prioritization after Qualitative Risk Estimation

Prioritization of qualitative risk estimates (here, denoted by letter **A** for comparison to other risk estimation schemes) is based on a 5×5 risk matrix (see Fig. 10.7). Aggregation of contaminants across space and time is not straightforward as discussed in Section 3.5.2. In general, each hazard is treated independent of each other. Due to the ordinal scale, it is possible to compare hazards of different contaminant types with each other.
To compare the hazards of Tab. 10.1, the hazards are classified according to the ordinal scale of the transfer function provided in Tab. 10.3. The failure rate λ of each hazard informs the risk managers about the probability of hazard failure. The mass being stored at each hazard location indicates the damage potential for the respective hazard. The transport aspect from the hazard location to the well is neglected, i.e., the impact is measured at the risk source

instead of at the receptor. An alternative qualitative approach that considers transport is the semi-quantitative method proposed by Haakh et al. (2013).

Doessing Overheu et al. (2013) proposed to use a log-scale for the mass discharge-based categorization. For illustration, I choose the categorization bounds for mass and probability in a subjective manner, as proposed by AS/NZS 4360:2004 (2004). Therefore, the rankings need adaption by the water stakeholder to their specific local problem. The categorization of the synthetic hazard test case follows the risk matrix as shown in Fig. 10.7.

The results are visualized by the hazard numbers in Fig.10.7. Apparently, Hazard 5 falls

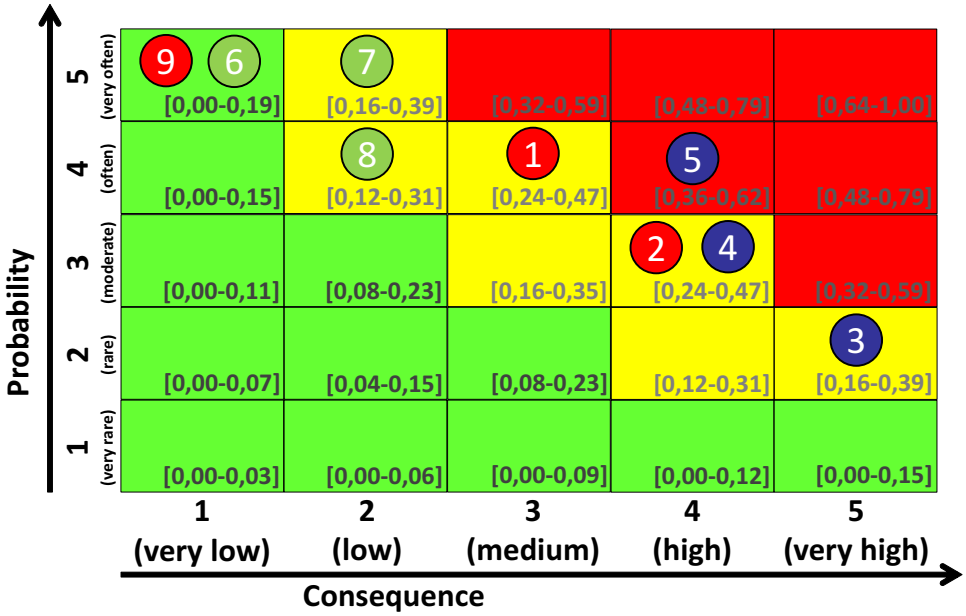


Figure 10.7.: 5×5 risk matrix, showing the categorization of hazards from the synthetic hazard test case. The nine hazards are colored according to the contaminant type (CT I=red, CT II=blue, CT III=green).

into the zone that is not acceptable and only hazards 9 and 6 are acceptable. In qualitative risk assessment, stakeholders often do the mistake to still interpret risk as probability times damage, although such a clear product rule can only apply quantitatively on cardinal scales as used in quantitative risk analysis approaches. Stakeholders tacitly use cardinal operators

Rank	Ranking bounds	Mass bounds	Probability bounds	Evaluation
1	$0.0 < 0.2$	$< 50\,kg$	$< 1 failure\, in\, 50\, years$	very small/rare
2	$0.2 < 0.4$	$< 200\,kg$	$\geq 1\, failure\, in\, 50\, years$	small/rare
3	$0.4 < 0.6$	$< 1.000\,kg$	$\geq 1\, failure\, in\, 20\, years$	medium
4	$0.6 < 0.8$	$< 5.000\,kg$	$\geq 1\, failure\, in\, 10\, years$	large/often
5	$0.9 \leq 1.0$	$\geq 5.000\,kg$	$\geq 1\, failure\, in\, 5\, years$	very large/often

Table 10.3.: Transfer function to assess hazard ranks for probability and damage.

with qualitative (ordinal) values, taking the computed outcome as a quantitative result. For example, the risk matrix is symmetric and delivers values from low „0“ to high „1“ risk values, after being normalized by the Neumann-Morgenstern utility function (see square brackets in Fig. 10.7). This transfer function normalizes the level of severity or probability by the maximum value of the categorization bound per axis. Using these numbers, one would still conclude that hazard 5 is the most severe risk source, followed by hazard 1, 2 and 4. For comparison with the following two other risk analysis schemes, the resulting ranking is also shown in Fig. 10.8.
Cox (2008) proposes to use the qualitative results only to distinguish between high and low risk, such that hazard 5 is ranked as unacceptable, whereas hazards 6 and 9 are acceptable. The other hazards would need further investigation. These hazards are potentially as severe as hazard 5 (see discussion on negative slopes, Section 3.5).
Instead of looking at individual hazards, one can also look at hazard groups, defined by contaminant types (see color coding of the hazard numbers in Fig. 10.7). Fig. 10.7 shows contamination type II (blue) to be most severe, and that hazards 1 and 2 (CT I, red) and hazards 7 and 8 (CT III, green) would need further investigation. In the previous section, hazard 7 is indeed a relevant risk source for assessing the maximum concentration ratio. As highlighted in Section 3.5, risk matrices are a valuable tool if used appropriate. Qualitative approaches can be adequately used for screening in order to find relevant hazards.

Prioritization of hazards based on a **qualitative** risk analysis (**A**):

$$\text{Common practice: } h_5 > h_1 = h_2 = h_4 > h_3 = h_7 > h_8 > h_6 = h_9$$

$$\text{Correct way: } h_5 > [h_9, h_6]; [h_1, h_2, h_3, h_4, h_7, h_8] = \text{ more information}$$

B: Impact-based Prioritization

Impact-based prioritization is based on deterministic risk (vulnerability) modeling results, where epistemic uncertainty is neglected or cannot be quantified, respectively. The proposed quantitative impact assessment approach is mass-discharge-based (see Fig. 10.1). Considering each hazard as individual risk sources (**B1**), the mean impact of hazard 2 ($WET = 1143\,d$) is most severe, followed by hazard 3 and hazard 4. This prioritization neglects aggregation of mass discharges and frequency aspects.
According to Tait et al. (2004) and Troldborg et al. (2008), cumulative effects across space and time have to be considered even in this type of vulnerability assessment. Here, the arrival and the concentration profile of contaminants of the same type at the well are assumed to occur perfectly simultaneous. This is similar to the aggregation of breakthrough curves in Fig. 3.9 (black line). The maximum mass release potential to the aquifer is assumed. Thus, the aggregation-based impact approach (**B2**) denotes a *worst case situation*. The well exposure time for contaminant type I equals $1529\,d$ and $722\,d$ for contaminant type II. The third contaminant type does not exceed the critical threshold value (see Fig. 10.6, CT III). The ranking follows Eq. (7.14) and is presented in Fig. 10.8 for comparison.

Please note, the total risk is influenced by an individual hazard based on the proportion of that hazard on the contaminant-specific breakthrough curve. Therefore, the ranking is only valid under the assumption that the ratio $\Delta R_{r,h}$ behaves linear across contaminant types, such that the contaminant-specific impact occurs consecutively without any overlap. The most severe hazard is hazard 2, followed by 3 and 4 (see Fig. 10.8).

Prioritization of hazards based on **vulnerability-based** risk analysis (**B2**):

$$\text{Impact-Ratio: } h_2 > h_3 > h_4 > h_1 > h_9 > h_5 > h_6 > h_7 > h_8$$

C: Risk-based Prioritization

A simplified risk ranking strategy is available, assuming impacts of individual hazards at the well to be independent of each other, such that contaminant BTCs do not overlap in time (**C1**). Following this simplification, hazard 2 is ranked most severe ($R_{a,2}(WET) = 32\,d/a$), followed by 4 ($R_{a,4}(WET) = 27\,d/a$) and 3 ($R_{a,3}(WET) = 21\,d/a$). Hazards 6, 7, 8 and 9 pose no risk.

Prioritization of hazards based on **probabilistic** risk analysis (**C1**, independent hazard):

$$R_a : h_2 > h_4 > h_3 > h_5 > h_1 > h6 > [h_7, h_8, h_9] = 0$$

This assumption of independence is strong, and is not true in most of the cases (see Fig. 7.6). Therefore, convolution of mass discharges across space, time and frequency is essential to assess the cumulative impact of contaminants of the same type (**C2**). There are two possible ways to interpret Eq. (7.14) for hazard ranking. First, the hazard impact is related to the total annual risk across all contaminant types $R_{a_{Tot},r}$. Second, it is related to the cumulative contaminant-specific total breakthrough curve $R_{a_{CT},r}$.

Prioritization of hazards based on **probabilistic** risk analysis (**C2**, aggregation-based):

$$R_{a_{Tot},r} : h_7 > h_2 > h_6 > h_8 > h_3 > h_4 > h_1 > h_9 > h_5$$
$$R_{a_{CT},r} : h_2 = h_7 > h_3 > h_4 > h_1 = h_6 > h_5 > h_8 = h_9$$

The results show that hazard 7 is one of the most severe ones. A similar observation has been made before for considering MCR. The reasons are given in Section 10.3.3. Hazard 2 is, as expected, a top risk source in both cases. Furthermore, both prioritization approaches within C2 lead to similar results, regarding the least important hazards, such as hazard 9

and 5. Contrary, hazard 6 is ranked third ($R_{a_{Tot},r}$), although being the least important hazard in the contaminant-specific ranking ($R_{a_{CT},r}$). The same is true, when compared to other rankings (e.g., vulnerability-based (B2), risk-based (C1)). Also hazard 8, which is ranked last for the contaminant-specific ratio ($R_{a_{CT},r}$), is ranked fourth, when assessing the relative change of risk in comparison to the total risk ($R_{a_{Tot},r}$).
Overall, the contaminant-specific approach ($R_{a_{CT},r}$) seems to be more valid than $R_{a_{Tot},r}$. This is due to the fact that for $R_{a_{Tot},r}$ the damage (here: well exposure time) is assumed to occur in a perfect concurrence, such that well exposure durations due not overlap across contaminant types. Contrary, the contaminant-specific approach ($R_{a_{CT},r}$) is limited in ranking hazards across contaminant types, here, showing hazard 2, 3 and 7 being each the top risk source per contaminant type. But how is it possible to compare the three top events per contaminant-type against each other? Therefore, hazard prioritization would need further investigation. A possible solution could be to consider the canceling out effect of damages as shown in Fig. 3.9. Due to this unresolved problem, ranking by **C1** is presented in Fig. 10.8.
The last option to prioritize hazards is based on annuality, using either a high return period (**D1** $= T_a = 10\,years$) or a low return period (**D2** $= T_a = 100\,years$). The ranking has been illustrated in Fig. 10.4. Here, hazard 2, 3, 4 and 5 are deemed to be the top three risk sources for both two return periods.

