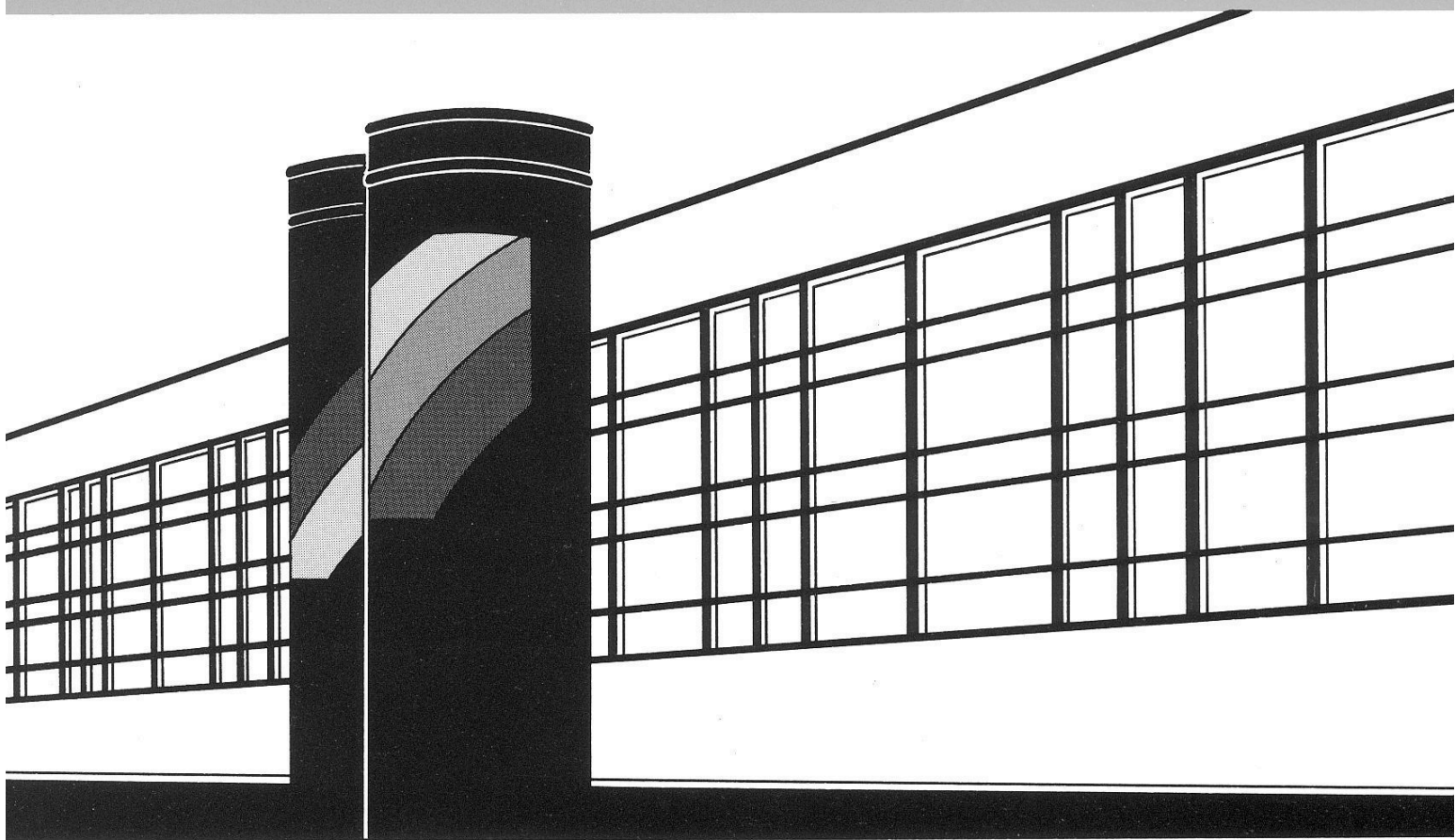


Institut für Wasserbau · Universität Stuttgart

Mitteilungen



Heft 206 Claus Haslauer

Analysis of Real-World Spatial
Dependence of Subsurface Hydraulic
Properties Using Copulas With a Focus
on Solute Transport Behaviour

Analysis of Real-World Spatial Dependence of Subsurface Hydraulic Properties Using Copulas With a Focus on Solute Transport Behaviour

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

Vorgelegt von
Claus Haslauer
aus Nürnberg, Deutschland

Hauptberichter: Prof. Dr. rer.nat. Dr.-Ing. András Bárdossy
Mitberichter: Prof. Dr. Edward Sudicky

Tag der mündlichen Prüfung: 16. September 2011

Institut für Wasserbau der Universität Stuttgart
2011

Heft 206 Analysis of Real-World
Spatial Dependence of
Subsurface Hydraulic Properties
Using Copulas
With a Focus on Solute Transport
Behaviour

von
Dr.-Ing.
Claus Haslauer

D93 Analysis of Real-World Spatial Dependence of Subsurface Hydraulic Properties Using Copulas With a Focus on Solute Transport Behaviour

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

Haslauer, Claus:

Analysis of Real-World Spatial Dependence of Subsurface Hydraulic Properties Using Copulas With a Focus on Solute Transport Behaviour / von Claus Haslauer. Institut für Wasserbau, Universität Stuttgart. - Stuttgart: Inst. für Wasserbau, 2011

(Mitteilungen / Institut für Wasserbau, Universität Stuttgart: H. 206)

Zugl.: Stuttgart, Univ., Diss., 2011

ISBN 978-3-942036-10-8

NE: Institut für Wasserbau <Stuttgart>: Mitteilungen

Gegen Vervielfältigung und Übersetzung bestehen keine Einwände, es wird lediglich um Quellenangabe gebeten.

Herausgegeben 2011 vom Eigenverlag des Instituts für Wasserbau
Druck: Document Center S. Kästl, Ostfildern

Acknowledgments

I'd like to thank the following people and institutions:

- András Bárdossy for being an unlimited source of magnificent ideas
- Ed Sudicky for his decisive numerical modelling expertise and his ability to bring stochastic hydrogeology “down to earth”
- Oliver Röhrle for acting as chair of the thesis committee at late notice and at an unthankful time
- students who have worked and discussed things with me: Philipp Guthke, Theresia Heisserer, Luise Mensch, Markus Rau, Maria Niedermaier, Sebastian Hörning, Michael Seeger
- colleagues at the department of hydrology and geohydrology and at the institute of hydraulic engineering – particularly Ferdi Beck, Thomas Pfaff, Dirk Schlabing, Maggie Eder, Jhan Rodríguez-Fernández, Felix Herma, Jochen Seidel, Takayuki Sugimoto, Jens Götzing, Jürgen Brommundt, Christian Ebert, Jan Bliefernicht, Astrid Lemp, Krista Uhrmann
- the German Research Foundation (DFG) and the State Institute for Environmental Protection Baden-Württemberg (LUBW) for funding
- my parents who have always believed in me
- Maïke Springmann who has supported me

Contents

List of Figures	V
List of Tables	V
Notation	VI
Kurzfassung	VII
Abstract	XIII
1 Introduction	1
2 Geostatistics	4
2.1 State of the Art	4
2.2 Critique of the State of the Art	8
2.2.1 Philosophy	9
2.2.2 Gaussian Assumption	9
2.2.3 Other Kriging-Issues	12
2.3 Copulas for Spatial Statistics	13
2.3.1 Basics of Copulas	13
2.3.2 Empirical Geostatistics using Copulas	15
2.3.3 Theoretical Spatial Copulas	19
2.3.3.1 Gaussian Copula	19
2.3.3.2 V-Copula	20
2.3.3.3 Maximum Gauss Copula	22
2.3.4 Parameter Estimation	24
2.3.4.1 Parameter Estimation for V-Copula	25
2.3.4.2 Parameter Estimation for Maximum Gauss Copulas	26
2.3.5 Properties of Spatial Copula Models	27
2.3.5.1 Interpolation	27
2.3.5.2 Simulation	28
3 Used Data-Sets and Their Geo-Statistical Description	31
3.1 Basic Description of the Data-Sets	31
3.1.1 The Borden Data-Set	31
3.1.2 The North Bay Data-Set	32
3.1.3 Why These Two Data-Sets	34
3.1.4 Marginal Distributions	36

3.2	Empirical Geostatistics on the Data-Sets	37
3.2.1	Empirical Geostatistics on the Borden Data-Set	37
3.2.2	Empirical Geostatistics on the North Bay Data-Set	39
3.3	Parameter Estimation	42
3.3.1	Gauss Copula	43
3.3.2	V-Copula	43
3.3.3	Maximum Gauss Copula	43
3.4	Comparison of Theoretical Copula Models With Data	48
3.4.1	Comparing Bivariate Copula Densities	49
3.4.2	Bootstrap Test of the Measure Symmetry	52
3.4.3	Fit in Higher Dimensions	52
3.5	Spatial Interpolation and Simulation	54
4	Implications on Solute Transport	57
4.1	Groundwater Flow and Solute Transport Model	58
4.2	Analysis of Spatial Moments of Solute Plume	61
4.3	Analysis of Solute Transport Behaviour Using Spatial Structures of the Borden Data-Set	61
4.3.1	Results of Solute Transport Behaviour in Two Dimensions	62
4.3.1.1	Test of Significance	67
4.3.1.2	Tracer Tests Based on Dependence Modelled by Maximum Gauss Copula	69
4.3.2	Results of Solute Transport Behaviour in Three Dimensions	70
4.4	Analysis of Solute Transport Behaviour Using Spatial Structures of the North Bay Data-Set	74
4.5	Structure Within the Borden Data-Set	75
4.5.1	Geostatistical Analysis of the Homogeneity of the Borden Data-Set	76
4.5.2	Effects of Structure in the Data	83
5	Conclusions and Outlook	93
	Bibliography	i

List of Figures

1.1	View on an artificial outcrop cut into the Borden aquifer	1
2.1	Spatial interpolation	5
2.2	Comparison of Ordinary Kriging and v -copula based interpolation and associated uncertainty estimates	11
2.3	Schematic sketch of processes that can be captured by spatial copulas	17
2.4	Empirical bivariate copula densities explained	17
2.5	Comparison between rank correlation and semivariogram	18
2.6	Sketch to explain a v -transformation	21
2.7	Distribution of extremes of two Gaussian processes	22
2.8	Some intermediate results of an analytical maximum Gauss copula	24
2.9	Rank correlations transformation	26
2.10	Differences in spatial dependence between Gaussian and v -copula models	30
3.1	Three-dimensional setup of measurement locations at Borden	32
3.2	Measured K-values at Borden contoured	33
3.3	Setup of the coring locations at North Bay	34
3.4	Measured K-values at North Bay contoured	34
3.5	Histograms of K-values at Borden and at North Bay	35
3.6	Marginal distributions of K-values at Borden and at North Bay	37
3.7	Measures of spatial dependence for K-values at Borden	39
3.8	Empirical bivariate copulas for K-values at Borden	40
3.9	Measures of spatial dependence for K-values at North Bay	41
3.10	Empirical bivariate copula densities for K-values at North Bay	42
3.11	Likelihood surface of v -copula model for Borden	44
3.12	Comparison of Semivariograms and rank correlations based on simulated spatial fields	44
3.13	Theoretical bivariate v -copula density fitted to Borden	50
3.14	Bivariate maximum Gauss copula density functions based on cross section AA' of the Borden data-set.	50
3.15	Bivariate maximum Gauss copula density functions based on the entire Borden data-set.	51
3.16	Bivariate maximum Gauss copula density functions based on cross-section Line 1 of the North Bay data-set.	51
3.17	Bivariate maximum Gauss copula density functions based on the entire North Bay data-set.	51
3.18	Bootstrap test of theoretical copula models	53

3.19	Fit of v-copula model in three dimensions evaluated via likelihood	54
3.20	Comparison of Kriging and v-copula based interpolation and associated uncertainty at Borden	55
3.21	Simulated K fields of Borden	56
4.1	Bivariate dependence with correlation coefficient $r = 0.88$	57
4.2	Sketch of domain of numerical model	60
4.3	Effective K estimated from two-dimensional Monte Carlo simulations (Borden)	63
4.4	Average spatial moments of concentration fields based on different spatially distributed two-dimensional K-fields at Borden	64
4.5	Mean and variance of concentration fields at $t = 300d$ based on spatially distributed K fields with Gaussian dependence	66
4.6	Mean and variance of concentration fields at $t = 300d$ based on spatially distributed K fields with v-copula dependence	66
4.7	Comparison of numerical results with analytical solution by Vomvoris and Gelhar (1990)	67
4.8	Average spatial moments of concentration fields based on different spatially distributed K-fields fitted to the Borden data-set	69
4.9	Average spatial moments of concentration fields based on different spatially distributed three-dimensional K-fields at Borden	71
4.10	Spatially distributed K fields with different marginal distributions	72
4.11	Effective K estimated from three-dimensional Monte Carlo simulations (Borden)	73
4.12	Ensemble of distribution functions using two types of resampling on the North Bay data-set	75
4.13	Cumulative average of K values along cores at Borden and picks of boundary layer elevations	77
4.14	Arithmetic means and standard deviations of K per core, separated for sections above and below the boundary layer.	78
4.15	Bootstrap test of standard deviation above and below boundary layer	79
4.16	Sketch illustrating how to calculate standard deviations across boundary layer	80
4.17	Distributions of standard deviations across boundary layer	80
4.18	Semivariogram of deviation from boundary plane	81
4.19	Likelihood surface of v-copula model for section below boundary layer	82
4.20	Significance of boundary layer	83
4.21	Four spatially distributed K-fields fitted to Borden (boundary layer)	85
4.22	Spatial ensemble rank correlations	86
4.23	Domain for numerical model including the boundary layer	87
4.24	Ensemble means and standard deviation of concentration based on Gaussian K fields	89
4.25	Ensemble means and standard deviation of concentration based on v-copula K fields	90
4.26	Ensemble averages breakthrough curves evaluated at slices	91
4.27	Ensembles of breakthrough curves for four scenarios evaluating the structure of the Borden data-set	92

List of Tables

3.1	Descriptive statistics of K and $\ln(K)$ of the Borden and of the North Bay data-sets.	35
3.2	V-copula dependence models fitted to Borden	45
3.3	List of transformation parameters for different scenarios using maximum Gauss copulas	46
3.4	Results of parameter optimization using a maximum Gauss copula based on the Borden data-set.	47
3.5	Akaike criteria for maximum Gauss copula scenarios based on the Borden data-set.	48
3.6	Optimized parameters for the maximum Gauss copula model for the K data-set at North Bay.	49
4.1	Parameters of the numerical model	59
4.2	Estimated dispersivities from two-dimensional Monte Carlo simulations (Borden)	63
4.3	Results of Wilcoxon rank-sum test to evaluate if dispersivities based on different types of spatial dependence are significantly different	68
4.4	Estimated dispersivities from two-dimensional Monte Carlo simulations based on maximum Gauss copulas (Borden)	70
4.5	Estimated dispersivities from three-dimensional Monte Carlo simulations (Borden)	72
4.6	Estimated dispersivities from two-dimensional Monte Carlo simulations based on maximum Gauss copulas (North Bay)	74
4.7	Distributions of dispersivities in resampling ensembles	75
4.8	Descriptive statistics of K-measurements above and below the boundary layer.	78
4.9	Parameters of the numerical model used to evaluate the effect of the boundary layer.	87

Notation

The following table shows the significant symbols used in this work. Local notations are explained in the text.

Symbol	Definition
a	geostatistical range
C	copula
c	copula density
F	distribution function
f	density function
h	(separation) distance
K	saturated hydraulic conductivity
K	Kernel for non-parametric marginal distribution
m, k	transformation parameters for v-Copula
P	probability
conc	concentration
cov	covariance
edf	empirical distribution function
\mathbf{x}	geographic location
$\mathbf{\Gamma}$	covariance matrix
\mathbf{I}	identity matrix
\mathbf{D}	dispersion tensor
I	Indicator
α	dispersivity
β	cut-off value (indicator Kriging)
γ	semivariance
η	porosity
ρ	Spearman's rank correlation coefficient
σ^2	variance
σ	standard deviation
τ	Kendall's rank correlation coefficient
Φ	standard normal distribution function
ϕ	standard normal density function
ln	logarithm to the base of e
log	logarithm to the base of 10
$E[. . .]$	expected value

Kurzfassung

Diese Arbeit beschäftigt sich mit der Beschreibung, Modellierung und Simulation räumlich verteilter Felder mit Copulas basierend auf gemessenen Daten hydraulischer Leitfähigkeit. Copulas sind mehrdimensionale Verteilungsfunktionen mit gleichverteilten Randverteilungen, die es erlauben den räumlichen Zusammenhang unabhängig von der Randverteilung, und somit einen unverfälschten Zusammenhang, zu beschreiben. Wird der räumliche Zusammenhang von in der Natur gemessenen Daten mit Hilfe von Copulas analysiert, werden Strukturen erkennbar, die nicht mit herkömmlichen Maßen (z.B. Semivariogramm) für räumliche Abhängigkeit beschrieben werden können. Theoretische Copula Modelle die nicht auf Gauß-Annahmen beruhen, insbesondere v-Copulas und maximum Gauß Copulas, können solche beobachteten nicht-Gauß Strukturen abbilden. Die Parameter dieser räumlichen Copula Modelle wurden an reale Daten angepasst und zur räumlichen Interpolation und Simulation verwendet.

Es wurde analysiert und quantifiziert, welchen Unterschied räumlich verteilte Felder mit Gauß-Zusammenhang und mit nicht-Gauß-Zusammenhang auf das Ausbreitungsverhalten von gelösten Stoffen im Grundwasser haben. Dabei haben alle Felder die gleichen Rangkorrelationen, sind also in ihren räumlichen Momenten zweiter Ordnung nicht unterscheidbar. Allerdings unterscheiden sie sich in ihren räumlichen Copulas. Das Ausbreitungsverhalten wurde in numerischen Tracertests analysiert, wobei alle Randbedingungen und Anfangsbedingungen eines stationären Grundwasser-Strömungsmodells und eines instationären Modells für den Transport von gelösten Stoffen gleich blieben, bis auf den Eingangsparameter der hydraulischen Leitfähigkeit.

Geostatistik

Die meisten geostatistischen Modelle, die heutzutage verwendet werden, beruhen auf mehreren Annahmen, die einige Nachteile mit sich führen: Der räumliche Zusammenhang muss normalverteilt sein, ein solches Modell ist anfällig auf Extremwerte in der Randverteilung und das Modell beruht lediglich auf Durchschnittswerten des Zusammenhanges in verschiedenen Abstandsklassen. Ein unterschiedlich starker Zusammenhang für verschiedene Quantile hydraulischer Leitfähigkeit und für verschiedene Abstandsklassen existiert in gemessenen Daten. Copulas können solch ein Verhalten abbilden und auch Felder mit entsprechenden Eigenschaften simulieren. Das ist besonders wichtig für sekundäre nicht-lineare Modelle, wie Schadstoffausbreitungsmodelle, die auf räumlich verteilten Feldern als Input beruhen. Das Maß der Symmetrie gibt an, in welchen Quantilen der stärkste Zusammenhang in der räumlichen Struktur besteht. Je positiver das Symmetrie-Maß ist, desto stärker ist der Zusammenhang in hohen Quantilen und dementsprechend bilden hohe

Werte hydraulischer Leitfähigkeit isolierte Gebiete in den räumlichen Feldern. Je negativer das Symmetrie-Maß ist, desto stärker ist der Zusammenhang in niedrigen Quantilen und dementsprechend bilden niedrige Werte hydraulischer Leitfähigkeit isolierte Gebiete in räumlichen Feldern.

Im ersten Kapitel werden Modelle räumlicher Abhängigkeit beschrieben, die auf Copulas basieren. "Gauß Copulas", "v-Copulas", und "maximum Gauß Copulas" sind geeignete Modelle um räumliche Abhängigkeiten zu beschreiben. Für die maximum Gauß Copula wurde eine analytische Formel gezeigt. Neben der Kovarianzfunktion und dem Semivariogramm, zwei Maße räumlicher Abhängigkeit beruhend auf Gauß-Annahmen, kann die räumliche Struktur mit Hilfe der Rangkorrelation und der Symmetrie beschrieben werden. Weiterhin wird beschrieben, wie die Parameter dieser theoretischen Modelle angepasst werden können und die Modelle für Interpolation und Simulation verwendet werden können.

Angewandte Geostatistik

Sowohl traditionelle Geostatistik die auf Gauß-Annahmen beruht, als auch Copula-basierte Geostatistik wurde auf zwei Datensätze hydraulischer Leitfähigkeit angewandt. Die beiden Datensätze umfassen jeweils ~ 1500 Messwerte, die in Permeameterversuchen aus Proben aus Bohrkernen bestimmt wurden. Die vertikale Auflösung beträgt 5 cm, die horizontale Auflösung 1 m bis 4 m. Der eine Datensatz stammt aus einem relativ homogenen Aquifer nahe der Stadt Borden, Ontario (Sudicky, 1986), der andere Datensatz aus einem Aquifer nahe der Stadt North Bay, Ontario (Sudicky et al., 2010).

Beide Datensätze unterscheiden sich in der Form ihrer Randverteilung, die im Fall von Borden relativ symmetrisch ist, im Fall von North Bay rechtsschief ist. Auch in der räumlichen Struktur unterscheiden sich beide Datensätze. Borden weist für kurze Entfernungen positive Symmetrie auf, also starken Zusammenhang in hohen Quantilen für kurze Abstandsklassen. Für größere Abstandsklassen schwächt sich der Zusammenhang ab. Der Datensatz von North Bay hingegen weist für kurze Abstandsklassen stärkeren Zusammenhang in kleinen Quantilen auf, der sich für größere Abstandsklassen umkehrt in starken Zusammenhang in hohen Quantilen.

Die Parameter der Copula-Modelle wurden an die Datensätze angepasst: die Parameter der Gauß-Copula, der v-Copula und der maximum Gauß Copula an den Borden Datensatz; die Parameter der Gauß Copula und der maximum Gauß Copula an den North Bay Datensatz. Es wurde in einem auf dem Bootstrap-Verfahren basierenden Test gezeigt, dass die räumliche Struktur beider Datensätze nicht mit einem Gauß-Modell auf einem Signifikanzniveau von 90% modelliert werden kann. Die verwendeten nicht-Gauß-Modelle bilden einen größeren Teil der räumlichen Struktur ab als die Gauß Copula.

Die angepassten Copula-Modelle können für räumliche Interpolation und Simulation verwendet werden. In der Interpolation spiegelt sich die Struktur der Copula vor allem in dem Unsicherheitsmaß wieder. An jedem Interpolationspunkt gibt eine Copula eine vollständige Verteilungsfunktion an, bedingt auf die Beobachtungen in der Umgebung. Mit dieser möglicherweise schiefen Verteilungsfunktion können beliebige Konfidenzintervalle angegeben

werden, die die Unsicherheit der Interpolation angeben. Diese Konfidenzintervalle hängen nicht nur von der Geometrie des Messnetzes ab, sondern auch von den gemessenen Werten. Die Unsicherheit ist am größten in isolierten kleinen Gebieten. Dagegen ist die Unsicherheit der Interpolation innerhalb von großen homogenen Gebieten klein.

Mit Copula-Modellen simulierte Felder spiegeln die Struktur der Copula wieder, d.h. es können Felder simuliert werden, in denen vorgegebene Bereiche von Messwerten einen vorgegebenen Zusammenhang haben. Das erlaubt z.B. verbundene Gebiete hoher hydraulischer Leitfähigkeit in simulierten räumlich verteilten Feldern. Räumliche Felder mit solchen Eigenschaften sind für sekundäre nicht-lineare Modelle von großer Wichtigkeit.

Effekte von räumlich verteilten Feldern mit nicht-Gauß Zusammenhang auf das Transportverhalten von gelösten Stoffen im Untergrund

Ziel dieses Kapitels war es herauszufinden, wie die räumliche Struktur, modelliert mit Gauß- oder mit nicht-Gauß-Zusammenhang, physikalische Eigenschaften wie das Transportverhalten gelöster Stoffe im Untergrund beeinflusst. Wichtig dabei ist, dass die räumlichen Felder so generiert wurden, dass sie identisch in ihrer Randverteilung und ihren zweiten räumlichen Momenten sind – also nicht unterscheidbar mit herkömmlicher auf Gauß-Annahmen beruhender Geostatistik. Die räumlichen Felder unterscheiden sich in ihren Copulas.

Um das Transportverhalten zu analysieren, wurden viele numerische Tracertests im Monte Carlo Verfahren durchgeführt. In diesem stochastischen Konzept wird die Makrodispersion als Eigenschaft gesehen, die mit der räumlichen Heterogenität der hydraulischen Leitfähigkeit in Verbindung steht. Dabei ist das numerische Modell eine realistische Repräsentation eines Aquifers, in den ein konservativer Tracer in sehr kurzer Zeit injiziert wird, der sich in einem stationären Grundwasser-Strömungsmodell ausbreitet, wobei der Transport des Tracers in einem instationären Transportmodell gelöster Stoffe berechnet wird. Die sich daraus ergebenden Konzentrationsfelder werden mit räumlichen Momenten analysiert. Die Steigung des linearen Teils des zweiten räumlichen Moments ist proportional zum Dispersionskoeffizienten. Je größer der Dispersionskoeffizient, berechnet aus dem mittleren Verhalten über das Ensemble der Monte Carlo Simulationen, desto mehr breitet sich der Tracer relativ zum Massenschwerpunkt seines Konzentrationsfeldes aus.

Analyse des Borden Datensatzes

Der Dispersionskoeffizient basierend auf isotropen Feldern ist signifikant kleiner, wenn die zugrundeliegenden Felder hydraulischer Leitfähigkeit die an den Datensatz angepasste v-Copula Strukturen besitzen, als wenn sie eine Gauß-Copula Struktur besitzen. Der gleiche Vergleich im dreidimensionalen Raum ergibt nur sehr kleine Dispersivitäts-Unterschiede. Anisotropie verbunden mit mehr verbundenen Fließwegen führt dazu, dass die Ausbreitung nur unwesentlich durch die Struktur und die Randverteilung von Borden beeinflusst wird. Wenn man jedoch die Randverteilung nur leicht ändert, indem man die Varianz leicht

erhöht, den Mittelwert jedoch beibehält, ergeben sich auch im dreidimensionalen Raum signifikante Unterschiede im Ausbreitungsverhalten.

Die kleinere Dispersivität, die sich ergibt, wenn die zugrundeliegenden Felder v-Copula Strukturen aufweisen, lässt sich so erklären: Die auf Borden basierende v-Copula weist eine positive Symmetrie auf, d.h. hohe hydraulische Leitfähigkeiten bilden kleinere isolierte Zonen. Verbundene Zonen ergeben sich für mittlere bis niedrige hydraulische Leitfähigkeiten. Es ergibt sich eine kleinere Ausbreitung der Schadstofffahne, was sich in einer kleineren Dispersivität widerspiegelt. Gleichzeitig ergibt sich basierend auf nicht-Gauß Feldern mit positiver Symmetrie eine höhere Varianz der Konzentration stromaufwärts vom Massenschwerpunkt als stromabwärts vom Massenschwerpunkt.

Es wird angenommen, dass nicht-Gauß Zusammenhänge in der räumlichen Struktur hydraulischer Leitfähigkeit und eine Randverteilung mit größerer Schiefe einen signifikanten Einfluss auf das Ausbreitungsverhalten in den anderen Aquiferen haben wird, die weitaus heterogener sind als der Borden Aquifer.

In einer Zone im Borden Aquifer wurden relativ hohe hydraulische Leitfähigkeiten gemessen, die im Mittel signifikant höher als im restlichen Aquifer sind. Der Datensatz wurde auf eine mögliche Struktur im Datensatz – im Gegensatz zu räumlicher Struktur innerhalb eines homogenen Datensatzes – untersucht. Es wurde eine Methode entwickelt, wie solch eine Struktur in den Daten in ein geostatistisches Modell für den gesamten Aquifer mit einbezogen werden kann. Die Berücksichtigung einer solchen Struktur kann das Ausbreitungsverhalten, gemessen zum Beispiel an der Ankunftszeit der maximalen Konzentration in einer Durchbruchkurve, mehr beeinflussen als die räumliche Struktur.

Analyse des North Bay Datensatzes

Der Aquifer bei North Bay gilt als heterogener als der Aquifer bei Borden, bedingt durch seine rechtsschiefe Randverteilung und durch den oben beschriebenen Vorzeichenwechsel im Symmetriemaß für ansteigende Abstandsklassen. Die maximum-Gauß Copula gliedert im Schnitt die negative und positive Symmetrie aus, was zu einer Struktur führte, die nicht wesentlich von einer Gauß-Struktur abwich. Durch die schiefe Randverteilung ergaben sich trotzdem Unterschiede in Dispersionskoeffizienten zwischen Feldern hydraulischer Leitfähigkeit basierend auf Gauß Copula und maximum Gauß Copula.

Anhand dieses Datensatzes wurde untersucht, welchen Einfluss neben der räumlichen Struktur die Unsicherheit in der Randverteilung auf das Ausbreitungsverhalten gelöster Stoffe hat. Aus den Messdaten wurden verschieden große Stichproben gezogen, um in jeder Monte Carlo Simulation eine andere Randverteilung zu verwenden. Dabei fiel zum einen auf, dass egal wie die Stichproben gezogen wurden, der Dispersionskoeffizient größer war, wenn die räumliche Struktur einen nicht-Gauß-Zusammenhang aufwies. Zum anderen wurde auch ein Unterschied in der Schiefe der Verteilung der Dispersivitäten festgestellt, je nachdem, ob die räumliche Struktur einen Gauß- oder einen nicht-Gauß-Zusammenhang aufwies. Das bedeutet, dass der Einfluss der Randverteilung anders ist, je nachdem welche räumliche Struktur zugrunde liegt.

Diese Arbeit zeigt, dass räumliche Momente zweiter Ordnung zur Beschreibung der räumlichen Struktur hydraulischer Leitfähigkeit nicht immer ausreichend sind, um das Ausbreitungsverhalten von gelösten Stoffen im Untergrund zu modellieren. Gerade bei einer geringen Schiefe der Randverteilung hat die räumliche Struktur einen grossen Einfluss auf das Ausbreitungsverhalten. Copulas haben sich zur Beschreibung und Modellierung von nicht-Gauß Strukturen als sehr hilfreich erwiesen.

Abstract

Copulas are a novel tool in geostatistics that allows modelling of pure spatial dependence independently of the marginal distribution and without an assumption of multivariate Gaussian dependence. By using a transformation via the marginal distribution, the effect of extreme values is substantially decreased compared to traditional Gaussian based geostatistical measures such as Kriging. Additionally, the dependence is not described as an average variance as in Kriging, but a different degree of dependence can be modelled for different quantiles of the marginal distribution.

Two data-sets from field sites at Borden and North Bay, both in Ontario, Canada, were used to test the performance of copulas as stochastic models for spatial dependence. Furthermore, this thesis explores possible effects of modelling spatial dependence using non-Gaussian copulas on physical properties that are based on such heterogeneous fields. For comparison, the effects of Gaussian structures are evaluated. The Gaussian- and non-Gaussian structures can not be distinguished by their variograms.

It was shown that neither of the two data-sets exhibits Gaussian dependence – despite the fact that the Borden aquifer is commonly thought of as a relatively homogeneous porous medium with a small variance of hydraulic conductivity. Two non-Gaussian copula models, v -copulas and maximum Gaussian copulas were fitted to the hydraulic conductivity data, to be compared with a Gaussian copula model. The theoretical copula models were subsequently used for spatial interpolation and simulation.

In addition to evaluating the spatial dependence structure of the hydraulic conductivity data-sets, fitting theoretical copula models and using them for interpolation and simulation, the goal of this thesis is to explore if the structure of the hydraulic conductivity field influences a physical property, such as plume evolution as evaluated by second central moments of concentration fields.

Despite the fact that Borden is a relatively homogeneous porous medium, and despite the fact that both types of spatial fields are not distinguishable by their variograms, the solute transport characteristics based on these two types of fields differ significantly in two-dimensional settings. The difference is less pronounced in three-dimensions. Non-Gaussian dependence can lead to a non-symmetric distribution of variance of concentration along the main direction of flow. Increasing the variance of a marginal distribution by a certain factor does not necessarily lead to a dispersivity increased by the same factor in the case of non-Gaussian fields. It is postulated that non-Gaussian spatial dependence of hydraulic conductivity and a more skewed marginal distribution of hydraulic conductivity will have significant implications in the other more heterogeneous aquifers.

1 Introduction

Heterogeneity exists in naturally occurring spatially distributed parameters, which vary in space, often in complicated ways. Heterogeneity exists in all geologic formations, even in formations that are regarded as a fairly homogeneous aquifers such as the Borden aquifer [Sudicky \(1986\)](#), whose hydraulic conductivity data-set is used in this thesis. Even this aquifer exhibits variations of sediment properties in space, such as different grain sizes, which become obvious when looking at the faces of a man-made ditch through the aquifer (Figure 1.1).

Heterogeneity must be treated adequately in numerical hydrogeologic models. Classical hydrogeological models were deterministic models. The realization that basic assumptions of deterministic models are not similar to what is encountered in practice has led to the development of stochastic hydrogeology as a discipline ([Freeze, 1975](#)). A deterministic approach of hydrogeology assumes (1) a full knowledge and understanding of the processes that occur within an investigated domain. Such processes can be a combination of physical,



Figure 1.1: View on an artificial outcrop cut into the Borden aquifer. Photo was taken by Dave Thomson.

chemical, biological or other processes. Even different processes can act at different parts of the domain or at different scales. (2) a mathematical model that correctly represents these processes and can be used to predict future responses to imposed excitations, and (3) the availability of information (measured or derived) on all parameters and coefficients that appear in the mathematical model, and on the domain's geometry, as well as on the initial and boundary conditions.

Stochastic hydrogeology's objective is to deal with uncertainty in groundwater modelling (Delhomme, 1979). One building block of all the uncertainties involved in the modelling process is parameter uncertainty. Geological and hydrological processes such as sedimentation or the climate during sedimentation effect naturally occurring parameters such as hydraulic conductivity (K). Unfortunately, these processes are not fully understood, and likely details of the processes cannot be unraveled. Most of the uncertainty associated with modelling may be attributed to the heterogeneity of subsurface domains. Hence, dealing with uncertainty related to heterogeneity is important, and ignoring heterogeneity might lead to large errors of prediction. Only partial incorporation of heterogeneity or incorporation of a wrong kind of heterogeneity might lead to substantial model errors. In contrast to deterministic approaches, stochastic approaches attempt to gain useful information about the behaviour of natural heterogeneous systems by treating them as if the hydrological parameters were random.

Heterogeneity in the hydraulic conductivity of geologic strata leads to a heterogeneous velocity field which in turn influences solute transport behaviour. Solute transport has played an important role for the development of hydrogeology in general. The recognition of the importance of heterogeneity and solute transport has led to many approaches of incorporating heterogeneous structures into models of solute transport (among many others: Dagan (1982, 1984, 1988, 1990); Gelhar and Axness (1983); Neuman and Zhang (1990); Rubin (1990); Burr et al. (1994)).

In all these approaches, the spatial variability of hydraulic conductivity has been modelled with a Gaussian dependence structure. Among others, Gómez-Hernández and Wen (1998) pointed out the importance of non-Gaussian spatial structures for evaluating the above mentioned effects of spatial variability, in their case on groundwater travel times. Approaches exist for modelling non-Gaussian structures, for example via training images (Strebelle, 2002) or transition probabilities (Carle and Fogg, 1997). Fiori et al. (2007) showed with artificial setups in semi-analytical approaches that for data-sets with a variance of the natural logarithms of K $\sigma^2(\ln(K)) > 0.5$ transport departs from Gaussianity and plumes become skewed. The main reason being tailing due to zones of low hydraulic conductivity. Zinn and Harvey (2003) were able to model one type of non-Gaussian spatial dependence using a Chi-Square transformation, not a full copula model. This paper presents an approach where the dependence model, a multidimensional spatial copula, is fitted to the spatial dependence structure of real-world data.

A statistician can never have enough good data. Particularly, in hydrogeology there is never enough good data, because it is difficult and expensive to collect. There are initiatives to work on novel tools that increase the possibility of having more and better data (Liu et al., 2009; Dietrich et al., 2008). However, such methods should allow to account for data mea-

sured at different scales, which copulas can. The data-sets used in this thesis are two of the most detailed data-sets of hydraulic conductivity available (Sudicky, 1986; Sudicky et al., 2010).

Essentially, this thesis tries to answer still the same questions that Allan Freeze posed in one of the earliest works of stochastic hydrogeology (Freeze, 1975), even though Freeze did not model the spatial dependence structure, but instead treated heterogeneity as purely random. The two questions are (1) how do the properties of the statistical distributions of the output variables arising from a deterministic model change when input parameters were specified as statistical distributions taken from known populations, and (2) what additional uncertainties or reductions in confidence are introduced into the solutions by the fact that input population distributions are not known but must be estimated from the (usually small) available statistical sample? Important novelties of this work are that nowhere in this analysis are assumptions made regarding a Gaussian distribution of any parameter. The marginal distribution is fitted to data and the spatial dependence structure is optimized with no prerequisite of a multivariate Gaussian model.

The goal of this thesis is to explore and quantify the effects of heterogeneity modelled by different spatial dependence structures defined by different copula models, such as Gaussian copulas, v -copulas, or maximum Gauss copulas. All copula models are fitted to real-world data. The heterogeneous parameter chosen is saturated hydraulic conductivity on a macroscopic scale where Darcy's law applies, that is bigger than pore scale and smaller than geologic scale. The effects on solute transport behaviour are analyzed using numerical tracer tests, where a slug of conservative tracer is injected instantaneously and subsequently is transported through a steady state groundwater flow field. The shape of the evolving solute plume is affected by the heterogeneous groundwater velocity field which in turn is dependent on the heterogeneous field of hydraulic conductivity.

Chapter 2 gives an overview of the state of the art of geostatistics, including a discussion of its limitations. As an alternative, the basics of spatial copulas as a stochastic model are introduced, including the theoretical model, parameter estimation, as well as interpolation and simulation. The data used in this thesis is explained and analyzed in Chapter 3, according to the methods established previously. At the end of this chapter, spatial fields of hydraulic conductivity can be simulated with different spatial dependence structures, which are used as input for numerical tracer tests. The numerical tracer tests form the key component of Chapter 4, where the effects of different types of spatial dependence are analyzed and quantified. The thesis closes with conclusions and an outlook in Chapter 5.