Comparing the Prioritization Results

The hazard ranking of three risk approaches (qualitative, impact, risk) are compared with each other in Fig. 10.8. Here, only showing one risk estimation approach each. None of the proposed methods coincide in the top three risk sources. Hazards 2 and 4 are the only hazards to be identified, that are listed within the top three risk sources of the three presented approaches. Hazard 3 is ranked second (B2) and third (C1) within the quantitative risk models, whereas ranked fifth in the qualitative approach. However, hazard 3 could easily be ranked higher, when considering the possible range of impact (Fig. 10.7, values within the squared brackets). Considering the results of C2 (cumulative risk), D1 and D2 (risk attitude-based) hazard 3 is ranked within the high range. Furthermore, the three risk models coincide within the group of less relevant hazards (hazards 6, 7, 8 and 9). Please note, that hazard 7 is ranked as a top hazard, when considering cumulative effects (**C2**).
Although the differences between the prioritization results seem to be somewhat arbitrary, it has to be noticed that mostly the ordering changes, but the relevant hazards are always within the top group. Due to these reasons, I would recommend to relax the assumption of consecutive ranking and take the top two to three or outstanding risk sources to tackle risk mitigation options until more accurate prioritization is considered in future research. This is in line with the proposed methodology by Cox (2008) for qualitative risk matrices.

10.4. STORM Summary

The goal of this chapter was to demonstrate the applicability of STORM, showing aspects of cumulative risk assessment and the impact of risk factors, such as stakeholder objectives,

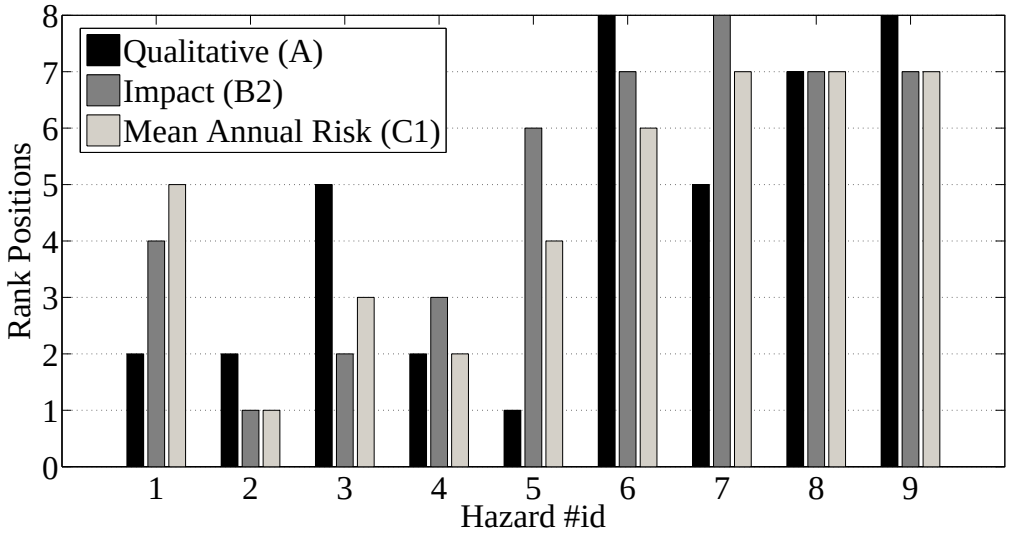


Figure 10.8.: Comparing three hazard rankings, including qualitative risk (A), impact (B2) and quantitative (C1) risk assessment using STORM.

on the prioritization. The STORM approach has been applied to a synthetic hazard test case with nine hazards of three different contaminant types. The hazards, including spatial and temporal variable release patterns, have been assumed to reside within the Burgberg catchment (see Chapter 9).

- The study shows that aggregation across space, time and frequency is necessary to assess cumulative risk. Summation of mass-discharge levels at the well is the only meaningful aggregation to perform risk analysis. Aggregation across type is not available by summation, but by statistical analysis.
- The framework shows flexibility in adapting to the multi-objective interests of stakeholders, by introducing well vulnerability criteria as intermediate risk levels.
- Two risk estimates are presented, the mean annual exposure time and the annuality curve of exposure, returning the expected well exposure for a given return period.
- Corresponding differences in risk prioritization have been discussed.
- Also, the impact using qualitative versus quantitative risk analysis methods has been investigated.
- The decision analysis framework shows that hazard prioritization and derived risk mitigation actions is dependent on the risk attitude (choice of return period), the stakeholder-objective risk measure (e.g., MCR, WET), the risk estimation equation (e.g., annual expected damage, risk ratio) and the choice of risk analysis method (e.g., qualitative, quantitative).

11. Conclusions and Outlook

The goal of this chapter is to summarize the main findings of this thesis and provide an overview of possible future research. Section 11.1 describes the overall findings in light of the pre-defined thesis as defined in Section 1.5. Section 11.2 presents a summary and conclusion regarding the VIP concept and Section 11.3 regarding STORM. The chapter concludes with an outlook on possible future research to further improve the field of risk quantification and management for drinking water supply systems (Section 11.4).

11.1. Overall Summary and Conclusions

At the beginning of this thesis, I provided a comprehensive overview, in-depth introduction and critical review of risk assessment and management. This summary enables risk managers and stakeholders to judge the suitability of available risk models to their actual problem setting. The advantages and limitations of qualitative and quantitative risk models have been intensively discussed, showing the supremacy of quantitative methods over qualitative ones. Also, requirements for quantitative risk models, regarding dispersion, risk aggregation and uncertainty have been postulated. Furthermore, a common risk terminology and risk management concept has been introduced to this thesis that spans across the three risk communities of qualitative, deterministic and probabilistic risk assessment.
Due to the findings of Chapter 3, it has been possible to state the *overarching goal of this thesis*: **to provide risk quantification and management concepts, guaranteeing safe drinking water supply from groundwater.** These concepts are designed to support risk-informed and robust decision making in the light of uncertainty quantification and state-of-the art transport-based models.

As a thesis I defined:

> *Risk assessment and management for drinking water supply systems has to be probabilistic, quantitative and physically-based. Only physically-based models enable the aggregation of mass discharges at the well, which is the only valid way to account for cumulative effects across hazards in space, temporal arrival and frequency. Aggregation across different contamination types is only available by statistical analysis on the level of risk estimation. The flow and transport calculations should avoid upscaled models that already include averaging procedures, because they would hinder a clean statistical analysis. Also, they should be set into a Bayesian and/or geostatistical probabilistic modeling framework to quantify and reduce epistemic uncertainty. That being said on a conceptual level, adopting the flexibility and modularity of graph-based risk methods will*

be a valuable step on the way from theory to practice. For practical application, it is also relevant to develop a unique concept that offers all desirable risk measures for multiple stakeholder views out of one method. Thus, stakeholder decisions can be supported by stakeholder-objective risk measures and by intuitive decisions analysis frameworks.

Conceptual advancements: Based on this thesis, I have successfully developed and presented two physically-based probabilistic risk quantification concepts. Both risk models fulfill the source-pathway-receptor concept as proposed by US EPA (1989), estimating risk at the receptor level (here: the drinking water well) and not at some intermediate level, such as at some downstream point from the source (e.g., at the groundwater table). The necessity to avoid upscaled flow and transport models, e.g., parameterization of macrodispersion by Fickian concepts, has been highlighted. I have shown that quantitative risk assessment is superior to qualitative (ordinal-scale based) risk analysis and that probabilistic and quantitative information on concentration breakthrough curves are a valuable tool for stakeholder-objective, robust, and risk-informed decision analysis. The adopted modular concept proved to be very useful. Thus, the thesis provided above is confirmed.

Model reduction: Both risk concepts resolve parameter uncertainty, such as hydraulic conductivity, via Monte Carlo simulations. As a counter measure to the high computational costs, triggered by the need for Monte Carlo simulations for probabilistic risk assessment, the flow field is reversed and moment-based transport calculations in combination with analytical and numerical breakthrough curve reconstruction techniques are adopted. The corresponding gain in computational efficiency and the resulting accessibility of probabilistic information at only a slightly reduced accuracy are much more valuable to my opinion.

Probabilistic vulnerability criteria: The contamination breakthrough curves at the well are aggregated based on the contamination type. From these cumulative breakthrough curves, well vulnerability criteria for each contaminant type are deduced, leading in the next step to exceedance-based cumulative distribution functions for each location across the domain. This probabilistic information on non-compliance is used to support stakeholder-objective decision analysis.

The issue of dispersion: Any transport-based risk measure is sensitive to the conceptual difference between uncertainty in plume location and actual dilution. For that reason, the flow and transport problem should be solved via finely resolved Monte Carlo simulations, where aquifer heterogeneity is resolved on and above the grid-scale in each realization. Thanks to this approach, the probability of peak dilution and the uncertainty in position and extent of protection zones can be assessed separately.

Catchment-wide risk aggregation: Contaminant loads from different release events at different times or location within the catchment are aggregated at the well, depending on the contaminant arrival of each hazard. This leads to cumulative concentration breakthrough curves per contaminant type. The aggregation of mass discharges can be done in

a fully quantitative manner by simple addition. Aggregation across contaminant type is only available by statistical analysis for non-linear risk measures. All threshold-based risk measures are non-linear, and risk dramatically changes depending on the chosen critical values. Therefore, simple summation of risk levels is prohibitive in a non-linear risk setting.

Uncertainty reduction and data worth: Within a Bayesian framework, the geostatistical models are calibrated to direct and indirect measurements. Uncertainty reduction is available by stochastic or conventional model calibration. The effect of conditioning on risk-aware delineation has been demonstrated. Uncertainty reduces with increasing amount and quality of data and depends on the sampling type. The trade-off between data quality, quantity and type has been demonstrated. The worth of data for improved well-head delineation has been set into an economical context, showing that higher reliability levels are available at no additional costs to guarantee safe drinking water supply, when uncertainty is priced via a utility function. These data worth analysis studies have been used to support stakeholders to take risk-aware decisions.

Stakeholder-objective risk measures: All presented decision analysis frameworks rely on the vulnerability concept and are flexible to adapt to the needs of stakeholders. The used probabilistic well vulnerability criteria allow stakeholders to address human-health, ecological, operational or economical risk-related questions. Thus, they serve as the basis for any transport-based risk measure. This is important, as different objectives lead to alternative decisions, such as demonstrated for hazard prioritization. Besides hazard ranking, the decision analysis framework is used to support stakeholders in their decision or risk mitigation scenario choice, respectively. Here, I have shown the use of peak arrival time and well exposure time in a risk-based decision analysis context.

Modularity and flexibility: The presented risk concepts are independent of the employed modeling software. Thus, they are flexible in their modeling tools and allow due to the modular conceptualization an easy adaption to current available software situation with the advantage that intrusion to the code is not necessary. Furthermore, the concept is independent of dimensionality, boundary conditions and (hydro)geological conditions. These properties make the approach flexible for any type of drinking water catchments and a wide range of applications.