2 Geostatistics

Geostatistics is the part of statistics that focuses on the variability of quantities in space with the goal to model heterogeneity. The biggest problem with spatial variability is that usually there is only one realization of the quantity under investigation, the one nature created. This means that there are no frequencies available that can describe spatial variability. There need to be some constructs in place that circumnavigate this problem. Typically, this is done by employing the concept of random functions. [de Marsily \(1986\)](#) visually describes this concept:

Suppose that we create in the laboratory several sand columns, each filled with the same type of sand. Each column represents the same porous medium, but is somehow different from the others. Each column is a “realization” of the same porous medium, defined as the ensemble of all possible realizations (infinite in number) of the same process

Additionally, when dealing with quantities in the subsurface, such as hydraulic conductivity, the one realization is not even visible, and measurements are difficult and/or expensive. Still, methods do exist to describe and model spatial variability. This chapter describes the most commonly used approaches, describes their limitations, and introduces the use of copulas for geostatistics.

2.1 State of the Art

Spatial variability is generally not completely random, there is some structure. This observation is one basic concept of geostatistics: spatial autocorrelation. Parameters measured in close vicinity are likely more similar than when they were measured at locations that are separated by a large distance ([Cliff and Ord, 1973](#)). Measurements in close vicinity are somehow more dependent than measurements separated by a large distance. This property has to be met by all models that try to describe spatial dependence, including spatial covariances and spatial copulas.

The concept of spatial autocorrelation also holds in a common task for anybody who works with spatially distributed data. This task of spatial interpolation is to estimate continuous functions based on observations made at discrete points in the vicinity. Continuous functions are necessary to estimate a parameter value at any desired location within the given domain. [Figure 2.1](#) illustrates this problem, where a parameter value (the value of the colour red representing concentration) is to be estimated at the location \hat{x} based on measurements at six locations x_i in the vicinity.

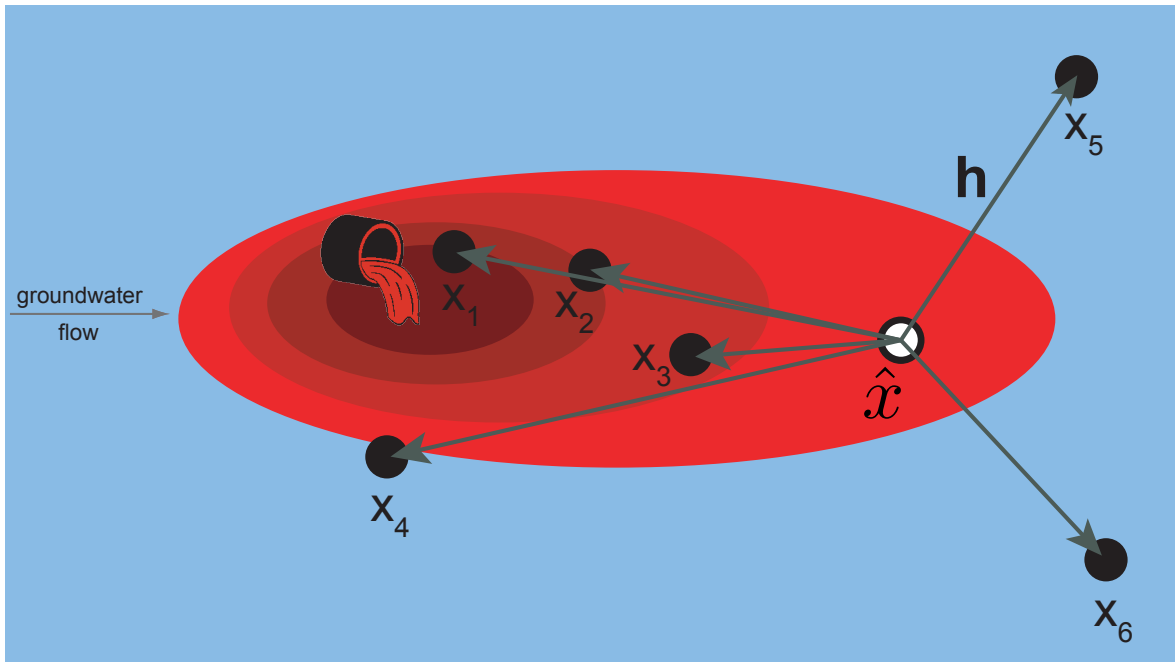


Figure 2.1: Sketch to explain interpolation. A parameter value is to be estimated at the location \hat{x} based on the n , in this example $n = 6$, surrounding locations x_n . The separation vectors \mathbf{h} between \hat{x} and the surrounding locations are indicated with arrows. The background concentration is indicated by the colour blue; whereas, higher concentration values, resulting from a point spill, are indicated by shades of red, the lighter the shading, the lighter the contamination

The simplest approach to interpolation is to implement deterministic models, which assume that the processes that generated the parameter-distribution are known in sufficient detail so that an accurate description of the entire parameter-distribution is known in sufficient detail so that an accurate description of the entire parameter-distribution can be made from only a few sample values (Isaaks and Srivastava, 1989). Examples for the methods used in deterministic models are inverse distance weighting or polynomial interpolation. Such interpolation schemes give “ready made” data weightings based on distances to known parameters at surrounding measurement locations, regardless of the physical aspects of the problem (Delhomme, 1978).

Probabilistic methods to model spatial dependence were developed while looking for ways to quantify the uncertainty associated with interpolation and for a possibility to generate multiple realizations of spatial fields whose mean spatial field has the same statistics as the data it is based on. Multiple realizations in spatial statistics are needed to evaluate uncertainties, because usually only one realization exists, the one that exists in nature.

Statistical methods for spatial interpolation spearheaded by Krige (1951) for mining. Consequently, the geostatistical interpolation method “Kriging” is named for Krige. The goal of Kriging is to find the set of weighting coefficients for a given parameter-set and parameter-configuration that minimizes the estimation variance. The method for finding the weighting

coefficients in such a way is commonly referred to as a best linear unbiased estimator (BLUE) (Brooker, 1979). BLUE is *linear* because its estimates are weighted linear combinations (least-squares regression algorithms) of the geometry of the locations from measured parameters (Goovaerts, 2001); it is *unbiased* because it tries to minimize the mean residual error (aim for zero mean error), and it is *best*, because it minimizes the variance of the estimation-errors (Isaaks and Srivastava, 1989). The last characteristic, the minimized error variance, is a key advantage of Kriging, compared to previously used methodologies. Using Kriging, it is possible to calculate a measure of the error associated with each parameter estimate. Unfortunately, that error can be characterized by a mean (zero) and a variance, which depends largely on the geometry of the observation network.

Matheron (1963, 1971) later significantly advanced and generalized Kriging methods and built the geostatistical framework of regionalized variables. Matheron (1963) described a phenomenon which is spread out in space and which exhibits a certain structure as “regionalized” and he called variables that describe such a phenomenon “regionalized variables” (ReV). The theory of ReVs fulfills two purposes. The theoretical aspects of ReVs express the structural properties of natural phenomenon in adequate mathematical form. The ReV theory also handles the practical problem of estimating a ReV from fragmentary sampling data. It is convenient to deal with ReVs by using the probabilistic theory of random functions (RFs). A ReV denoted by $z(x)$ is thus interpreted as a realization of a RF denoted by $Z(x)$. Random Variables (RVs), the outcomes of RFs, are described by their distribution function F_Z :

$$F_Z(z) = P(Z \leq z) \quad (2.1)$$

In a spatial context, a RF could theoretically be described by its multidimensional distribution functions by assigning the cumulative distribution function, F_{x_1, \dots, x_n} , at each set of points, x_1, \dots, x_n , in a chosen sample population, A . These functions could be evaluated for each set of possible values, w_1, \dots, w_n , making it possible to find the joint probability P that could be used for the estimation of global averages:

$$F_{x_1, \dots, x_n}(w_1, \dots, w_n) = P(Z(x_1) < w_1, \dots, Z(x_n) < w_n). \quad (2.2)$$

In practice, there is at most one measurement available at each point in A . One single realization is not enough to estimate the above given joint distribution function (Equation 2.2). Therefore, it is necessary to obtain the essential replication for statistical evaluations by different means, commonly by the assumption of strong stationarity.

Strong stationarity implies that the multi-variate probability law does not depend on the location x but only on the separation vector \mathbf{h} and hence it does depend on the geometric configuration of the points where the parameter in question was sampled. (“translation invariance”). All pairs of RVs separated by a particular \mathbf{h} have the same joint probability distribution function:

$$P(Z(x_1) < w_1, \dots, Z(x_n) < w_n) = P(Z(x_1 + \mathbf{h}) < w_1, \dots, Z(x_n + \mathbf{h}) < w_n). \quad (2.3)$$

The assumption of strong stationarity is further simplified to the assumption of second order stationarity where the expected value m is constant everywhere within the considered do-

main (Equation 2.5) and the covariance (Cov) of two RVs is dependent on \mathbf{h} (Equation 2.5).

$$E[Z(x)] = m \quad (2.4)$$

$$Cov(\mathbf{h}) = E[(Z(x + \mathbf{h}) - m) \cdot (Z(x) - m)] \quad (2.5)$$

The assumption of second order stationarity is even further relaxed to the intrinsic hypothesis, where the variance (not the covariance) of the increment corresponding to two points in A depends only on the vector separating them. This results in the semivariogram $\gamma(\mathbf{h})$, a measure for spatial variability (Equation 2.6).

$$\gamma(\mathbf{h}) = \frac{1}{2} \cdot var[Z(x + \mathbf{h}) - Z(x)] \quad (2.6)$$

The intrinsic hypothesis is commonly used for geostatistical evaluations because it does neither require a finite variance nor the knowledge or setting of the expectation. Additionally, semivariograms and covariance functions can be constructed quite easily. Equations 2.7 and 2.8 show how this is done numerically. Both $\gamma(\mathbf{h})$ and $Cov(\mathbf{h})$ represent averages based on the measurement values for all pairs of points which are separated by a distance of \mathbf{h} .

$$\gamma(\mathbf{h}) = \frac{1}{n(\mathbf{h})} \cdot \sum_{i=1}^{n(\mathbf{h})} [Z(\mathbf{x}_i) - Z(\mathbf{x}_i + \mathbf{h})]^2 \quad (2.7)$$

$$Cov(\mathbf{h}) = \frac{1}{n(\mathbf{h})} \cdot \sum_{i=1}^{n(\mathbf{h})} [(Z(\mathbf{x}_i) - \bar{x}) \cdot (Z(\mathbf{x}_i + \mathbf{h}) - \bar{x})] \quad (2.8)$$

Kriging follows a four-step approach (Journel, 1989):

1. The sample population A is defined, which is homogeneous enough to warrant statistical averaging within it.
2. All available data within A are scanned and the experimental \mathbf{h} characteristics of spatial variability, expressed in the experimental semi-variogram, are calculated.
3. A function, the theoretical semi-variogram $\gamma(\mathbf{h})$, is fitted to the experimental variogram.
4. This function is used for regression type techniques in spatial interpolation.

With a theoretical semivariogram, a property Z can be estimated at an unsampled location \bar{x} (Equation 2.9), by one of several Kriging methods, Ordinary Kriging, f.ex. In current stochastic hydrogeological practice, it is commonly assumed that the natural logarithm of K values is normally distributed, $Z = \ln(K)$.

$$\hat{Z}(\mathbf{x}) - \bar{Z}(\mathbf{x}) = \sum_{i=1}^n w_i(\mathbf{x}) [Z(\mathbf{x}_i) - \bar{Z}(\mathbf{x}_i)] \quad (2.9)$$

The weights w are usually chosen such that variance of the residual gets minimized

$$\sigma_{\hat{Z}}^2(\mathbf{x}) = \overline{\left[\hat{Z}(\mathbf{x}) - Z(\mathbf{x}) \right]^2} \rightarrow \min \quad (2.10)$$

This approach allows for two important things. First, a linear combination of the weights w_n of the n measurements in the vicinity of the location where a parameter value is to be estimated is used for the interpolation, such that the average error is zero and the modelled error variance is minimized. It can be shown that using Ordinary Kriging ($\bar{Z}(\mathbf{x}_i) = \bar{Z}(\mathbf{x})$), the requirement for an unbiased estimate is that the sum of the weights is unity (Equation 2.11). Second, the probability model is used to calculate the estimation variance as a function of n variables, namely the weights w_1, \dots, w_n .

$$\sum_{i=1}^n w_i(\mathbf{x}) = 1 \quad (2.11)$$

Thus the estimate reduces to

$$\hat{Z}_{ok}(\mathbf{x}) = \sum_{i=1}^n w_i(\mathbf{x}) Z(x_i) \quad (2.12)$$

The set of Kriging equations (Equation 2.13) can be solved with the help of Equation 2.11 for the unknown weights w_i and the Lagrange multiplier μ

$$\sum_{j=1}^n w_j(\mathbf{x}) cov_Z(x_i - x_j) + \mu(\mathbf{x}) = cov_Z(x_i - \mathbf{x}); i = 1, \dots, n \quad (2.13)$$

The estimation variance σ_{ok}^2 can be obtained with Equation 2.14

$$\sigma_{ok}^2 = cov_Z(0) + \sum_{j=1}^n \sum_{i=1}^n w_i w_j cov_Z(x_i - x_j) - 2 \sum_{i=1}^n w_i cov_Z(x_i - \mathbf{x}) \quad (2.14)$$

2.2 Critique of the State of the Art

Kriging is a well established geostatistical method, and is advantageous because it minimizes the error variance. Kriging takes into account several factors: (1) the spatial structure of the parameter through the (semi-) variogram, and (2) the distance between the point where the estimation is to take place and the surrounding data-points. But the Kriging-weights do not take into account the values of the parameters in the vicinity. There are additional critical comments related to Kriging that are discussed in the following sections.

2.2.1 Philosophy

In nature, the spatial distribution of a parameter is determined by the processes that created that parameter. An example of such a process is the infiltration and related spreading of a contaminant from a point source into an aquifer. The spatial distribution of the contaminant concentration is not the result of a random event but rather is shaped in a pattern that is induced by the process of the spill. Hence, the spatial dependence structure of the high concentration values relating to the spill is different than the dependence structure of the low background concentrations, which were induced by a different process, perhaps a geologic process.

Average measures of spatial dependence (Equations 2.7 and 2.8) describe the spatial dependence as an integral over the whole distribution of the parameter values within a certain separation distance class. It has been long recognized (Journel and Alabert, 1989) that different percentile values, for example extremes, can have a different spatial dependence structure. A measure for spatial dependence is needed that is able to describe a varying dependence structure that depends on any different processes that were acting on the parameter in various areas within the study site and that depends on the magnitude of the observed values (extremes). Such a measure needs to have non-symmetric characteristics because it needs to have the capability to express different degrees of dependence. A purely Gaussian interpolation such as Kriging can not describe and model a variable dependence structure.

Indicator Kriging was developed to describe a variable dependence structure by looking at different thresholds, which is partly considered in an indicator variogram.

Indicator variables I_β are defined using a cutoff value β in Equation 2.15. The form of an indicator variogram is given in Equation 2.16.

$$I_\beta = \begin{cases} 1 & \text{if } Z(\mathbf{x}) < \beta \\ 0 & \text{else} \end{cases} \quad (2.15)$$

$$\gamma_\beta = \frac{1}{2} \cdot E \left[(I_\beta(\mathbf{x} + \mathbf{h}) - I_\beta(\mathbf{x}))^2 \right] \quad (2.16)$$

The advantage of indicator Kriging is that extreme values and skewed marginal distributions have no major impact on the variogram, leading to a more stable variogram estimation and less influence of extremes on an interpolated field. Indicator Kriging however, does not provide a reasonable uncertainty estimate but is rather an expression of the local heterogeneity. Additionally, Indicator Kriging is not well based on a theoretical foundation. The indicator variograms are fitted for each threshold separately, thus their compatibility is not ensured. This can lead to problems with the monotonicity of the estimation.

2.2.2 Gaussian Assumption

Many parameters occurring in nature, such as hydraulic conductivity, exhibit a skewed distribution rather than a symmetrical distribution. It is not an underlying requirement for the

Kriging-parameter to be normally distributed. However, Kriging only works well if the parameter that is being Kriged is normally distributed. Often the necessity of an underlying Gaussian dependence is compensated for by “forcing” the one-point marginal distribution to fit into a normal or log-normal (parametric) model by transforming the raw data by means of a normal score– or a box cox transform (Goovaerts, 1999; Goovaerts et al., 2005). However, forcing the one-point marginal distribution to fit into a normal or log-normal model is not sufficient; the two-point through the n th-point marginal distributions need to fit into the chosen parametric distribution. Problems arise when even the log-transformed one-point distribution is skewed. In such a case, when the parameter values are back transformed into the original space after Kriging, the expected value can very well be different compared to the mean in the data. On the other hand, if the data is not transformed and left in the original space, it could very well be possible to obtain negative values in the interpolation due to negative Kriging weights. Despite the fact that these issues are known, current practice commonly involves transforming the one-point marginal distributions only to fit into a Gaussian model because applied Kriging facilitates this process.

Using Kriging, the obtained estimation variance is symmetric and can be fully described by mean and variance. Additionally, it depends heavily on the geometry of the observation network and not on the magnitude of the observations. For two large-scale precipitation events over the catchment of the river Neckar, Germany, Figure 2.2 shows both the estimation standard deviation (bottom panel) and the 60% confidence intervals associated with v-copula based interpolation. While the shape of the Ordinary Kriging estimation standard deviation is virtually identical for both events, the copula based confidence interval show high uncertainty in areas where isolated strong (south-east boundary of catchment in 1992) or weak precipitation (northern boundary of the catchment in 1992) is occurring and little uncertainty in areas of homogeneous precipitation (south/central region of the catchment in 1982).

In addition to the problems inherent to the methodology, a physical problem arises when using a Gaussian model for interpolation or simulation purposes. Multi-Gaussian models enforce maximum entropy (Gómez-Hernández and Wen, 1998) resulting in minimum connectivity of extreme values. Extreme values are dominating processes, for example in hydrogeology, zones of low conductivity act as flow and contamination barriers, and zones of high conductivity act as “funnels” for flow and contamination. Such barriers and funnels effect the contaminant breakthrough characteristics at the receptors, typically a water supply well, of a contaminant pathway significantly (Frind et al., 2002). A spatial method that predicts the parameter distribution in space correctly would lead to realistic and conservative estimates in groundwater flow– and transport models.

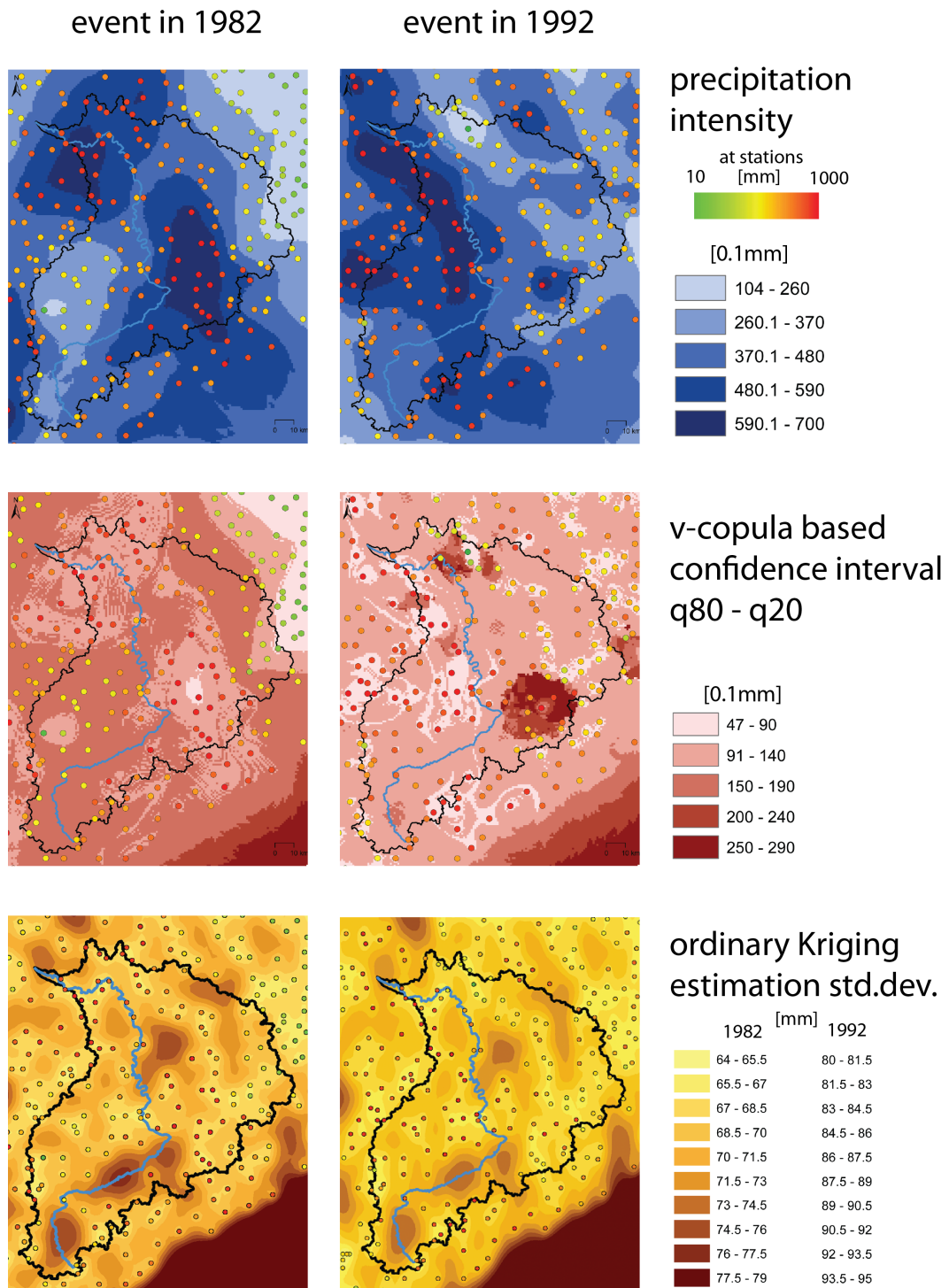


Figure 2.2: Comparison of Ordinary Kriging and v-copula based interpolation and associated uncertainty estimates for two large-scale precipitation events over the catchment of the river Neckar, Germany. The shape of the Kriging based estimation-variance is very similar, while the confidence intervals associated with v-copula based interpolation reflect also the magnitudes and the gradient of the measurement values.

Some approaches exist to model non-Gaussian dependence, most notably multiple point geostatistics (Strebelle, 2002) and approaches based on sedimentological descriptions, using for example transition probabilities (Carle and Fogg, 1997). The problem of both approaches can be illustrated with an example. Say the goal is to simulate the numbers that are drawn by the German lottery, and say we did not know that the possibilities are in the interval $[1, 49]$. If for our model we'd use the boundaries given by the maximum and the minimum of last Saturday's draw, our chances of winning would be fairly small, because very likely we would not cover in our model the entire range of numbers between $[1, 49]$. A training image, or a sedimentological record gives only the equivalent of last Saturday's draw. On the other hand, if we'd use a very large and detailed training image we would get exactly a replicate of last Saturday's draw, while we would like to simulate at least a little outside the range of last Saturday's draw – at least as long as it did not include the numbers 1 and 49.

2.2.3 Other Kriging-Issues

Kriging, essentially a regression method, smoothes out small-scale variability and extreme values (Delhomme, 1978) and assumes that the model of spatial variability (the semi-variogram) is known. However, the small sample sizes of measurements on which the empirical semi-variogram is based and the less than optimal fits of the theoretical semi-variogram can result in errors.

The procedure of fitting a theoretical variogram model to an empirical semivariogram was recently automated for meteorological applications by (Hinterding and Streit, 2002), and can be fitted using a maximum likelihood approach in some special cases (Todini and Pellegrini, 1999; Kitanidis, 1997). In any case, fitting a theoretical to an experimental semivariogram is not an unequivocally clear procedure.

Applications for predictions of parameter-distributions in space started in the late 1970's (Delhomme, 1978; Volpi and Gambolati, 1978; Gambolati and Volpi, 1979). Since then, various case studies for many parameters were conducted and the methodology was advanced. For instance, non-stationary methods (Universal Kriging, External Drift Kriging, Intrinsic Random Functions of order k), interpolations with categorical parameters (Indicator Kriging), and additional soft information was included in the prediction (Co-Kriging with a linear relationship between additional data and Kriged data, Markov-Bayes Kriging with a non-linear relationship).

Most of these approaches do not use a statistical model as the backbone of the method. If they do use a statistical model as a basis for the method, they use only a simple multivariate distribution which is similar to empiricism. The reason why Multi-Gaussian models are commonly used is “*the lack of instruments to handle other multivariate models*” (Gómez-Hernández and Wen, 1998). Spatial copulas are one such instrument.

2.3 Copulas for Spatial Statistics

Copulas are a novel tool in general, and particularly for modelling spatial dependence. Copulas originated in financial mathematics where they are used frequently (Genest et al., 2009). Some people claim that the Gaussian copula has been used unsuccessfully (Salmon, 2009). Recently, copulas were introduced in bi- or higher variate hydrology (Dupuis, 2007; Evin and Favre, 2008; Genest et al., 2007; Grimaldi and Serinaldi, 2006; Salvadori and de Michele, 2007; Serinaldi and Grimaldi, 2007; Zhang and Singh, 2006).

Bárdossy (2006) spearheaded the use of copulas in spatial statistics, followed by Bárdossy and Li (2008). Other uses of copulas in spatial statistics remain limited to date: Kazianka and Pilz (2009), Emery (2002), Marcotte and Gloaguen (2008).

Instead of averaged measures as semivariograms, the structure of spatially distributed parameters can be analyzed and modelled using copulas. Nelsen (1999) defines copulas as “functions that join or couple multivariate distribution functions to their one-dimensional marginal distribution functions” and as “distribution functions whose one-dimensional margins are uniform”.

Using copulas, similar steps as in a classical geostatistical workflow are necessary. First, descriptive copula based geostatistics are performed using the data under investigation (Section 2.3.2). The next steps requires to choose a theoretical model (Section 2.3.3) and fit its parameters to the properties of the data (Section 2.3.4). Finally, this model can be used for interpolation– and simulation purposes.

2.3.1 Basics of Copulas

Copulas are multivariate distributions defined on the n -dimensional hypercube

$$C : [0, 1]^n \rightarrow [0, 1] \quad (2.17)$$

$$C(\mathbf{u}^i) = u_i \text{ if } \mathbf{u}^i = (1, \dots, 1, u_i, 1, \dots, 1) \quad (2.18)$$

For any n -dimensional hypercube within the unit hypercube the corresponding probability has to be non-negative:

$$\sum_{i=0}^{2^n-1} (-1)^{n-\sum_{i=1}^n j_i} C(u_i + j_i \Delta_1, \dots, u_n + j_n \Delta_n) \geq 0$$

if $0 \leq u_i \leq u_i + \Delta_i \leq 1$ and $i = \sum_{k=0}^{n-1} j_k 2^k$

Nelsen (1999) gives an example in two dimensions of what copulas are:

Consider for a moment a pair of random variables X and Y , with distribution functions $F(x) = P[X \leq x]$ and $G(y) = P[Y \leq y]$, respectively, and a joint distribution function $H(x, y) = P[X \leq x, Y \leq y]$. To each pair of real numbers (x, y) we can associate three numbers: $F(x)$, $G(y)$, and $H(x, y)$. Note that each of these numbers lies in the interval $[0, 1]$. In other words, each pair (x, y) of real numbers

leads to a point $(F(x), G(y))$ in the unit square $[0, 1] \times [0, 1]$, and this ordered pair in turn corresponds to a number $H(x, y)$ in $[0, 1]$. [...] This correspondence, which assigns the value of the joint distribution function to each ordered pair of values of the individual distribution functions, is indeed a function. Such functions are copulas.

In the following, key mathematical concepts of copulas are introduced. The interested reader is referred to [Nelsen \(1999\)](#) and [Joe \(1997\)](#) for more detailed mathematical background and to [Bárdossy \(2006\)](#); [Bárdossy and Li \(2008\)](#); [Li \(2010\)](#) for more details regarding copulas in a geostatistical sense.

Any multivariate joint distribution function $F(t_1, \dots, t_n)$ with margins $F_{t_1}(t_1), \dots, F_{t_n}(t_n)$ can be represented with a copula C ([Sklar, 1959](#)):

$$F(t_1, \dots, t_n) = C(F_{t_1}(t_1), \dots, F_{t_n}(t_n)) \quad (2.19)$$

If the margins are continuous, then the C is unique. Conversely, if C is a copula and F and G are distribution functions, then the function $F(t_1, \dots, t_n)$ defined by Equation 2.19 is a multivariate joint distribution function with margins $F_{t_1}(t_1), \dots, F_{t_n}(t_n)$. Two limits of copulas, the Fréchet-Hoeffding bounds W and M can be defined using $u_i = F(x_i)$:

$$W(u_i) \leq C(u_i) \leq M(u_i) \quad (2.20)$$

where

$$M(u_i) = \min(u_i) \quad (2.21)$$

$$W(u_i) = \max(u_1 + \dots + u_n - n + 1) \quad (2.22)$$

M can also be regarded as the copula where the marginal distributions are fully positively dependent. A third important copula is the copula where the marginal distributions are independent, Π .

$$\Pi = \prod_{i=1}^n u_i \quad (2.23)$$

[Schweizer and Wolff \(1981\)](#) were the first to relate copulas with the study of dependence among random variables. They wrote:

... under almost surely increasing transformations of (the random variables), the copula is invariant while the margins may be changed at will, it follows that it is precisely the copula which captures those properties of the joint distribution which are invariant under almost surely strictly increasing transformations. Hence the study of rank statistics – insofar as it is the study of properties invariant under such transformations – may be characterized as the study of copulas and copula-invariant properties.

In a geostatistical sense, this property of copulas brings the advantage that frequently used transformations of the marginal do not influence the dependence structure as modelled by the copula. Frequent discussions among geostatisticians using Gaussian based methods if the marginal is symmetric or if it is symmetric after a transformation become obsolete when using copulas. In a sense, a dependence structure that is independent of its margin can be called “pure” dependence.

Assuming that C is continuous, then the copula density $c(u_1, \dots, u_n)$ can be written as

$$c(u_1, \dots, u_n) = \frac{\partial^n C(u_1, \dots, u_n)}{\partial u_1 \dots \partial u_n}. \quad (2.24)$$

A conditional copula is given by

$$C(u|U_1 = u_1, \dots, U_n = u_n) = \frac{\partial^n C(u, u_1, \dots, u_n)}{\partial u_1, \dots, \partial u_n} \cdot \frac{1}{c(u_1, \dots, u_n)} \quad (2.25)$$

2.3.2 Empirical Geostatistics using Copulas

Empirical bivariate copula densities are an option to visualize the spatial dependence structure of a given data-set. The visualization is only in two dimensions, and the entire spatial dependence structure is in much higher dimensions, still the impression of the spatial dependence structure can be useful. Empirical bivariate copula densities can be evaluated for different directions and different angles between pairs of points. The assessment of the bivariate spatial copulas from measured data $z(x_1), \dots, z(x_n)$ can be done by first calculating the empirical distribution function $F_n(z)$. Using this distribution function for any given vector \mathbf{h} , the set of pairs $S(\mathbf{h})$, consisting of distribution function values corresponding to the parameter at locations \mathbf{x} separated by the vector \mathbf{h} , can be calculated.

$$S(\mathbf{h}) = \{F_n(z(\mathbf{x}_i)), F_n(z(\mathbf{x}_j)) \mid (\mathbf{x}_i - \mathbf{x}_j \approx \mathbf{h}) \text{ or } (\mathbf{x}_j - \mathbf{x}_i \approx \mathbf{h})\} \quad (2.26)$$

$S(\mathbf{h})$ is thus a set of points in the unit square. Note that $S(\mathbf{h})$ is by definition symmetrical regarding the major axis $u_1 = u_2$ of the unit square, namely, if $(u_1, u_2) \in S(\mathbf{h})$, then $(u_2, u_1) \in S(\mathbf{h})$.

Empirical copula densities for pairs of points separated by a $S(\mathbf{h})$ can be plotted and give an insight into the spatial dependence structure. On such plots, points with low measurements are plotted close to the origin, and points where the measured value is high are plotted far from the origin. If the empirical copula density for a certain quantile is high, then there are a lot of pairs of points separated by the given distance which have the corresponding quantile values.

To get an impression of the spatial dependence structure of a data-set with empirical copula densities, either the scatterplots of Equation 2.26 or contours of those scatterplots (details in Li (2010)) could be constructed for different separation distances and different angles of separation.

The advantages that empirical bivariate copula densities offer shall be illustrated by the following example of two spills into an aquifer with negligible background contamination (Figure 2.3). Subfigures 2.3a and 2.3b show pairs of points where the concentration of the water is to be sampled. The pairs of points in both subfigures are separated by different separation distances in each subfigure. On Subfigure 2.3a the separation distance is small enough such that a pair of observation points can lie entirely within the highly concentrated plume area. For such pairs the spatial dependence is high, and hence the empirical bivariate copula density is also high, indicated by the dark red shading in the top right corner corresponding to high quantiles on Subfigure 2.4a. For the small separation distance, there are also pairs which are both within areas of background concentration, or pairs of which one point is within the plume area, the other outside the plume area, which lead to some significant degree of spatial dependence in the medium- and small quantiles, but not as strong – and hence different – as in the high quantiles. An average measure could not model such a typical dependence structure.

If the separation distance between a pair of points is increased beyond the characteristic length of the contamination process (as indicated on Subfigure 2.3b), any significant magnitude of a copula density must originate from two separate plumes. If there is only one plume, then for a separation distance bigger than the characteristic length of this plume, there will be no more strong dependence for high quantiles (as indicated on Subfigure 2.4b). Copulas offer the possibility for such improved process understanding.

Theoretically, copulas should be evaluated at arbitrarily high dimensions. Empirical bivariate copula densities give an useful impression of the spatial dependence, at least in two dimensions. Still the process of sorting through empirical bivariate copula densities is more time consuming than looking at empirical semivariograms. To facilitate this procedure, two measures that summarize information from the bivariate copula densities over multiple separation distances are the copula-based rank correlation (“Rank”, Equation 2.27), and the asymmetry of an empirical bivariate copula density (“Sym”, Equation 2.29).

“Rank” is a measure for the degree of the spatial dependence. For a pair of points with a short separation distance, the rank correlation is expected to be very high, that is close to unity. Its value will decrease to a value close to zero for large separation distances. The copula rank correlation is a measure comparable to the semi-variogram or the spatial covariance function. It expresses the correlation coefficient of empirical bivariate copula densities for varying separation distances. The distance where the copula rank correlation reaches zero is effectively the range of traditional geostatistics (Bárdossy and Li, 2008).

The measure rank can have advantages, when the variogram is deteriorated, for example by extreme values (Figure 2.5). In this example, the semivariogram is increasing with bigger separation distance, while the rank correlation reaches a value of zero at ~ 20 m. An alternative measure for the rank correlation is Kendall’s tau given in Equation 2.28.

“Sym” is a measure for the symmetry of the empirical copula density function representing which range of quantiles the density is strongest (Equation 2.29). High positive symmetry values indicate strong dependence for high quantiles and high negative symmetry values indicate strong dependence for low quantiles. A Gaussian-type dependence structure is fully symmetric, and its measure of symmetry $Sym = 0.0$. Gaussian dependence is most pro-

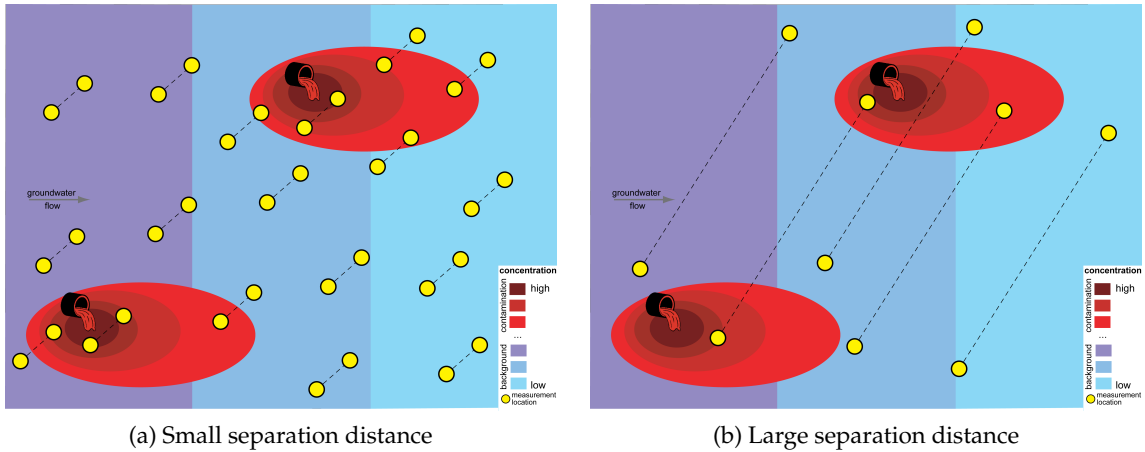


Figure 2.3: Conceptual drawing of two spills and their plumes within an aquifer. High concentrations related to the spills are indicated by red colours, the relatively small background concentration of the aquifer is indicated by blue colours. Indicated by yellow dots are fictious observation locations which are separated by a small distance on Subfigure 2.3a and by a large distance on Subfigure 2.3b.

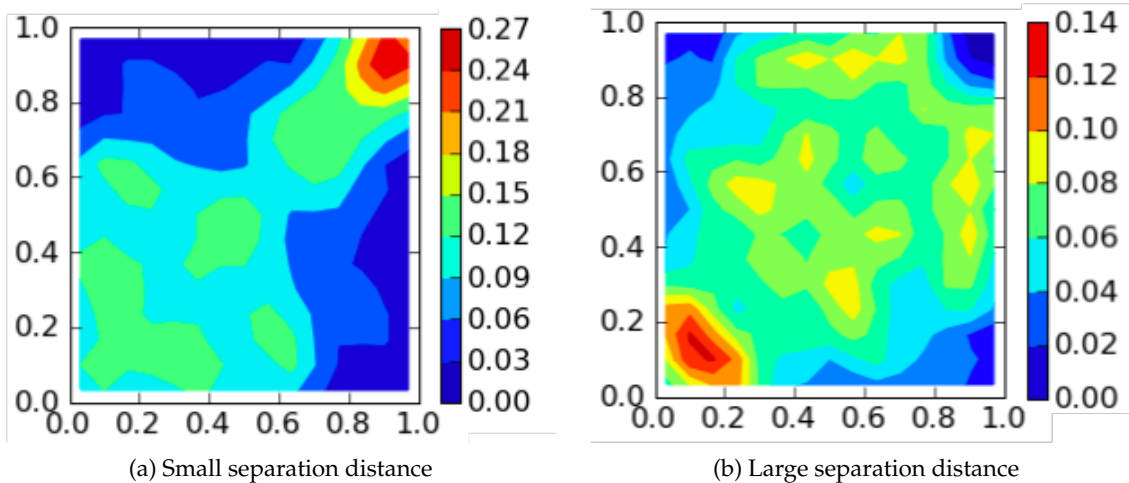


Figure 2.4: Schematic empirical bivariate copula density plots show the degree of dependence between pairs of observation points. High values are scaled to be at high quantiles (close to 1), low values are scaled to be at small quantiles (close to 0). The closer the shading is to a dark red, the higher the Copula density and the stronger the degree of dependence. The empirical bivariate copula density plots are shown for a small separation distance on Subfigure 2.4a corresponding to the sketch on Subfigure 2.3a; for a large separation distance on Subfigure 2.4b corresponding to the sketch on Subfigure 2.3b.