Two new risk assessment frameworks: The two introduced physically-based probabilistic risk quantification concepts are called VIP and STORM. Although VIP is used for the transport and intermediate vulnerability assessment step of STORM, the VIP concept can be used on its own to quantify and manage catchment risk as shown. As VIP is a backward approach, it does not consider the aggregation of contaminant mass so far. Instead, the focus is on pixel-wise concentration-based information, that can be flexibly adapted to the given risk source situation. STORM uses the backward transport information from VIP in a forward risk context, adding the aggregation over any given hazard inventory towards overall quantitative risk measures.

11.2. Summary and Conclusions for the Proposed VIP Concept

Concept of methods: The VIP concept is based on probabilistic well vulnerability criteria that are used for delineating *V*ulnerability *I*so*P*ercentile lines. These VIP lines quantify the risk of non-compliance that drinking water is exposed to contamination above a pre-defined or regulatory-based threshold level. On a methodological level, VIP consists of nine modular steps. VIP is a backward risk assessment approach.

Key contributions for improved management: Compared to existing approaches, the presented VIP concept provides valuable additional probabilistic and advective-dispersive information, such as

1. The probability distribution of peak arrival travel time from a potential spill location to the well;
2. The possible levels of peak concentration arriving at the well, while accounting for dilution effects through diffusion and dispersion;
3. The probability distribution of the time window available to react after a spill event until a critical concentration level is exceeded in the well (e.g., a drinking water standard); and
4. The probability that the well has to be shut down for more than a given duration, which is the information required to estimate economic damage and consider alternative risk treatment measures.

The vulnerability isopercentile maps are easy to understand, such as showing the level of exposure risk for all locations within the well catchment by displaying areas with lower or higher probability to not comply with pre-set safety standards. By visualizing these zones prioritization of contaminated sites and location of protection zones are available. Thus, VIP maps deliver information beyond the classical deterministic wellhead delineation and vulnerability approaches.

VIP - modularity and flexibility: For demonstrating the modularity and flexibility two modeling tool sets for VIP analysis have been presented. These tool sets acknowledge state-of-the-art flow and transport models and is applicable to available and known software.

VIP is ready for practice: As demonstration for practical application, the VIP concept has been applied to a real catchment operated by the „Zweckverband Landeswasserversorgung“. The resulting VIP maps provide options for risk-aware delineation of wellhead protection zones. I have presented three management options that allow water managers to (a) quantify the current safety level of existing delineations, (b) newly delineate protection areas in an area-neutral fashion to achieve higher consistent safety levels, (c) control the costs for larger risk-aware delineations and other risk mitigation measures, (d) render the decision making process more robust against uncertainties and (e) improve transparency towards consumers and stakeholders.

11.3. Summary and Conclusions for the Proposed STORM Concept

Concept: The STORM-based forward and cumulative risk model quantifies *ST*akeholder-*O*bjective *R*isk *M*easure within a forward risk mode, accounting for multiple risk sources of different contaminant types and hazard failure frequencies across the well catchment. The STORM framework has been applied to the real test case, Burgberg, with a synthetic hazard data set.

Proper risk aggregation: Meaningful hazard ranking (prioritization) has to be mass-discharge-based. Hazards that are considered independent of each other may often pose no risk to the drinking water well. Only in a cumulative and realistic setting, these hazards threaten safe water supply, as concentration loads may exceed critical concentration levels, when overlapping from different failure events at different points in space or time. Risk in a cumulative problem setting is not the sum of risks, but the sum of hazard loads per contaminant type evaluated in a statistical framework. Furthermore, risk is highly non-linear due to threshold-based impact estimates.

Decision support with STORM: The STORM framework allows to smartly investigate the most effective risk mitigation alternative to reduce risk the most. Mitigation options are based on probability and impact reduction or both simultaneously. The risk mitigation is highly dependent on the risk attitude, hazard location, and the transport-specific catchment properties. Therefore, the choice of mitigation option (mass, probability) depends on the above mentioned factors and the local situation.

Hazard prioritization: Prioritization is either based on mean annual risk or on the impact for a given return period. Both risk estimates are related to each other, as the integral area under the annuality curve equals the annual mean risk estimate. These frequency-based risk estimates provide additional information to risk managers in comparison to the classical risk interpretation that risk is the expected impact or the $90\,\%$-confidence level given a damage. The annuality curves provide additional information to risk managers that are even more valuable than the annual mean risk measure. The curves transparently reflect, which hazards are most severe for a given return period.

Influence of stakeholder-objective risk measures: Hazard ranking is highly dependent on the risk attitude of the stakeholder. Furthermore, they allow to pinpoint hazards that contribute to severity levels that are not acceptable. Prioritization also depends on the chosen stakeholder-objective risk measure. This has been demonstrated by comparing two hazard ranking lists, one showing the ranking based on maximum concentration ratio (MCR), the other on well exposure time (WET). Furthermore, prioritization depends on the risk estimation model, here comparing three risk assessment approaches, varying from qualitative, via deterministic „risk“ (impact) to probabilistic risk assessment.

Benefits of total risk aggregation: The STORM framework assesses the total risk of supply failure, taking all hazards into account and still allows prioritization of individual hazards based on their risk contribution. Due to a single overall risk estimate, risk can be mapped over time, such that trends can be quickly visualized and recognized. Furthermore, different catchments can be compared based on the overall risk, leading to catchment prioritization for investments regarding risk reduction.
In conclusion, the conceptual and mathematical-stochastic framework behind STORM provides valuable information in a cumulative and thus complex problem setting:

1. It allows prioritization and aggregation of risks from several hazards within a well catchment across different contaminant types and failure frequency.
2. It delivers the overall probability of drinking water contamination and thus supply failure as a whole, if there are two or more hazards in the catchment, but keeps the information on individual hazards to prioritize them.
3. STORM allows to distinguish between the need of aggregation and when hazards can be treated independently of each other.
4. The framework maps the trend over time in the overall risk.
5. STORM is flexible to adapt to the objective of stakeholders (e.g., human-health, ecology, economy).
6. It transparently reflects the decisions made by stakeholders, and supports stakeholders in decision alternatives, when considering different stakeholder preferences (objectives), risk attitudes, risk estimates and measures.

11.4. Outlook

Risk quantification and management concepts for well catchment need further general improvements:

1. Risk quantification in this thesis and in most other works is based on steady-state flow fields. Festger and Walter (2002) showed, that steady-state assumptions (e.g., bounding streamlines) break down under transient flow conditions. The groundwater flow field and thus contaminant transport is subject to seasonal variations, time-dependent pumping regimes and extreme hydrological situations such as droughts and floods. Therefore, the significance of flow transients within a risk management context need more attention, such as to investigate the interplay between spatial uncertainty and flow transients or the influence of uncertain transients on risk measures and related risk-informed decisions.
2. Conceptual model error (beyond considering the Màtern correlation function to encode geostatistical model uncertainty) has been neglected in this thesis and can be incorporated by Bayesian model averaging (e.g., Hoeting et al., 1999). Taking conceptual model error into account would resolve uncertainty to a next level. Sousa et al.

(2012) transparently disclosed uncertainty in wellhead delineation by utilizing three different flow and transport models.

3. The risk framework quantifies uncertainty on the top level, such as the uncertainty in the risk measure. Using Bayesian inference methods (e.g., Perl, 2009) it is possible to trace uncertainty from element to element and thus identify the individual modeling step that introduces the largest parts of uncertainty. There are methods available that transparently reflect the degree of uncertainty per component, such as ANOVA or Bayesian networks (e.g., Thomsen et al., 2012). This enables stakeholders to smartly invest available resources to reduce epistemic uncertainty most effectively.

4. The choice of sampling designs to reduce epistemic uncertainty most effectively or to find the highest available safety level has been performed in a scenario-based manner in this thesis. To do so in a rational and more efficient manner, one can use optimal design techniques in combination with a worth of data framework, based on formal optimization methods (e.g., Feyen et al., 2003; Nowak et al., 2010). Such techniques provide the information, which types of data should be collected where and in which quality, in order to achieve the largest uncertainty reduction or highest reliability level for a given limited budget. In this area, several open questions remain. For example, Bakr and Butler (2004) investigated only the required amount of data to delineate a wellhead protection area up to a certain accuracy level without considering limited investigation budgets or quality and type of possible data types.

5. Optimization techniques could also be applied to choose the best available risk mitigation scenario for a given budget from a continuous space of scenarios. Scenarios to choose from can include premium payments, costs for data acquisition and alternative risk treatment methods.

6. I often encountered the wish to combine the four well vulnerability criteria to one unified risk measure. In this thesis, I proposed to consider a risk analysis approach, mainly based on one risk measure that is meaningful regarding the pre-defined risk objective. Nevertheless, in light of finding an optimal solution for all participating stakeholders within a risk assessment and management process, it would be necessary to find a solution between the multi-objective interests. One way to encounter this dilemma is by using multi-objective optimization (e.g., Marler and Arora, 2004).

Future Research in VIP

In specific, VIP could be extended in the following fields:

1. The reaction time denotes the available time until a contaminant load is above a certain critical threshold value. Therefore, the third vulnerability criterion is suitable to assess the reliability of groundwater monitoring networks in the context of an early warning system. By setting the problem of monitoring network design into an optimization framework, monitoring systems can be designed such that both the detection reliability and the available time for counter measures between first detection and impact at the well is maximized. It is also possible to find a cost-optimal monitoring design solution. The main challenge is to set up a suitable set of objective functions in light of costs, available time, and detection reliability within a multi-objective optimization framework.
2. The utility value of areal error assumed the effects of excess delineated area and wrongly undelineated area to be equal. In fact, the error of delineating area that is not part of the actual catchment outline is only a financial risk, whereas the error of missing land that should have been protected is more severe. Due to the second type of error, customers are unwillingly and unknowingly exposed to contaminated water that may lead over time to severe health effects. Therefore, I suggest to find a more realistic economically-based utility function that provides a more realistic view on data worth.

Future Research in STORM

In specific, STORM needs extension in the following fields:

1. For the event model, I assume independent time sequences between two events that follow an exponential waiting time distribution. This has been justified by only considering the useful life period of the risk source. The exponential function could be replaced by a beta or a Weibull distribution, such that the failure rate is used as a time-dependent variable. Due to the Monte Carlo framework, one could even randomize the time-dependent failure rate in order to account for the uncertainty in parameterization of the hazard rate.
2. The aspect of source geometry (e.g., Thomsen et al., 2012) is neglected so far. As contaminant source architecture influences the model prediction, the module on mass release could be extended by a model that accounts for randomized or scenario-based spatial mass release patterns.
3. The STORM model could be extended to account for the time-lag and mass discharge rates to the aquifer. The flow and transport through the unsaturated zone could be estimated by analytical transfer functions, such as pedotransfer functions (e.g., Wang et al., 2009), a cascade leachate model (e.g., Troldborg et al., 2008) or numerical solutions. The relevance of using time-lag models has been discussed by Sousa et al. (2012).

Bibliography

A/64/L.63/Rev.1, U., 2010. The Human Right to water and sanitation. United Nations Sixty-fourth General Assembly, Plenary, 108th Meeting, GA/10967.

Abramowitz, M., Stegun, I., 1964. Handbook of mathematical functions with formulas, graphs, and mathematical tables. Dover publications, Mineola, New York, U.S., Chapter 25.4 Gaussian Quadrature.

Alcolea, A., Carrera, J., Medina, A., 2006. Pilot points method incorporating prior information for solving the groundwater flow inverse problem. Advances in Water Resources 29 (11), 1678–1689.