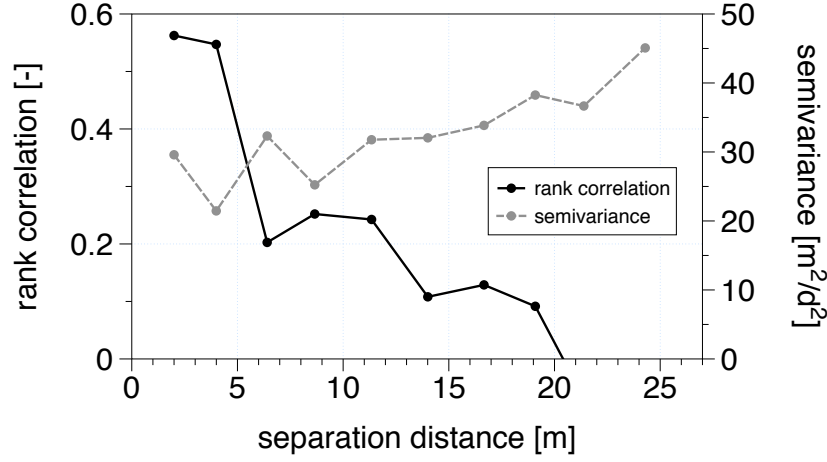


Figure 2.5: Comparison between rank correlation and semivariogram. The rank correlation is less influenced by extreme values and shape of the distribution of the data, hence often it is more useful to determine the “range” of the data-set.

nounced for extremes and equally strong in the extreme high and in the extreme low values. In a spatially distributed K field with Gaussian dependence structure, both the highest K and lowest K values have the strongest dependence and hence form isolated patches. The more positive the copula symmetry is, the higher the degree of dependence for high quantiles, the more patches of high K zones form, and the higher the likelihood that connected and continuous zones of low K are present. Similarly, the more negative the copula symmetry, the higher the degree of dependence for low quantiles, and the higher the chance for presence of connected high- K zones that could rapidly advect solutes.

Each of these measures is calculated for a given value and/or direction of the separation vector \mathbf{h} . The number of pairs of points for each \mathbf{h} is denoted by $n(\mathbf{h})$; $F_n(z(\mathbf{x}_i))$ is the value of the empirical distribution of the measurements z at location x_i .

$$\text{Rank}(\mathbf{h}) = \frac{1}{12n(\mathbf{h})} \cdot \sum_{\mathbf{x}_i - \mathbf{x}_j \approx \mathbf{h}} \left(F_n(z(\mathbf{x}_i)) - \frac{1}{2} \right) \cdot \left(F_n(z(\mathbf{x}_j)) - \frac{1}{2} \right) \quad (2.27)$$

$$\begin{aligned} \tau(\mathbf{h}) &= \frac{1}{N(\mathbf{h}) \cdot (N(\mathbf{h}) - 1)} \cdot \\ &\cdot \sum_{\mathbf{x}_i - \mathbf{x}_j \approx \mathbf{h}} \sum_{\mathbf{x}_k - \mathbf{x}_l \approx \mathbf{h}} \text{sign}((Z(x_i) - Z(x_k)) \cdot (Z(x_k) - Z(x_l))) \end{aligned} \quad (2.28)$$

$$\begin{aligned} \text{Sym}(\mathbf{h}) &= \frac{1}{n(\mathbf{h})} \cdot \sum_{\mathbf{x}_i - \mathbf{x}_j \approx \mathbf{h}} \left(F_n(z(\mathbf{x}_i)) - \frac{1}{2} \right)^2 \left(F_n(z(\mathbf{x}_j)) - \frac{1}{2} \right) + \\ &+ \left(F_n(z(\mathbf{x}_i)) - \frac{1}{2} \right) \left(F_n(z(\mathbf{x}_j)) - \frac{1}{2} \right)^2 \end{aligned} \quad (2.29)$$

2.3.3 Theoretical Spatial Copulas

This section describes first the theoretical models available to model non-Gaussian spatial dependence with copulas. Later, it will be explored how to evaluate the fit of a model to data.

A variety of copula models exist, among others: Elliptical copulas (Durante and Sempi, 2010), Archimedean copulas (Genest and Rivest, 1993; Frees and Valdez, 1998) including Frank copulas (Frank, 1978; Genest, 1987), or Eyraud-Farlie-Gumbel-Morgenstern (EFGM) copulas (Morgenstern, 1956; Farlie, 1960). Overviews are given by Nelsen (1999); Durante and Sempi (2010); Joe (1997). However, to be useful in geostatistics, a copula has to fulfill the following conditions:

1. Similar to the variogram- or covariance functions, the bivariate spatial copula of the random variable $Z(x)$ corresponding to two locations separated by the vector \mathbf{h} is assumed to be only dependent on \mathbf{h} . The marginal distribution of $Z(x)$ is supposed to be the same everywhere.
2. It must be possible to build a multivariate copula with any dimension $n > 2$ based on its bivariate marginals.
3. The parameterization of the copula should enable any n -dimensional copula corresponding to any selected n points to reflect their spatial configurations.
4. The parameterization of the copula should allow arbitrarily strong dependence.

Most of the above mentioned copula models do not meet the above criteria. However, the Gaussian copula does. The drawback of a Gaussian copula is that it can model only Gaussian, hence symmetric, dependence, which does not occur in the two real-world data-sets that this thesis is dealing with, for the reasons discussed (Section 3.1.3). There are two groups of copulas that are capable of modelling non-Gaussian dependence and that meet the above criteria, and they are used in this thesis. Both build on the Gaussian copula. The one group is called “v-copulas” (Bárdossy and Li, 2008) due to the shape of the non-monotonic transformations applied to Gaussian variables. The other group is called “maximum Gauss copula”, because it is a copula of the maximum of a number of Gaussian processes.

Despite containing more information than traditional average geostatistical measures, copulas are still related to classical approaches. For any cut-off value β , the indicator variogram γ_β is related to the diagonal of the spatial copula (Bárdossy, 2006).

2.3.3.1 Gaussian Copula

A Gaussian copula can be described with mean vector and covariance matrix Γ , Φ denotes the univariate standard normal distribution function and $\Phi_{\Gamma,n}$ a n -dimensional Gaussian distribution, then the corresponding Gaussian copula is given in Equation 2.30.

$$C_\Gamma(u_1, \dots, u_n) = \Phi_{\Gamma,n}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n)) \quad (2.30)$$

A Gaussian copula density is then given by Equation 2.31 with $x_i = \Phi^{-1}(u_i)$.

$$c_n(u_1, \dots, u_n) = \frac{1}{\sqrt{|\mathbf{\Gamma}|}} \cdot \left(-\frac{1}{2} \mathbf{x}^T (\mathbf{\Gamma}^{-1} - \mathbf{I}) \mathbf{x} \right) \quad (2.31)$$

2.3.3.2 V-Copula

The basis of v-copulas is a non-monotonic transformation of a standard-normally distributed function $\mathbf{Y} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Gamma})$ into a non-symmetric distribution \mathbf{X} using two parameters $m, k \in \mathbb{R}^+$ (Equation 2.32).

$$X_j = \begin{cases} k(Y_j - m), & \text{if } Y_j \geq m \\ m - Y_j, & \text{else} \end{cases} \quad (2.32)$$

This means the slope of the left “arm” of the v stays always unity. The parameters m and k are limited to be positive. If these two conditions were not imposed, two identical copulas with different sets of parameters could be created. For $|m| \rightarrow \infty$ the v-copula approaches the Gaussian copula, which for practical purposes is the case for $|m| > 3$.

For any one-dimensional marginal, its distributions $H(x)$ can be established by calculating the area under the standard normal density between the bounds given by the shape of the v (Equation 2.33).

$$H(x) = P(X \leq x) = \Phi\left(\frac{x}{k} + m\right) - \Phi(-x - m) \quad (2.33)$$

A marginal density $h(x)$ is obtained by differentiating $H(x)$ (Equation 2.34)

$$h(x) = \frac{1}{k} \phi\left(\frac{x}{k} + m\right) + \phi(-x - m) \quad (2.34)$$

In two dimensions, the approach of how the copula (distribution) is calculated, is visualized on Figure 2.6. The volume above the grey shaded square has to be integrated, between the integration bounds that are defined by the shape of the v-transformation, as given in Equation 2.33. For orientation purposes, the area where the multivariate standard normal density function intersects the x- and y-axis are shaded in light blue and light orange, respectively.

The multivariate distribution- and density functions in v-transformed space are given in Equations 2.35 and 2.36, respectively.

$$H_n(x_1, \dots, x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n) = \sum_{i=0}^{2^n-1} (-1)^{n-\sum_{j=0}^{n-1} i_j} \Phi(\zeta_i + \mathbf{m}) \quad (2.35)$$

$$h_n(x_1, \dots, x_n) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{\Gamma}|}} \cdot \sum_{i=0}^{2^n-1} \left[\frac{1}{k^{n-\sum_{j=0}^{n-1} i_j}} \cdot \exp\left(-\frac{1}{2} (\zeta_i + \mathbf{m})^T \mathbf{\Gamma}^{-1} (\zeta_i + \mathbf{m})\right) \right] \quad (2.36)$$

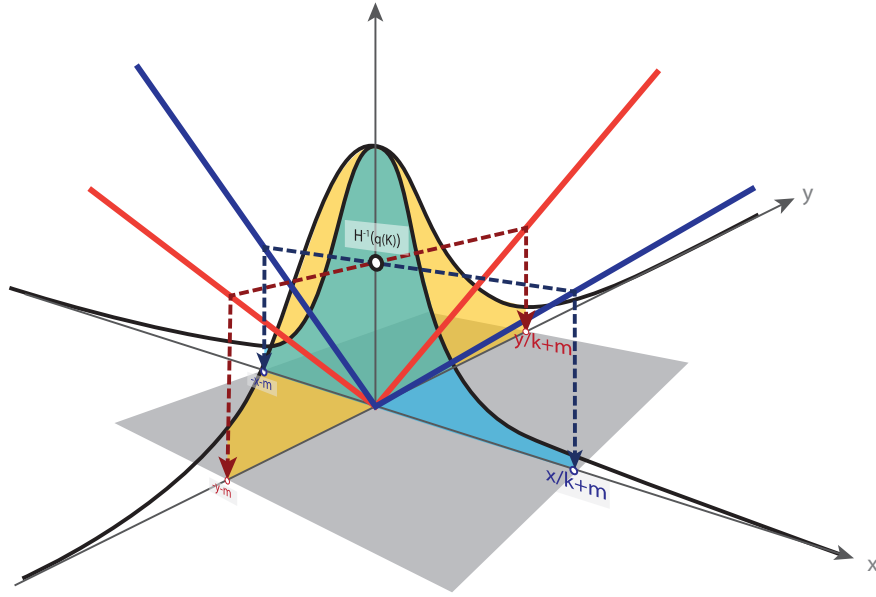


Figure 2.6: Sketch that explains how a copula (distribution) can be calculated in two dimensions. The volume bounded by the grey rectangular area on the bottom, and the bivariate standard normal density function has to be calculated. The slices where the standard normal distribution intersects the x- and y-axis are shaded in light blue and light orange, respectively. The integration bounds are defined by the shape of the v-transformation.

In these equations, the following substitutions are necessary:

$$\zeta_i^T = (b((-1)^{i_1}) \cdot x_1, \dots, b((-1)^{i_n}) \cdot x_n) \quad (2.37)$$

$$i = \sum_{j=0}^{n-1} i_j 2^j \quad (2.38)$$

$$b = \begin{cases} -1 & \text{if } (-1)^{i_j} = -1 \\ \frac{1}{k} & \text{if } (-1)^{i_j} = +1 \end{cases} \quad (2.39)$$

Equation 2.38 is a binary representation that facilitates adding the correct parts of the transformed distribution, as depicted on Figure 2.6. The dimension is n , which is equivalent to the number of points in the parameter estimation. As an example, say there are 5 points per tile, then there are $2^4 = 16$ numbers to be represented in 5 digit long binary representation. From all these possibilities, say $i = 5$, then $i_j = 00101$, and $\sum_{j=0}^{n-1} i_j = 2$.

Finally, the the v-copula density is the joint density divided by the marginal densities (Equation 2.40).

$$c_n(u_1, \dots, u_n) = \frac{h_n(x_1, \dots, x_n)}{\prod_{i=1}^n h_i(x_i)} \quad (2.40)$$

2.3.3.3 Maximum Gauss Copula

Let F and G be two random variables and let H be a third random variable corresponding to the maximum of F and G : $Z = \max(F, G)$. The distribution function of H is the distribution function where x is statistically bigger than F and at the same time bigger than G . Then the distribution function of H , $F(H)$, is the product of the two underlying distribution functions, representing the events where both F and G are smaller than x :

$$F_H(x) = P(F \leq x \wedge G \leq x) = P(F \leq x) \cdot P(G \leq x) = F_F(x) \cdot F_G(x) \quad (2.41)$$

This can be generalized such that the distribution of the maxima H of a set of random variables is distributed as the product of the distribution function of each of the random variables F_k (Equation 2.42, Coles (2001)).

$$H(x_1, \dots, x_n) = \prod_{k=1}^l F_k(x_1, \dots, x_n) \quad (2.42)$$

In one dimension, the distribution functions F , G , and H are shown on Figure 2.7, where the two processes are $F \sim \mathcal{N}(1, 1)$ and $G \sim \mathcal{N}(2, 1)$. The distribution of the maxima is $H = F \cdot G$ and is in this case slightly not symmetric.

For this thesis the general foundations are shown which could be expanded in the future:

1. The dimension of the problem is 2. Theoretically, the dimension could be arbitrarily high. The number of terms in the derivative of the distribution of the maxima would grow quickly as the dimension increases.

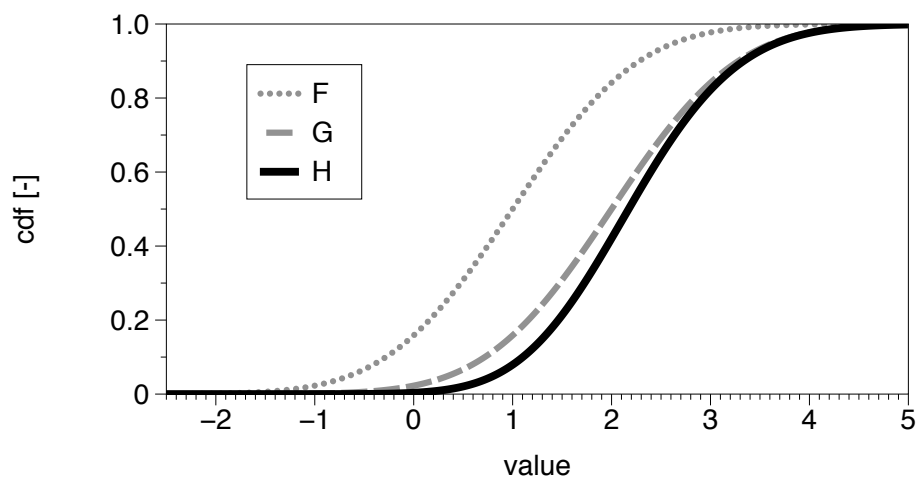


Figure 2.7: Sketch of a non-Gaussian distribution function H which is the maximum of two Gaussian processes F and G .

2. For reasons of parsimony, the number of processes that contribute to the distribution of the maxima is 2. Theoretically, there could be an infinite number of contributing processes. Geologically, a much larger number than 2 seems rare.
3. Both of the underlying processes are Gaussian processes, such that they are mathematically easier to handle, because then both processes can be described by means and covariance matrix. This constraint does not imply that the resulting maximum process H is necessary Gaussian. To facilitate the solution to fit the mean and the covariance, one of the underlying processes can be limited to be standard normal.

The following steps are necessary to calculate a copula density from the multivariate distribution function of the maxima of two underlying processes:

$$H(x_1, x_2) = F(x_1, x_2) \cdot G(x_1, x_2) \quad (2.43)$$

$$\begin{aligned} h(x_1, x_2) &= \frac{\partial^2 H(x_1, x_2)}{\partial x_1 \cdot \partial x_2} \\ &= \frac{\partial(F(x_1, x_2) \cdot G(x_1, x_2))}{\partial x_1 \cdot \partial x_2} \\ &= \underbrace{\frac{\partial^2 F}{\partial x_1 \cdot \partial x_2}}_f \cdot G + \underbrace{\frac{\partial F}{\partial x_1} \cdot \frac{\partial G}{\partial x_2} + \frac{\partial F}{\partial x_2} \cdot \frac{\partial G}{\partial x_1}}_{\text{"mixed terms"}} + F \cdot \underbrace{\frac{\partial^2 G}{\partial x_1 \cdot \partial x_2}}_g \end{aligned} \quad (2.44)$$

$$\begin{aligned} h(x_1) &= \frac{\partial H(x_1, x_2)}{\partial x_1} = \frac{\partial(F(x_1, x_2) \cdot G(x_1, x_2))}{\partial x_1} \\ &= F(x_1, x_2) \cdot \frac{\partial G(x_1, x_2)}{\partial x_1} + \frac{\partial F(x_1, x_2)}{\partial x_1} \cdot G(x_1, x_2) \\ &= F(x_1, x_2) \cdot g(x_1) + f(x_1) \cdot G(x_1, x_2) \end{aligned} \quad (2.45)$$

$$\begin{aligned} h(x_2) &= \frac{\partial H(x_1, x_2)}{\partial x_2} = \frac{\partial(F(x_1, x_2) \cdot G(x_1, x_2))}{\partial x_2} \\ &= F(x_1, x_2) \cdot \frac{\partial G(x_1, x_2)}{\partial x_2} + \frac{\partial F(x_1, x_2)}{\partial x_2} \cdot G(x_1, x_2) \\ &= F(x_1, x_2) \cdot g(x_2) + f(x_2) \cdot G(x_1, x_2) \end{aligned} \quad (2.46)$$

$$H(x_i) = F(x_i) \cdot G(x_i) \quad (2.47)$$

$$c(H^{-1}(x_1), H^{-1}(x_2)) = \frac{h(H^{-1}(x_1), H^{-1}(x_2))}{h(H^{-1}(x_1)) \cdot (H^{-1}(x_2))} \quad (2.48)$$

Each of the “mixed terms” in Equation 2.44 contain conditional distributions and the marginal density. Figure 2.8 shows bivariate results of Equations 2.43-2.48, using some example processes F and G .

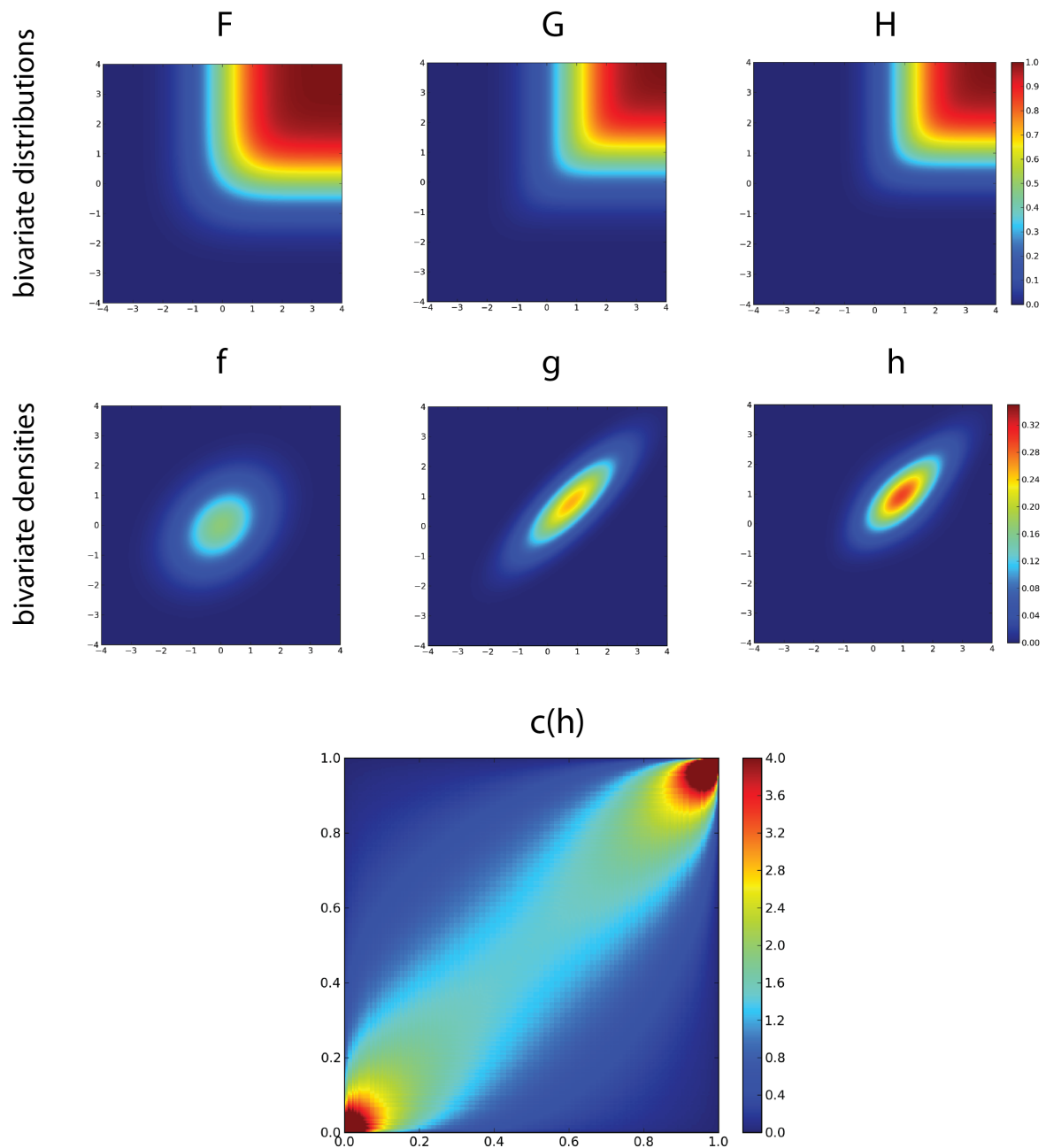


Figure 2.8: Some intermediate results (Equations 2.43-2.48) of an analytical maximum Gauss copula. Chosen parameters: $F \sim \mathcal{N}([0.0, 0.0], [1.0, 1.0])$; $G \sim \mathcal{N}([0.9, 0.9], [1.1, 0.82])$.

2.3.4 Parameter Estimation

The parameters of a given theoretical copula model need to be chosen such that the model fits to the observed data. "Fit" means foremost that the chosen theoretical copula model

is to capture the spatial dependence structure of the data-set. The dependence-structure exists in high dimensions, essentially as high as the number of measurement locations. And theoretically, the fit should be good over all dimensions. The margins of a multidimensional copula model should fit to the bivariate empirical copula densities. On the other hand, the fit does not necessarily need to be “perfect” on the margins, because the fit should be good in all dimensions.

In traditional geostatistics, a theoretical variogram is typically fitted “by eye” to an empirical variogram, sometimes with (semi-)automated methods. The approach chosen here offers a more quantitative approach for parameter estimation, which is adopted from [Bárdossy and Li \(2008\)](#). In this approach the set of observation points is divided into “tiles” representing subsets of the set of observations of size m . Each observation point is allowed to occur only once in a tile. Additionally, the points in each subset are chosen such that the histogram of separation distances over all subsets is fairly uniform, meaning each subset should contain pairs of points with a big separation distance and also small separation distances. Initial parameters are chosen, for each tile the copula density c is calculated, and the $\ln(c)$ of all subsets added. The resulting sum is to be maximized by varying the parameter values. This parameter optimization approach has been adopted for all three types of copula models.

One question that often arises is the amount of data that is needed for a spatial copula. A simple answer is that if there is enough data for a good variogram estimation, and that means foremost at least enough pairs of points in a variety of distance classes, then this data-set is likely good enough for spatial copulas. It is important to stress, that the parameter estimation approach employed does not use binned distance classes but is maximum likelihood based. This means, to estimate a Gaussian copula, less data-points might be sufficient than for variogram estimation. Furthermore, the more parameters a theoretical copula model has, the more data-points are needed.

2.3.4.1 Parameter Estimation for V-Copula

The above described parameter estimation approach estimates m and k . It also optimized the hidden covariance function model. For the purpose of comparing solute transport properties based on Gaussian and non-Gaussian spatial fields (Chapter 4) it was necessary to ensure identical second order moments for all the spatially distributed fields. Hence, the covariance structure of the underlying Gaussian field R_{ug} had to be determined such that the covariance structure after the v-transformation R_v is identical to the given or pre-determined covariance structure R (Figure 2.9).

Such an involved procedure of generating fields whose spatial rank correlations are identical was deemed necessary because the results of the solute transport analysis are then comparable between the different types of spatial dependence structures. The correlation structure Γ of the “underlying” (index “ug”) field $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \Gamma)$ needs to be such that the new correlation structure after the transformation R_v fits to the empirical correlation structure of the Borden data-set (Equation 2.49). With the help of Equation 2.50 correlations ρ can be converted to

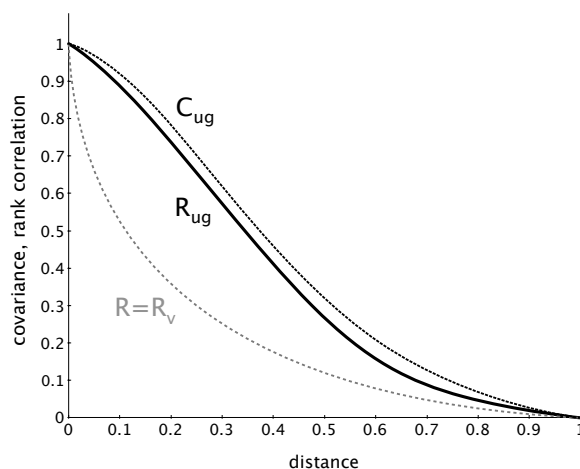


Figure 2.9: Sketch to visualize typical properties of correlation structures of an “underlying” (index ug) Gaussian spatial field that is to be transformed to a non-Gaussian field with a certain given or known rank correlation $R = R_v$.

rank correlations R (Joe (1997)).

$$R_v = f(R_{ug}|m, k) \quad (2.49)$$

$$R(h) = \frac{6}{\pi} \arcsin\left(\frac{\rho(h)}{2}\right) \quad (2.50)$$

2.3.4.2 Parameter Estimation for Maximum Gauss Copulas

A bivariate maximum Gauss copula density function based on two bivariate normal Gaussian distribution functions $F(x_1, x_2)$ and $G(x_1, x_2)$ depends on a number of parameters. Mean values ($\mu_{F1}, \mu_{F2}, \mu_{G1}, \mu_{G2}$), variances ($\sigma_{F1}, \sigma_{F2}, \sigma_{G1}, \sigma_{G2}$) and correlation coefficients (ρ_{F12}, ρ_{G12}) add up to a total of ten parameters. This seemed like a complex optimization to start with, hence the number of parameters was decreased as much as possible for initial optimization, and later slowly increased. The options to reduce the number of parameters include:

- have identical margins for each process: $\mu_{F1} = \mu_{F2}, \mu_{G1} = \mu_{G2}, \sigma_{F1} = \sigma_{F2}, \sigma_{G1} = \sigma_{G2}$;
- leave one of the two processes as a “standard normal process”: $\mu_{F1} = \mu_{F2} = 0$ and $\sigma_{F1} = \sigma_{F2} = 1$;

These simplifications have led to a total of four parameters $\mu_G, \sigma_G, \rho_{F12}, \rho_{G12}$.

The correlation coefficients depend on the covariance matrix, which in a spatial setting varies with the separation distance between points considered. As in the estimation procedure for the v -copula model, here is also a hidden covariance model employed, using

a superposition of nugget and spherical variogram types. The variance is standardized, hence for each process only the range (a) and one additional value, either nugget or sill of the spherical model have to be optimized.

Including the covariance structure, there are a minimum of six parameters needed for a spatial maximum Gauss copula model: $\mu_G, \sigma_G, Nug_F, a_F, Nug_G$ and a_G . The optimization tried to maximize the variance of one process. Usually, the margins of one process were fixed, hence the variance of the other process was maximized. In scenario 1cc for the Borden aquifer, the variance of both processes was allowed to vary during the optimization, then the same thing happened, one variance was maximized.

The maximum Gauss copula model is quite flexible and can incorporate anisotropy, even different anisotropy for each of the underlying Gaussian processes. These parameters of anisotropy can be included in the optimization procedure, with the drawback that the number of parameters increases. These parameters include stretch factors s , possibly one in each direction of each process and rotating factors, indicated by small case greek letters, possibly one less than the number of dimensions involved. Table 3.3 lists the four sets of parameters that were at different stages incorporated into the parameter estimation of maximum Gauss copulas.

2.3.5 Properties of Spatial Copula Models

Spatial copulas are a full stochastic model and can be used directly for interpolation and simulation purposes. How does the copula approach compare with commonly used Kriging procedures? – There are two key points:

The first relates to how the interpolation and simulation is mathematically executed: No system of linear equations has to be solved using copulas. Instead, at the point where a value is to be interpolated or simulated, a conditional distribution function is built, conditioned on the surrounding observations. This distribution function can be not symmetric. For interpolation purposes, a representative value of the distribution function is chosen (the expected value, f.ex.). A major advantage of the copula approach is that because a full distribution function is available, confidence intervals can be easily evaluated, a $q_{90} - q_{10}$ width, f.ex.

The second point relates to the uncertainty estimates related to interpolation: Using Ordinary Kriging, the estimation weights depend on the observation network's density, geometry, and selected variogram model – but not on the magnitude of the measured values. A copula model takes the magnitude of measurements into account. Being able to quantify uncertainty in a much more meaningful way using spatial copulas is one of the method's largest advantages. This is also a prerequisite for meaningful simulation.

2.3.5.1 Interpolation

The typical goal of an interpolation method is to estimate a random variable at unsampled locations x' . This section describes the interpolation algorithm:

1. The observation network consists of n locations x_1, \dots, x_n . At each location there are observations available, z_1, \dots, z_n , which are transformed to u_1, \dots, u_n by $F(z_i) = u_i$.
2. In the neighbourhood of a x' , m observation points are selected.
3. The copula density value corresponding to those m locations and their observation values is calculated: $c_m(u_1, \dots, u_m)$.
4. For the point x' , for a quantile v , the $m + 1$ dimensional copula density $c_{m+1}(u_1, \dots, u_m, v)$ is calculated.
5. The density function corresponding to x' conditioned on the n observations in the vicinity is calculated:

$$c^*(v) = c(v_k | u_1, \dots, u_n) = \frac{c_{n+1}(u_1, \dots, u_n, v)}{c_{n+1}(u_1, \dots, u_n)} \quad (2.51)$$

6. The conditional copula C^* is calculated from its density c^* by summing the density up cumulatively.

Depending on which theoretical copula model is used, different forms of c on the right hand side of Equation 2.51 are used: Equation 2.31 when using Gaussian copulas, Equation 2.40 when using v-copulas, Equation 2.48 when using maximum Gauss copulas. Interpolation using copulas is exact, meaning that estimates at measured locations return the measured value.

From the conditional copula, arbitrary quantiles can be evaluated. For interpolation purposes, this typically is the 50% quantile q_{50} or the expected value. Confidence intervals can also be evaluated, a $q_{90} - q_{10}$ width, f.ex. It should be stated that all these values are still in copula space and can be transformed with the inverse of the marginal distribution to the space in which the variable of interest was originally measured.

2.3.5.2 Simulation

Spatial uncertainty is a measure of the joint uncertainty about attribute values at several locations taken together, for example the probability of occurrence of a string of large or small values. Such spatial uncertainty can be modelled by generating multiple realizations of the joint distribution of attribute values in space, a process known as stochastic simulation [Goovaerts \(1997\)](#). Subsequently these realizations can be used as input parameters for physical models which return a distribution of responses, which can be analyzed to evaluate the effects of the spatial uncertainty, which is done in Chapter 4. The input for subsequent linear models should be simulated spatial fields and not interpolated spatial fields. Interpolated fields tend to smooth out spatial variation, which decreases extreme values which are important for physical models.

The procedure of generating unconditionally simulated fields with v-copula structure, a multi-Gaussian realization is generated with one of the existing approaches (sequential

Gaussian simulation, LU decomposition algorithm, turning bands algorithm). The simulated values are transformed using the optimized v -copula parameters, and then the marginal distribution is exchanged with a random sample from the marginal distribution. The procedure of generating unconditionally simulated fields with maximum Gauss copula structure is similar to, except that two multi-Gaussian fields are simulated with their sets of optimized parameters. At every location, the maximum of the simulated multi-Gaussian variables is selected, before the marginal distribution is exchanged.

Some properties of simulated Gaussian and non-Gaussian fields are visualized on Figure 2.10. The top panel shows the bivariate copula densities used to simulate the spatial fields shown in the second row. The 10% highest values of the simulated field are shown in white in the third row, the rest is censored in black shading. The bottom panel shows the 10% smallest values. The structure of the 10% highest values is very similar for both fields, since in both copulas have high densities in high quantiles. However, the v -copula has no high density in the low quantiles, hence the structure of the smallest 10% quantiles is degraded (bottom right panel) compared to the Gaussian copula.

Using conditional simulation, the spatial structure and the variability are simulated as prescribed with the additional feature that values at measurement locations can be incorporated and remain unaltered in the simulated field. The method chosen is a sequential simulation algorithm, which involves these steps:

1. The observation network consists of n locations x_1, \dots, x_n . At each location there are observations available, z_1, \dots, z_n , which are transformed to u_1, \dots, u_n by $F(z_i) = u_i$.
2. The locations where values are to be simulated needs to be chosen (typically a raster), as well as a random path that defines in which order (“sequence”) the locations are visited.
3. For each of those locations, in the previously defined order, the copula density needs to be calculated according to Equation 2.51, as well as its distribution. From this cdf a random number is drawn using its inverse, and the result added to the set of conditioning points.
4. These simulated quantiles are back-transformed into the space of the measurement values.

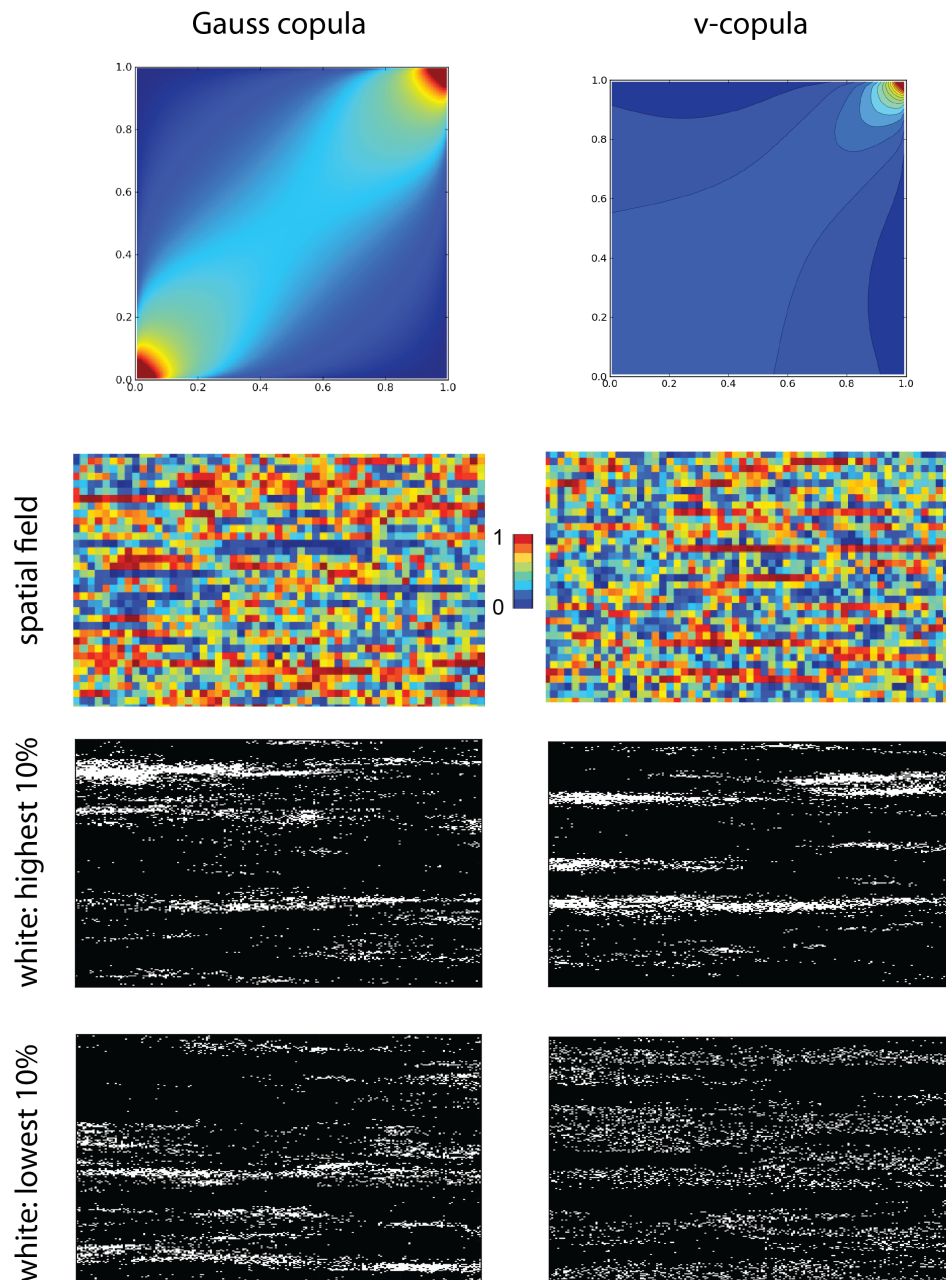


Figure 2.10: Differences between Gaussian and non-Gaussian spatial fields. The top panel shows the bivariate copula densities used to simulate the spatial fields shown in the second row. The 10% highest values of the simulated field are shown in white in the third row, the rest is censored in black shading. The bottom panel shows the 10% smallest values. The structure of the 10% highest values is very similar for both fields, since in both copulas have high densities in high quantiles. However, the v-copula has no high density in the low quantiles, hence the structure of the smallest 10% quantiles is degraded (bottom right panel) compared to the Gaussian copula.

3 Used Data-Sets and Their Geo-Statistical Description

The previous chapter described some classical and copula-related theory on geostatistics. In this chapter, the previously described theory is applied on two data-sets of hydraulic conductivity (Section 3.1). The analysis starts with a descriptive geostatistical analysis of the data in Section 3.2. Subsequently, the parameters of theoretical copula models are fitted to the data-sets (Section 3.3). These fitted models will be compared to the data (Section 3.4), and used for Interpolation and Simulation purposes (Section 3.5).

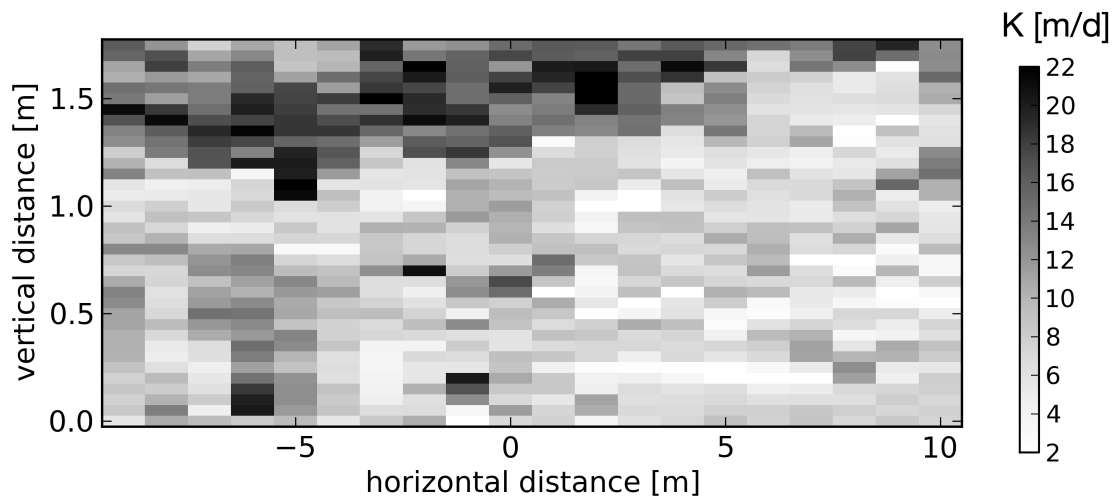
3.1 Basic Description of the Data-Sets

This thesis applies spatial copulas, novel geostatistical methods, to two real world data-sets of hydraulic conductivity (K). The one data-set is from a well characterized field site located at the C.F.B. Borden, Canada (Sudicky, 1986), the other from an aquifer located in the vicinity of the city of North Bay, Canada (Sudicky et al., 2010). Collecting data is generally time- and money- consuming. Each data-set comprises more than one thousand values of measured K, which makes the two sites two of best characterized aquifers in the world.

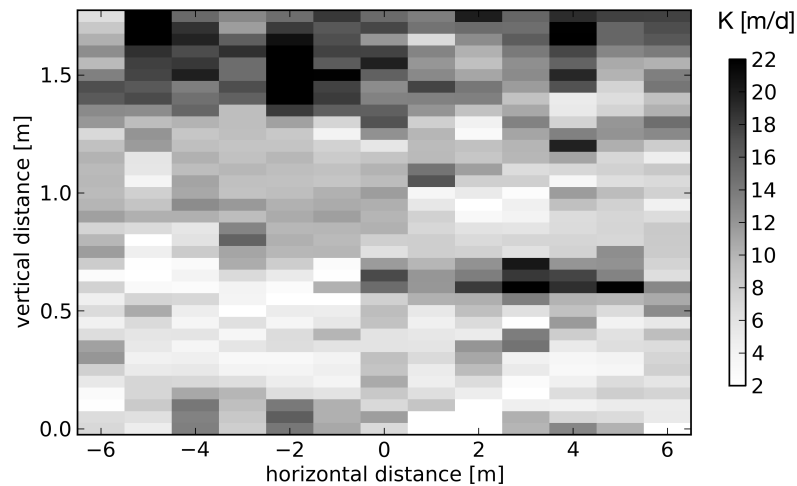
3.1.1 The Borden Data-Set

The Borden site is investigated hydrogeologically by scientists from the University of Waterloo since 1978 (Sudicky, 1983) when a salt tracer experiment for studies of dispersion under natural flow conditions was conducted in a shallow sandy unconfined aquifer which originated as beach sand along the shore of a glacial meltwater lake that existed at the end of Pleistocene time (Cherry et al., 1996). The experiment showed strong influence of aquifer heterogeneity on flow and transport phenomena, particularly dispersion, even though the aquifer is relatively homogeneous in its hydraulic behaviour compared with most other aquifers (Cherry et al., 1996).

The Borden aquifer hydraulic conductivity data-set has been inspiration for many hydrogeological studies, particularly related to geostatistics and solute transport behaviour, including Woodbury and Sudicky (1992); Dagan (1990); Zhang and Neuman (1990); Wang and Kitanidis (1999); Sun et al. (2008); Ritzi and Allen-King (2007); Rajaram and Gelhar (1991) among many others.



(a) AA'



(b) BB'

Figure 3.2: Contours of the measured K values on the two orthogonal cross sections AA' and BB' at the Borden site. In the contour plots, each measurement value occupies the geometric space equally distributed between measurement locations.

conductivity measurements along the two cross-sections at North Bay in a similar manner as at Borden (c.f. Figure 3.2). Here, the extent of a rectangle in x -direction was assumed to be 2 m, the closest horizontal distance between two cores. Areas where no measurement was taken are shaded in grey. The structure of K is less clearly dominated by lenses compared to Borden. This is the case particularly at large x -coordinates along line 1 where mostly low K values were measured. The geologic setting at North Bay is more complex than at Borden.

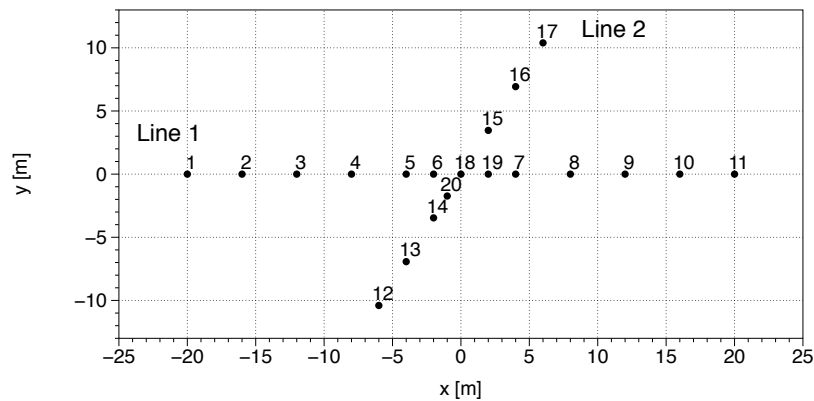


Figure 3.3: Setup of the coring locations at North Bay.

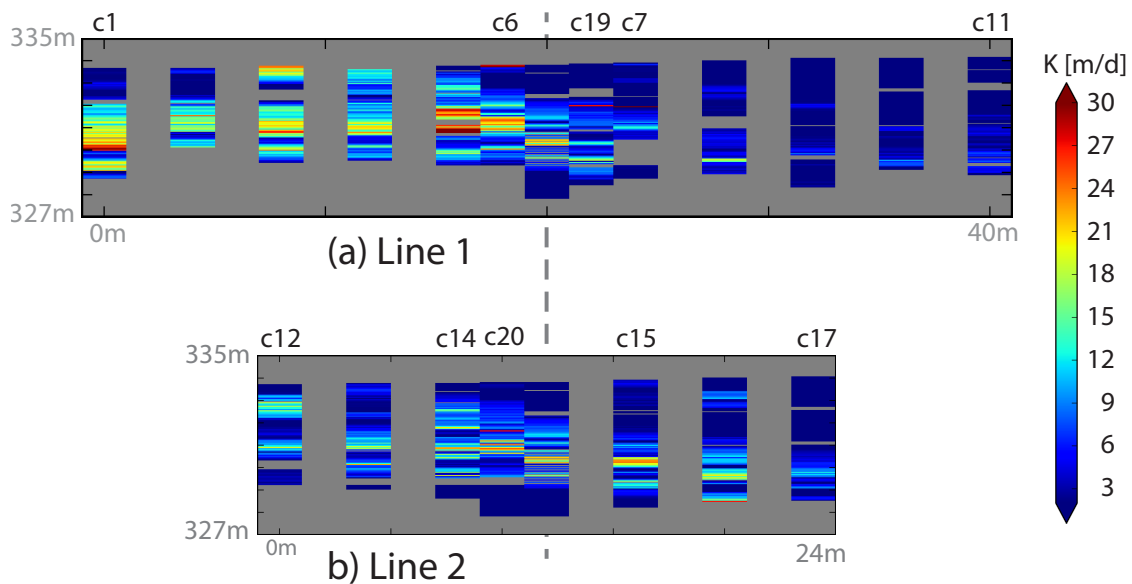


Figure 3.4: The two cross sections, Line 1 and Line 2, of the North Bay dataset. Inverse distance interpolation to get an impression of the distribution of K values at North Bay.

3.1.3 Why These Two Data-Sets

The Borden data-set is one of the most homogeneous aquifers studied, with a low variance of K of $\text{var}(\ln(K)) = 0.39$. The marginal distribution of K is fairly symmetric with $\text{skew}(K[\text{m/d}]) = 0.81$ (Table 3.1, Figure 3.5). Generally, the range between smallest and largest K value is ~ 30 m/d and is considered to be a fairly small portion of hydraulic conductivities encountered in naturally occurring geologic media (Table 3.1, Figure 3.5). Both aquifers fall in the category of “clean sand” porous medium aquifers. For comparison, Gelhar (1992) suggests the variance of \ln -hydraulic conductivity might range in naturally oc-

curing porous media from 0.16 to 16, and spatially distributed K fields used by Zinn and Harvey (2003) exhibit a $\ln(K)$ -variance of 9.

The North Bay data-set is geologically more complex, which might have led to its larger range of hydraulic conductivity (~ 40 m/d compared to ~ 30 m/d at Borden), a larger variance of $\text{var}(\ln(K)) = 1.79$ compared to $\text{var}(\ln(K)) = 0.39$ at Borden, and a bigger skewness of $\text{skew}(K[\text{m/d}]) = 1.62$ compared to $\text{skew}(K[\text{m/d}]) = 0.80$ at Borden (Table 3.1, Figure 3.5). Despite being geologically more complex than the fairly homogeneous aquifer at Borden, the aquifer at North Bay represents a more common geologic setting.

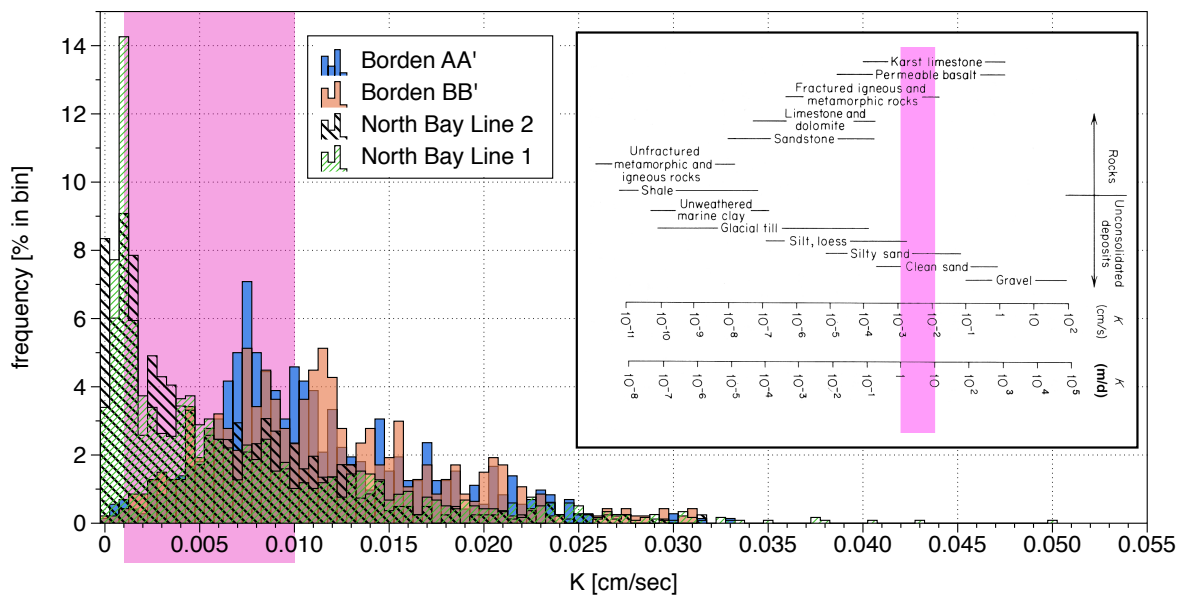


Figure 3.5: Histograms of the Borden and the North Bay K data-sets. Shaded in pink is one order-of-magnitude range, which corresponds to the range shaded in the inset that depicts the range of naturally occurring K (after Freeze and Cherry (1979)).

Table 3.1: Descriptive statistics of K and $\ln(K)$ of the Borden and of the North Bay data-sets.

	Borden		North Bay	
	K [m/d]	$\ln(K)$	K [m/d]	$\ln(K)$
min	0.05	-3.00	0.01	-4.61
max	28.44	3.35	43.16	3.76
range	28.39	6.34	43.15	8.37
arithm. mean	9.62	2.11	5.90	1.12
geom. mean	8.24		3.06	
variance	24.26	0.39	37.28	1.79
std deviation	4.93	0.62	6.11	1.34
skewness	0.80	-1.65	1.62	-0.73
curtosis	0.48	7.33	2.93	0.37

3.1.4 Marginal Distributions

One advantage of using copulas for spatial statistics is that the type and shape of the marginal distribution can be chosen such that it fits best to the observed data. There is no requirement, as there is with Ordinary Kriging, that the marginal distribution must be symmetric (Section 2.2.2).

In the case of the Borden aquifer, whose skewness is not very non-symmetric (Section 3.1, Table 3.1, Figure 3.5), the best fit using a standard maximum likelihood based approach among Weibull-, Gumbel-, Log-Normal- and Normal-Distributions was a two-parametric Weibull distribution (Equation 3.1) with parameters $\alpha = 0.0072$ and $\beta = 2.0659$ for K to be in units of m/d.

$$F(x) = 1 - e^{-\alpha x^\beta} \quad (3.1)$$

The North Bay data-set has a much higher skewness than the Borden data-set. No parametric distribution function would fit to a satisfactory degree, hence a non-parametric approach was chosen. The width of moving classes for a density estimation was optimized using a kernel K with a leave-one-out maximum likelihood based optimization.

$$f(x) = \frac{1}{n} \sum_{i=1}^n K_u(x - x_i) \quad (3.2)$$

The kernel K needs to fulfill three conditions (Equations 3.4, 3.5, and 3.5).

$$K(x) \geq 0 \quad (3.3)$$

$$\int K(x) dx = 1 \quad (3.4)$$

$$\int xK(x) dx = 0 \quad (3.5)$$

A Gaussian Kernel K_G (Equation 3.6) fulfills these three conditions and was used. The Gaussian Kernel is symmetric and defined in the interval $]-\infty, +\infty[$. Hydraulic conductivities are defined in the interval $]0, +\infty[$. This problem is avoided if the measurements are transformed into positive space, by taking the natural logarithm f.ex., before the kernel is fitted, and back-transformed after successful fitting.

$$K_G(x) = \frac{1}{\sqrt{2\pi}d} \cdot e^{-\frac{x^2}{2d^2}} \quad (3.6)$$

The optimized kernel widths in ln-space were $d = 0.1498$ for the entire North Bay data-set, $d = 0.2190$ for the bottom part of the Borden data-set, and $d = 0.2888$ for the top part of the Borden data-set.

It is a common approach to estimate one marginal distribution based on the given data. Given the high quantity and quality of the Borden- and the North Bay data-sets, the

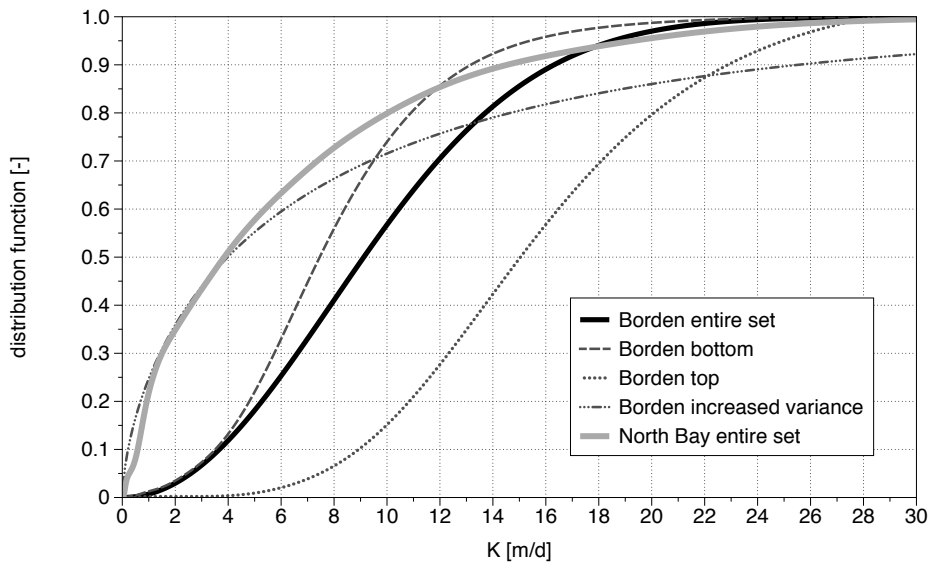


Figure 3.6: Marginal distributions fitted to the measurements of K for different (sub-) data-sets used in this thesis. The thick solid lines – black for Borden and grey for North Bay – are related to the data presented thus for in this thesis. The other distributions will be needed later. The distributions for the entire Borden data-set and for the Borden data-set with increased variance is a two-parametric Weibull distribution, the others are parameter-free distributions with optimized Kernel widths.

marginal distributions are fairly well characterized. Still, there remains an uncertainty related to the shape and position of the marginal distribution. This uncertainty is accounted for if for every Monte Carlo realization the marginal distribution is fitted to a subsample of the entire data-set (Section 4.4).

3.2 Empirical Geostatistics on the Data-Sets

In this section empirical geostatistical measures of the two data-sets are constructed and analyzed. These include copula-based measures as “Sym” (Equation 2.29) and “Rank” (Equation 2.27) over varying separation distances and empirical bivariate copula densities (Equation 2.26) as well as traditional measures as semi-variograms (Equation 2.7) and covariance functions (Equation 2.8).

3.2.1 Empirical Geostatistics on the Borden Data-Set

The above mentioned measures of spatial dependence for the Borden data-set are plotted on Figures 3.7b and 3.7a in horizontal and in vertical direction, respectively.

The semi-variograms compare well with the previously published ones (Woodbury and Sudicky, 1991). For the Borden data-set, there is no visible difference in structure between the covariance function and rank correlation, which is attributable to the fairly symmetric marginal distribution. The rank correlation is slightly higher for short separation distances in the vertical direction (highest value for shorted measured separation distance ~ 0.7 m) than in horizontal direction (highest value for shorted measured separation distance ~ 0.6 m). From the rank correlations, the anisotropy ratio, given by the range in horizontal direction divided by the range in vertical direction, is $11 \text{ m}/0.55 \text{ m} \sim 20$, which is slightly smaller than previously estimated value of ~ 24 (Woodbury and Sudicky (1991)).

For short separation distances, the measure “Sym” is positive in both directions, reaching zero, and becoming negative in the horizontal direction. This means that for separation distances larger than ~ 0.3 m in the vertical direction, the strongest dependence is in the low quantiles. This property had to be ignored, because to date no copula model exists that can describe a behaviour corresponding to a sign-change in the measure “Sym”. However, the error introduced is considered to be small because the important factor is the fit between theoretical- and empirical copula for short separation distances.

From analyzing these measures of spatial dependence separately for both cross-sections AA' and BB', and jointly (red, blue, and green respectively on Figures 3.7b and 3.7a), it is evident that the joint behaviour is a mixture of results from both cross-sections. This phenomenon has been observed in the past (Woodbury and Sudicky, 1991) and is most pronounced in the measure “Sym”.

Copulas are capable of modelling a different degree of dependence for different quantiles, and thus contain more information compared to an average variance for a given lag-distance, as in semi-variograms. The different degree of dependence can be visualized best by examining bivariate copula densities. Figure 3.8 shows two bivariate copula densities for two different separation distances both in horizontal (panel a) and in vertical direction (panel b). For larger separation distances, the structure of the bivariate copulas flattens out, eventually reaching a flat copula density, which is reached at a separation distance corresponding to the “range” in traditional geostatistics – the same distance where the rank correlation reaches zero. With shorter separation distances of empirical bivariate copula densities, it is evident that the highest values of K which are scaled to be close to unity in the uniform copula-space unit square do not have the same degree of dependence as the lowest values of K which would be expected if the spatial dependence structure was Gaussian. Rather, the highest degree of dependence occurs in the highest quantiles, both in the vertical and horizontal directions. This means, for a pair of points separated by a short separation distance, that the probability of encountering another high K value at the second point is high, if also a high K value was measured at the first point. This property will lead to patches or “blobs” of high-K values.

There are two main differences between the horizontal and the vertical direction: First, the required separation distance to reach a flat copula structure is longer in the horizontal direction (anisotropy). The distance at which the copula density becomes flat corresponds with the measure “range” of traditional variogram analysis. Secondly, the maximum correlation is weaker in horizontal direction.

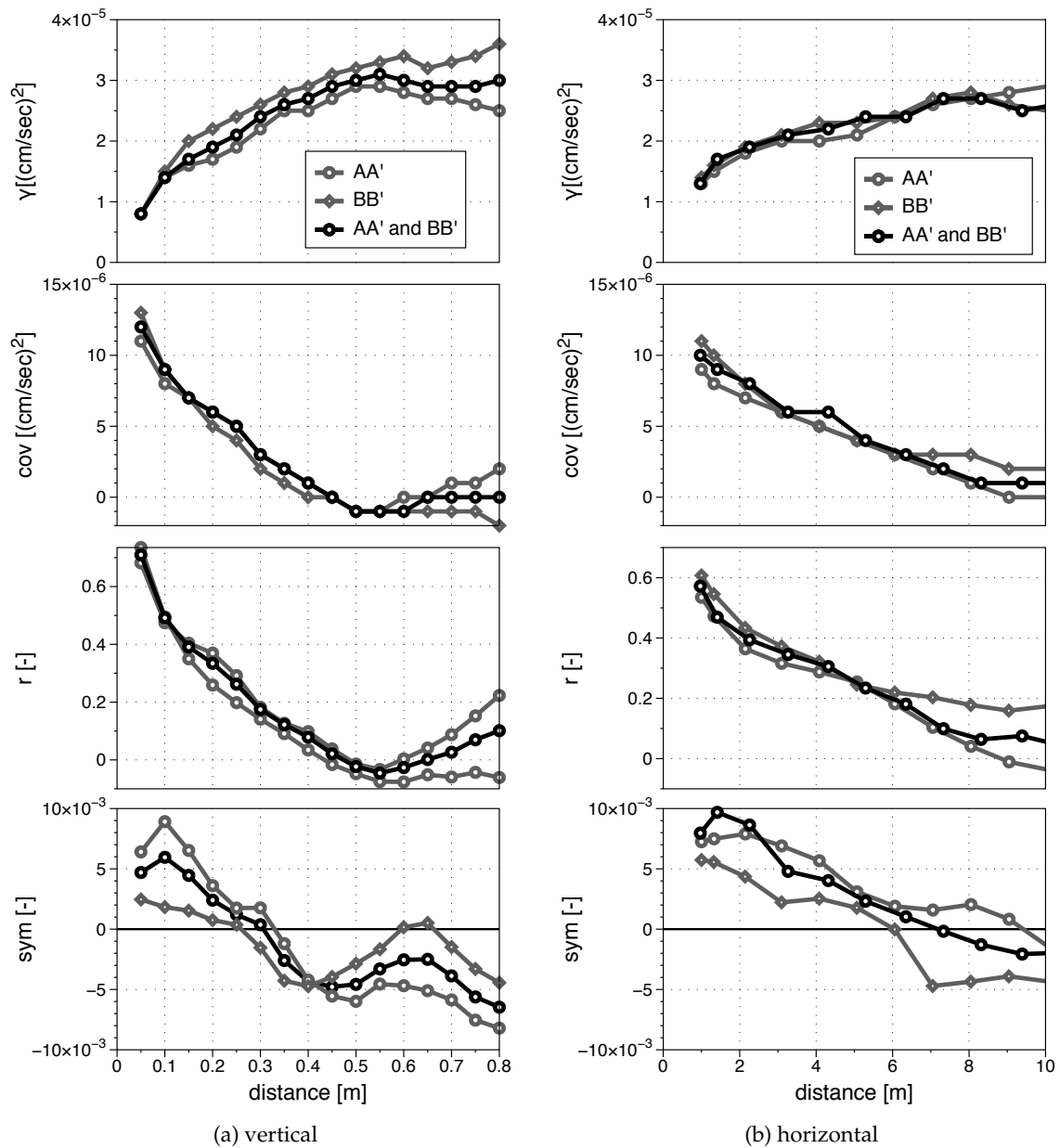


Figure 3.7: Measures of spatial dependence (semi-variance γ [(cm/sec)²], covariance "cov" [(cm/sec)²], rank correlation r [-], and symmetry "sym" [-]) over separation distance for the **Borden** hydraulic conductivity field.

3.2.2 Empirical Geostatistics on the North Bay Data-Set

The above mentioned measures of spatial dependence for the North Bay data-set are plotted on Figures 3.9b and 3.9a in horizontal and in vertical direction, respectively.

It would be relatively difficult to estimate a theoretical semivariogram to the North Bay data-set based on not transformed marginals. In horizontal direction, the range would be difficult

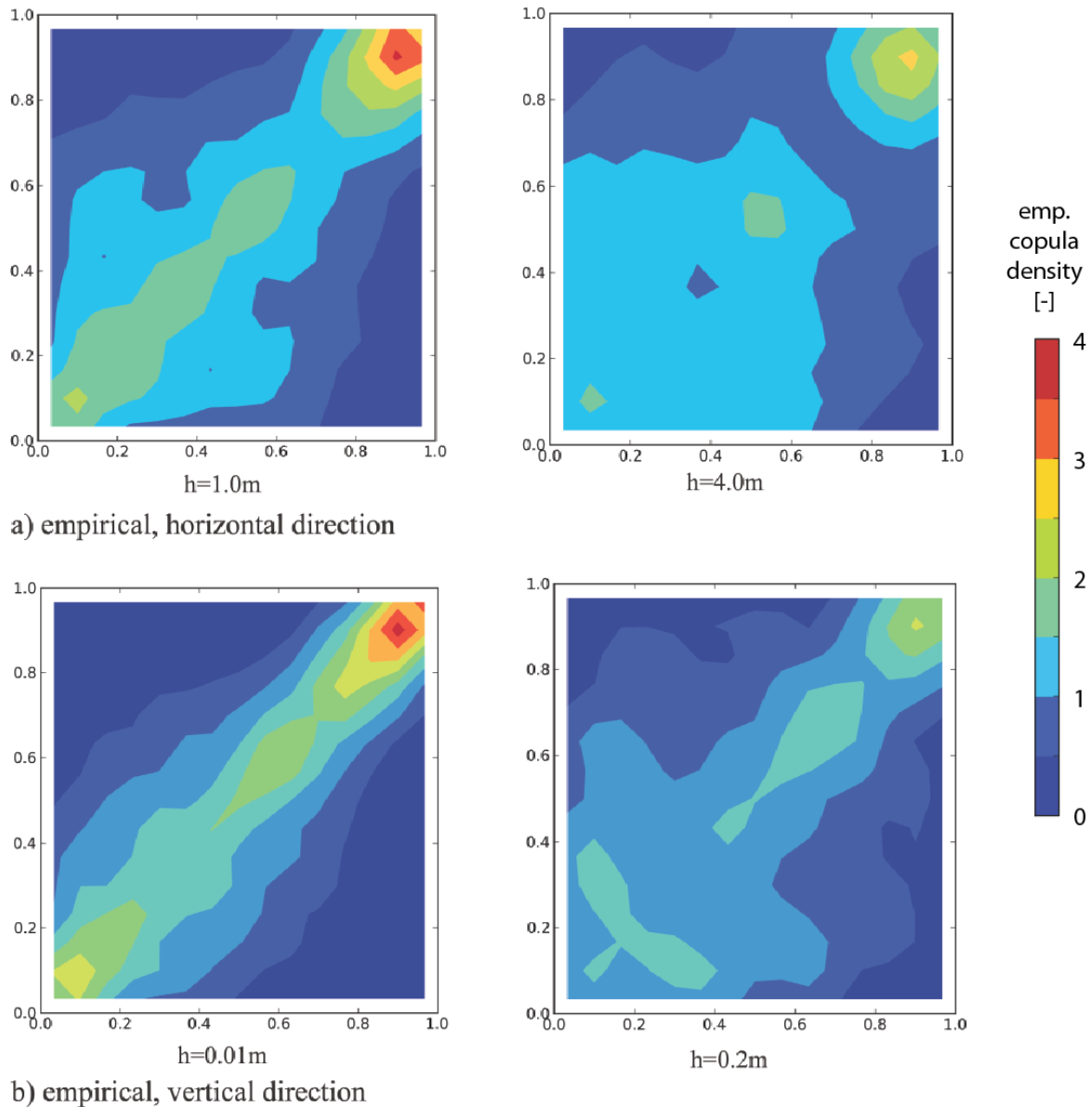


Figure 3.8: Bivariate empirical copula densities for different separation distances based on the Borden data-set of hydraulic conductivity. Panels show a) horizontal, b) vertical densities.

to estimate because the empirical semivariogram is increasing. The vertical semivariogram reaches two plateaus.

The empirical rank correlation are different for Line 1 and for Line 2, similar as in the Borden data-set, but more pronounced. This property could be modelled by using a different anisotropy ratio for both directions, however it was ignored for this study. The anisotropy ratio of range in horizontal direction divided by range in vertical direction based on rank correlation is $14 \text{ m} / 2.5 \text{ m} = 5.6$.

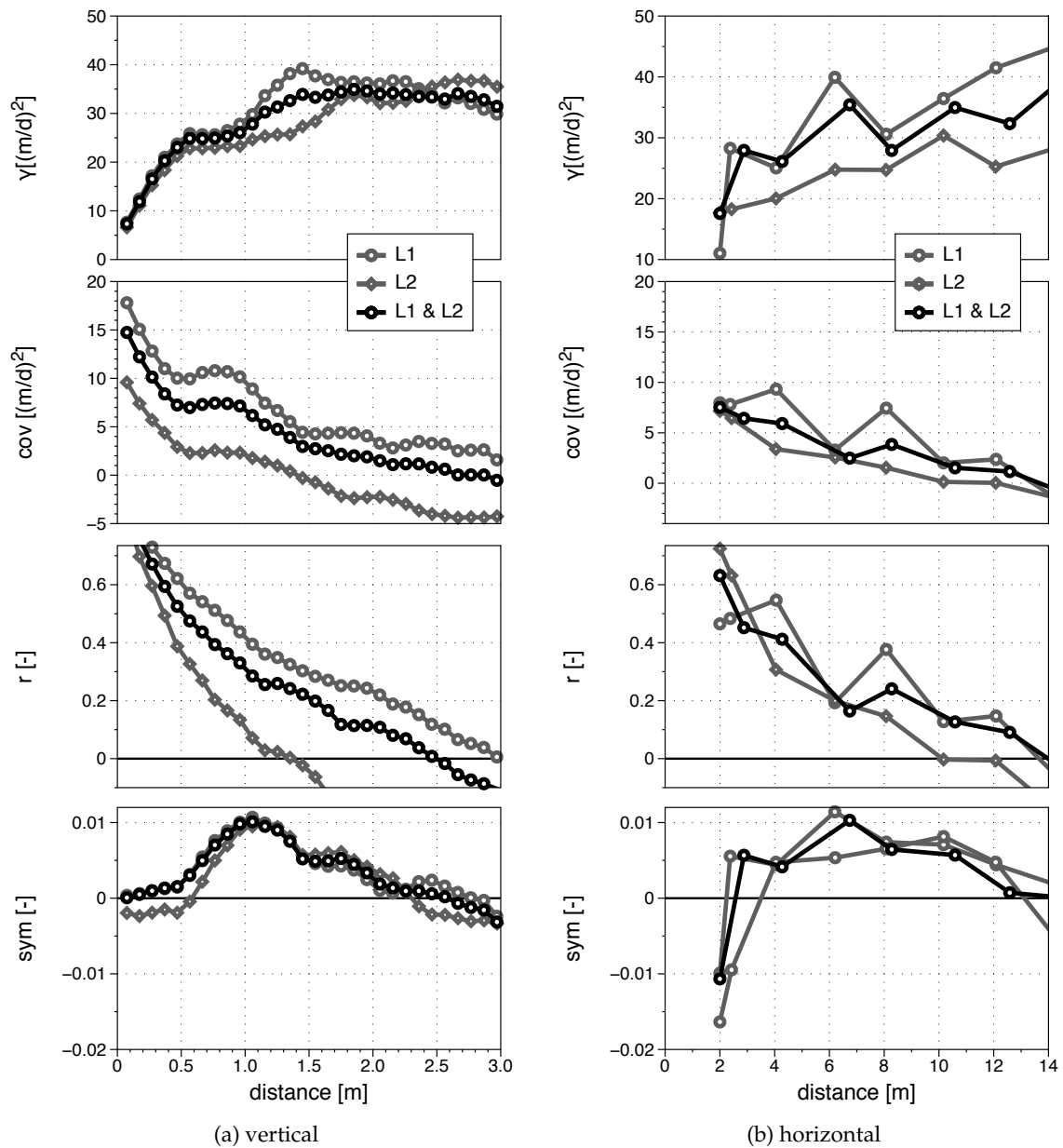


Figure 3.9: Measures of spatial dependence (semi-variance $\gamma[(m/d)^2]$, covariance “cov” $[(m/d)^2]$, rank correlation $r[-]$, and symmetry “sym” $[-]$) over separation distance for the **North Bay** hydraulic conductivity field.

The measure of symmetry starts to be negative for small separation distances and continues to be positive for larger separation distances. This behaviour is more pronounced in horizontal direction than in vertical direction, and it is much more pronounced in the data-set based at North Bay than in the one based at Borden. This property is not clearly visible in the smoothed contour plots of the empirical bivariate copula densities (Figure 3.10)

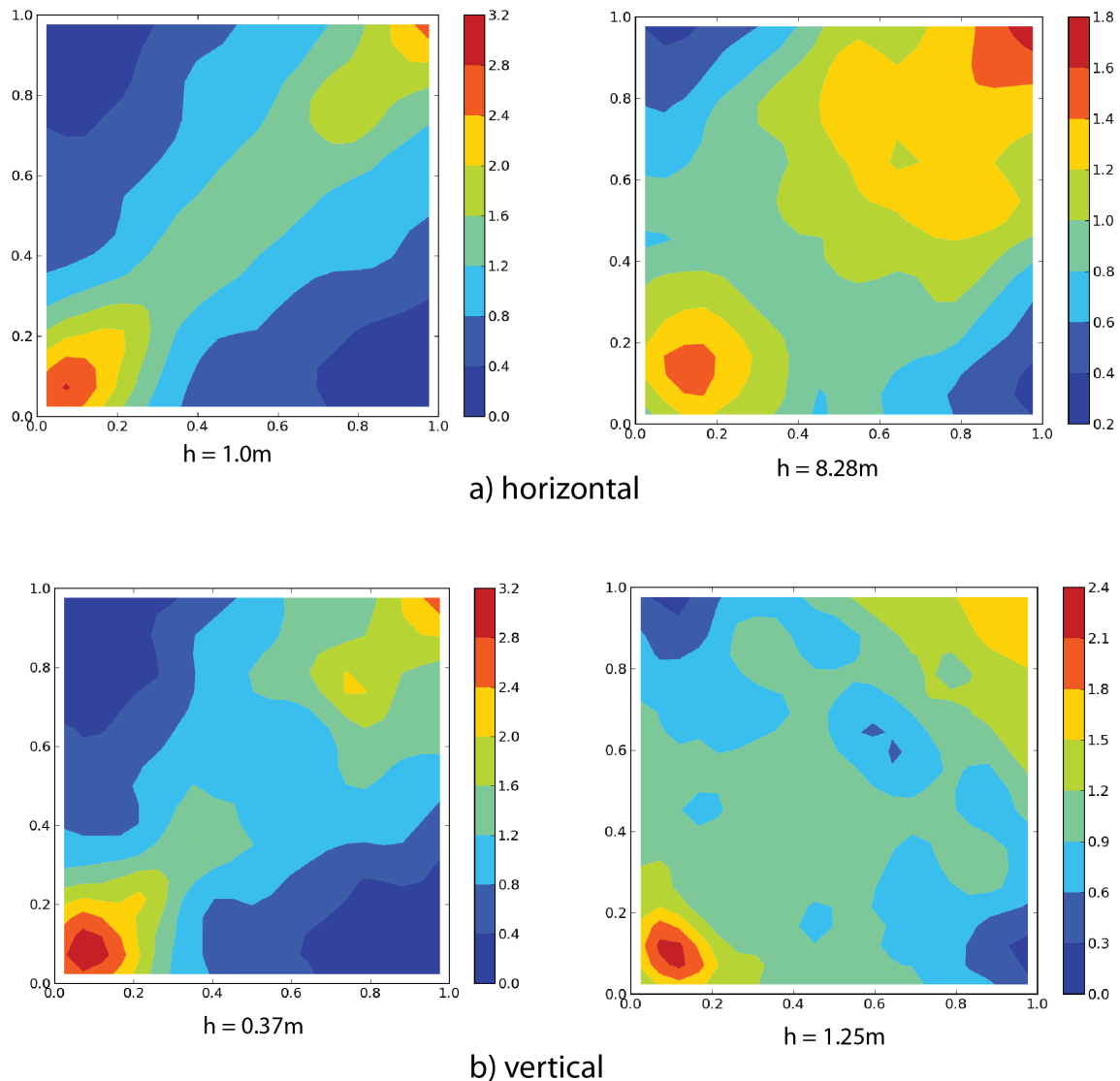


Figure 3.10: Empirical bivariate copula densities in a) horizontal and b) vertical direction for different separation distances based on the North Bay data-set of hydraulic conductivity.