Aller, L., Lehr, J., Petty, R., Bennett, T., 1987. DRASTIC - A standardized system to evaluate groundwater pollution potential using hydrogeologic setting. Journal of the Geological Society of India 29 (1), 23–37.

Andersson, C., Destouni, G., 2001. Risk-cost analysis in ground water contaminant transport: The role of random spatial variability and sorption kinetics. Groundwater 39 (1), 35–48.

Andricevic, R., Cvetkovic, V., 1998. Relative dispersion for solute flux in aquifers. Journal of Fluid Mechanics 361 (1), 145–174.

ASME, 2006. RAMCAP: Risk analysis and management for critical asset protection - The framework. American Society of Mechanical Engineers (ASME), Innovative Technologies Institute, LLC, Version 2.0.

AS/NZS 4360:2004, 2004. Risk Management. Standards Australia/Standards New Zealand, Sydney, Australia, ISBN 0 7337 5904 1.

Aven, T., 2010. On how to define, understand and describe risk. Reliability Engineering & System Safety 95 (6), 623–631.

Aven, T., 2011. On the new ISO guide on risk management terminology. Reliability Engineering & System Safety 96 (7), 719–726.

Aven, T., 2012. On the link between risk and exposure. Reliability Engineering & System Safety 106, 191–199.

Aven, T., Kørte, J., 2003. On the use of risk and decision analysis to support decision-making. Reliability Engineering & System Safety 79 (3), 289–299.

Aven, T., Renn, O., 2009. On risk defined as an event where the outcome is uncertain. Journal of Risk Research 12 (1), 1–11.

Bakr, M., Butler, A., 2004. Worth of head data in well-capture zone design: Deterministic and stochastic analysis. Journal of Hydrology 290 (3), 202–216.

Barrenblatt, G., Zheltov, I., Kochina, I., 1960. Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks. Journal of Applied Mathematics and Mechanics 24 (5), 1286–1303.

Bedford, T., Cooke, R., 2001. Probabilistic risk analysis: Foundations and methods. Cambridge University Press, Cambridge, UK.

Berkowitz, B., Scher, H., 1995. On characterization of anomalous dispersion in porous and fractured media. Water Resources Research 31 (6), 1461–1466.

Blokker, M., Ruijg, K., de Kater, H., 2005. Introduction of a substandard supply minutes performance indicator. Water Asset Management International 1, 19–22.

Bolster, D., Barahona, M., Dentz, M., Fernandez-Garcia, D., Sanchez-Vila, X., Trinchero, P., Valhondo, C., Tartakovsky, D., 2009. Probabilistic risk analysis of groundwater remediation strategies. Water Resources Research 45 (6), W06413.

CalTox, 1994. A multimedia total exposure model for hazardous-waste sites. [last accessed on 22nd march 2013].
URL `http://energy.lbl.gov/ied/era/caltox/index.html`

Chambon, J., Binning, P., Jørgensen, P., Bjerg, P., 2011. A risk assessment tool for contaminated sites in low-permeability fractured media. Journal of Contaminant Hydrology 124 (1), 82–98.

Christakos, G., Olea, R., 1992. Sampling design for spatially distributed hydrogeologic and environmental processes. Advances in Water Resources 15 (4), 219–237.

Cirpka, O., Frind, E., Helmig, R., 1999. Numerical simulation of biodegradation controlled by transverse mixing. Journal of Contaminant Hydrology 40 (2), 159–182.

Cirpka, O., Kitanidis, P., 2000a. An advective-dispersive stream tube approach for the transfer of conservative-tracer data to reactive transport. Water Resources Research 36 (5), 1209–1220.

Cirpka, O., Kitanidis, P., MAY 2000b. Characterization of mixing and dilution in heterogeneous aquifers by means of local temporal moments. Water Resources Research 36 (5), 1221–1236.

Cole, B., Silliman, S., 2000. Utility of simple models for capture zone delineation in heterogeneous unconfined aquifers. Groundwater 38 (5), 665–672.

Cornaton, F., 2003. Deterministic models of groundwater age, life expectancy and transit time distributions in advective dispersive systems. Phd thesis, University of Neuchâtel.

Cox, A., 2008. What's wrong with risk matrices? Risk Analysis 28 (2), 497–512.

Cushman, D., Driver, K., 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 (Supplement 1), 155–162.

Cvetkovic, V., Shapiro, A., Dagan, G., 1992. A solute flux approach to transport in heterogeneous formations: 2. Uncertainty analysis. Water Resources Research 28 (5), 1377–1388.

Dagan, G., 1984. Solute transport in heterogeneous porous formations. Journal of Fluid Mechanics 145, 151–177.

Dagan, G., Nguyen, V., 1989. A comparison of travel time and concentration approaches to modeling transport by groundwater. Journal of Contaminant Hydrology 4 (1), 79–91.

Davison, A., Howard, G., Stevens, M., Callan, P., Deere, D., Bartram, J., 2005. Water safety plans - Managing drinking-water quality from catchment to consumer. WHO/SDE/WSH/05.06.

de Barros, F., Bolster, D., Sanchez-Vila, X., Nowak, W., 2011. A divide and conquer approach to cope with uncertainty, human health risk, and decision making in contaminant hydrology. Water Resources Research 47 (5), W05508.

de Barros, F., Ezzedine, S., Rubin, Y., 2012. Impact of hydrogeological data on measures of uncertainty, site characterization and environmental performance metrics. Advances in Water Resources 36, 51–63.

de Barros, F., Guadagnini, A., Fernàndez-Garcia, D., Riva, M., Sanchez-Vila, X., 2013. Controlling scaling metrics for improved characterization of well-head protection regions. Journal of Hydrology 494, 107–115.

de Barros, F., Rubin, Y., 2008. A risk-driven approach for subsurface site characterization. Water Resources Research 44 (14), W01414.

de Barros, F., Rubin, Y., Maxwell, R., 2009. The concept of comparative information yield curves and its application to risk-based site characterization. Water Resources Research 45 (W06401), doi:10.1029/2008WR007324.

De Marsily, G., 1986. Quantitative hydrogeology: Groundwater hydrology for engineers. Academic Press San Diego, California, iSBN-10: 0122089162.

Delhomme, J., 1979. Spatial variability and uncertainty in groundwater-flow parameters - geostatistical approach. Water Resources Research 15 (2), 269–280.

Dentz, M., Kinzelbach, H., Attinger, S., Kinzelbach, W., 2000. Temporal behavior of a solute cloud in a heterogeneous porous medium: 1. Point-like injection. Water Resources Research 36 (12), 3591–3604.

DIN, 2000. Zentrale Trinkwasserversorgung - Leitsätze für Anforderungen an Trinkwasser, Planung, Bau, Betrieb und Instandhaltung der Versorgungsanlagen. Beuth Verlag, Deutsches Institut für Normung, Berlin, Germany.

Doessing Overheu, N., Tuxen, N., Flyvbjerg, J., Ansger Aabling, J., Andersen, J., Pedersen, J., Thyregod, T., Binning, P., Bjerg, P., 2013. Risk-based prioritisation of ground water threatening point sources at catchment and regional scale. In: AquaCon Soil. Barcelona.

Doherty, J., Hunt, R., 2010. Approaches to highly parameterized inversion: A guide to using PEST for groundwater-model calibration. No. Report 20105169. US Department of the Interior, US Geological Survey, Reston, Virgina, USA.

DVGW, 2000. Technisches Sicherheitsmanagement. Deutscher Verein des Gas- und Wasserfaches e.V. (DVGW), Bonn, Germany.

DVGW, 2006. Arbeitsblatt W101: Richtlinien für Trinkwasserschutzgebiete; Teil 1: Schutzgebiete für Grundwasser. Deutscher Verein des Gas- und Wasserfaches e.V. (DVGW), Bonn, Germany.

DVGW, 2009. Arbeitsblatt W 1001: Sicherheit in der Trinkwasserversorgung - Risikomanagement im Normalbetrieb. Deutscher Verein des Gas- und Wasserfaches e.V. (DVGW), Bonn, Germany.

Einarson, M., Mackay, D., 2001. Peer reviewed: Predicting impacts of groundwater contamination. Environmental Science & Technology 35 (3), 66–73.

Enzenhoefer, R., Bunk, T., Nowak, W., 2013. Nine-steps to risk-informed wellhead protection and management: The Burgberg case study. Groundwater (submitted).

Enzenhoefer, R., Nowak, W., Helmig, R., 2011. Optimal risk management for drinking water catchments by means of probabilistic well vulnerability criteria. In: Alegre, H., Merkel, W. (Eds.), Strategic Asset Management of Water and Wastewater Infrastructure. IWA, Mühlheim an der Ruhr, Germany.

Enzenhoefer, R., Nowak, W., Helmig, R., 2012. Probabilistic exposure risk assessment with advective-dispersive well vulnerability criteria. Advances in Water Resources 36, 121–132.

European Commission, 2003. Report from the Commission to the European Parliament and the Council on the application and effectiveness of the EIA Directive (Directive 85/337/EEC as amended by Directive 97/11/EC).

Evers, S., Lerner, D., 1998. How uncertain is our estimate of a wellhead protection zone? Groundwater 36 (1), 49–57.

Fernàndez-Garcia, D., Bolster, D., Sanchez-Vila, X., Tartakovsky, D. M., 2012. A Bayesian approach to integrate temporal data into probabilistic risk analysis of monitored NAPL remediation. Advances in Water Resources 36, 108–120.

Fernàndez-Garcia, D., Sanchez-Vila, X., 2011. Optimal reconstruction of concentrations, gradients and reaction rates from particle distributions. Journal of Contaminant Hydrology 120, 99–114.

Festger, A., Walter, G., 2002. The capture efficiency map: The capture zone under time-varying flow. Groundwater 40 (6), 619–628.

Feyen, L., Beven, K. J., De Smedt, F., Freer, J., 2001. Stochastic capture zone delineation within the generalized likelihood uncertainty estimation methodology: conditioning on head observations. Water Resources Research 37 (3), 625–638.

Feyen, L., Gorelick, S., 2005. Framework to evaluate the worth of hydraulic conductivity data for optimal groundwater resources management in ecologically sensitive areas. Water Resources Research 41 (3), W03019.

Feyen, L., Ribeiro, P., Gomez-Hernandez, J., Beven, K., De Smedt, F., 2003. Bayesian methodology for stochastic capture zone delineation incorporating transmissivity measurements and hydraulic head observations. Journal of Hydrology 271 (1-4), 156–170.

Fienen, M., Doherty, J., Hunt, R., Reeves, H., 2010. Using prediction uncertainty analysis to design hydrologic monitoring networks: Example applications from the great lakes water availability pilot project. Tech. rep., U. S. Geological Survey, Middleton, Wisconsin, U.S.

Fienen, M., Masterson, J., Plant, N., Gutierrez, B., Thieler, E., 2012. Emulation modeling with Bayesian networks for efficient decision support. AGU Fall Meeting Abstract-ID H32F-01.

Fischhoff, B., 1990. Understanding long-term environmental risks. Journal of Risk and Uncertainty 3 (4), 315–330.

Fishburn, P. C., 1970. Utility theory for decision making. John Wiley & Sons, Inc., New York, USA, ISBN: 471 26060 6.

Folks, J., Chhikara, R., 1978. The inverse Gaussian distribution and its statistical application–a review. Journal of the Royal Statistical Society. Series B (Methodological), 263–289.