3.3 Parameter Estimation

The parameters of a given theoretical copula model need to be chosen such that the model fits to the observed data. “Fit” means foremost that the chosen theoretical copula model is to capture the spatial dependence structure of the data-set. The dependence-structure exists in high dimensions, essentially as high as the number of measurement locations. And theoretically, the fit should be good over all dimensions. The margins of a multidimensional copula model should fit to the bivariate empirical copula densities (Section 3.2.1). On the other hand, the fit does not necessarily need to be “perfect” on the margins, because the fit

should be good in all dimensions.

A Gauss copula was fitted to both data-sets for comparison purposes. A v-copula was fitted to the Borden data-set, while a fit to the North Bay data-set gave no satisfying results. A maximum Gauss copula model was fitted to the North Bay data-set, as well as to the Borden data-set. In the following, variograms and superpositions of variograms are given in the form "*sill · type (range)*"

3.3.1 Gauss Copula

The best fitting correlation structure for isotropic Gaussian fields of Borden was found to be $0.35 \cdot \text{nugget}(0.0) + 0.33 \cdot \text{exponential}(3.4) + 0.32 \cdot \text{gaussian}(9.4)$, which corresponds well with the model used by [Sudicky \(1986\)](#) and has an exponential component as stipulated by [Ritzi and Allen-King \(2007\)](#).

The isotropic Gaussian model for North Bay was determined to be $0.2 \cdot \text{nugget}(0.0) + 0.8 \cdot \text{sph}(14.0)$. Anisotropy was modelled geometrically by scaling the vertical coordinates with the previously determined factors of anisotropy (Section 3.2).

3.3.2 V-Copula

The v-copula model was fitted to the Borden data-set successfully. From its parameters, only the variogram parameters were optimized, while the copula density is calculated for a variety of m, k combinations. The resulting likelihood surface can be visualized (Figure 3.11). The optimal parameters were found to be $m = 0.75, k = 1.1$, corresponding to the location of the maximum of the contoured surface.

[p]

The underlying Gaussian structure of the hydraulic conductivity field before the v-transformation according to Equations 2.49 and 2.50 was found to be $0.2 \cdot \text{nugget}(0.0) + 0.8 \cdot \text{spherical}(20.3)$. Figure 3.12 shows that the second order moments are identical for both the Gaussian and v-copula based simulated fields.

For comparison, additional transformations were evaluated: 1) a transformation that corresponds to a chi-shaped transformation ("*chi*", $m = 0.0, k = 1.0$), 2) an extreme transformation ("*extr*", $m = 0.75, k = 10.0$), and 3) for each transformation, the inverse transformation (indicated by a minus sign). In all transformations, the final fields are not distinguishable by their first two moments due to the above described transformations of the rank correlations ([Guthke, 2010](#)). Table 3.2 gives an overview of the v-transformation parameters used.

3.3.3 Maximum Gauss Copula

A bivariate maximum Gauss copula density function based on two bivariate normal Gaussian distribution functions $F(x_1, x_2)$ and $G(x_1, x_2)$ depends on a number of parameters.

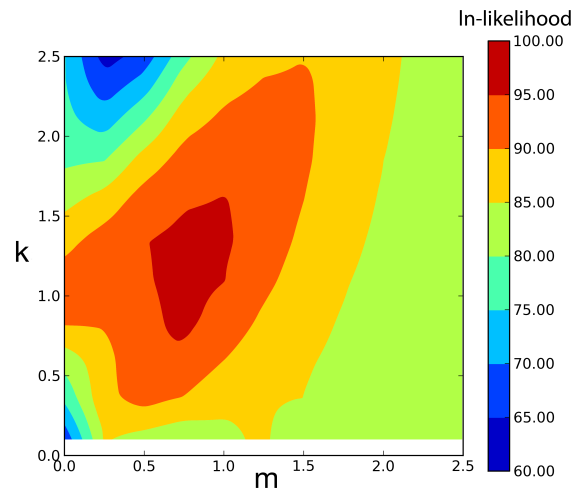


Figure 3.11: Contour plot of the ln-likelihood function L that was optimized for various parameter-sets of m, k .

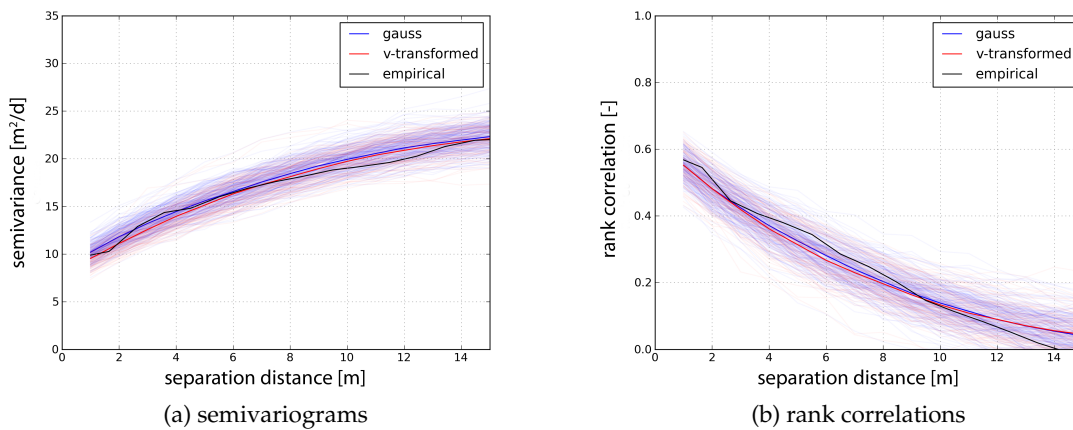


Figure 3.12: Semivariograms and rank correlations for the simulated spatial fields with Gaussian and non-Gaussian dependence.

Table 3.2: List of the v-copula dependence models, including their hidden covariance structure. A superposition of multiple variograms is given in the form “*sill · type (range)*”.

model	m	k	underlying Gaussian model
g	$\rightarrow \infty$	–	0.35 Nug(0) + 0.33 Exp(3.4) + 0.32 Gau(9.4)
v	0.75	1.1	0.20 Nug(0) + 0.80 Sph(20.3)
chi	0.0	1.0	0.13 Nug(0) + 0.22 Exp(4.9) + 0.66 Gau(14.9)
extr	0.75	10.0	0.35 Nug(0) + 0.13 Exp(3.5) + 0.77 Gau(14.7)

Mean values ($\mu_{F1}, \mu_{F2}, \mu_{G1}, \mu_{G2}$), variances ($\sigma_{F1}, \sigma_{F2}, \sigma_{G1}, \sigma_{G2}$) and correlation coefficients (ρ_{F12}, ρ_{G12}) add up to a total of ten parameters. This seemed like a complex optimization to start with, hence the number of parameters was decreased as much as possible for initial optimization, and later slowly increased. Viable options to reduce the number of parameters include:

- identical margins for each process: $\mu_{F1} = \mu_{F2}, \mu_{G1} = \mu_{G2}, \sigma_{F1} = \sigma_{F2}, \sigma_{G1} = \sigma_{G2}$;
- leave one of the two processes as a “standard normal process”, f.ex.: $\mu_{F1} = \mu_{F2} = 0$ and $\sigma_{F1} = \sigma_{F2} = 1$;

These simplifications have led to a total of four parameters $\mu_G, \sigma_G, \rho_{F12}, \rho_{G12}$.

The correlation coefficients depend on the covariance matrix, which in a spatial setting varies with the separation distance between points considered. As in the estimation procedure for the v-copula model, here is also a hidden covariance model using a combination of nugget (Nug) and spherical variogram is employed. The variance is standardized, hence only one the range (a) and one additional value, either nugget or sill of the spherical model have to be optimized.

The final smallest number of parameters with which the maximum Gauss copula model works, including the hidden covariance model, is six: $\mu_G, \sigma_G, Nug_F, a_F, Nug_G$ and a_G . The optimization tried to maximize the variance of one process. Usually, the margins of one process were fixed, hence the variance of the other process was maximized. Optimization results were considered only then as good results, if the bounds were not reached.

The maximum Gauss copula model is quite flexible and can incorporate anisotropy, even different anisotropy for each of the underlying Gaussian processes. These parameters of anisotropy can be included in the optimization procedure, with the drawback that the number of parameters increases. These parameters include stretch factors s , possibly one in each direction of each process and rotating factors, indicated by small case greek letters, possibly one less than the number of dimensions involved. Table 3.3 lists the four sets of parameters that were at different stages incorporated into the parameter estimation of maximum Gauss copulas.

Table 3.4 gives an overview of the different parameter optimizations performed on the Borden data-set using a maximum Gauss copula (Rau, 2011). Optimizations 1a), 1b) and 1c)

Table 3.3: List of transformation parameters for different scenarios using maximum Gauss copulas

dimension	F	G	number parameters
AA' & BB'	$s_x, s_y, s_z, \alpha, \beta$	$s_x, s_y, s_z, \alpha, \beta$	10
AA'	s_x, s_z, α	s_x, s_z, β	6
AA'	—	s_x, s_z, α	3
AA'	—	s_z, α	2

were performed on only cross-section AA', whereas optimization 1d) was performed on the entire data-set

- a) all possible twelve parameters used, unstable;
- b) limited number (2) of transformation parameters used; stable solution; good fit to empirical rank correlation
- c), cc) no coordinate transformations. In cc) also σ_F optimized. The optimized parameter set for scenario 1cc) was further refined using a brute-force approach, which yielded no further increase in likelihood;
- d) all data of the Borden data-set used, no anisotropy;

	L	means, std.devs		nuggets		ranges		stretch factors			rotations			
		σ_F	μ_G	σ_G	n_F	n_G	a_F	a_G	$s_{x,F}$	$s_{z,F}$	$s_{x,G}$	$s_{z,G}$	α_F	α_G
1a)	-11.70	-	-0.65	50.	0.12	7.42	8.5	16.80	0.09	2.91	0.01	2.33	0.34°	9.17°
1b)	-10.46	-	-1.09	50.	0.24	7.56	10.45	15.15	-	-	-	19.61	-	9.00°
1c)	-7.02	-	-1.13	40.	0.17	5.96	10.16	16.5	-	-	-	-	-	-
1cc)	-7.25	35.75	1.5	0.23	4.72	0.025	16.6	8.3	-	-	-	-	-	-
AA' & BB'	-13.75	-	0.49	8.64	0.2	1.73	12.6	13.5	-	-	-	-	-	-

Table 3.4: Results of parameter optimization using a maximum Gauss copula based on the Borden data-set. Optimization 1a)-1c) are based on data of cross section AA' only while optimization 1d) is based on all available data.

The optimized parameters are listed in Table 3.4. In scenario 1cc), the variance of both processes was allowed to vary during the optimization, then also one variance, σ_F , was maximized while the other variance, σ_G , was optimized to a small magnitude. With the maximum number of parameters used in scenario 1a) no satisfying solution could be found. Scenarios 1b) and 1c)/1cc) resulted in reliable solutions, where scenario 1b) includes some anisotropy.

The Akaike information criterion (AIC, Akaike (1974)) can provide support for the decision, which model should be preferred. The AIC is calculated using Equation 3.7. There exists a version of the AIC that attempts to correct for sample size k (AICc, Equation 3.8).

$$\text{AIC} = 2k - 2 \ln(L) \quad (3.7)$$

$$\text{AICc} = \text{AIC} + \frac{2k(k+1)}{n-k-1} \quad (3.8)$$

Generally, an increase in parameters leads to a better fit of the model. However, the most parsimonious approach should be chosen, for which an increased number of parameters does not lead to an increase in the model fit. The AIC does not provide a test. The preferred model is the one with the smallest AIC or AICc value. For the scenarios that were fit to set of data from cross-section AA', the sample size remained constant ($k = 239$). For these scenarios, the AIC and AICc values are listed in Table 3.5. Both the AIC and the AICc indicate the fitted model of scenario 1b) to be optimal, while 1c) and 1d) are close second. The model of scenario 1a) with all the possible parameters for anisotropy using the maximum Gauss copula performed worst by far.

scenario	L	k	AIC	AICc
1a	120572	12	0.60	1.98
1b	34892	8	-4.92	-20.92
1c	1119	6	-2.04	-14.04
1cc	1408	7	-0.50	-14.50

Table 3.5: Akaike criteria for maximum Gauss copula scenarios based on the Borden data-set.

The optimized parameters of the maximum Gauss copula model for the North Bay data-set are given in Table 3.6. Two optimizations were performed using the most parsimonious model, one for the hydraulic conductivity data based on Line 1 only (scenario 2a) and one for the entire data-set (scenario 2b).

3.4 Comparison of Theoretical Copula Models With Data

The parameters of a copula model are determined via the maximum likelihood approach outlined in Section 2.3.4. The goal of this section is to show possibilities to evaluate the fit of

Table 3.6: Optimized parameters for the maximum Gauss copula model for the K data-set at North Bay.

data-base	scenario	μ_G	σ_G	n_F	n_G	a_F	a_G
Line 1	2a)	4.03	25.00	0.10	2.50	30.00	21.63
Lines 1 & 2	2b)	3.44	40.00	0.00	1.15	5.65	15.80

the theoretical copula models with the optimized parameters to the empirical data.

One basic approach is to compare visually the empirical bivariate copula densities based on the optimized parameters with empirical bivariate copula densities (Section 3.4.1). A second possibility is to compare the measure of symmetry. Using a bootstrap analysis, confidence intervals can be calculated within which the measure of symmetry should fall for a given copula model (Section 3.4.2).

The parameter estimation approach is based on high dimensions, hence the fit of the theoretical models should be evaluated in higher dimensions than two as is the case in the previous two possibilities. The fit in higher dimensions can be evaluated by calculating likelihood values based on sets of data of dimensions higher than two (triplets instead of pairs, f.ex) which is shown in Section 3.4.3.

3.4.1 Comparing Bivariate Copula Densities

The fit of the v-copula model to the Borden data-set based on bivariate copula densities is exemplarily shown on Figure 3.13. Shown are the bivariate v-copula densities for two correlations that correspond to separation distances of the empirical bivariate copula densities shown on Figure 3.8. The dependence of the v-copula is non-symmetric with strongest dependence for high quantiles. The dependence structure flattens the bigger the separation distance is. From visual inspection it can be seen, that the v-copula with its non-symmetric structure fits better to the empirical bivariate copula density than a symmetric Gaussian bivariate density would.

The fit of the maximum Gauss copula model to the Borden data-set is shown on Figures 3.14 and 3.15, for the subset based on cross-section AA' only, and from both cross-sections, respectively. One of the most interesting features of the theoretical maximum Gauss copulas fitted to the subset from cross-section AA' is that they all, independent of the number of parameters used in the optimization, exhibit three zones of high dependence, one in each corner of the bivariate copula density corresponding to the extremes, and a third area around the 0.7 quantile, leading to non-symmetric dependence. Such a third area of high dependence is not visible on the theoretical copula density fitted to the entire Borden data-set (Figure 3.15), which is the one that is closest to be symmetric among the fitted non-Gaussian copulas.

The maximum Gauss bivariate copula densities fitted to the North Bay data-set are shown on Figures 3.16 and 3.17. Both exhibits peaks of dependence in high and in low quantiles, however the peak in high quantiles is more pronounced and thus non-symmetric dependence is modelled corresponding also to the observed measure of symmetry.

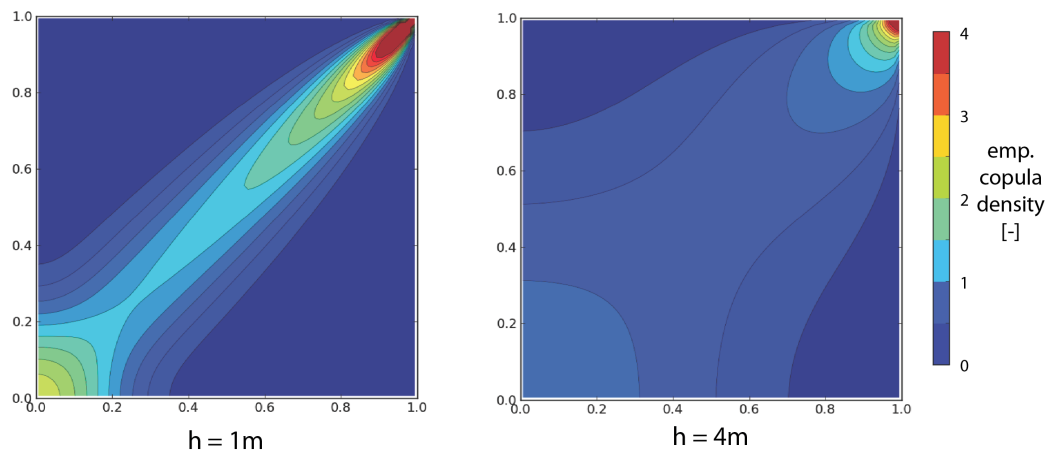


Figure 3.13: Theoretical bivariate v -copula density fitted to the Borden data-set. Shown are the densities whose correlations correspond to the empirical bivariate copula densities shown on Figure 3.8.

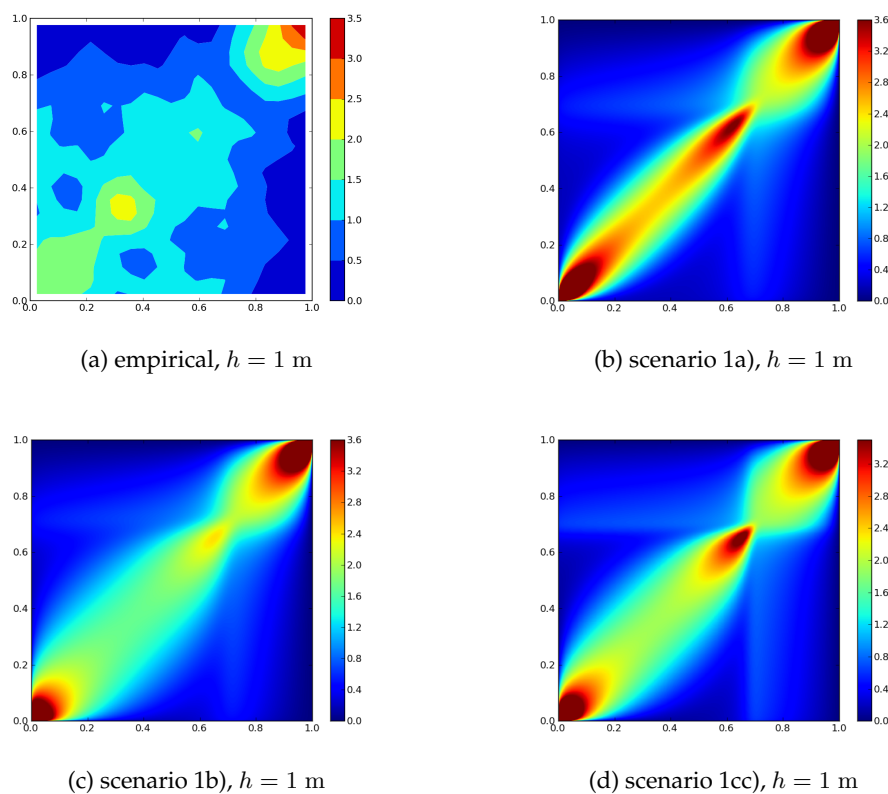


Figure 3.14: Bivariate maximum Gauss copula density functions based on cross section AA' of the Borden data-set.

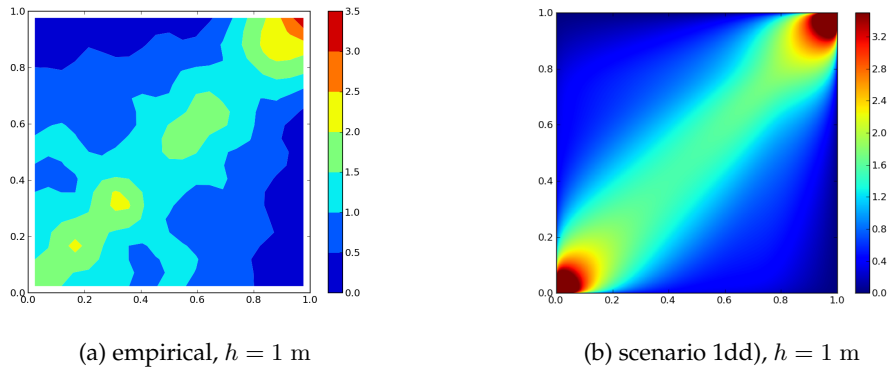


Figure 3.15: Bivariate maximum Gauss copula density functions based on the entire Borden data-set.

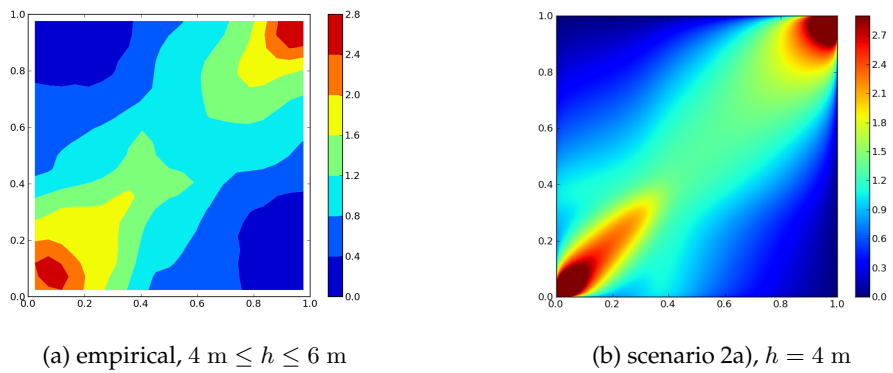


Figure 3.16: Bivariate maximum Gauss copula density functions based on cross-section Line 1 of the North Bay data-set.

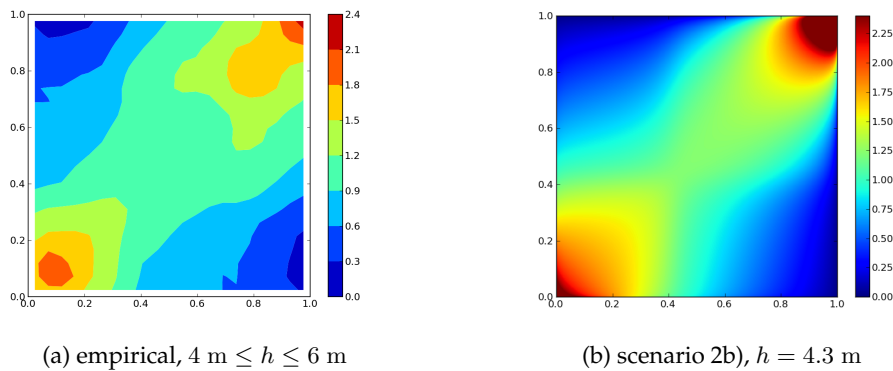


Figure 3.17: Bivariate maximum Gauss copula density functions based on the entire North Bay data-set.

3.4.2 Bootstrap Test of the Measure Symmetry

If the spatial dependence structures of the studied data-sets were purely Gaussian, the symmetry measure would be required to be zero, which was not observed. However, this deviation from zero could be by chance. Hence, a simple bootstrap method was employed to test if the symmetry observed in the one real realization can occur by chance with a Gaussian model. Two hundred sets of hydraulic conductivity values were simulated at locations where a measurement exists. For each simulated field, the symmetry was calculated for a given set of separation distances, both in vertical and in horizontal directions, resulting in an empirical distribution of calculated symmetry values for each class of separation distances. For this empirical distribution, the 90% confidence interval ($q_{95} - q_5$) was calculated, which is shown as the shaded areas in orange on Figure 3.18.

For both data-sets, both in horizontal- and vertical direction, at least for some separation distances, the empirical symmetry values lie outside the 90% confidence interval, within which they should lie if the spatial dependence structure was Gaussian.

The procedure was repeated for the fitted non-Gaussian structures, resulting in 90% confidence intervals shaded in blue for the v-copula model and in green for the maximum Gaussian copula models on Figure 3.18. In all cases, the confidence intervals based on non-Gaussian structures include bigger portions of the empirical symmetries than if based on Gaussian structures. Yet, none of the non-Gaussian models includes the entire empirical symmetries for all separation distances. There is still room for improvement.

3.4.3 Fit in Higher Dimensions

The goal of this subsection is to evaluate the fit of a theoretical copula model in dimensions higher than two. Instead of pairs of points, triplets of points are considered. The entire data-set is divided into triplets, for which the corresponding copula density is calculated using the optimized parameters. The log copula densities for each tile are summed up. The procedure is repeated using a different copula model. The model with the higher sum of log-copula densities can be regarded as the model that fits better to the data.

This procedure was performed comparing the v-copula and the Gaussian copula model on the full set of hydraulic conductivity measurements at Borden. In detail, it involved the following steps:

1. choose which direction is analyzed (horizontal or vertical)
2. start with one core (vertical direction) or three adjacent cores (horizontal direction)
3. select three adjacent measurements of K
4. convert these K values to uniform space
5. calculate the three-dimensional copula density (either v-copula or Gaussian copula based) that corresponds to this triplet of measurements

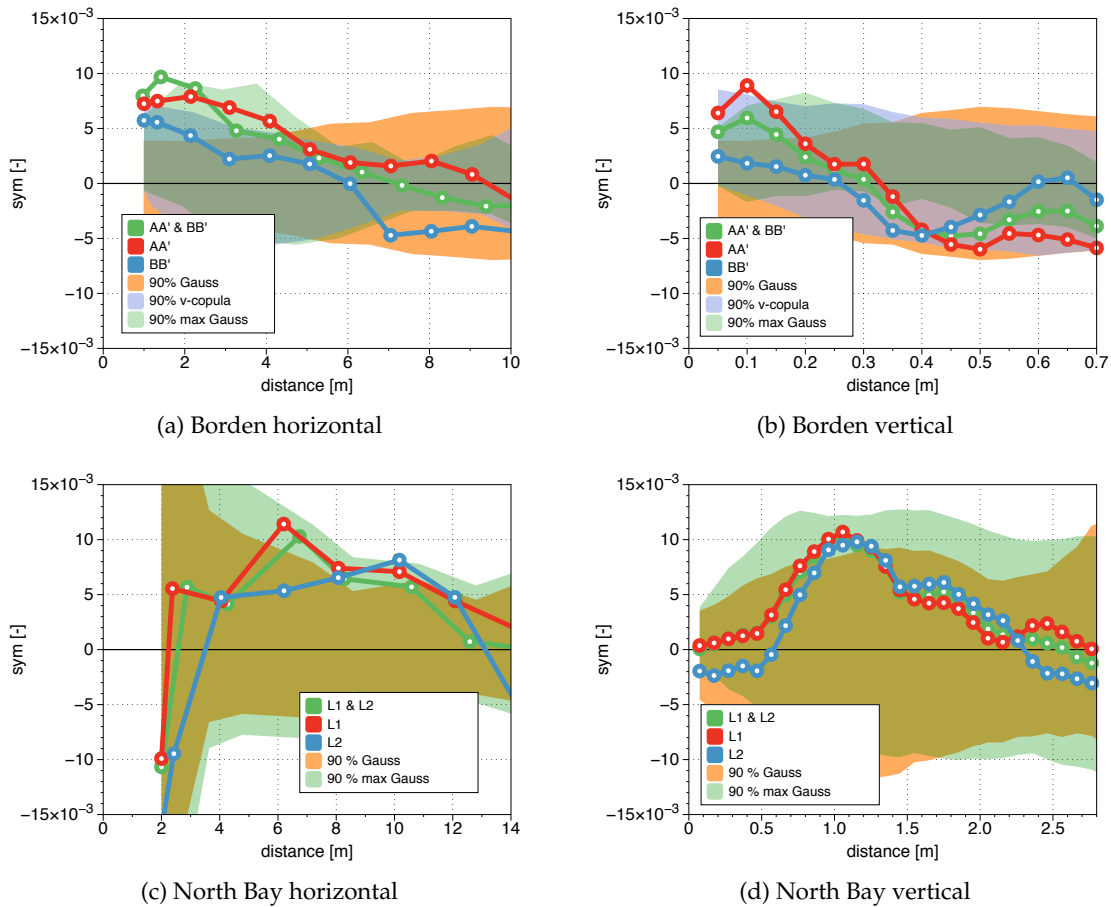


Figure 3.18: Bootstrap test comparing empirical symmetries of the Borden and North Bay data-sets with 95% confidence intervals of symmetries based on simulations using different copula models.

6. take the logarithm and store
7. proceed with the next triplet of K measurements in the given direction and repeat

The analysis was repeated for different separation distances. In this approach the distance between measurements that are used for the triplets was stepwise increased by leaving more and more adjacent measurements out. The number of measurements left out is referred to as “triplets separation”. This approach neglects that the distance between all points in the triplets should be identical.

The results of this evaluation are shown on Figure 3.19. The likelihood is higher for the v -copula model both in horizontal and vertical direction compared to the Gaussian copula model, especially for triplets where the three points are close together. This means, that also in three dimensions, from where on the copula densities are difficult to contour, the v -copula model fits better to the Borden data than the Gaussian model.

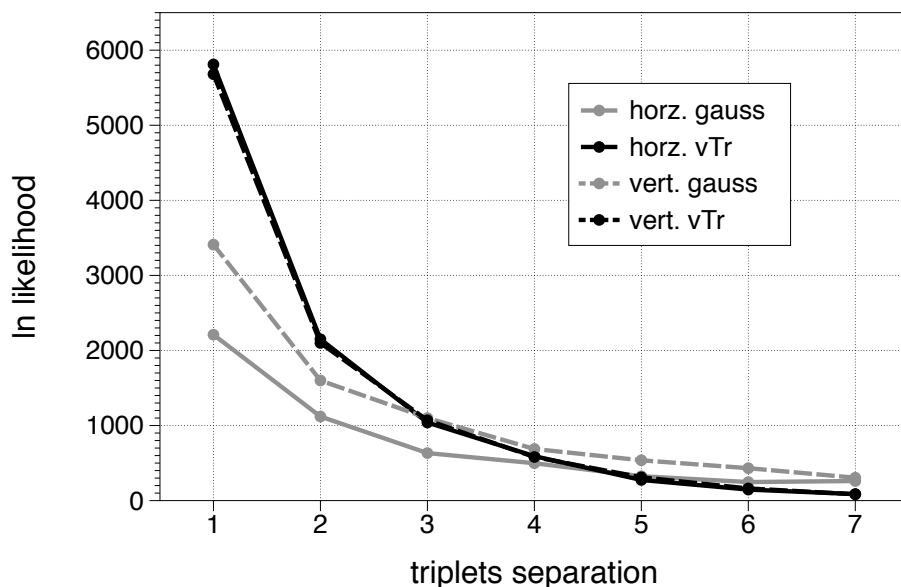


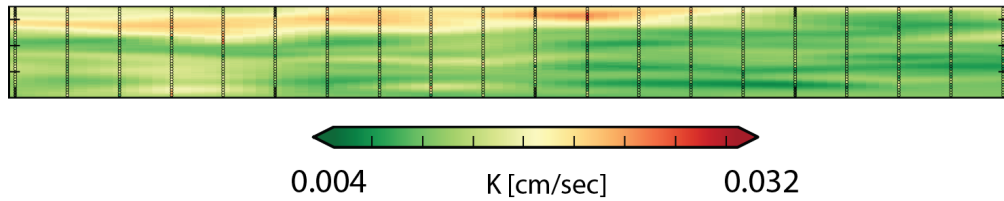
Figure 3.19: Likelihood obtained for the Gaussian and the v-copula model fitted to Borden. The likelihood is calculated using triplets in either horizontal or vertical direction, according to the geometrical setup of the Borden data-set.

3.5 Spatial Interpolation and Simulation

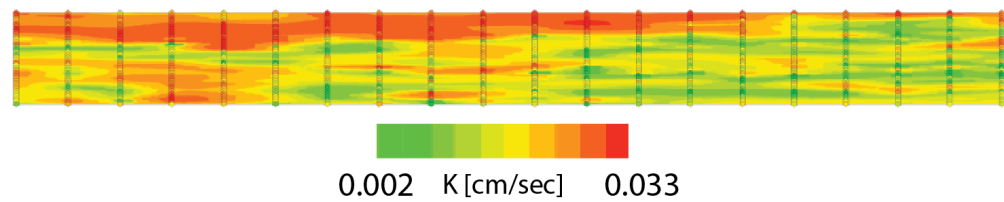
The properties of copula-based spatial interpolation are demonstrated along cross-section AA' of the Borden data-set (Figure 3.20). Copula based interpolation is shown on panel *a* and can be compared with interpolation using Ordinary Kriging on panel *b*. For the copula based interpolation, the arithmetic average of the distribution at every estimation point is used. There are no significant differences between the two interpolation schemes visible. [Bárdossy and Li \(2008\)](#) have shown, that cross-validation results tend to be favourable for copula based interpolation. One major advantage of using copulas for interpolation purposes is visible from the uncertainty estimates of the interpolation. The estimation standard deviation associated with Ordinary Kriging reflects mostly the geometry of the observation network (panel *c* on Figure 3.20), while the 90% confidence interval based on the v-copula interpolation (panel *d* on Figure 3.20) is largest and hence has the highest uncertainty in areas where the gradient is highest. In this case this is the zone of high hydraulic conductivity values in the top portion of the cross-section.

With the same theoretical copula models that were fitted to data and that were used for interpolation, fields of spatially distributed parameters can be simulated. Realization of such fields based on Gaussian, v-copula, and maximum Gauss copula on the entire Borden data-set are shown on the left hand side of Figure [reffig:SimFieldsBorden](#). Bivariate copula densities of the copulas responsible for each field for a separation distance of 1 m are shown adjacent to the right of each field on Figure 3.21. The Gaussian field (panel *a*) shows the typical isolated patches (“blobs”) of high- and low- K values. Such blobs exist for high- K

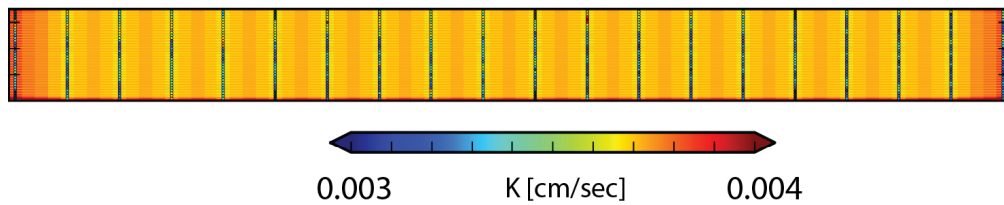
a) interpolation: ordinary Kriging



b) interpolation: using v-copula



c) ordinary Kriging estimation standard deviation



d) confidence interval q90-q10 using v-copula

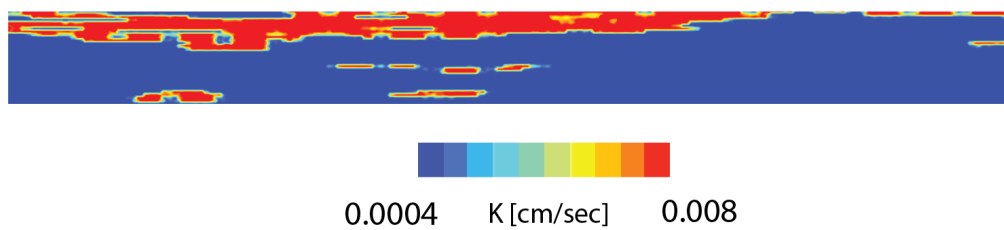


Figure 3.20: Comparison of Kriging and v-copula based interpolation and associated uncertainty along cross-section AA' at Borden.

values also in the v -copula based fields (panel *b*) and the maximum Gauss copula based fields (panel *c*). However, all three types of fields differ most significantly by the degree of dependence in low- and medium quantiles.

All three types of fields were designed such that their rank correlations are identical. The remainder of this thesis explores and tries to quantify if this different spatial dependence as modelled in the spatial copula has an effect on a physical property that is based on the spatially distributed fields of hydraulic conductivity.

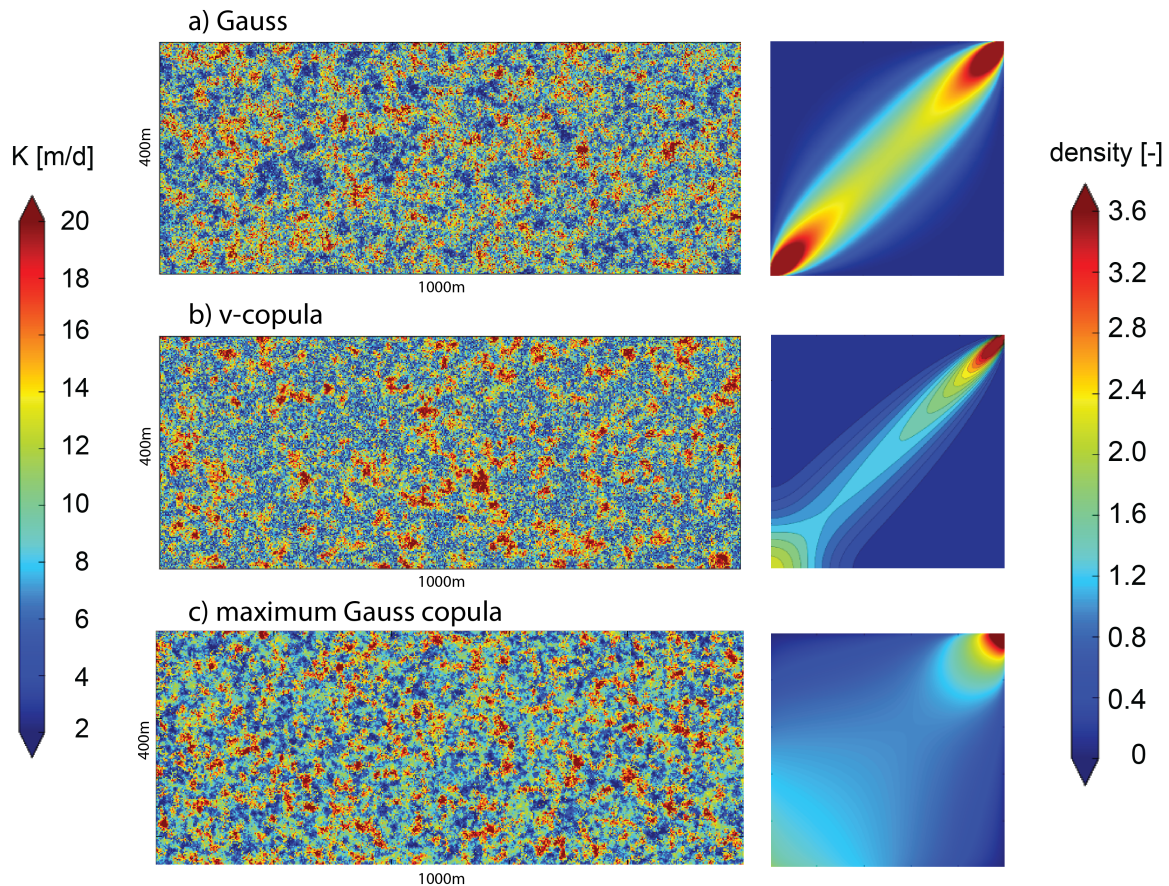


Figure 3.21: Three simulated spatially distributed K fields fitted to the Borden data-set. All have the same marginal distribution and rank correlation. They differ in their spatial copulas. The bivariate copula densities for a separation distance of 1 m is show on the right panel, adjacent to the spatially distributed field.

4 Implications on Solute Transport

One goal of this thesis is to explore if the spatial dependence structure of a hydraulic conductivity field as modelled by Gaussian and by non-Gaussian copulas influences a physical property, such as plume evolution as evaluated by second central moments of concentration fields. Two types of hydraulic conductivity fields, non-Gaussian copula models and the Gaussian copula model, were constructed such that their second-order spatial moments are identical, and hence they can not be distinguished by second-order based geostatistics.

Despite identical second order moments, the copulas of the spatial fields are not identical. Copulas are so flexible, that for a given correlation, different kinds of copulas can exist, for example a Gaussian copula and non-Gaussian copulas. And different copulas imply a different type of dependence, even though the correlation is identical. This property is illustrated on the three panels of Figure 4.1 which show different types of bivariate dependence, even though all three types have the identical correlation of $r = 0.88$.

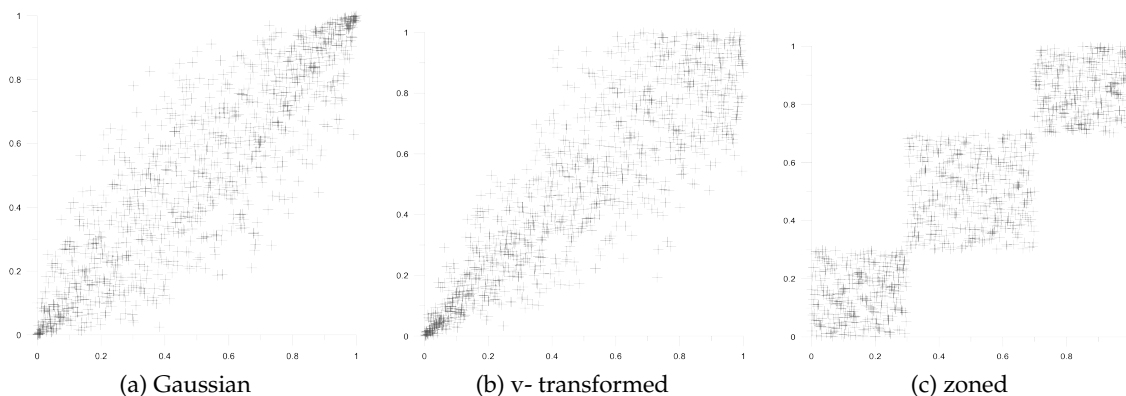


Figure 4.1: Different type of dependence as modelled by three different copulas. All three types of dependence have identical correlation coefficient of $r = 0.88$.

The previous chapter has shown that copulas are a full stochastic model. Their parameters can be optimized such that the copula fits well to a multidimensional dependence-structure and such that they fit better to observations than using a Gaussian dependence structure. The copula approach offers two key advantages. The first advantage is that the shape of the marginal distribution does not influence the dependence structure. This means that neither possible extreme values influence the dependence structure. The second key advantage is that a full conditional distribution function can be calculated at every location. This conditional distribution function can be non-symmetric and is a very powerful tool for evaluating uncertainties of interpolation or in the data-set. This can be very useful for sampling network design (Li et al., 2011).

The spatial fields in this thesis were designed to have identical (rank-) correlations, hence the eyes of an observer who only knows about second order moments based geostatistics can not detect these differences in spatial dependence as expressed by copulas.

The migration of solutes in heterogeneous aquifers is evaluated using numerical tracer tests (Frind et al., 1987; Smith and Schwartz, 1980; Naff et al., 1998). This stochastic modelling concept casts macroscopic dispersion as a process related to the spatial heterogeneity in hydraulic conductivity. The numeric model represents a realistic analog of the actual process and does not require the use of large dispersivity values to generate macroscopic dispersion (Davis, 1986). A slug of conservative tracer is instantaneously injected and transported within a steady-state flow field. The resulting concentration field is recorded and its spatial moments analyzed.

By classical theory (Bear, 1972; Gelhar and Axness, 1983), the slope at the linear proportion of the second central spatial moment of a concentration field can be related to the dispersion coefficient. The advective spreading of a solute migrating in the subsurface is influenced by the groundwater velocity field, which in turn is directly impacted by the spatial distribution of K . Due to this spreading, portions of the plume advance more rapidly than the average velocity in some zones, while in other zones migration rates are slower than the average velocity. This spreading phenomenon is commonly referred to as macrodispersion. Here, solute spreading will be examined in both two- and three-dimensional high-resolution K fields.

The evaluation of the solute migration is performed in a Monte Carlo style, taking the spatial variability of the flow parameters into account and the inability to describe a heterogeneous conductivity distribution exactly, thus enabling a statistical and hence more conclusive evaluation of the spreading process. Many realizations of K fields were generated for each type of spatial dependence of K , and for each K field, a numerical tracer experiment was conducted. For the purpose of this paper, local dispersion is reduced as much as numerically possible and the modelling domain is finely discretized in order to avoid numerical dispersion and oscillations.

The following two sections describe the setup of the numerical model necessary for the numerical tracer tests (Section 4.1), and the procedures necessary to evaluate the spreading behaviour of solute mass (Section 4.2). These two steps are employed in the sections describing solute transport behaviour based on Gaussian and non-Gaussian K -fields in the aquifers at Borden (Section 4.3) and at North Bay (Section 4.4). Section 4.5 takes another look at the Borden data-set trying to assess its homogeneity.

4.1 Groundwater Flow and Solute Transport Model

Saturated groundwater flow through the constructed K fields was simulated under steady-state conditions and the transport of a conservative solute body from an instantaneous release is used to contrast the spreading behaviour for both the Gaussian and non-Gaussian structured K fields. The robust numerical model HydroGeoSphere (Therrien and Sudicky,

1996) is used to solve standard flow and advection-dispersion equations with careful attention paid to space and time discretization issues to ensure solution accuracy.

The domain size was chosen to be long enough in all dimensions such that the plume analysis, especially the calculation of the macrodispersion coefficient at late time (Section 4.2) is done before the plume touches a boundary of the domain (allowed a maximum of 1% of current maximum concentration at any boundary node). The grid spacing needed to fulfill two criteria. First, it is fine enough to avoid numerical dispersion and oscillation which was ensured by grid refinement tests. Second, one spatial correlation length contains at least five grid elements in all dimensions. The analysis was performed in two dimensions ($\sim 200,000$ nodes, 200 Monte Carlo runs for each type of field) and three dimensions (~ 3.75 million nodes, 100 Monte Carlo runs for each type of field). With the help of several test runs it was found that the domain size had to be at least ten, preferably twenty times the correlation length in the main direction of flow and at least six times the correlation length perpendicular to the main direction of flow. Three-dimensional analyses of groundwater flow was considered necessary, especially in non-Gaussian K fields, due to the possibility of connected zones of high K or low K barriers to flow. Basic model parameters are listed for the two- and three-dimensional simulations in Table 4.1, the setup of the domain with boundary conditions is shown on Figure 4.2.

Table 4.1: Parameters of the numerical models used for two- and for three dimensions.

		2D	3D
domain length [m]	lx	1000	75
	ly	400	10
	lz	-	4
number elements [-]	nx	500	375
	ny	200	50
	nz		200
grid spacing [m]	dx	2	0.2
	dy	2	0.2
	dz	-	0.01
hydraulic gradient [-]	$i = dh/dl$	0.1	0.0045
local dispersivities [m]	$\alpha_{l,h}$	2	0.21
	$\alpha_{t,h}$	0.2	0.021
	$\alpha_{t,v}$	0.02	0.0021

The steady-state flow model is a steady-state system with constant head boundary conditions at the left and right boundary. All other domain boundaries were chosen to be Neumann boundaries with zero flow conditions. The key component of the flow model is the spatially-distributed K field.

In the solute transport model, the initial concentration at all nodes was set to zero, except at one node at the left side of the domain which has a unity concentration assigned to represent

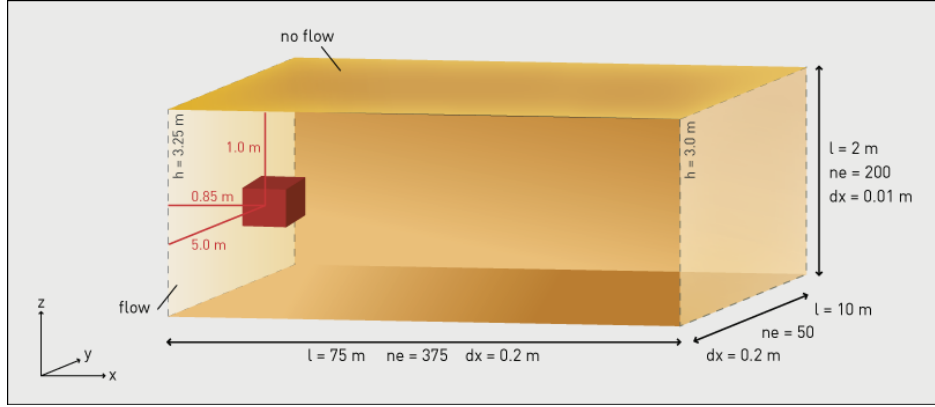


Figure 4.2: Sketch showing the dimensions (length l), the discretization (number of elements ne and grid spacing dx), the initial and boundary conditions for the three-dimensional domain for the numerical tracer tests.

a slug injection of solute. The K values within a radius of 5 grid lengths around the injection node were set to the geometric mean of all K values of the domain, to avoid that the early time plume behaviour significantly influences absolute travel times of the solute plume.

The pore geometry in real porous media varies in space, which results in a heterogeneous flow field. Variability in the flow field causes an additional deformation of the transported solute plume. This spreading due to velocity variability at different scales is called dispersion. Two dispersion processes can be distinguished: hydrodynamic dispersion (mechanical dispersion plus diffusion) and macrodispersion.

Structures in the conductivity field that are smaller than the grid cell size are not explicitly resolved from a numerical model. Therefore, dispersive effects on this Darcy scale are considered within the mathematical model. Dispersion processes are modelled to have the same influences on a solute plume as molecular diffusion due to brownian motion, only on another scale: A spreading of the solute plume in every direction. Therefore, both processes are summarized in the hydrodynamical dispersion/diffusion coefficient D_{local} (Equation 4.1), where α is a characteristic length scale (dispersivity), $v = \frac{q}{\phi}$ is the groundwater velocity in flow direction with ϕ being the porosity, and D^* is the molecular diffusion coefficient ($D^* \rightarrow 0$).

$$D_{local} = \alpha v + D^* \quad (4.1)$$

Structures larger than the grid size are resolved within the numerical model. Therefore, such structures will result in a heterogeneous flow field. This leads to dispersion on a greater scale, or macroscale, due to the same reasons as explained above. The total dispersion in a groundwater flow and solute transport model with a heterogeneous K structure is therefore the sum of macrodispersion and hydrodynamic dispersion (Equation 4.2). While being interested in the effects of heterogeneity on D_{macro} , D_{local} is decreased as much as numerically possible.

$$D^l = D_{macro} + D_{local} \quad (4.2)$$

4.2 Analysis of Spatial Moments of Solute Plume

This paper adapts the notation of [Freyberg \(1986\)](#) to describe the zeroth, first, and higher spatial moments, and their relation to plume evolution and spreading rates. Spatial moments are integrated measures to describe complex plume behaviour. In three dimensions, the n -th moment is defined in Equation 4.3, where $n = i + j + k$ and $conc(x, y, z, t)$ is the concentration at a location x, y, z and at time t ; η is the porosity.

$$m_{i,j,k}(t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \eta \cdot conc(x, y, z, t) \cdot x^i y^j z^k dx dy dz; \quad i, j, k = 1, 2, 3 \quad (4.3)$$

The zeroth moment $m_{0,0,0}$ corresponds to the total solute mass in the system, and the first central moment in x-direction $m_{c1,0,0} = \frac{m_{1,0,0}}{m_{0,0,0}}$ corresponds to the x-coordinate of the center of mass. The ratio between second central moments and the zeroth moment is a measure for the bulk spatial spreading of the plume in each direction. The third and fourth standardized central moments are measures for the skewness and the kurtosis of the solute plume.

The analysis of the spatial moments is critical to this study, because the second central spatial moments are linked to the macrodispersion tensor as described in Equation 4.4. As the conservative tracer migrates through the K field, driven by a head gradient, it is expected that the second central spatial moment initially grows at a non-linear rate but reaches a linear behaviour at later time when the macrodispersion coefficient reaches an asymptotic value.

$$D_{i,j,k} = \frac{1}{2} \cdot \frac{d}{dt} [m_{i,j,k}(t) - m_{i-1,j-1,k-1}(t)] / m_{0,0,0}; \quad (i, j, k) = (2, 0, 0), (0, 2, 0), (0, 0, 2) \quad (4.4)$$

For every Monte Carlo simulation and for every time-step, the spatial moments were calculated and later averaged to eliminate dispersive artefacts related to the “dispersion” in plume centroid positions across realizations as pointed out by [Rajaram and Gelhar \(1993\)](#).

The set of Monte Carlo realizations for each type of K structure was analyzed produce an average concentration plume, and its uncertainty as described by the concentration variance. For multi-Gaussian fields, a symmetric variance plume is expected ([Vomvoris and Gelhar, 1990](#)).

4.3 Analysis of Solute Transport Behaviour Using Spatial Structures of the Borden Data-Set

In this work, two copula models were fitted to the Borden data-set: a Gaussian- and a non-Gaussian v-transformed copula. Both are designed such that they are not distinguishable by second-order moments (their spatial covariance functions) and they have identical marginal distributions. The impacts of these two types of spatial dependence structures of K on solute

transport behaviour will be tested in a series of detailed numerical tracer tests, evaluated using a Monte Carlo approach.

If a non-Gaussian spatial dependence structure of K , as modelled by a fitted copula, leads to a different solute transport behaviour than when it does based on Gaussian K or $\ln(K)$ assumptions, then this will have implications for solute transport in most other aquifers. Furthermore, if such deviations are found, other hydrogeological parameters that depend on the spatial structure of K could be impacted.

In the two-dimensional case, in addition to the Gaussian and the v -transformed K -fields, other transformations were performed. As before, the marginal distribution and the first two spatial moments remain what they were, but different copulas were used to evaluate the effect of different and partly more extreme spatial dependence on solute transport. For comparison, a field that corresponds to a chi-shaped transformation (“chi”) and an extreme transformation (“extr”) were analyzed. For each transformation, the inverse transformation (indicated by a minus sign) was analyzed.

4.3.1 Results of Solute Transport Behaviour in Two Dimensions

First it has to be pointed out that the velocity of the center of mass of all plumes modelled in two dimensions moves on average with the same velocity, independent of the type of spatial dependence structure that was used for simulating the K fields (Figure 4.4, bottom panel; m_{100}). If the average first central moments of a conservative plume are identical, the average K that represents the spatial K fields, should be identical. It was thought that the best measure of an average K in the given case is the K obtained when considering each Monte Carlo simulation as a Darcy experiment. The resulting empirical distributions of K , evaluated along the principal direction of flow, based on the Gaussian, the v -transformed, and the inversely v -transformed dependence structure is shown on Figure 4.3. Those distributions are fairly similar. However, the distribution of K values based on the v -transformed K field, the one that fits best to the Borden data, is slightly shifted towards higher K values for medium quantiles compared to the Gaussian K structure.

The spreading of the plumes as measured by the second central spatial moment over time exhibits a typical behaviour with non-linear growth in early times and linear growth in late time – independent of which spatial structure was used for the simulation of the K field (Figure 4.4, central panel, m_{200} ; Freyberg (1986)). However, the slope of the linear portions of the graphs is different. Most importantly, the Gaussian (thick dark blue) and the non-Gaussian v -copula slopes (thick dark green) are different. This difference occurs despite the fact that both of them are not distinguishable by second-order moments (their semi-variograms) and despite the fact that the average velocity of their center of masses are identical. The slope of second central moment over time is milder for the v -transformed non-Gaussian spatial dependence structure of K than for the Gaussian structure. Low to medium K values are connected in the non-Gaussian v -copula model, leading to blocking of the spreading solute, and a smaller spreading in general, as indicated by a smaller value of macrodispersion derived from the slope of the linear proportion (Equation 4.4): $D_{macro,Gauss}^l = 2.62 \frac{m^2}{d}$ and a

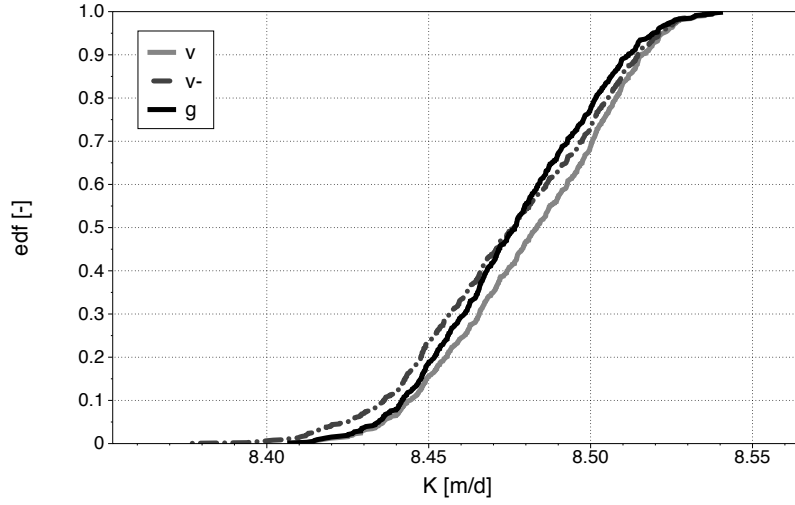


Figure 4.3: Empirical distributions of K obtained from running each two-dimensional Monte Carlo simulation as a Darcy experiment.

$D_{macro,v-trans}^l = 1.79 \frac{m^2}{d}$, which equates to dispersivities of $1.08m$ and $0.74m$ (Table 4.2). The dispersivities based on dependence models with strongest dependence in high quantiles (v , g , $extr$, chi) are smaller than the dispersivity based on the Gaussian dependence structure ($< 1.08m$). Similarly, the dispersivities based on dependence models with strongest dependence for low quantiles ($v-$, $g-$, $extr-$, $chi-$) are $> 1.08m$.

Table 4.2: Average second spatial moments for different dependence models and their inverted structures (indicated by “-”), for two-dimensional simulations. In all cases, $D_{local} = \alpha \cdot v_x \approx 4.85 \frac{m^2}{d}$.

	g	v	v-	chi	chi-	extr	extr-
$D_l^{macro} = D^l - D_{local}^l [\frac{m^2}{d}]$	2.62	1.79	3.27	1.28	3.32	1.47	3.48
difference to Gaussian $[\frac{m^2}{d}]$	0	-0.83	0.65	-1.34	0.7	-1.15	0.86
$\alpha_{macro}^l = D_{macro}^l / v_{bulk}^l [m]$	1.08	0.74	1.35	0.53	1.37	0.61	1.44

Table 4.2 gives the macrodispersion values calculated based on Gaussian and v -transformed dependence structure of K , and for two other types of non-Gaussian spatial dependence (“ $chi-$ ” and “ $extr-$ ”) which have also highest dependence in high quantiles of K , and are shown here for comparison purposes (Guthke, 2010). Additionally, for all non-Gaussian dependence structures, their inverses were calculated, which exhibit the strongest dependence in low quantiles of K . In these inverse structures, the ranks are exchanged: the location with the highest K in the original field will be assigned to the lowest K in the inverted case. This means that the inverted fields have more and on average bigger zones of connected high K , which act as “channels” or “windows” and transport some parts of the solute faster than average. The result are larger macrodispersion coefficients. The slopes of their second cen-

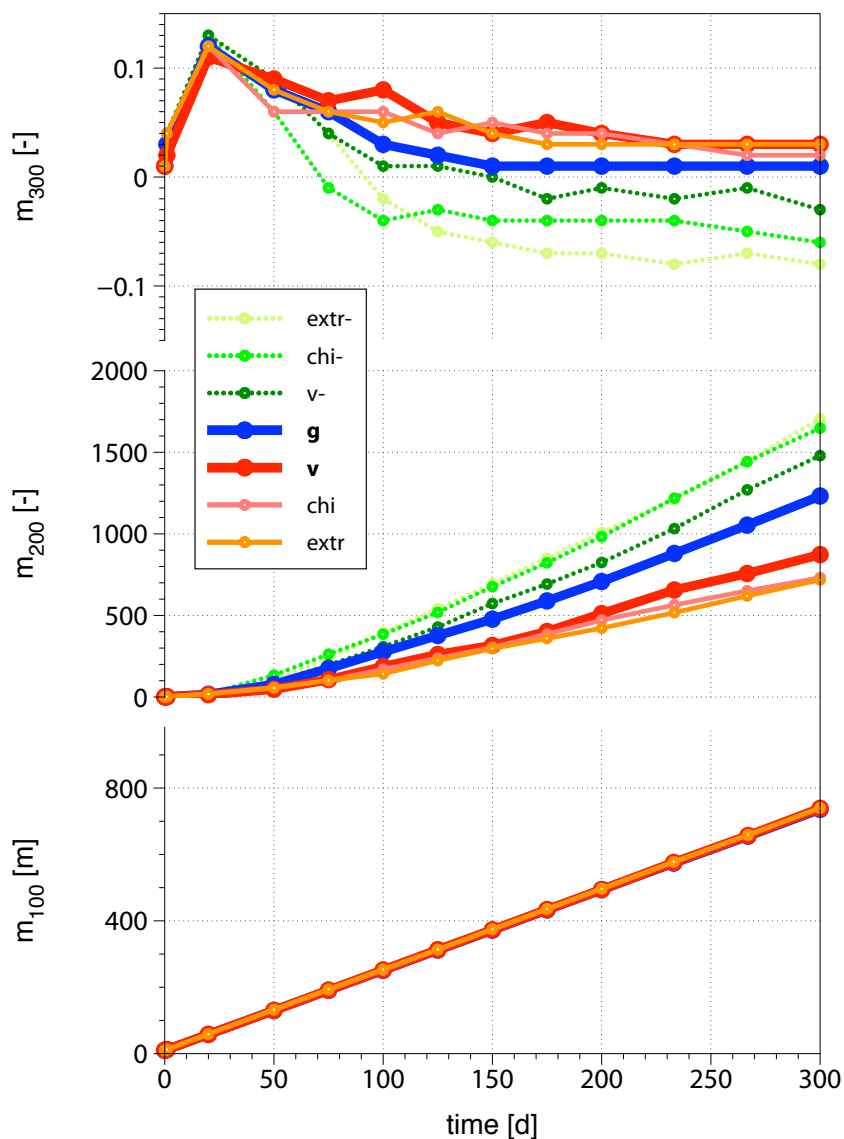


Figure 4.4: First (m_{100} , bottom panel), second (m_{200} , central panel), and third (m_{300} , top panel) spatial moments in principal flow direction for two-dimensional simulations. The first moment indicates the position of the center of mass. The second moment indicates the spreading of the plume. The third moment indicates the skewness of the plume.

tral moments are symmetric around the Gaussian graph for a type of dependence model (Figure 4.4).

The top panel of Figure 4.4 shows the third central moment in main direction of flow, normalized by the second central moment. All curves flatten out after ~ 250 d, but at a different level, which means that the mean concentration plume is somehow distorted in a different manner for every underlying dependence model of K . The seven dependence models can be distinguished into three groups, similarly as was the case for the second central mo-

ments. The third spatial moment with an underlying Gaussian K field reaches at large time a value of ~ 0 . The third spatial moments with an underlying non-Gaussian structure with strongest dependence in high quantiles flatten out to a positive third spatial moment, and the ones with an underlying inverted non-Gaussian structure with strongest dependence in low quantiles flatten out to a negative third spatial moment.

A negative skewness indicates that parts of the solute are temporary trapped in low conductivity zones while the bulk of the solute plume moves further, enhanced by connected zones of high K (inverted non-Gaussian structures). On the contrary, for the underlying non-Gaussian structures where the solute is forced to flow through low conductivity areas, the skewness tends to be positive. Only some small parts of the solute mass advect faster than the bulk solute.

The same phenomena can be observed looking at average solute plumes at a given time and at plumes of variance of concentration at a given time. It becomes evident, that the average solute mass after $t = 300d$ is very similar, independent of the type of the underlying spatial dependence structure of K. Figure 4.5a shows the average plume based on a Gaussian K field at $t = 300d$ and Figure 4.6a shows the average plume based on a non-Gaussian v-copula K field at $t = 300d$. There are only minor differences visible, such as a slightly higher peak concentration at the very center of the plume in the case of the non-Gaussian v-copula-based underlying structure. However, the difference becomes evident in the variance “plumes” (Figures 4.5b and 4.6b).

For the Gaussian dependence structure, the variance plume is symmetric around the center of mass, corresponding to zero third spatial moment. The background values are quite certain, the location of the center of mass is a little less certain, but the location of the edges of the plume where the highest concentration gradients occur is least certain. In principal, this scheme is also true for the underlying non-Gaussian v-copula based K field, however the location of the upstream edge of the plume is again less certain than the location of the edge in downstream direction.

The analytical solution for concentration variance of Vomvoris and Gelhar (1990) was used to check the numerical solutions (Equation 4.5). To calculate the concentration variance $\sigma_{conc}^2(x, y, z, t)$ the following parameters were used: σ_f^2 , the variance of the $\ln(K)$ data (0.39); γ , the flow factor of Gelhar and Axness (1983); $\frac{\partial \overline{conc}}{\partial x_1}$, the mean concentration gradient along the longitudinal axis of the plume where $conc(x, y, z, t)$ was calculated using standard three-dimensional solution of the advection dispersion equation (Sudicky (1983), $v_l = 2.45 \frac{m}{d}$ from Figure 4.4, $D_l^{macro} = 2.62 \frac{m^2}{d}$ (Table 4.2); α_T , transverse dispersivity, was augmented with local diffusion as described in Burr et al. (1994); l_i , the integral scale of the $\ln(K)$ field in the i th direction when the field is described by a hole exponential covariance function. Here was an error introduced since the the effective range of an exponential covariance function was used and not a length fitted to a hole exponential covariance function. However, the sensitivity of the lengths scales was within reasonable ranges was not found to influence the results significantly.

$$\sigma_{conc}^2(x, y, z, t) = \sigma_f^2 \frac{1}{\gamma^2} \left(\frac{\partial \overline{conc}}{\partial x_1} \right)^2 \frac{2}{3} \frac{1}{\alpha_T} l_1 l_1 l_3 \quad (4.5)$$

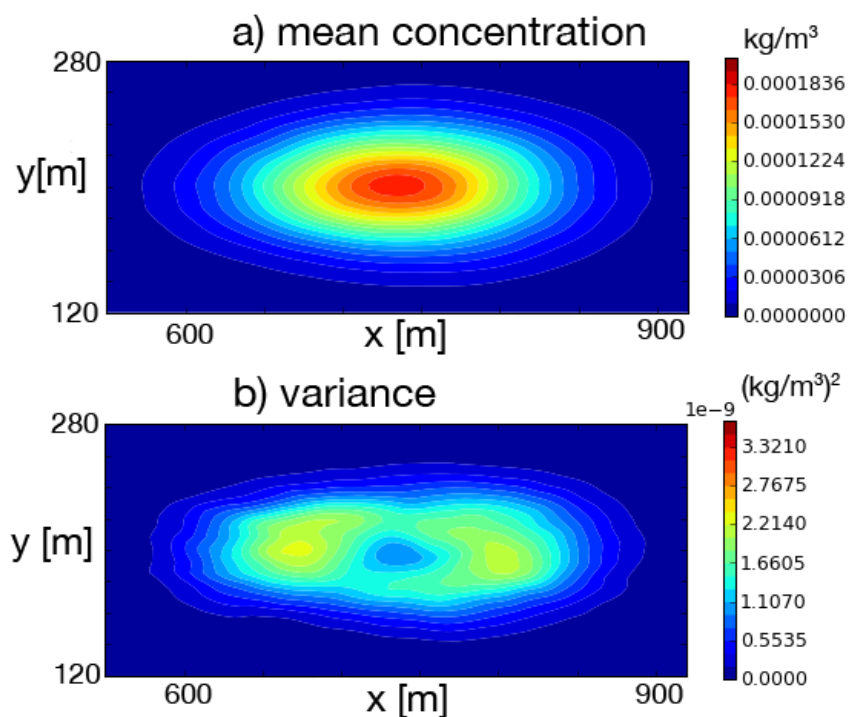


Figure 4.5: Mean and variance of concentrations calculated at time $t = 300d$ in each grid cell over 200 realizations for Gaussian dependence structure in the hydraulic conductivity fields.

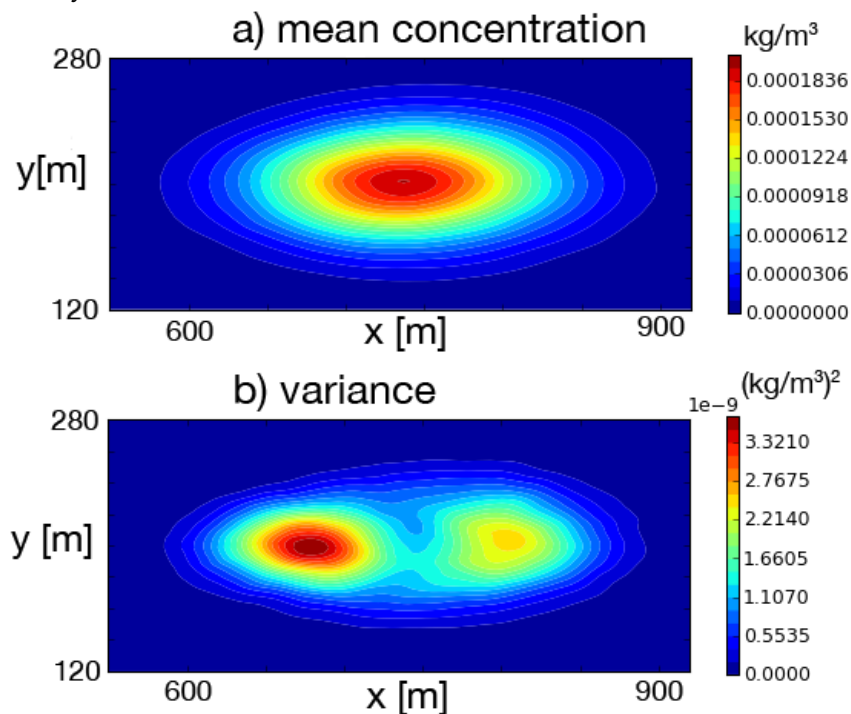


Figure 4.6: Mean and variance of concentrations calculated at time $t = 300d$ in each grid cell over 200 realizations for a non-Gaussian dependence structure in the hydraulic conductivity fields.

Generally the fit of the analytical solution of [Vomvoris and Gelhar \(1990\)](#) to average modelled concentration variances based on Gaussian K fields was good. However, the observed tailing in concentration variance when the spatial structure of K is non-Gaussian can not be reproduced. This tailing is pronounced at $\sim 200d$ and weakens for larger times (Figure 4.7).

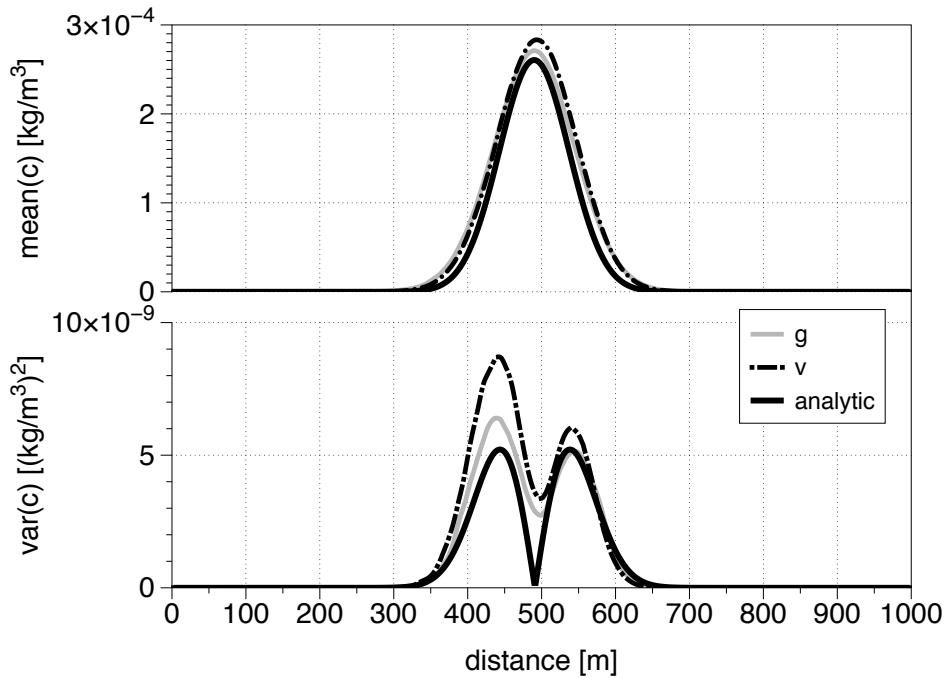


Figure 4.7: Concentration and concentration variance along main direction of flow at $t = 200d$. Comparison between numerical results based on averages of 200 simulations based on spatially distributed K fields with “g”, “v”, and “chi” structure compared with analytical solutions of concentration ([Sudicky, 1983](#)) and concentration variance ([Vomvoris and Gelhar, 1990](#)).

4.3.1.1 Test of Significance

The results so far have indicated that there is a difference in macrodispersivity depending on the type of model used to simulate the spatial dependence of K: a Gaussian model and a non-Gaussian v-copula based model were fitted to the Borden data-set. Especially for the second moment in the interval $200d \leq t \leq 300d$, which was used for the calculation of the dispersion coefficients, confirmation that the differences were significant and not an artifact of chance, was necessary. At any given time-step, the spatial moments exhibit a positive skewness and it is supposed, that they are not normally distributed. Therefore, a Wilcoxon Rank-Sum Test was applied to the second spatial moments in x-direction of the Gaussian model and all v-transformed models. The test is a non-parametrical statistical procedure that tests the significance whether two distributions $F1$ and $F2$ are based on the

same population or differ from each other by a shift. The null and alternative hypotheses can be formulated in three ways:

1. $H_{01} : F1(x) \leq F2(x)$ against $H_{11} : F1(x) \geq F2(x)$, '>' for at least one x
2. $H_{02} : F1(x) = F2(x)$ against $H_{12} : F1(x) \neq F2(x)$
3. $H_{03} : F1(x) \geq F2(x)$ against $H_{13} : F1(x) \leq F2(x)$, '<' for at least one x

All tests showed similar and clear results. Therefore, just one example for the comparison of the Gaussian and the v -transformed model that was fitted to the Borden data-set is given here. Table 4.3 shows the critical p-values for the null hypotheses H_{01} with $F1$ representing spatial moments of the Gaussian model and $F2$ the spatial moments of the v -copula model. The p-values (significance level) can be understood as the probability of making an error if the null hypothesis is rejected. Therefore, a low p-value indicates that the null hypothesis has to be rejected. In this case, H_{01} is rejected with low probability error for all four time-steps in the relevant interval $200d \leq t \leq 300d$. It can also be seen that for very small time-steps, it is not very likely to reject the null hypothesis. That is consistent with the data, as for smaller time-steps the spatial moments are more or less the same.

To conclude, it can be stated that the spatial moments of the non-Gaussian v -copula model are systematically smaller than the spatial moments of the Gaussian model. The tests showed, that the spatial moments of the v , chi, and extr models are indeed significantly smaller than the spatial moments of the Gaussian model. Also the spatial moments of the v -, chi-, and extr- models are significantly bigger than the spatial moments of the Gaussian model.

Table 4.3: Critical p-values of the Wilcoxon rank-sum test for spatial second moments at different times.

time [days]	critical p-value for H_{01}
1	0.283744661
20	0.382815922
50	0.017914135
75	0.004437013
100	0.026335570
125	0.001146688
150	0.000006460
175	0.000001279
200	0.000000299
233	0.000000071
267	0.000000029
300	0.000000869

4.3.1.2 Tracer Tests Based on Dependence Modelled by Maximum Gauss Copula

The effect of a maximum Gauss spatial dependence structure was analyzed with numerical tracer tests in two dimensions using the same setup as before. Scenario 1d, which is based on all of the hydraulic conductivity measurements was analyzed, as well as scenarios 1b and 1c which are based on the data of cross-section AA' only. Scenario 1a was excluded from the analysis, because the fit was not deemed to be good enough, and it performed worst by far in the Akaike test. The evolution of second spatial central moments is shown on Figure 4.8. The graph of the second central moment based on scenario 1b departs in that respect significantly from the graphs based on the other scenarios. This is due to the fact that it includes significant anisotropy, hence it is not really comparable with previous isotropic fields. The graphs of the ensemble averages based on scenarios 1cc and 1d fall in the range of the Gaussian and the v -copula fitted fields. Scenarios 1d of the parameter optimization for the maximum Gauss copula is directly comparable to previous results using v -copula structures, because it is also based on the entire data-set of hydraulic conductivities at Borden, not just on data from cross-section AA'. The resulting dispersivity value of 0.95 m is higher than the v -copula based estimate of 0.74 m and below the Gaussian based estimate of 1.08 m. Using maximum Gaussian copulas the solute transport behaviour of Borden is modelled similar to Gaussian behaviour.