Franzetti, S., Guadagnini, A., 1996. Probabilistic estimation of well catchments in heterogeneous aquifers. Journal of Hydrology 174 (1-2), 149–171.

Freeze, R. A., Gorelick, S. M., 1999. Convergence of stochastic optimization and decision analysis in the engineering design of aquifer remediation. Groundwater 37 (6), 934–954.

Freeze, R. A., Massmann, J., Smith, L., Sperling, T., James, B., 1990. Hydrogeological decision analysis: 1. A framework. Groundwater 28 (5), 738–766.

Freeze, R. A., McWhorter, D. B., 1997. A framework for assessing risk reduction due to DNAPL mass removal from low-permeability soils. Groundwater 35 (1), 111–123.

Frind, E., Molson, J., Rudolph, D., 2006. Well vulnerability: A quantitative approach for source water protection. Groundwater 44 (5), 732–742.

Fritz, J., Neuweiler, I., Nowak, W., 2009. Application of FFT-based algorithms for large-scale universal kriging problems. Mathematical Geosciences 41 (5), 509–533.

Gelhar, L., Axness, C., 1983. Three-dimensional stochastic analysis of macrodispersion in aquifers. Water Resources Research 19 (1), 161–180.

Goldscheider, N., 2005. Karst groundwater vulnerability mapping: Application of a new method in the Swabian Alb, Germany. Hydrogeology Journal 13 (4), 555–564.

Guadagnini, A., Franzetti, S., 1999. Time-related capture zones for contaminants in randomly heterogeneous formations. Groundwater 37 (2), 253–260.

Haakh, F., Bunk, T., Emmert, M., Pfaff, S., 2013. Risikomanagement für Wasserschutzgebiete. DVGW energie | wasser-praxis 3, 43–49.

Haggerty, R., Gorelick, S., 1995. Multiple-rate mass transfer for modeling diffusion and surface reactions in media with pore-scale heterogeneity. Water Resources Research 31 (10), 2383–2400.

Handcock, M., Stein, M., 1993. A Bayesian analysis of kriging. Technometrics 35 (4), 403–410.

Harbaugh, A., Banta, E., Hill, M., 2000. MODFLOW-2000, the U.S. Geological Survey modular ground-water model: User guide to modularization concepts and the ground-water flow process. No. Report 00-92. U.S. Geological Survey, Reston, Virgina, USA.

Harvey, C., Gorelick, S., 1995. Temporal moment-generating equations - modeling transport and mass-transfer in heterogeneous aquifers. Water Resources Research 31 (8), 1895–1911.

Hassan, A., Andricevic, R., Cvetkovic, V., 2002. Evaluation of analytical solute discharge moments using numerical modeling in absolute and relative dispersion frameworks. Water Resources Research 38 (2), W1009.

Hassan, A., Mohamed, M., 2003. On using particle tracking methods to simulate transport in single-continuum and dual continua porous media. Journal of Hydrology 275 (3-4), 242–260.

Hathhorn, W., 1996. A second look at the method of random walks. Stochastic Hydrology and Hydraulics 10 (4), 319–329.

Hathhorn, W., Charbeneau, R., 1994. Stochastic fluid travel times in heterogeneous porous media. Journal of Hydraulic Engineering 120 (2), 134–146.

Hendricks Franssen, H.J.and Stauffer, F., Kinzelbach, W., 2004. Joint estimation of transmissivities and recharges–application: Stochastic characterization of well capture zones. Journal of Hydrology 294 (1-3), 87–102.

Hendricks Franssen, H., Alcolea, A., Riva, M., Bakr, M., Van Der Wiel, N., Stauffer, F., Guadagnini, A., 2009. A comparison of seven methods for the inverse modelling of groundwater flow. Application to the characterisation of well catchments. Advances in Water Resources 32 (6), 851–872.

Herrera, G. S., Pinder, G. F., 2005. Space-time optimization of groundwater quality sampling networks. Water Resources Research 41 (12), W12407.

Hodgson, E., 2012. Chapter one - Human environments: Definition, scope, and the role of toxicology. Progress in molecular biology and translational science 112, 1–10.

Hoeting, J., Madigan, D., Raftery, A., Volinsky, C., 1999. Bayesian model averaging: A tutorial. Statistical science 14 (4), 382–401.

Hoeting, J. A., Davis, R. A., Merton, A. A., Thompson, S. E., 2006. Model selection for geostatistical models. Ecological Applications 16 (1), 87–98.

Hokstad, P., Roestum, J., Sklet, S., Rosn, L., Pettersson, T., Lindhe, A., Sturm, S., Beuken, R., Kirchner, D., Niewersch, C., 2009. Methods for risk analysis of drinking water systems from source to tap - Guidance report on Risk Analysis. TECHNEAU, Integrated Proj ect Funded by the European Commission under the Sixth Framework Programme, Sustainable Development, Global Change and Ecosystems Thematic Priority Area, D 4.2.4.

Indelman, P., Fiori, A., Dagan, G., 1996. Steady flow toward wells in heterogeneous formations: Mean head and equivalent conductivity. Water Resources Research 32 (7), 1975–1983.

ISO, 2009. Risk management - Principles and guidelines. Tech. Rep. ISO 31000:2009, International Organization for Standardization, Geneva, Switzerland.

IWA, et al., 2004. The Bonn charter - For safe drinking water. International Water Association, IWA publishing, London, UK.

Jacobson, E., Andricevic, R., Morrice, J., 2005. Probabilistic capture zone delineation based on an analytic solution. Groundwater 40 (1), 85–95.

Jaime Gómez-Hernánez, J., Sahuquillo, A., Capilla, J., 1997. Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data – I. Theory. Journal of Hydrology 203 (1), 162–174.

James, B. R., Freeze, R. A., 1993. The worth of data in predicting aquitard continuity in hydrogeological design. Water Resources Research 29 (7), 2049–2065.

James, B. R., Gorelick, S. M., 1994. When enough is enough: The worth of monitoring data in aquifer remediation design. Water Resources Research 30 (12), 3499–3513.

Jamin, P., Dollé, F., Chisala, B., Orban, P., Popescu, I., Hérivaux, C., Dassargues, A., Brouyère, S., 2012. A regional flux-based risk assessment approach for multiple contaminated sites on groundwater bodies. Journal of Contaminant Hydrology 127 (1), 65–75.

Jaynes, E., 1957. Information theory and statistical mechanics. Physical Review 106 (4), 620–630.

Jaynes, E., 2003. Probability theory: The logic of science. Cambridge Univ Press, Cambridge, UK.

Jeffries, J., 2009. CLEA software handbook. Tech. rep., Environment Agency, ISBN: 978-1-84911-105-8.
URL `http://www.environment-agency.gov.uk/research/planning/33732.aspx`

Kaplan, S., Garrick, B. J., 1981. On the quantitative definition of risk. Risk analysis 1 (1), 11–27.

Kinzelbach, W., 1988. The random walk method in pollutant transport simulation. Vol. 224 of Groundwater Flow and Quality Modelling. Springer, Berlin, Germany.

Kitanidis, P., 1994. The concept of the dilution index. Water Resources Research 30 (7), 2011–2026.

Kitanidis, P. K., 1986. Parameter uncertainty in estimation of spatial functions: Bayesian analysis. Water Resources Research 22 (4), 499–507.

Kitanidis, P. K., 1988. Prediction by the method of moments of transport in a heterogeneous formation. Journal of Hydrology 102 (1), 453–473.

Kitanidis, P. K., 1995. Quasi-linear geostatistical theory for inversing. Water Resources Research 31 (10), 2411–2419.

Kolluru, R., Bartell, S., Pitblado, R., Stricoff, S., 1996. Risk assessment and management handbook for environmental, health, and safety professionals. McGraw-Hill, New York, USA.

Kourakos, G., Klein, F., Cortis, A., Harter, T., 2012. A groundwater nonpoint source pollution modeling framework to evaluate long-term dynamics of pollutant exceedance probabilities in wells and other discharge locations. Water Resources Research 48 (6), W00L13.

Kunstmann, H., 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.

Lang, U., Justiz, J., 2009. Grundwassermodell Burgberg: Modellaufbau, Modelleichung und Transportberechnung. Tech. rep., Ingenieursgesellschaft Prof. Kobus und Partner GmbH, Stuttgart.

Lee, Y. W., Dahab, M. F., Bogardi, I., 1992. Nitrate risk management under uncertainty. Journal of Water Resources Planning and Management 118 (2), 151–165.

Leube, P., Nowak, W., Schneider, G., 2012. Temporal moments revisited: Why there is no better way for physically based model reduction in time. Water Resources Research 48 (11), W02501.

Lindhe, A., Rosén, L., Norberg, T., Bergstedt, O., 2009. Fault tree analysis for integrated and probabilistic risk analysis of drinking water systems. Water Research 43 (6), 1641–1653.

Lindhe, A., Rosen, L., Norberg, T., Bergstedt, O., Pettersson, T., 2011. Cost-effectiveness analysis of risk-reduction measures to reach water safety targets. Water Research 45 (1), 241–253.

MacGillivray, B., Hamilton, P., Strutt, J., Pollard, S., 2006. Risk analysis strategies in the water utility sector: An inventory of applications for better and more credible decision making. Critical Reviews in Environmental Science and Technology 36 (2), 85–139.

Marler, R. T., Arora, J. S., 2004. Survey of multi-objective optimization methods for engineering. Structural and multidisciplinary optimization 26 (6), 369–395.

Massmann, J., Freeze, R. A., 1987. 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.

Maxwell, R. M., Kastenberg, W. E., Rubin, Y., 1999. A methodology to integrate site characterization information into groundwater-driven health risk assessment. Water Resources Research 35 (9), 2841–2855.

McKnight, U., Rasmussen, J., Kronvang, B., Bjerg, P., Binning, P., 2012. Integrated assessment of the impact of chemical stressors on surface water ecosystems. Science of the Total Environment 427–428, 319–331.

Meyer, P. D., Valocchi, A. J., Eheart, J. W., 1994. Monitoring network design to provide initial detection of groundwater contamination. Water Resources Research 30 (9), 2647–2659.

Mohammad-Djafari, A., 2001. A Matlab program to calculate the maximum entropy distributions. In: Smith, C., Erickson, G., Neudorfer, P. (Eds.), Maximum Entropy and Bayesian Methods. Vol. 50 of Fundamental Theories of Physics. Eleventh International Workshop on Maximum Entropy and Bayesian Methods of Statistical Analysis, Kluwer Academic Publishers, Alphen aan den Rijn, The Netherlands, pp. 221–234.

Moutsopoulos, K., Gemitzi, A., Tsihrintzis, V., 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.

Müller, W., 2007. Collecting spatial data: Optimum design of experiments for random fields. Springer, Berlin, Germany.

Neukum, C., Azzam, R., 2009. Quantitative assessment of intrinsic groundwater vulnerability to contamination using numerical simulations. Science of the Total Environment 408 (2), 245–254.

Neuman, S., 2003. Maximum likelihood Bayesian averaging of uncertain model predictions. Stochastic Environmental Research and Risk Assessment 17 (5), 291–305.

Neupauer, R., Wilson, J., 2001. Adjoint-derived location and travel time probabilities for a multidimensional groundwater system. Water Resources Research 37 (6), 1657–1668.