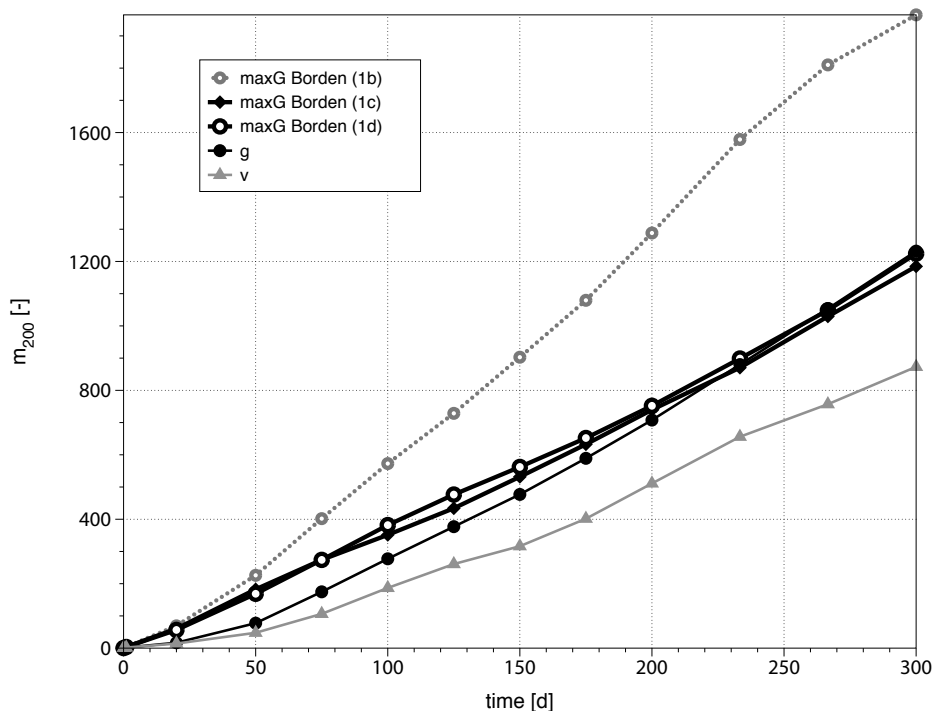


Figure 4.8: Ensemble averages of second central spatial moments in principal flow direction for two-dimensional simulations. The second moment is corrected for local dispersion and indicates the spreading of the plume.

Table 4.4: Estimated dispersivities from two-dimensional Monte Carlo simulations based on maximum Gauss copulas. $D_{local} = \alpha \cdot v_x \approx 4.86 \frac{m^2}{d}$.

scenario	1b	1cc	1d
$D_l^{macro} = D^l - D_{local}^l [\frac{m^2}{d}]$	5.09	2.13	2.29
$\alpha_{macro}^l = D_{macro}^l / v_{bulk}^l [m]$	1.99	0.88	0.95

4.3.2 Results of Solute Transport Behaviour in Three Dimensions

The two-dimensional analysis so far presents a depth-averaged model. Three-dimensional systems are more realistic, flow paths can be more connected. Similar analyses as before in two dimensions were performed in three dimensions, with the following adaptations: The hydraulic gradient in principal flow direction was adjusted to a value of 0.0045 which is the observed gradient at the Borden site. Additionally, the local dispersivities were decreased even further by one order of magnitude, the number of cells was increased significantly and the dimensions of the domain slightly decreased (Table 4.1). Four types of K fields were analyzed in four dimensions (“g”, “v”, “g10”, “v10”):

1. “base case”: Two types of fields, one with a Gaussian (“g”) and one with a non-Gaussian v-copula (“v”) dependence structure. These two types are the exact same spatial fields as the “g” and the “v” fields in the two-dimensional simulations.
2. “increased variance”: Two types of fields, again with the same “g” and “v” spatial dependence structure, and also with the same mean K ($\sim 9.62m/d$), labelled “g10” and “v10”. The only difference is in the variance of the marginal distribution, which is increased by a factor of ten. This resulted in a variance of $\sim 240(m/d)^2$ instead of $\sim 24(m/d)^2$ and parameters of the marginal Weibull distribution in Equation 3.1 of $\alpha \sim 0.284$ and $\beta \sim 0.646$. The distribution function with the increased variance is shown on the chart of the marginal distributions (Figure 3.6). This is a significant increase in the variance, but not unrealistic. It is a nice feature of copulas that they are capable of treating the marginal distribution independently of the spatial structure.

In the three-dimensional setting, each containing the same mass of solute, the average center of mass in the base case for both “g” and “v” K-fields moves with identical velocity. However, this is a different velocity than the velocity with which the plumes in the scenario with the increased variance move (Figure 4.9 bottom panel). The increased variance with a larger skewness towards small K values leads for both the Gaussian and non-Gaussian dependence structure to a slightly slower movement of the plume.

The spreading of the plumes as measured by second spatial moments in the base case is much less pronounced in three dimensions compared to the spreading in two dimensions. Like in two dimensions, the dispersion based on the v-transformed K field is smaller than the dispersion based on the Gaussian field (compare top panel of Figure 4.9 with central panel of Figure 4.4). For both “g” and “v” in the base case scenario, $D_l^{macro} \sim 0.022m^2/d$, which equals a dispersivity value of 0.36m, which is exactly the value obtained by Freyberg

(1986). However, the macrodispersion in the scenario with the increased variance lead to significantly increased macrodispersion coefficients, namely $D_l^{macro} \sim 0.207m^2/d$ ($\alpha = 3.94m$) in the case of the Gaussian K field, and to $D_l^{macro} \sim 0.088m^2/d$ ($\alpha = 1.52m$) in the case of the v-copula based field. Like in two dimensions, the dispersivity is also in three dimensions smaller for a v-copula based structure than for a Gaussian based structure. The dispersion coefficients and dispersivity values are summarized for the three-dimensional simulations on Table 4.5.

The increased dispersion with the increased variance of the marginal distribution in the three-dimensional setup is attributed to the more tortuous pathways of groundwater flow, induced by more blocking of flow by more small-K values. Figure 4.10 visualizes these effects by showing a cross-section of a simulated three-dimensional block with the same copula-based K-field but with the two different marginal distributions.

A Darcy experiment was also conducted in three dimensions. The empirical distributions of the effective K values obtained for each simulation run are plotted on Figure 4.11. With a more skewed marginal distribution (dotted lines), the effective K values are smaller than with a more symmetric marginal distribution (solid lines) – despite the fact that all four types of fields have the same mean K.

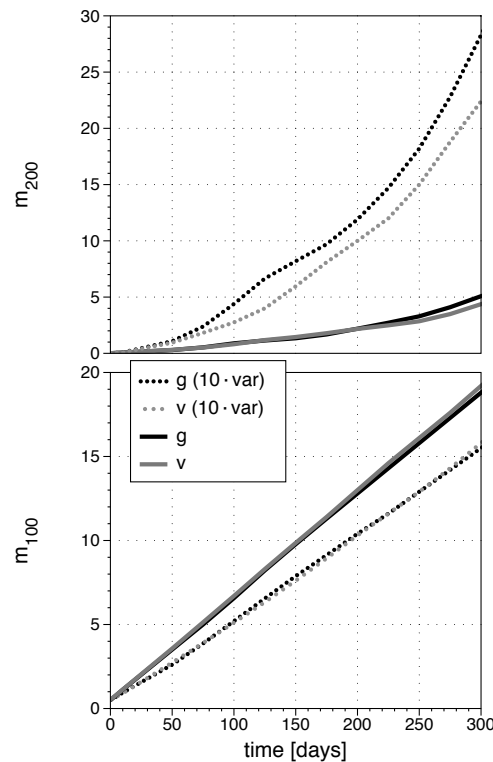


Figure 4.9: First (m_{100} , bottom panel), and second (m_{200} , top panel) spatial moments in principal flow direction for three-dimensional simulations. The first moment indicates the position of the center of mass. The second moment is corrected for local dispersion and indicates the spreading of the plume.

Table 4.5: Average second spatial moments for different dependence models and their inverted structures (indicated by “-”), for three-dimensional simulations. In all cases, $D_{local} = \alpha \cdot v_x \approx 0.012 \frac{m^2}{d}$.

	g	v	g10	v10
$D_l^{macro} = D^l - D_{local}^l [\frac{m^2}{d}]$	0.022	0.023	0.207	0.088
difference to Gaussian $[\frac{m^2}{d}]$	0	0.001	0.185	0.066
$A^l = D_{macro}^l / v_{bulk}^l [m]$	0.36	0.35	3.94	1.52

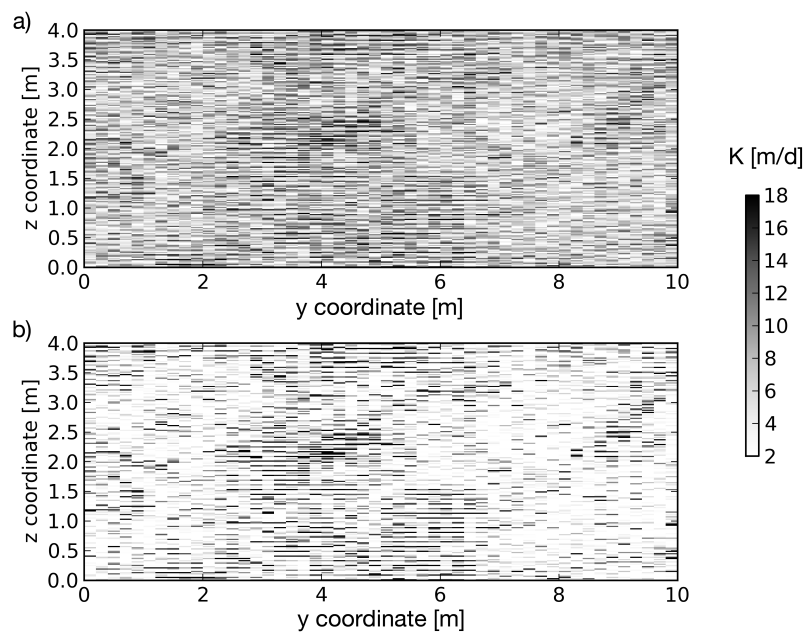


Figure 4.10: Two types of spatially distributed K fields. Both with the same structure as modelled by the v-copula fitted to the Borden data-set. The marginal distributions are different. The field shown on panel a) has the marginal distribution fitted to the Borden data-set, the field shown on panel b) has the same spatial structure, but a ten times increased variance. The different marginal distribution functions are shown on Figure 3.6

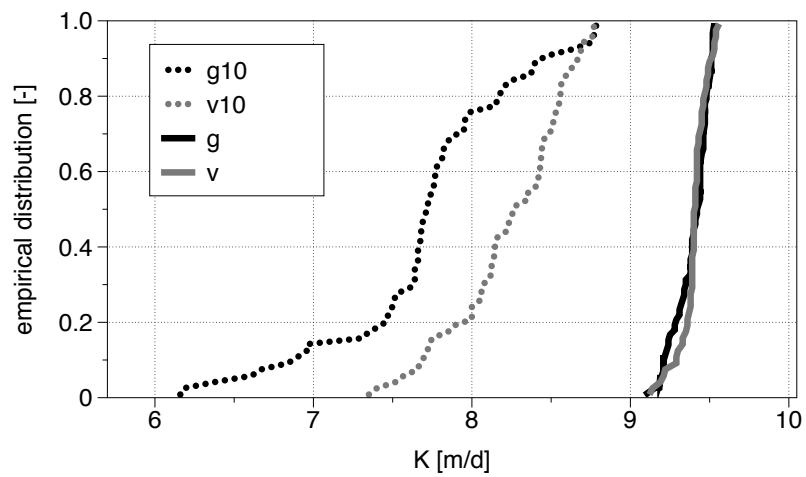


Figure 4.11: Empirical distributions of K obtained from running each three-dimensional Monte Carlo simulation as a Darcy experiment.

4.4 Analysis of Solute Transport Behaviour Using Spatial Structures of the North Bay Data-Set

Gaussian copula and maximum Gauss copula were fitted to the North Bay data-set. The solute transport properties were analyzed in two-dimensional numerical tracer tests, with the identical setup as for the two-dimensional analysis using the data from the Borden data-set. The resulting dispersivities are significantly larger than the ones determined based on the Borden data-set. Dispersivities of 4.20 m using the fitted Gaussian copula and 4.97 m using the fitted maximum Gauss copula were determined based on 200 simulations. These values are slightly below the value of 5.31 m reported by [Sudicky et al. \(2010\)](#). The overall fit of the maximum Gauss copula to the North Bay data-set equals out the negative symmetry for short separation distances and the positive symmetry for large separation distances. The overall result is very similar to Gaussian. For the maximum Gauss copula fitted to data from line 1 only, a dispersivity value of 7.14 m was obtained (Table 4.6).

The influence of the uncertainty in the marginal distribution is analyzed using the North Bay data-set by two different resampling approaches. From the 1878 measured K values were either 250 samples drawn without replacement or 1878 samples with replacement. The latter approach minimizes variability, which can be seen in the ensemble of distribution functions for 200 draws for each resampling approach (Figure 4.12). Independently of the resampling technique, the dispersivity obtained from the average spatial moments is always bigger if the spatial structure of K was modelled using the maximum Gauss copula and not the Gaussian copula.

At first glance, the results of the dispersivities obtained from the ensemble averages of the resampling procedure might seem counterintuitive. When the marginal distribution is based on a small subsample, the dispersivities are smaller than when they are based on a constant marginal distribution. On the contrary, when the marginal distribution is based on a big sample with replacement, the dispersivities are bigger than when they are based on a constant marginal distribution. One aspect that sheds some light into the issue is the observation that if only a small sample is drawn, then the chances are fairly high that less extreme values were included in that sample and consequently in the marginal distribution.

Table 4.6: Estimated dispersivities from two-dimensional Monte Carlo simulations based on maximum Gauss copulas based on the North Bay data-set. “G” stands for Gaussian spatial dependence, “mG” for dependence modelled using the maximum Gauss copula. The four right most columns summarize dispersivities based on a resampled marginal distribution, based on different sample size (250 or all).

$$D_{local} = \alpha \cdot v_x \approx 2.30 \frac{m^2}{d}.$$

scenario	constant marginal			resampled, based on 2b			
	G	2a	2b	maxG 250	G 250	maxG all	G all
$D_l^{macro} = D^l - D_{local}^l [\frac{m^2}{d}]$	4.71	2.33	1.64	5.34	4.53	5.79	4.69
$\alpha_{macro}^l = D_{macro}^l / v_{bulk}^l [m]$	4.20	7.14	4.97	4.69	3.92	5.70	4.42

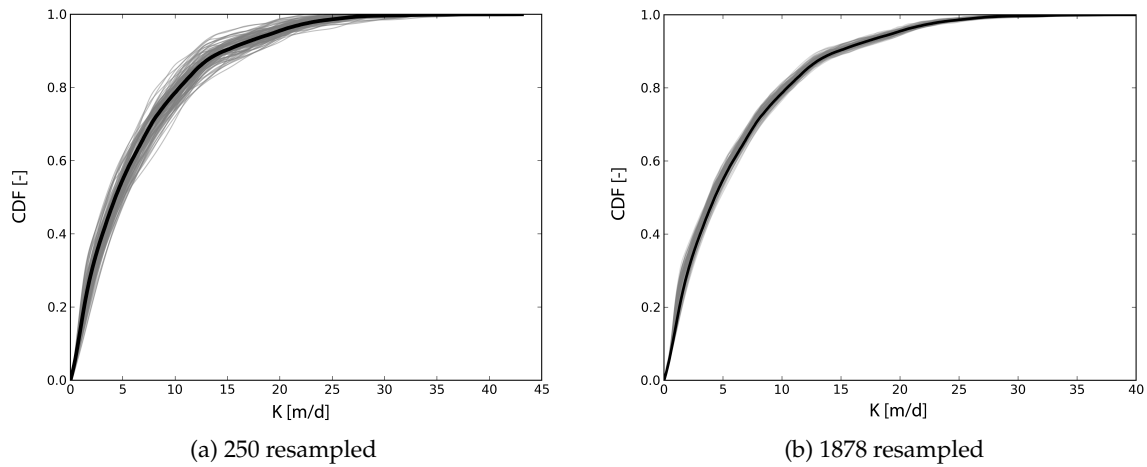


Figure 4.12: Ensemble of distribution functions using two types of resampling on the North Bay data-set: drawing 250 samples without replacement (Figure 4.12a) or drawing 1978 samples with replacement (Figure 4.12a).

The issue can be even more resolved by looking at some statistics of the distribution of the dispersivities, and not the ensemble averages (Table 4.7). The skewness of the distribution of dispersivities calculated based on Gaussian dependence structure is zero, indicating a Gaussian symmetry, independently if the sample was small or big. However the skewness of the dispersivities differs significantly from zero if the underlying structure is non-Gaussian. In this case, the structure was maximum Gaussian of North Bay type, which is only slightly non-Gaussian – but still the difference in skewness in the dispersivities occurred. This means that the marginal distribution does have an influence on the solute transport behaviour as well, not only the spatial structure. But the influence of the marginal distribution is different, depending on the spatial structure

4.5 Structure Within the Borden Data-Set

The previous section has shown, that non-Gaussian structures can lead to a significant difference in solute transport behaviour – despite identical properties in terms of second or-

Table 4.7: Distributions of dispersivities in resampling ensembles. “G” and “maxG” indicates Gaussian and maximum Gaussian dependence structure. The integer stands for the number of samples taken in the resampling approach.

	G 250	G 1878	maxG 250	maxG 1878
mean	2.99	4.47	4.80	5.76
std.dev	1.99	1.96	2.75	2.83
skewness	0.01	-0.15	0.86	0.99

der moments. The Borden aquifer is generally considered to be fairly homogeneous, the marginal distribution to be fairly symmetric, and its structure to be well represented with a traditional Gaussian variogram model. Two different models that are not distinguishable by classical methods, neither by their marginal distribution nor by their second order moments, were fit to the Borden K data-set. The only difference was in their spatial copula, and the difference in the spatial copula was not very large. There exist other pairs of v-transformation parameters m, k which lead to much more non-Gaussian dependence (m closer to zero and/or k closer to 0 or closer to larger numbers). Still, under some conditions, significant differences in solute transport behaviour were observed.

This section does not only look at the model of spatial dependence but takes a closer look at structure within the data. What if we did not just assume that Borden is one homogeneous data-set? Doesn't it appear as if there are generally more high-K values in the upper portion of the cross-sections, indicated by darker shading on Figure 3.2? This section explores if this could be the case and how this observation could be incorporated into a geostatistical model.

4.5.1 Geostatistical Analysis of the Homogeneity of the Borden Data-Set

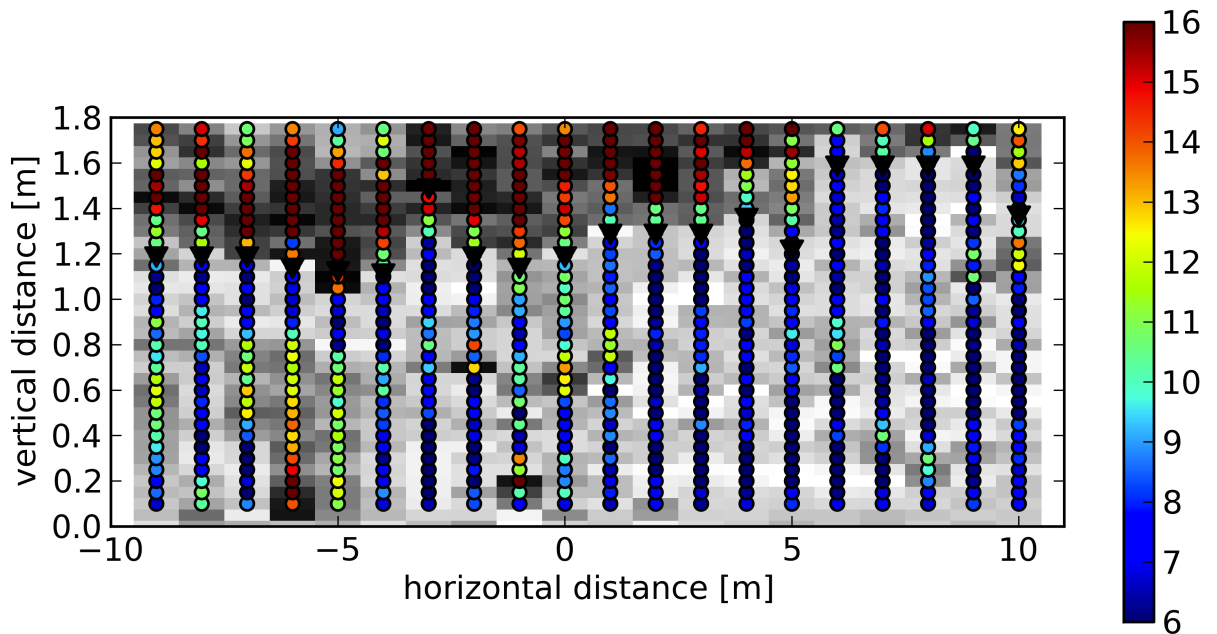
Based on the above noted observation, vertically cumulative averages were calculated for each core and plotted on Figure 4.13. The cumulative average confirms that there are generally higher K values at the top, just below ground surface, along both cross sections (Haslauer et al., 2008).

If we see that there is a part at the top of the aquifer that behaves statistically different from the bottom part of the aquifer, then there must be a layer separating the top and the bottom. This layer was picked manually at each core at a depth where the gradient of the cumulative mean is sharpest with the additional constraint of a maximum geologically possible slope of 20° (Louis and Fischer, 1960). The picks of the boundary layer are shown as black triangles on Figure 4.13.

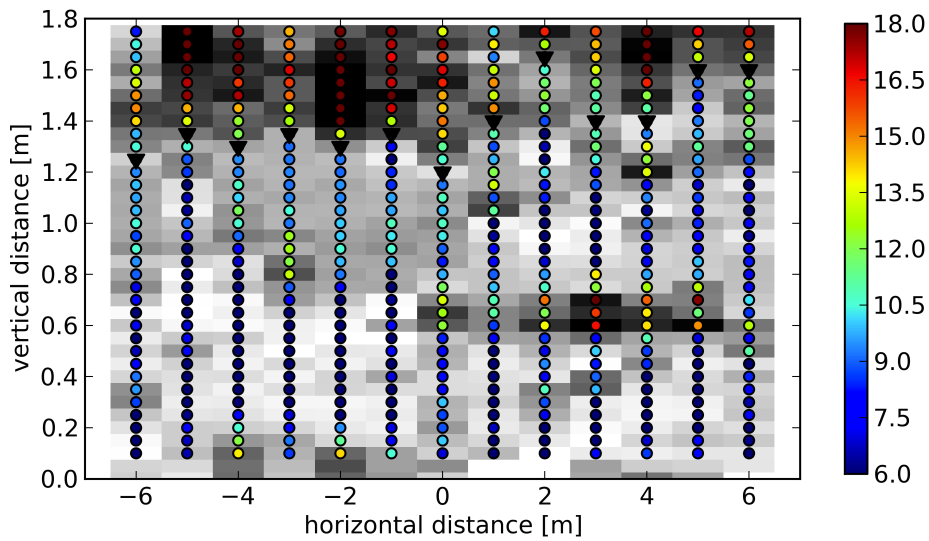
The assumed boundary layer separates the aquifer into a top part and into a bottom part. The arithmetic mean calculated at each core for the top and the bottom part is clearly distinguishable (Figure 4.14a), whereas there doesn't seem to be a clear difference in the standard deviation at each core (Figure 4.14b). The mean K is ~ 2 times higher in the top section than in the bottom section, whereas the standard deviation differs by a factor of ~ 1.13 m/d only (Table 4.8).

The fitted parameter-free distribution functions (Equation 3.2) for the entire top and the entire bottom part of the aquifer are shown on Figure 3.6. Both curves have a similar shape, whereas the distribution of the K values in the top section is shifted by ~ 8 m/d towards higher K values.

The shift in mean K between the bottom and the top section is fairly substantial. It seems like it is more substantial than the difference in standard deviation. Still, the question arises if such a difference in variance between the bottom and the top section of the Borden aquifer could occur just by chance? To answer this question, a bootstrap-test was conducted. After



(a) cross-section AA'



(b) cross-section BB'

Figure 4.13: Moving average of K along cores at both cross-sections at Borden. Locations of picks of elevation of the boundary layer are indicated by black triangles. The moving average of three K measurements was calculated, starting at $z = 0$, and the resulting value plotted at the top location of the three values (coloured shading). The background shows the same contours of K as on Figure 3.2.

having the boundary layer established, there are 257 measurements of the data-set in the top

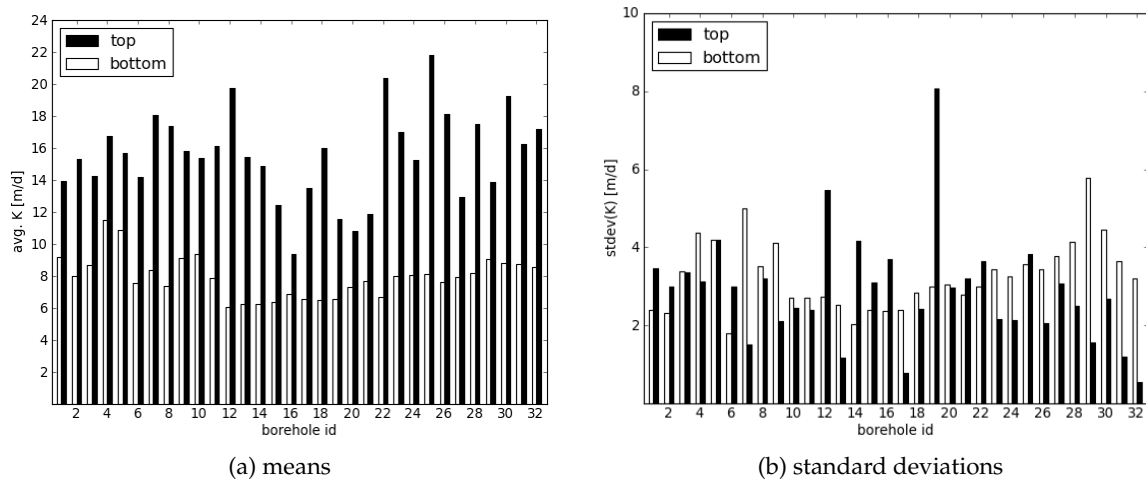


Figure 4.14: Arithmetic means and standard deviations of K per core, separated for sections above and below the boundary layer.

Table 4.8: Descriptive statistics of K-measurements above and below the boundary layer.

	bot	top	bot & top	units
count	895	257	1152	-
mean	7.89	15.65	9.62	m/d
variance	13.00	16.72	24.26	$(m/d)^2$
std. dev	3.61	4.09	4.93	m/d

section of the aquifer and 895 in the bottom section of the aquifer. For both sections, 1000 times random samples were drawn from the entire set of K measurements and converted to a value of standard deviation by subtracting the mean of either the top or the bottom section. Every time, an empirical distribution was created from the random sample of standard deviations. The mean distribution of these empirical distributions is plotted on Figure 4.15 – once for the top section and once for the bottom section. Assuming the boundary layer exists, its standard deviation of the top section exceeds the standard deviation that can occur by chance in about 8 cases of 100. Additionally, and much more meaningful, the standard deviation of the top section of the Borden aquifer is so large, that it is in 99 cases of 100 bigger than the standard deviation that can occur by chance in the bottom section.

So far we have seen that in terms of the marginal distribution, a top and bottom section as it was defined at the beginning of this section can hardly occur by chance. In addition to the marginal distributions, does the boundary layer influence the spatial structure of the dataset? This is a difficult question to answer. However we will introduce a measure, we call it the standard deviation across the boundary layer, which can be calculated based on the measurements, and which needs to be reproduced by the statistical model that we are going to use later to model the boundary layer and the sections of the aquifer above and below.

Vertically along each core, a standard deviation can be calculated for different separa-

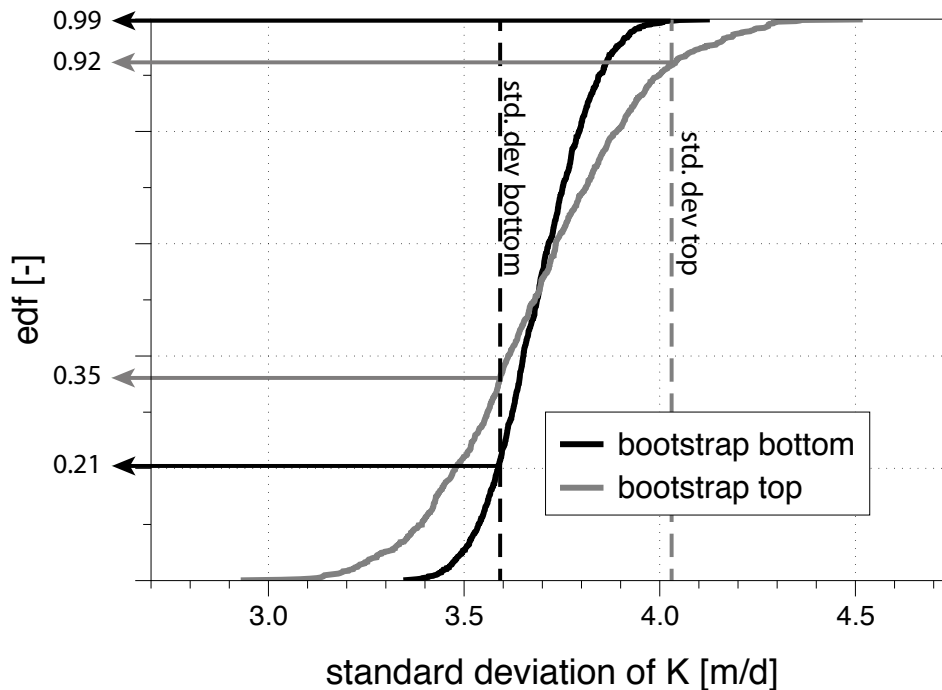


Figure 4.15: Shown are the empirical distributions of standard deviation of K in the sections above and below the boundary layer. The dashed lines represent the value of the standard deviation of the measurements above and below the boundary layer, respectively.

tion distances above (green), below (blue), and across (red) the boundary layer as shown schematically on Figure 4.16. The bigger the separation distance, the lighter the colour.

Employing this scheme on the measurements results in the set of curves shown on Figure 4.17a. Generally, the distribution of the standard deviation is similar in the top section and in the bottom section, and increases with bigger separation distance (lighter colour). However, it is a slightly smaller in the bottom section, because the blue distribution functions are a little bit shifted towards the top left compared to the green curves. The standard deviation across the boundary layer is much bigger than within the sections above and below the boundary layer.

Two components of the geostatistical model of the aquifer considering a boundary layer need to be established before the model can be used for simulation purposes.

The first component includes the geostatistical model of the boundary layer. The boundary layer is modelled as a plane with some spatially distributed deviation from the plane. The plane is modelled by minimizing the Euclidean distance to each location where the boundary layer was picked (triangles on Figure 4.13). The z -coordinate of the plane is given by $1.75 - z = (d - by - ax)/c$, where $a = 0.000635$, $b = 0.000785$, $c = -0.0307$, $d = -0.0414$, with the coordinate system as shown on Figure 3.1. The deviation from the plane is modelled using a classical variogram of the form $0.015 \cdot \text{Sph}(2.5)$. The empirical semivariogram and the fitted theoretical variogram are shown on Figure 4.18. Due to limited availability of data

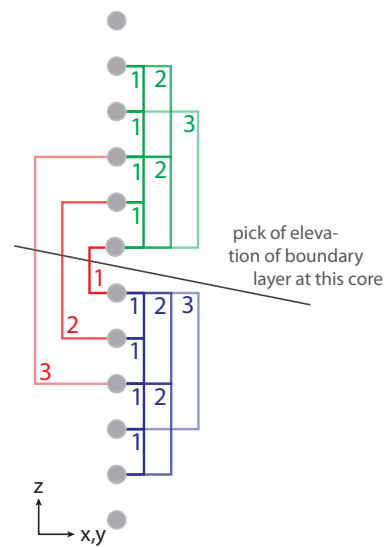


Figure 4.16: Sketch that illustrates how a standard deviation above (green), below (blue), and across (red) the boundary layer can be calculated. These standard deviations serve as a measure that needs to be met by a statistical model that incorporates the boundary layer.

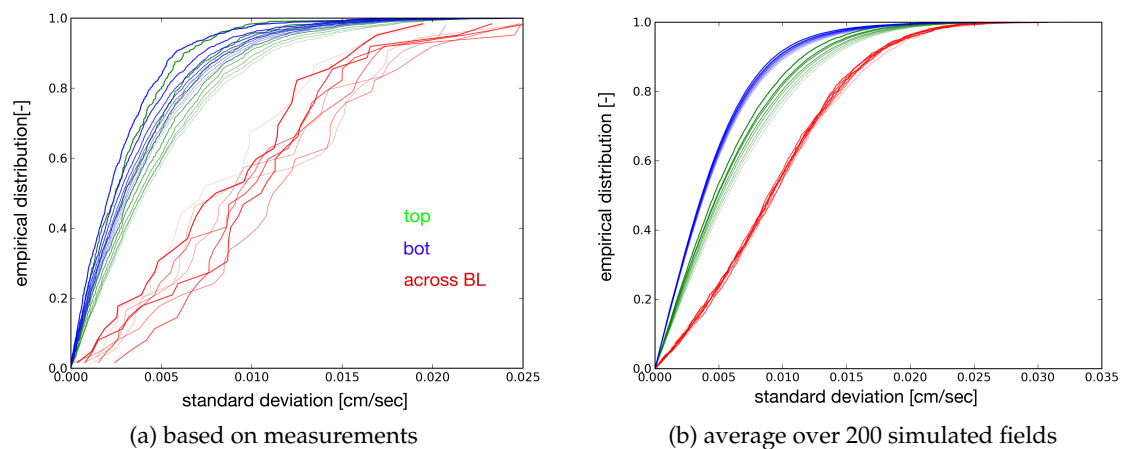


Figure 4.17: Distributions of standard deviation of pairs of points separated vertically by a certain distance (Figure 4.16). The lighter the colour the bigger the separation distance. Blue and green coloured lines are based on points that are exclusively below or above the boundary layer. The red lines are based exclusively on pairs of points which are separated by the boundary layer. Figure 4.17a shows the distributions of standard deviations of K based on measured data, Figure 4.17b shows the average over 200 simulated data-sets at the original measurement points. The model of the underlying Gaussian model for the spatial structure above and below the boundary layer is $0.17 \cdot \text{Nug}(0.0) + 0.83 \cdot \text{Gau}(7)$. For the v -transformation the optimized parameters $m = 0.75$, $k = 2.35$ were used.

for the location of the boundary layer – only at the (x, y) -coordinates of the 20 cores, a simple Gaussian approach was chosen.

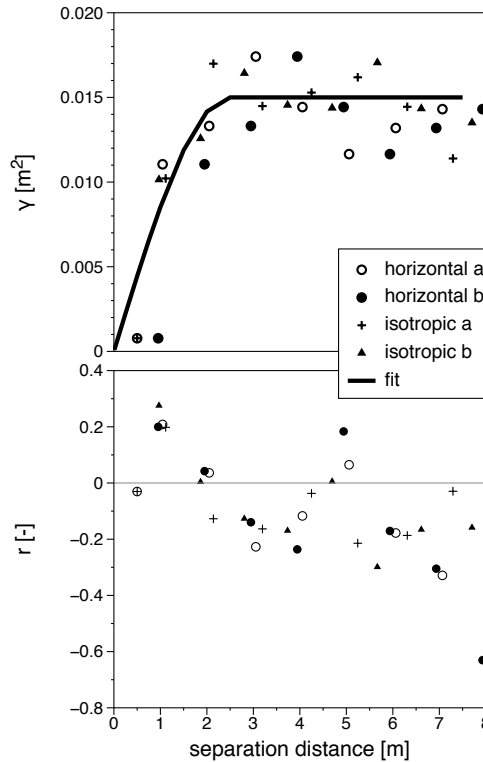


Figure 4.18: Semivariogram (top panel) and rank correlation of the deviation from the mean of the boundary layer at the Borden site.

The second component includes the spatial model of the K-fields in the top and the bottom section. Conceptually, this model assumes that the processes that led to the spatially distributed K fields in the top section and in the bottom section are the same. There is no evidence that suggests otherwise. The shift in mean K values in the top part might suggest that material with slightly larger K values was deposited at later times. The fit of a v-copula model was performed for the bottom section analogous to the procedure described in Section 2.3.4 for the entire data-set of Borden, resulting in the parameters $m = 0.75$, $k = 2.35$ and in the underlying Gaussian model of $0.17 \cdot \text{Nug}(0.0) + 0.83 \cdot \text{Gau}(7)$. Figure 4.19 shows the surface of ln-likelihood values evaluated for different combinations of m, k . The optimum is marked with a cross.

The process to simulate spatial fields of K involves the following steps

- establish the set of points within the domain at which a K value is to be simulated;
- for every combination of x, y coordinates within the domain, calculate the z coordinate of the boundary plane. At each of those x, y coordinates, simulate a value of the deviation from the boundary plane. The deviation has zero mean, hence its value can be added to the z -coordinate of the plane to obtain the elevation of the boundary layer at each x, y coordinate;

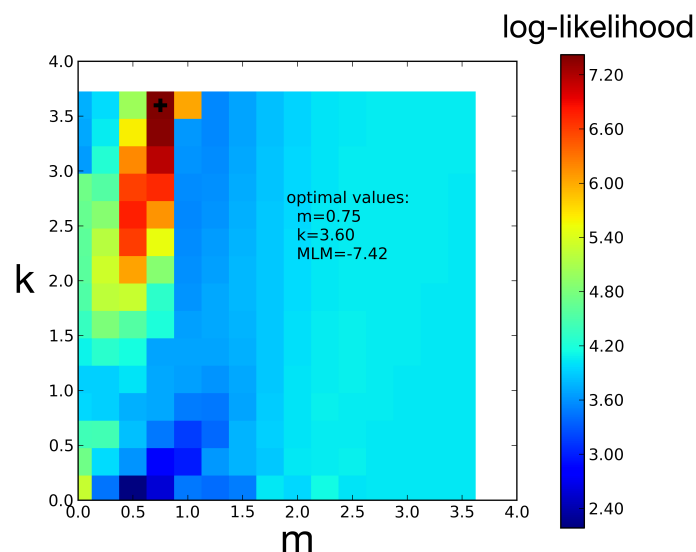


Figure 4.19: Contour plot of the log-likelihood function L that was optimized for various parameter-sets of m, k for the bottom section of the Borden K data-set.