Nilsen, T., Aven, T., 2003. Models and model uncertainty in the context of risk analysis. Reliability Engineering & System Safety 79 (3), 309–317.

Nowak, W., 2005. Geostatistical methods for the identification of flow and transport parameters in the subsurface. Phd thesis, University of Stuttgart, Institute for Modelling Hydraulic and Environmental Systems, Stuttgart, Germany.

Nowak, W., Cirpka, O. A., 2004. A modified Levenberg–Marquardt algorithm for quasi-linear geostatistical inversing. Advances in Water Resources 27 (7), 737–750.

Nowak, W., de Barros, F., Rubin, Y., 2010. Bayesian geostatistical design: Task-driven optimal site investigation when the geostatistical model is uncertain. Water Resources Research 46 (3), W03535.

Nowak, W., Schwede, R. L., Cirpka, O. A., Neuweiler, I., 2008. Probability density functions of hydraulic head and velocity in three-dimensional heterogeneous porous media. Water Resources Research 44 (8), W08452.

Öberg, T., Bergbäck, B., 2005. A review of probabilistic risk assessment of contaminated land. Journal of Soils and Sediments 5 (4), 213–224.

O'Connor, D. R., 2002. Report of the Walkerton Commission of Inquiry: A strategy for safe drinking water (Part Two). Ontario Ministry of the Attorney General, ISBN: 0-7794-2621-5.

Page, R., Scheidler, S., Polat, E., Svoboda, P., Huggenberger, P., 2012. Faecal indicator bacteria: Groundwater dynamics and transport following precipitation and river water infiltration. Water, Air, & Soil Pollution 223 (5), 2771–2782.

Pappenberger, F., Beven, K., 2006. Ignorance is bliss: Or seven reasons not to use uncertainty analysis. Water Resources Research 42 (5), W05302.

Pardo-Igúzquiza, E., 1999. Bayesian inference of spatial covariance parameters. Mathematical Geology 31 (1), 47–65.

Park, I., Amarchinta, H. K., Grandhi, R. V., 2010. A Bayesian approach for quantification of model uncertainty. Reliability Engineering & System Safety 95 (7), 777–785.

Perl, J., 2009. Causality: Models, reasoning and inference, 2nd Edition. Cambridge University Press, Camebridge, UK.

Pollard, S., Davies, G., Coley, F., Lemon, M., 2008. Better environmental decision making - Recent progress and future trends. Science of the Total Environment 400 (1), 20–31.

Pollock, D., 1994. User's guide for MODPATH/MODPATH-Plot, Version 3: A Particle Tracking Post-processing Package for MODFLOW, the US: Geological Survey Finite-difference. U.S. Geological Survey, Reston, Virigina.

Pollock, D., Salama, R., Kookana, R., 2002. A study of atrazine transport through a soil profile on the Gnangara Mound, Western Australia, using LEACHP and Monte Carlo techniques. Soil Research 40 (3), 455–464.

Pollock, D. W., 1988. Semianalytical computation of path lines for finite-difference models. Groundwater 26 (6), 743–750.

Prickett, T., Naymik, T., Lonnquist, C., 1981. A 'random walk' solute transport model for selected groundwater quality evaluations. Illinois State Water Survey, Bulletin 65 (103), 68ff.

Pukelsheim, F., 2006. Optimal design of experiments. Vol. 50 of Classics in Applied Mathematics. John Wiley & Sons, Inc., New York, U.S., society for Industrial and Applied Mathematics.

Refsgaard, J., van der Sluijs, J., Højberg, A., Vanrolleghem, P., 2007. Uncertainty in the environmental modelling process–A framework and guidance. Environmental Modelling & Software 22 (11), 1543–1556.

Reilly, T., Pollock, D., NOV-DEC 1996. Sources of water to wells for transient cyclic systems. Groundwater 34 (6), 979–988.

Renard, P., 2007. Stochastic hydrogeology: What professionals really need? Groundwater 45 (5), 531–541.

Riva, M., Guadagnini, L., Guadagnini, A., Ptak, T., Martac, E., 2006. Probabilistic study of well capture zones distribution at the Lauswiesen field site. Journal of Contaminant Hydrology 88 (1-2), 92–118.

Riva, M., Willmann, M., 2009. Impact of log-transmissivity variogram structure on groundwater flow and transport predictions. Advances in Water Resources 32 (8), 1311–1322.

Rodak, C., Silliman, S., 2012. Probabilistic risk analysis and fault trees: Initial discussion of application to identification of risk at a wellhead. Advances in Water Resources 36, 133–145.

Rubin, Y., 2003. Applied stochastic hydrogeology. Oxford University Press, Inc., New York, USA, ISBN: 0-19-513804-X.

Rubin, Y., Sun, A., Maxwell, R., Bellin, A., 1999. The concept of block-effective macrodispersivity and a unified approach for grid-scale-and plume-scale-dependent transport. Journal of Mathematical Fluid Mechanics 395, 161–180.

Sadiq, R., Kleiner, Y., Rajani, B., 2007. Water quality failures in distribution networks - Risk analysis using fuzzy logic and evidential reasoning. Risk Analysis 27 (5), 1381–1394.

Salamon, P., Fernàndez-Garcia, D., Gómez-Hernández, J., 2006. A review and numerical assessment of the random walk particle tracking method. Journal of Contaminant Hydrology 87 (3-4), 277–305.

Sander, P., Öberg, T., 2006. Comparing deterministic and probabilistic risk assessments. A case study at a closed steel mill in southern Sweden. Journal of Soils and Sediments 6 (1), 55–61.

Scheidegger, A., 1954. Statistical hydrodynamics in porous media. Journal of Applied Physics 25 (8), 994–1001.

Schloz, W., Armbruster, V., Prestel, R., Weinzierl, W., 2007. Hydrogeologisches Abschlussgutachten zur Neubegrenzung des Wasserschutzgebietes Donauried-Hübe für die für die Fassungen des Zweckverbandes Landeswasserversorgung im württembergischen Donauried und bei Giengen-Burgberg. Tech. rep., Regierungspräsidium Freiburg, Landesamt für Geologie, Rohstoffe und Bergbau.

Schöniger, A., Nowak, W., Hendricks Franssen, H., 2012. Parameter estimation by ensemble Kalman filters with transformed data: Approach and application to hydraulic tomography. Water Resources Research 48 (4), W04502.

Shao, J., Tu, D., 1996. The jackknife and bootstrap. Springer Series in Statistics. Springer Science + Business Media New York, Inc., New York, U.S., iSBN-10:0387945156.

Shrader-Frechette, K. S., 1991. Risk and rationality: Philosophical foundations for populist reforms. University of California Press, Berkley and Los Angeles, USA, ISBN: 978-0520072879.

Sousa, M., Jones, J., Frind, E., Rudolph, D., 2012. A simple method to assess unsaturated zone time lag in the travel time from ground surface to receptor. Journal of Contaminant Hydrology 144 (1), 138–151.

Srinivasan, G., Keating, E., Moulton, J., Dash, Z., Robinson, B., 2012. Convolution-based particle tracking method for transient flow. Computational Geosciences 16 (3), 551–563.

Stauffer, F., Guadagnini, A., Butler, A., Hendricks Franssen, H., Van den Wiel, N., Bakr, M., Riva, M., Guadagnini, L., 2005. Delineation of source protection zones using statistical methods. Water Resources Management 19 (2), 163–185.

Stauffer, F., Hendricks Franssen, H., Kinzelbach, W., 2004. Semianalytical uncertainty estimation of well catchments: Conditioning by head and transmissivity data. Water Resources Research 40 (8), W08305.

Tait, N., Davison, R., Whittaker, J., Leharne, S., Lerner, D., 2004. Borehole optimisation system (BOS) - a GIS based risk analysis tool for optimising the use of urban groundwater. Environmental Modelling & Software 19 (12), 1111–1124.

Tarantola, A., 2004. Inverse Problem Theory and Model Parameter Estimation. SIAM: Society for Industrial and Applied Mathematics, Paris, France, iSBN-10: 0898715725.

Tartakovsky, D., 2013. Assessment and management of risk in subsurface hydrology: A review and perspective. Advances in Water Resources 51, 247–260, 35th Year Anniversary Issue.

Tartakovsky, D. M., 2007. Probabilistic risk analysis in subsurface hydrology. Geophysical Research Letters 34 (5), L05404.

Tesfamichael, A. A., Caplan, A. J., Kaluarachchi, J. J., 2005. Risk-cost-benefit analysis of atrazine in drinking water from agricultural activities and policy implications. Water Resources Research 41 (5), W05015.

Thomsen, N., Troldborg, M., Mcknight, U., Binning, P., Bjerg, P., 2012. Uncertainty of mass discharge estimation from contaminated sites at screening level. In: AGU Fall Meeting. San Francisco, USA, Abstract-ID: H51A-1327.

Tiemann, M., 2010. Safe Drinking Water Act (SDWA): A summary of the act and its major requirements. Congressional Research Service, CRS Report RL31243.

TrinkwV, 2001. Verordnung über die Qualität von Wasser für den menschlichen Gebrauch. Viega Verlag, Attendorn, Germany, federal regulation, last amended: 28.11.2011.

Troldborg, M., 2010. Risk assessment models and uncertainty estimation of groundwater contamination from point sources. Phd thesis, DTU Environment, Technical University of Denmark, Lyngby, Denmark.

Troldborg, M., Lemming, G., Binning, P., Tuxen, N., Bjerg, P., 2008. Risk assessment and prioritisation of contaminated sites on the catchment scale. Journal of Contaminant Hydrology 101 (1), 14–28.

Turner, K., Georgiou, S., Clark, R., Brouwer, R., 2004. Economic valuation of water resources in agriculture. Vol. 27 of FAO Water Reports. FAO Land and Water Development Division, Rome, Italy, ISBN 92-5-105190-9.
URL `http://www.fao.org/docrep/007/y5582e/y5582e08.htm`

UBA, 2013. Accessed on 20th February 2013.
URL `http://www.umweltbundesamt.de/wasser/themen/grundwasser/index.htm`

Uffink, G., 1989. Application of Kolmogorov's backward equation in random walk simulations of groundwater contaminant transport. In: Kobus, H., Kinzelbach, W. (Eds.), Contaminant Transport in Groundwater: Proceedings of the International Symposium on Contaminant Transport in Groundwater. A. A. Balkema, Brookfield, Vt., pp. 283–298.

US EPA, 1989. Risk assessment guidance for superfund. Volume I: Human health evaluation: Manual (Part A). U.S. Environmental Protection Agency, Washington, D.C., USA.

US EPA, 1993. Guidelines for delineation of wellhead protection areas. U.S. Environmental Protection Agency, Washington, D.C., USA, EPA/440/5-93/001.

US EPA, 1997. Guiding principles for Monte Carlo analysis. U.S. Environmental Protection Agency, Washington, D.C., USA, EPA/630/R-97/001.

US EPA, 2000. Supplementary guidance for conducting health risk assessment of chemical mixtures. U.S. Environmental Protection Agency, Washington, D.C., USA, EPA/630/R-00/002.

US EPA, 2001. Risk Assessment Guidance for Superfund: Volume III - Part A, Process for Conducting Probabilistic Risk Assessment. U.S. Environmental Protection Agency, Washington, D.C., USA, EPA/540/R-02/002.

US EPA, 2007. Concepts, methods, and data sources for cumulative health risk assessment of multiple chemicals, exposures and effects: A resource document. U.S. Environmental Protection Agency, Washington, D.C., USA, EPA/600/R-06/013F.

US EPA, 2009. National Primary Drinking Water Regulations. U.S. Environmental Protection Agency, Washington, D.C., USA, EPA/816/F-09/004.

US EPA, 2012. Setting standards for safe drinking water. Last updated May 21, 2012.
URL `http://www.epa.gov/safewater/standard/setting.html`

U.S. Nuclear Regulatory Commission, 1974. Reactor safety study: An assessment of accident risks in US commercial nuclear power plants. US Government Printing Office, Washington D.C., USA.

van Leeuwen, M., te Stroet, C. B., Butler, A. P., Tompkins, J. A., 1998. Stochastic determination of well capture zones. Water Resources Research 34 (9), 2215–2223.

Vargas-Guzmán, J., Yeh, T., 2002. The successive linear estimator: A revisit. Advances in Water Resources 25 (7), 773–781.

Varljen, M., Shafer, J., 1991. Assessment of uncertainty in time-related capture zones using conditional simulation of hydraulic conductivity. Groundwater 29 (5), 737–748.

Vassolo, S., Kinzelbach, W., Schäfer, W., 1998. Determination of a well head protection zone by stochastic inverse modelling. Journal of Hydrology 206 (3), 268–280.

Verreydt, G., Van Keer, I., Bronders, J., Diels, L., Vanderauwera, P., 2012. Flux-based risk management strategy of groundwater pollutions: The CMF approach. Environmental Geochemistry and Health 34 (6), 1–12.

Wackerly, D., Mendenhall, W., Scheaffer, R., October 10,2007 2002. Mathematical statistics with applications, 7th Edition. Duxbury Press, ISBN: 0495110817.

Walker, W. E., Harremoës, P., Rotmans, J., Van der Sluijs, J., Van Asselt, M., Janssen, P., Von Krauss, M. K., 2003. Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. Integrated Assessment 4 (1), 5–17.

Wang, T., Zlotnik, V. A., Šimunek, J., Schaap, M. G., 2009. Using pedotransfer functions in vadose zone models for estimating groundwater recharge in semiarid regions. Water Resources Research 45 (4), W04412.

WHO, 2004. Guidelines for Drinking-Water Quality, 3rd Edition. WHO publications, Geneva, Switzerland, ISBN: 978 92 4 154761 1.

Wilkins, D., 2002. The bathtub curve and product failure behavior. Reliability HotWire 21, 17.

Woodbury, A., Ulrych, T., 2000. A full-Bayesian approach to the groundwater inverse problem for steady state flow. Water Resources Research 36 (8), 2081–2093.

Zanini, A., Kitanidis, P. K., 2009. Geostatistical inversing for large-contrast transmissivity fields. Stochastic Environmental Research and Risk Assessment 23 (5), 565–577.

Zwahlen, F., 2004. Vulnerability and risk mapping for the protection of carbonate (karst) aquifers. Final Report (COST Action 620), European Commission, Brussels.

A. Obtaining Vulnerability Criteria by Moment-based Breakthrough Curve Reconstruction

The goal of this chapter is to introduce the mathematical concept to reconstruct contaminant breakthrough curves from statistical moments. The information on the breakthrough curve history is needed for the VIP concept as introduced in Chapter 4 to assess the four well vulnerability criteria (step 5, Fig. 4.1). Due to model reduction techniques, the transport model delivers only information on the zeroth to fourth temporal moments. Based on these moments, it is possible to obtain a non-linear polynomial function that represents a skewed breakthrough curve. The aim is to find the so-called Lagrangian parameters of the polynomial function with the help of an optimization scheme, such as Newton-Raphson iteration. The methodology presented here in this chapter is applied in the synthetic test case (Chapter 8).
First, I give a brief problem description and present the Equations to be solved (see Section A.1). As a next step, I introduce the Gauss-Hermite Integration, as it was adapted to this specific problem setting (see Section A.2). The third part introduces how to calculate analytically the third and fourth well vulnerability criteria by the MATLAB „root“ function with the previously obtained Lagrangian parameters (see Section A.3). This analytical approach helps to save computational demand, making the transport code faster.

A.1. Maximum Entropy in Log-Time

The entropy E of a probability density function $p(t)$ is defined as

$$E = -\int_{-\infty}^{\infty} p(t) \ln p(t) dt \tag{A.1}$$

In the maximum entropy framework, a probability density is found that maximizes entropy, E, which is subject to side conditions that a set of moments $E[\Xi_k]$ is given, where Ξ_k denotes the temporal moment. $\Xi_k(t)$ is the polynomial function for which the Langrangian parameters have to be found. It has been shown by Mohammad-Djafari (2001) that the resulting probability density function $p(t)$ must have the form:

$$p(t) = \exp\left[-\sum_{\ell=0}^{n_\ell} \lambda_\ell \cdot \Xi_\ell(t)\right] \tag{A.2}$$

with n_ℓ being the number of side conditions and λ_ℓ the respective ℓ-th Lagrangian parameter. The moments are calculated by the general form

$$E[\Xi_k] = \int_{-\infty}^{\infty} \Xi_k \, p(t) dt = \mu_k \tag{A.3}$$

with μ_k being the k-th moment. In the current context, the random variable t is time, and μ_k are temporal moments obtained from the transport code, and the probability density function $p(t)$ represents the breakthrough curve, $c(t)$. Taking Eq. (A.2) and inserting into Eq. (A.1) leads to

$$G_k(\lambda) = \int \Xi_k(t) \exp\left[-\sum_{\ell=0}^{n_\ell} \lambda_\ell \cdot \Xi_\ell(t)\right] dt = \mu_k \tag{A.4}$$

with G_k being the integral function of the k-th moment (see Eq. A.3). It is now possible to solve Eq. (A.4) by any optimization scheme for the Lagrangian parameters λ_ℓ. Following the conceptual idea by Harvey and Gorelick (1995) to solve the contaminant breakthrough curve in log-time, reads for $p(t)$:

$$p(t) = c(t) = \frac{1}{t} \exp\left[-\sum_{\ell=0}^{n_\ell} \lambda_\ell \cdot \ln(t)^\ell\right] \tag{A.5}$$

Substituting Eq. (A.5) into Eq. (A.4) leads to:

$$G_k(\lambda) = \int_{-\infty}^{\infty} t^k \frac{1}{t} \exp\left[-\sum_{\ell=1}^{n_\ell} \lambda_\ell \cdot t^\ell\right] \mathrm{dt} = \mu_k \tag{A.6}$$

with

$$\sum_{\ell=1}^{n_\ell} \lambda_\ell \cdot t^\ell = p(t) = \left(\lambda_4 \ln(t)^4 + \lambda_3 \ln(t)^3 + \lambda_2 \ln(t)^2 + \lambda_1 \ln(t) + \lambda_0\right) \tag{A.7}$$

with $\ell = 0, \ldots 4$, when solving for up to four temporal moments, leading to five Lagrangian parameters λ_ℓ.

Back-Transformation to Maximum Entropy in Time

While maximizing entropy in log-time has been shown to produce good results in Harvey and Gorelick (1995), it is mathematically more convenient to solve the problem without the logarithm. Thus, I apply a transformation of the form

$$\begin{aligned} t &= \exp(s) \\ \ln(t) &= s \end{aligned} \tag{A.8}$$

to Eq. (A.6). Thus, $dt = \exp(s) ds$ changes Eq. (A.6) to:

$$G_k(s) = \int_{-\infty}^{\infty} \exp\left[ks + P(s)\right] \mathrm{ds} = \mu_k \tag{A.9}$$

with

$$P(s) = -\left(\lambda_4 s^4 + \lambda_3 s^3 + \lambda_2 s^2 + \lambda_1 s + \lambda_0\right) \tag{A.10}$$

The only remaining difference to the standard problem of finding a Maximum Entropy distribution is the appearance of ks in Eq. (A.9).

A.2. Gauss-Hermite Integration

The integration in Eq. (A.9) has to be performed many times, for $k = 0, \ldots, K$ (when K is the largest order of temporal moments considered), and for many cycles of iteration within he optimization to find the values of λ_k. One of the best methods for 1-dimensional integration is Gauss-Hermite quadrature. Gauss-Hermite quadrature (Abramowitz and Stegun, 1964) approximates the value of integrals the following form:

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{dx} = \int_{-\infty}^{\infty} \omega(x)\,\exp(x^2)\cdot f(x)\,\mathrm{dx} \approx \sum_{i=1}^{n} \omega_i\,\exp(x_i^2)\,f(x_i) \tag{A.11}$$

with n being the number of Gauss-Integration points x_i, with corresponding weights ω_i and a weighting function $\omega(x)$. Depending on the probabilistic or physicist interpretation of the Hermite polynomial (Eq. A.13), the probabilists' weighting function $\omega(x)$ follows a standard normal distribution

$$\omega(x) = \exp\left[-\frac{1}{2}\,x^2\right], \tag{A.12}$$

The values x_i are found as the roots of the Hermite polynomials $H_n(x)$, and the corresponding weights can be calculated

$$\omega_i = \frac{2^{n-1}n!\sqrt{\pi}}{n^2\left[H_{n-1}(x_i)\right]^2} \tag{A.13}$$

The explicit form of the probabilists' Hermite polynomials of order n, i.e., for any total number of Gauss-Iteration-Points n is

$$H_n(x) = n!\sum_{m=0}^{n/2}\frac{(-1)^m}{m!\,(n-2m)!}\frac{x^{n-2m}}{2^m} \tag{A.14}$$

Adaption to the Maximum Entropy Problem

In order to evaluate Eq. (A.9) with Gauss Hermite quadrature, I extend Eq. (A.9) by:

$$1 = \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]\cdot\exp\left[+\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] = \frac{\omega(x)}{\omega(x)}, \tag{A.15}$$

and define:

$$\frac{f(x)}{\omega(x)} = \exp\left[ks + P(s)\right] \tag{A.16}$$

As the weighting function $\omega(x)$ follows a standard normal distribution (see Eq. A.12), now I have to find suitable values for μ and σ. The weights have to be shifted to the location of the mean μ_s and stretched by the variance σ_s^2 of the integral function (see Eq. A.9). The weights change through the rescaling proportional to σ_s. The transformed weighting function follows the distribution $w' \sim N\left(\mu_s, \sigma_s\right)$ and reads

$$\omega' = \exp\left[-\frac{1}{2}\left(\frac{x-\mu_s}{\sigma_s}\right)^2\right] \tag{A.17}$$

Thus Eq. (A.9) becomes

$$\mu_k = \int_{-\infty}^{\infty} \exp\left[ks + p(s)\right] \cdot w' \cdot \exp\left[+\frac{1}{2}\left(\frac{s-\mu_s}{\sigma_s}\right)^2\right] \text{ ds} \tag{A.18}$$

with μ_s being the mean value of the integral function of Eq. (A.9). The corresponding shifted Gauss-Integration-Points are:

$$s_i = x_i \cdot \sqrt{2\sigma_s^2} + \mu_s \tag{A.19}$$

and the corresponding weights ω_i':

$$\omega_i' = w_i \cdot \sqrt{2\sigma_s^2} \tag{A.20}$$

This leads to the final equation that can be numerically integrated at the shifted Gauss-Integration-Points s_i and applied with the transformed weights ω_i:

$$\mu_k = \int_{-\infty}^{\infty} \omega_i' \cdot \exp\left[ks_i + P(s_i) + \frac{1}{2\sigma_s^2}\left(s_i - \mu_s\right)^2\right] \text{ ds} \approx \sum_{i=1}^{n} w_i' \cdot f(s_i) \tag{A.21}$$

with n being the number of Gauss-Integration-Points. This is the equation on which the optimization for finding the Lagrangian parameters λ_k will actually be applied.

Getting μ_s and σ_s

The mean, variance and standard deviation of the log-normal distribution is generally given as follows

$$\begin{aligned} E[X] = \mu &= \exp\left[\mu_s + \frac{1}{2}\sigma_s^2\right] \\ VAR[X] = \sigma^2 &= \left(\exp\left[\sigma_s^2\right] - 1\right) \cdot \exp\left[2\mu_s + \sigma_s^2\right] \\ STD[X] = \sigma &= \sqrt{\exp\left[\sigma_s^2\right] - 1} \cdot \exp\left[\mu_s + \frac{1}{2}\sigma_s^2\right] \end{aligned} \tag{A.22}$$

μ_s and σ_s are the mean and standard deviation of the logarithmic distribution, whereas μ_ℓ and σ_ℓ are the ℓ-th temporal moments calculated by Eq. (5.33). Eq. (A.22) is rearranged in order to receive μ_s and σ_s^2.

$$\begin{aligned} \log(\mu) &= \mu_s + \frac{1}{2}\sigma_s^2 \\ \log(\sigma^2) &= \log\left\{\left(\exp\left[\sigma_s^2\right] - 1\right) \cdot \exp\left[2\mu_s + \sigma_s^2\right]\right\} \end{aligned} \tag{A.23}$$

The logarithmic mean is

$$\mu_s = \log(\mu) - \frac{1}{2}\sigma_s^2 \tag{A.24}$$

Substituting μ_s from Eq. (A.24) in Eq. (A.23) leads for σ_s^2 to

$$\sigma_s^2 = \log\left(\frac{\sigma^2}{\mu^2} + 1\right) = \log\left(\frac{m_{2c}}{m_1^2} + 1\right) \tag{A.25}$$

with m_{2c} being the second central moment (variance) and m_1 being the first temporal moment. Substituting Eq. (A.25) into Eq. (A.24) leads to

$$\mu_s = \log(\mu) - \frac{1}{2}\log\left(\frac{\sigma^2}{\mu^2}+1\right) = \log(m_1) - \frac{1}{2}\log\left(\frac{m_{2c}}{m_1^2}+1\right) \tag{A.26}$$

Using Eq. (A.21) in combination with Eq. (A.26) and Eq. (A.25) within an optimization routine directly leads to the Lagrangian parameters λ_ℓ.

A.3. Reconstructing Well Vulnerability Criteria by the MATLAB 'roots'-Function

Here, I show how to assess the points in time, when the concentration rises above (t_1) and falls again below (t_2) a given critical level c_{crit}. t_1 equals the third well vulnerability criterion t_{react}. Well down-time is calculated by $t_{exp} = t_2 - t_1$.
Eq. (A.5) (concentration profile in log-time) is set equal to c_{crit} in order to receive the roots t_1 and t_2:

$$c(t) = \frac{1}{t}\exp\left[-\lambda_4\ln(t)^4 - \lambda_3\ln(t)^3 - \lambda_2\ln(t)^2 - \lambda_1\ln(t) - \lambda_0\right] = c_{crit} \tag{A.27}$$

Due to substitution (see Eq. A.8), Eq. (A.27) reads:

$$c(s) = \exp(-s)\,\exp\left[-\lambda_4 s^4 - \lambda_3 s^3 - \lambda_2 s^2 - \lambda_1 s - \lambda_0\right] = c_{crit} \tag{A.28}$$

Eq. (A.28) is rearranged to

$$\lambda_4 s^4 + \lambda_3 s^3 + \lambda_2 s^2 + (\lambda_1 + 1)\,s + \lambda_0 + \ln(c_{crit}) = 0 \tag{A.29}$$

The polynomial coefficients c_i to be passed to the MATLAB „roots" function are:

$$\begin{aligned} c_0 &= \lambda_0 + \ln(c_{crit}) \\ c_1 &= \lambda_1 + 1 \\ c_2 &= \lambda_2 \\ c_3 &= \lambda_3 \\ c_4 &= \lambda_4 \end{aligned} \tag{A.30}$$

Instead of using the MATLAB „roots"-function, one can calculate according to the „roots"-algorithm the eigenvalues of the companion matrix A:

$$A = \begin{bmatrix} \frac{c_0}{c_4} & \frac{c_1}{c_4} & \frac{c_2}{c_4} & \frac{c_3}{c_4} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{A.31}$$

With the coefficients given in Eq. (A.30) and the Lagrangian parameters λ_k attained from Eq. (A.21), it is now possible to calculate for the polynomial of Eq. (A.29) the corresponding roots, which are the eigenvalues of matrix A (Eq. A.31). In order to calculate the third and fourth well vulnerability criterion, the roots $\hat{s}_i$ received here (Eq. A.29) are re-substituted by $t_i = \exp(\hat{s}_i)$ with $i = 1, 2$.

B. Appendix - Figures

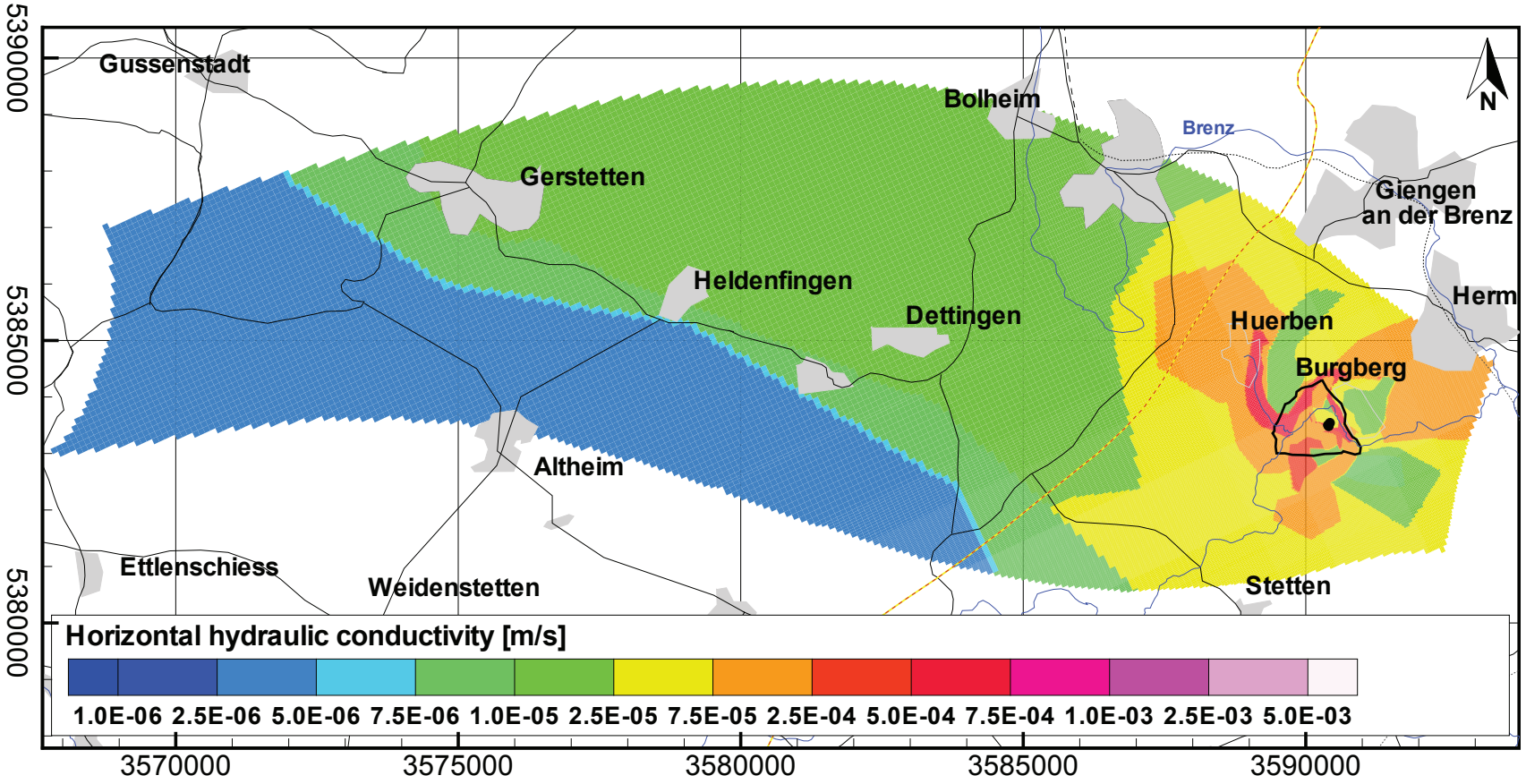


Figure B.1.: Zonation for the lower Burgberg aquifer, showing layers 3 of the MODFLOW model by Lang and Justiz (2009). The position of the wells are marked by the black dot, black lines show the road system, blue lines mark creeks and rivers and the gray areas show cities and villages.

C. Appendix - Tables

	Max. Concentration Ratio	Contaminant Load Exposure	Well Exposure Time	Response Time	Required Blending Ratio
Abbrev.	MCR	CLE	WET	RT	RBR
WVC	c_{peak}	m'_0	t_{exp}	t_{react}	c_{peak}, t_{exp}
Spatial type	point, line, area	point, line, area	point, line, area	point, line, area	point, line, area
Temporal type	spill, contin.	spill, contin.	spill	spill, contin.	spill, contin.
Contaminant type	all	all	e.g., microbiological	all	all
Risk type	acute health	chronic health	technical, economic	technical	health, technical
Risk transfer	TU	HQ	CML	-	-

Table C.1.: Possible (intermediate) risk measures by means of well vulnerability criteria.

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

1 Röhnisch, Arthur: *Die Bemühungen um eine Wasserbauliche Versuchsanstalt an der Technischen Hochschule Stuttgart*, und
Fattah Abouleid, Abdel: *Beitrag zur Berechnung einer in lockeren Sand gerammten, zweifach verankerten Spundwand*, 1963

2 Marotz, Günter: *Beitrag zur Frage der Standfestigkeit von dichten Asphaltbelägen im Großwasserbau*, 1964

3 Gurr, Siegfried: *Beitrag zur Berechnung zusammengesetzter ebener Flächentragwerke unter besonderer Berücksichtigung ebener Stauwände, mit Hilfe von Randwert- und Lastwertmatrizen*, 1965

4 Plica, Peter: *Ein Beitrag zur Anwendung von Schalenkonstruktionen im Stahlwasserbau*, und Petrikat, Kurt: *Möglichkeiten und Grenzen des wasserbaulichen Versuchswesens*, 1966

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: *Versucheinrichtung 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, ISBN 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: Upscaling, 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, 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 Hochwasserrückhaltebecken-systemen,* 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 3-933761-57-3

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

Die Mitteilungshefte ab der Nr. 134 (Jg. 2005) stehen als pdf-Datei über die Homepage des Instituts: www.iws.uni-stuttgart.de zur Verfügung.