- separate the points where a K value is to be simulated in the ones below and the ones above the simulated boundary layer;
- for each of the two sub-sets, simulate K values independently. This can be done with a Gaussian or the v-copula model;
- exchange the marginal distribution, using the marginal distributions shown on Figure 3.6.

The above described model can be used to simulate K values. An initial test involved simulating K values at the same locations where K measurements were taken, in order to evaluate the distribution of standard deviation across the boundary layer, as was done for the measurements. The average distributions obtained from 200 simulations are shown on Figure 4.17b. These average distributions based on simulations are similar to the distribution from the one observable set (Figure 4.17a). It can be concluded, that the geostatistical model of the boundary layer can mimic a behaviour that is seen in the data, if a boundary layer is assumed at the previously defined locations in the data. What if those locations were not ideal?

To assess this question 200 boundary layers were drawn at random such that their slope would nowhere be bigger than 20° . For each of the 200 boundary layers, the above outlined procedure of simulating spatially distributed K fields was followed, and the distribution of standard deviations across the boundary layer calculated, similarly to the distributions shown on Figure 4.17. It was then evaluated for each separation distance, where the distribution of the standard deviations based on the purposefully picked boundary layer (Figure 4.17a) lie relative to the 200 simulated distributions. If they lie completely outside the range of the 200 distributions, they were assigned a “significance” value of 1, if they lie in the middle of the 200 distributions they were assigned a value of 0. Two scenarios were

evaluated: One where the 200 boundary layers were drawn fully at random within the domain of the measurements at Borden (Figure 4.20a), and one where each of the 200 random boundary layers had to fall within the top and bottom z-coordinate of the picked boundary layer (Figure 4.20b). In both scenarios, the distributions based on the picked boundary layer are outside the range dictated by 200 randomly drawn boundary layers (values of significance of or close to unity). When the random boundary layer is bound to somewhat more realistic limits, in some cases the the distribution based on the picked boundary layer further inside the range of possible distributions based on the 200 simulations – some points on Figure 4.20b are farther away from unity than on Figure 4.20a. Still, it can be concluded that with fairly high significance, the picked boundary layer led to a distribution of standard deviations of K across itself, that vary rarely could occur by chance.

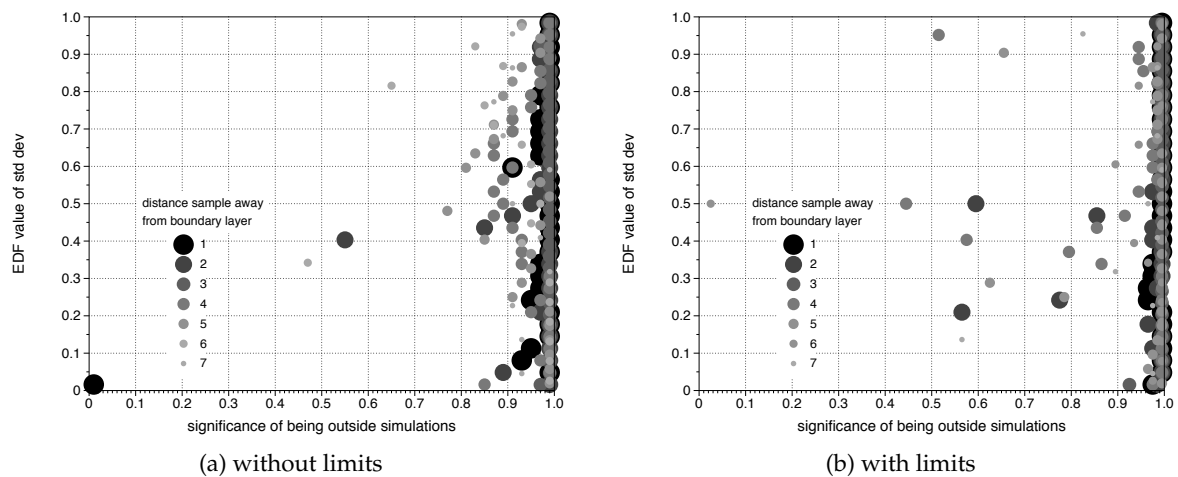


Figure 4.20: Comparison of where the distribution of standard deviations across the picked boundary layer are, relative to the possible range given by 200 boundary layers drawn at random. If the distributions across the picked boundary layer lie completely outside the range given by the 200 distributions, a “significance” value of 1 was assigned to that empirical distribution value; if they lie in the middle of the 200 distributions they were assigned a value of 0. The boundary layers drawn at random were unbounded for Figure 4.20a and were bounded by the top and bottom z-elevation of the picked boundary layer on Figure 4.20b.

4.5.2 Effects of Structure in the Data

The previous Sections presented four new possibilities for the spatial structure of Borden, all of which could have some advocates. The options include all combinations of the following: with or without a boundary layer, with Gaussian structure, and with v-copula structure. This means there are two choices: the choice between a Gaussian or a v-copula spatial dependence structure and the choice between two subsets of data which are treated independently. Both choices are based on “geostatistical understanding”, but they are different. The first choice is related to the geostatistical model used, the second choice is related to

careful analysis of the existing data. This section explores if there are differences in physical properties, such as solute transport behaviour, depending on which choice is made.

One simulated field of each of the four options is shown on Figure 4.21, along a typical cross-section. Anisotropy was modelled according to the anisotropy ratio of 20 based on the ratio of horizontal to vertical rank correlation, leading to horizontal layering. Using the v-copula model, zones of high hydraulic conductivity, which in an isotropic field yielded isolated patches, form now lenticular layers of high conductivity (panel c) of Figure 4.21), typical of Borden [Sudicky \(1986\)](#).

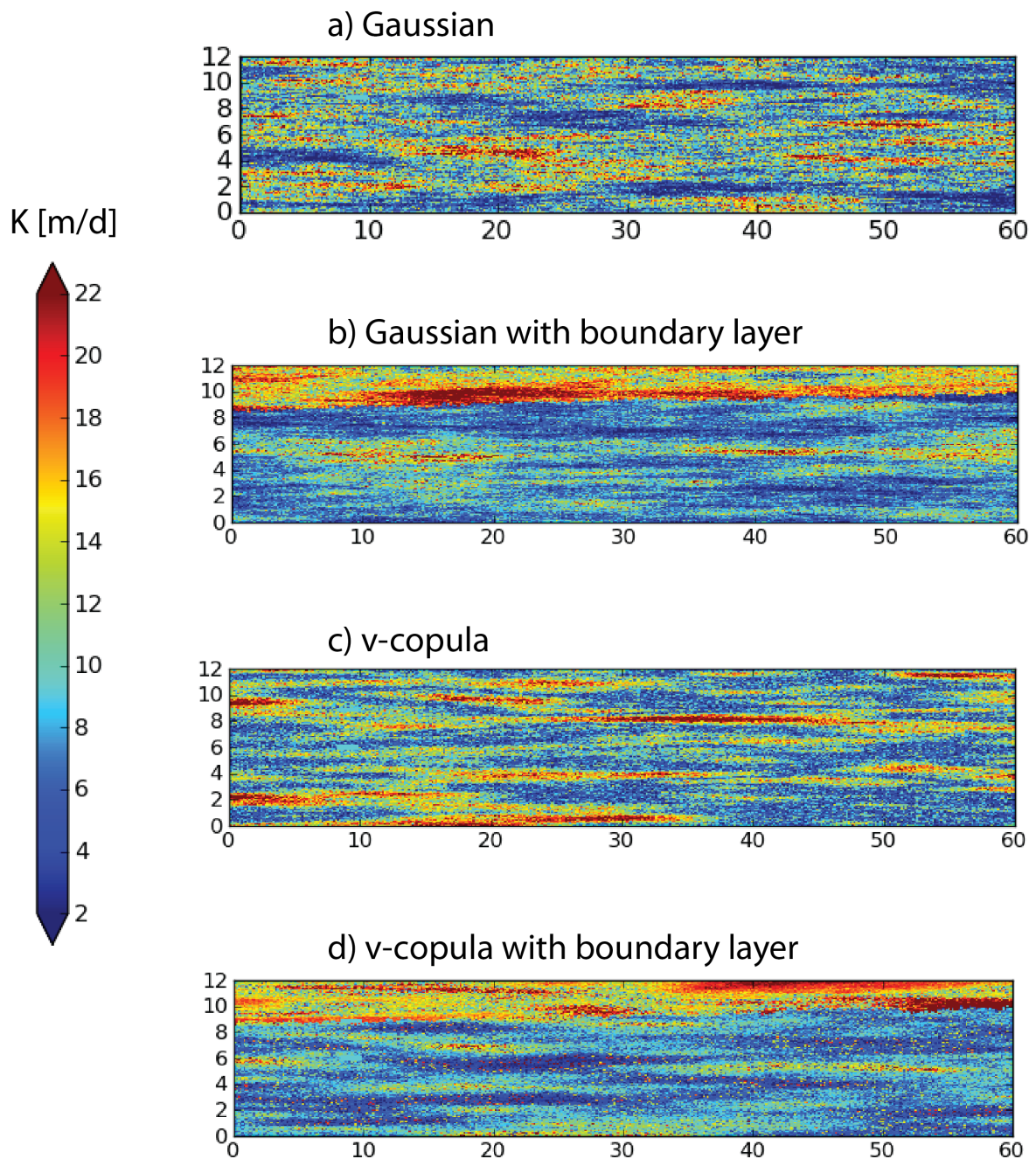


Figure 4.21: Overview over the four scenarios of spatial structures evaluated for the boundary layer problem. Spatial extent in m.

To make results comparable, the rank correlations of all four types of fields are designed to be very similar. The ensembles of rank correlations of the 200 simulated fields for each option are shown on Figure 4.22 together with the ensemble averages indicated by the thick lines and with the empirical rank correlation. The averages of the ensembles and the empirical rank correlation are very similar.

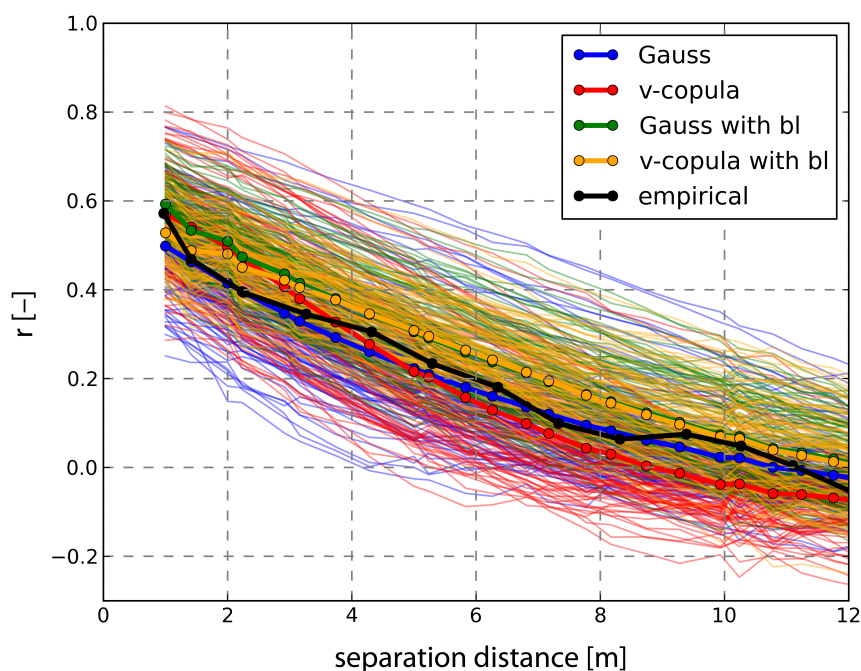


Figure 4.22: The ensembles of rank correlations for four scenarios: Gaussian- and v-copula modelled spatial dependence and including a boundary layer (bl) or not. The ensemble averages are plotted with a thick line with plotting positions marked. For comparison the empirical rank correlation is shown. The ensembles of each scenario are based on 200 simulations.

The evaluation of solute transport behaviour was to be analyzed in a similar manner as before. A solute tracer is transported over time within a steady state flow field imposed by constant head boundary conditions on the left-hand and right-hand boundary of the domain and no-flow boundary conditions everywhere else. However, some modifications were necessary: A slug injection at one defined small location was thought to lead to representative spreading for the ensemble of plumes. The relative position of that location relative to the boundary layer would influence the spreading of the plume. Instead, over the entire thickness of the domain on the left-hand boundary, a constant concentration of $conc = 1 \text{ kg/m}^3$ boundary condition was implemented (Figure 4.23), which was active for ~ 30 min. Grid refinement tests were performed leading to the optimized parameters listed in Table 4.9.

The movement of the plume through the domain is visualized on Figures 4.24 for Gaussian structure and on Figure 4.25 for non-Gaussian structure. Both figures are structured identical: the left-hand part of the panel shows the average concentration fields over 200 simulations at $t = 200$ d, $t = 400$ d, and $t = 600$ d. The right-hand side shows the standard

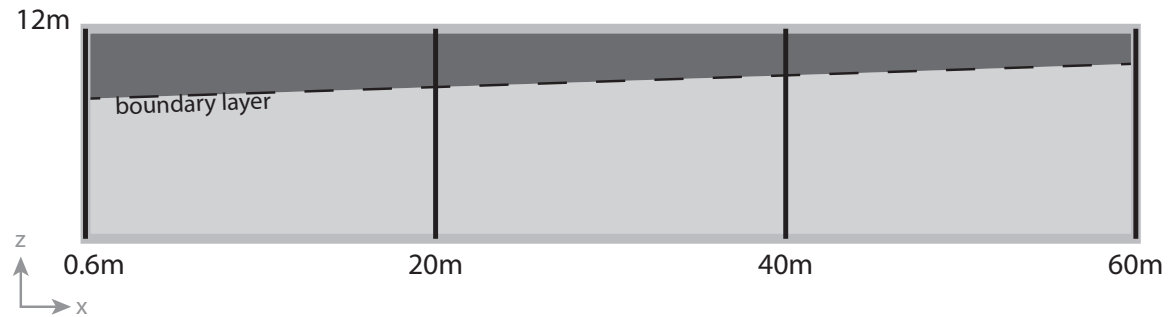


Figure 4.23: Domain for numerical model including the boundary layer.

Table 4.9: Parameters of the numerical model used to evaluate the effect of the boundary layer.

domain length [m]	lx	60
	ly	-
	lz	12
number elements [-]	nx	400
	ny	-
	nz	400
grid spacing [m]	dx	0.15
	dy	-
	dz	0.03
hydraulic gradient [-]	$i = dh/dl$	0.0033
local dispersivities [m]	$\alpha_{l;h}$	0.15
	$\alpha_{t;h}$	0.015
	$\alpha_{t;v}$	0.015

deviation of concentration within the domain, calculated based on the ensemble of 200 simulations, for the same times as the average concentration. The top half of the figure is based on spatial fields without a boundary layer, the bottom half includes a boundary layer.

The following aspects are notable:

- Due to the larger hydraulic conductivity values in the section above the boundary layer, the plume moves faster in that section. For late time, the top part of the plume moves so fast, that it almost separates from the plume in the bottom section.
- At early times the maximum concentration is identical, no matter if the boundary layer is included or not. Due to the slower movement in the section below the boundary layer, peak concentration are higher at late times, if a boundary layer is included.
- The pattern of the standard deviation is symmetric around the center of the plume in principal flow direction, independent of incorporating the boundary layer or not.

The top two characteristics hold true in the case of v-copula based structure. However, the pattern of the standard deviation is no longer symmetric around the center of mass in the principal direction of flow. Higher standard deviations are observable upstream than downstream, a phenomenon that was also observed in the analysis of solute transport in the Borden aquifer (Section [4.3](#)).

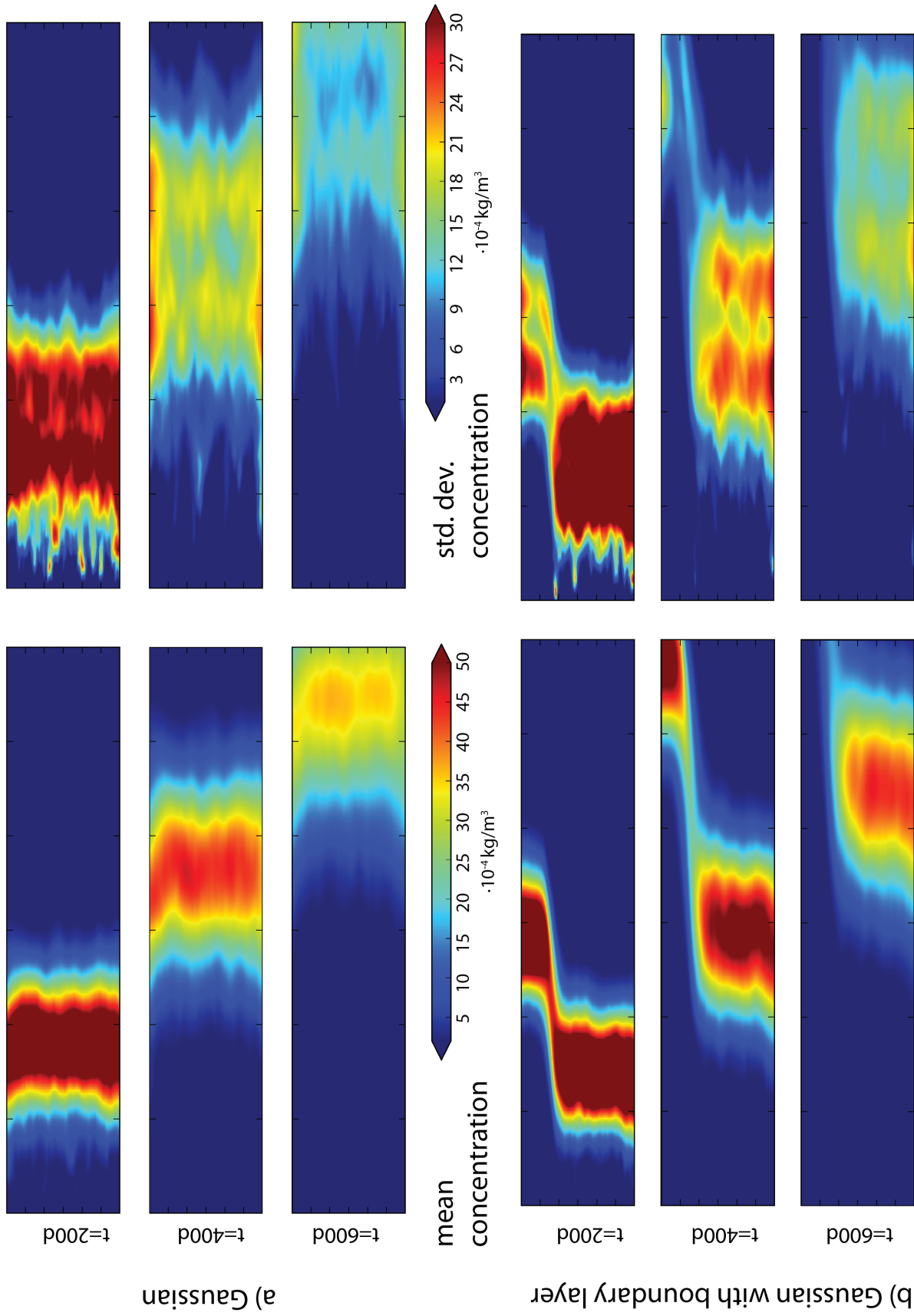


Figure 4.24: Ensemble means and standard deviation of concentration based on Gaussian K fields

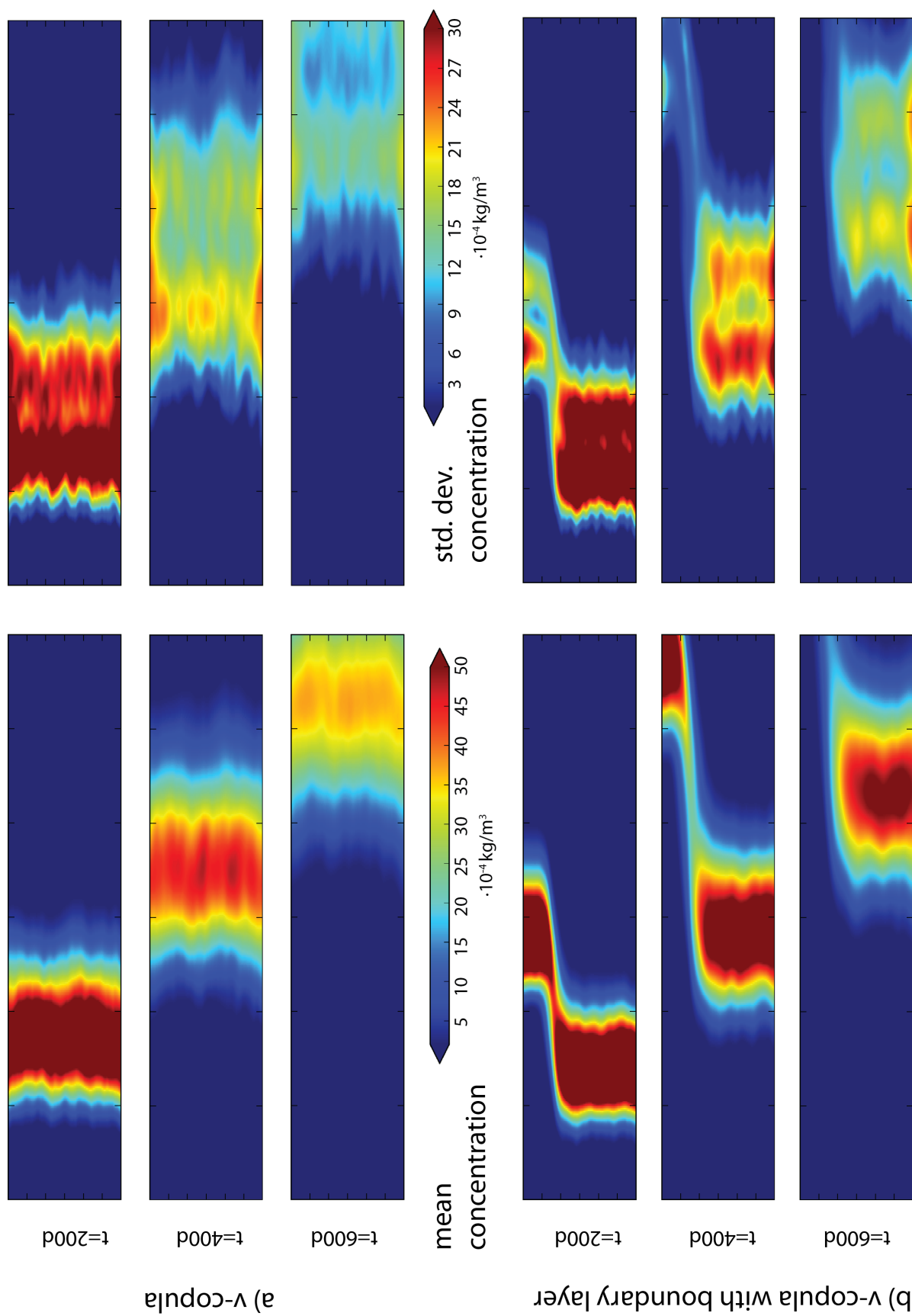


Figure 4.25: Ensemble means and standard deviation of concentration based on v-copula K fields

Differences in peak concentration and timing are summarized on Figure 4.26. The differences between including the boundary layer or not are larger than the differences between Gaussian and v-copula based structure. Peak breakthrough occurs at earlier times, if the boundary layer is not included. However, due to the faster movement of the plume in the section above the boundary layer, concentration starts to increase above zero at earlier times when a boundary layer is included, resulting in breakthrough curves with two peaks. The ensemble of the breakthrough curves through each plane for the four different scenarios is shown on Figure 4.27.

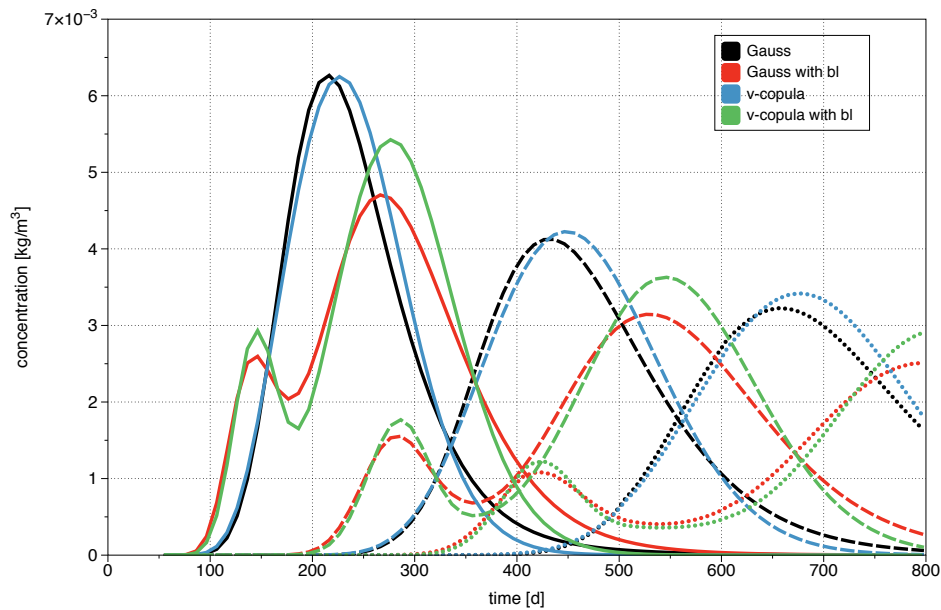


Figure 4.26: Average breakthrough curves for the Monte Carlo based numerical tracer tests based on the four types of spatially distributed K fields. The four types of distribute either include a boundary layer or not and are based on Gaussian dependence or v-copula dependence. For each type of field, three ensemble breakthrough curves averaged over 200 simulations are shown, indicated by different dashing. These concentration profiles are evaluated at cross-sections through the numerical model at $x = 20$ m (solid lines), $x = 40$ m (dashed lines), and $x = 60$ m (dotted lines).

This section has shown that there can be an internal structure with reasonable significance within a data-set, which can be detected and accounted for in the setup of the geostatistical model. In a certain sense, detecting and representing such a structure within the data is more important than the model of the spatial dependence structure. The difference in breakthrough curves is more pronounced between cases where the boundary layer is modelled explicitly or not than between models of different spatial structures.

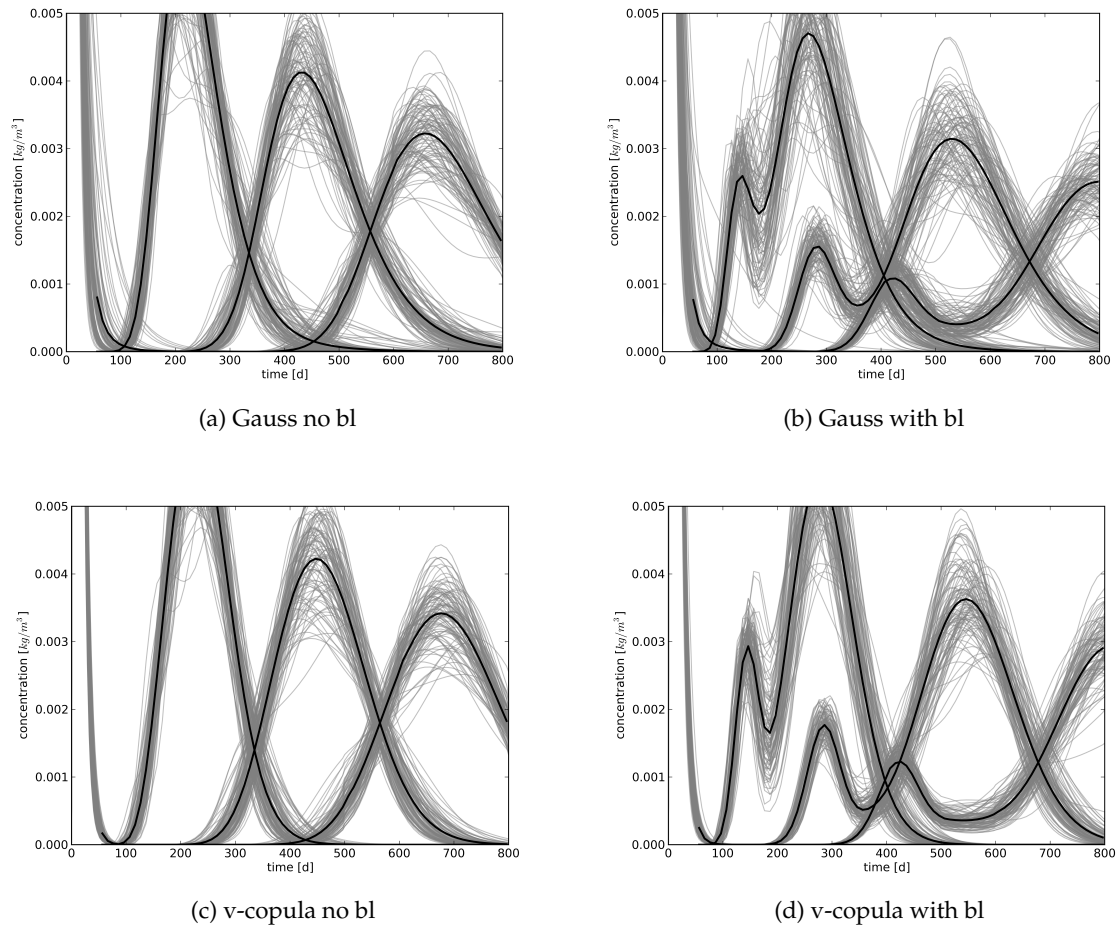


Figure 4.27: Breakthrough curves based on different spatial structures of the hydraulic conductivity field (Gaussian and v-copula based) and on different treatment of the homogeneity of the data-set (one block and with boundary layer). Shown are the breakthroughs for each Monte Carlo simulation (grey curves) and the ensemble average (thick black line) through slices at $x = 20$ m, $x = 40$ m, and $x = 60$ m (from left to right in each plot).

5 Conclusions and Outlook

Copulas offer a flexible tool to model a variety of spatial dependence structures, Gaussian structures and especially non-Gaussian structures. Using non-Gaussian copulas, a wider range of spatial dependence structures can be parameterized than using Gaussian copulas. The theoretical non-Gaussian v -copula and maximum Gauss copula models used in this thesis, are flexible stochastic models that were fitted to successfully to two data-sets of hydraulic conductivity, the data-sets of Borden and North Bay. For the maximum Gauss copula a two-dimensional analytical expression was derived that can be expanded into higher dimensions.

It was shown that within 90% confidence the spatial structure of neither of the two data-sets can be modelled with a Gaussian structure of spatial dependence, particularly for the important short separation distances. This conclusion is drawn from the results of a bootstrap algorithm, which is based on evaluating the measure of symmetry of simulated fitted Gaussian spatially distributed fields of K . It was also shown that non-Gaussian copula models lead to improved fits of the models. Using the fitted non-Gaussian models, the 90% confidence intervals were expanded such that a bigger portion of the symmetry measure falls into their bounds. This result is even more important considering that the Borden aquifer is quite homogeneous and has a small variance of K .

These fitted copula models were used for spatial interpolation and to simulate spatially distributed fields of hydraulic conductivity.

When using copula-based methods for interpolation, the full distribution function of estimated values at every location can be calculated. This distribution function is not dependent on the geometry of the observation network only, but also on the magnitude of the measurements. Hence, realistic uncertainty estimates of the interpolation can be given.

The simulated fields of hydraulic conductivity were designed such that they are not distinguishable by their first two spatial moments, but they are distinguishable by their spatial copulas and hence have a different dependence structure. A geostatistician who knows only about multi-Gaussian dependence could not distinguish these two types of spatial fields. One objective of this thesis was to explore if the spatial structure as modelled by copulas of a hydraulic conductivity field influences plume evolution as evaluated by second central moments. Plume evolution is largely influenced by velocity variations which were introduced by the finely discretized heterogeneous hydraulic conductivity field.

Different spreading behaviour of solute plumes in numerical tracer tests based on those different types of K fields was observed – despite the fact that the different types of spatial K fields are not distinguishable by classical geostatistical methods, and despite the fairly homogeneous setting of Borden with its small variance of K . In two dimensions, the dispersion coefficient obtained from Monte Carlo second central spatial moment analysis of the solute

plumes was shown to be statistically significantly different when K fields with a Gaussian spatial dependence were used as input compared to when K fields with the fitted non-Gaussian spatial dependence were used. In three dimensions, the dispersion coefficients were shown to be less different. Three-dimensional flow paths combined with anisotropy in vertical direction compensate much of the non-Gaussian dependence structure of Borden when analyzing solute transport behaviour. However, only a slight change in the variance of the marginal distribution while keeping the mean unchanged, such that the marginal distribution is different from Borden but still within a possibly naturally occurring range, leads also in three dimensions to significantly different transport behaviour between the same Gaussian and non-Gaussian spatial dependence structures of Borden that were also used in two dimensions. This indicates that for a “well behaved” marginal distribution with small variance and small skewness, the spatial dependence structure might not be the most critical factor influencing solute transport behaviour. However, the less well behaved a marginal distribution is, the more important becomes the type of the spatial dependence structure of K for predicting solute transport behaviour, and even minor deviations from a Gaussian dependence structure such as the Borden spatial dependence matter.

The maximum Gauss copula fitted to the entire data-set of Borden showed properties very similar to the Gaussian copula, however when the data of both cross-section was fitted individually, it was shown that the maximum Gauss copula is capable of modelling non-Gaussian dependence.

The North Bay data-set has a much more skewed marginal distribution than Borden and it has a complicated spatial dependence structure. It was found that its type of dependence changes from strong dependence in low quantiles at short separation distances to strong dependence in high quantiles at larger separation distances. The overall fit of maximum Gauss copula to North Bay equals out the negative symmetry for short separation distances and the positive symmetry for large separation distances. The overall result in terms of ensemble average dispersion is not very different to Gaussian, even though slightly increased dispersivity values compared to the case of Gaussian dependence were found.

Within the process of geostatistical analysis it is important to treat data carefully. It was detected that a region at the of the Borden data-set exhibits elevated hydraulic conductivity measurements compared to the rest of the data-set. A methodology was shown that showed with reasonable statistical significance that the upper region should be treated differently than the rest of the aquifer. Furthermore this methodology shows how this recognition can be incorporated into the the geostatistical workflow leading to simulated hydraulic conductivity fields that represent both the top region and the rest of the aquifer. It was shown that solute transport behaviour, for example peak concentration, is more sensitive to the way the internal structure of the data is modelled than to the model of spatial dependence structure.

This work has proven that the spatial structure of hydraulic conductivity as modelled by using second order moments may not be sufficient in all cases to describe solute transport behaviour. Structure within the data can be detected and accounted for in reasonable ways.

This thesis has shown that spatial copulas can be fitted to data that has been measured in the environment. However, that data is based from only two sites and it remains to be tested if copulas are also useful used with data from other field sites. Such data could be more com-

plex, for example not sandy-aquifers requiring more complex models of spatial dependence. The phenomenon of changing dependence from strongest dependence in low quantiles to strongest dependence in high quantiles was observed in the North Bay data-set, but could not be represented explicitly in any current spatial copula model. Incorporating the property of changing symmetry over increasing separation distances would be one particular improvement to a copula model. It would reflect different processes acting at different length scales, which is probably a very realistic feature.

All the aquifers that were analyzed with copulas to date are sandy porous medium aquifers. They can be modelled well using v -copulas or maximum Gauss copulas. Copulas can be seen as the representation of the process that lead to the formation of the parameter under investigation. It is stipulated that copulas could be used as a proxy for the process. Maybe, eventually, the process or processes could be inferred that created an aquifer or that created a contamination if the spatial structure of hydraulic conductivity or solute concentration is analyzed. One application could be to distinguish between the age of contamination, because a contaminant source that has been in place for a long time (geogenic concentration) tends to lead to Gaussian dependence in their concentration-copula, whereas recent contamination (anthropogenic) tends to lead to non-Gaussian dependence in copulas of related concentration fields.

As pointed out in this thesis, environmental data and particularly hydrogeological data is difficult and expensive to measure. Hence the number of flow and transport parameters that can be measured at any given site is usually quite limited. It is therefore important to augment such data with other hard or soft information which is related to it quantitatively or qualitatively. Among the most promising developments in this area are geophysical methods (such as seismic and electromagnetic geotomography) which remotely image the subsurface in a way that correlates with permeability, porosity, and fluid or salt content [Neuman and Di Federico \(2003\)](#). The further development of such methods of field investigation, and of computational algorithms which take advantage of such correlations to help reduce uncertainty about subsurface flow and transport conditions is an important area of active research. Yet, most of these measurements are based on different support volumes. Copulas as a full stochastic model are thought to be effective in incorporating varying sample support.

It was shown that non-Gaussian and Gaussian dependence structures can lead to significantly different solute transport behaviour. From a hydrogeological standpoint it would be interesting to explore if the dispersivities that were obtained in the numerical tracer tests and / or the hydraulic conductivity modelled by non-Gaussian copulas could lead to a better calibration of solute transport models that try to represent observed concentrations.

Bibliography

- Akaike, H., 1974. A New Look at the Statistical Model Identification. *Automatic Control, IEEE Transactions on* 19 (6), 716–723. doi:10.1109/TAC.1974.1100705.
- Bárdossy, A., 2006. Copula-Based Geostatistical Models for Groundwater Quality Parameters. *Water Resources Research* 42 (W11416). doi:10.1029/2005WR004754.
- Bárdossy, A., Li, J., 2008. Geostatistical Interpolation Using Copulas. *Water Resources Research* 44 (W07412). doi:10.1029/2007WR006115.
- Bear, J., 1972. *Dynamics of Fluids in Porous Media*. Elsevier.
- Brooker, P. I., 1979. Kriging. *Engineering and Mining Journal* 180 (7-12), 148–153.
- Burr, D. T., Sudicky, E. A., Naff, R. L., Mar. 1994. Nonreactive and Reactive Solute Transport in Three-Dimensional Heterogeneous Porous Media: Mean Displacement, Plume Spreading and Uncertainty. *Water Resources Research* 30 (3), 791–815.
- Carle, S. F., Fogg, G. E., 1997. Modeling Spatial Variability with One and Multidimensional Continuous-Lag Markov Chains. *Mathematical Geology* 29 (7), 891–918.
- Cherry, J. A., Barker, J. F., Feenstra, S., Gillham, R. W., MacKay, D. M., Smyth, D. J. A., 1996. The Borden Site for Groundwater Contamination Experiments: 1978–1995. In: Kobus, H., Barczewski, B., Koschitzky, H.-P. (Eds.), *Groundwater and Subsurface Remediation - Research Strategies for In-Situ Remediation*. Springer Verlag, pp. 101–127.
- Cliff, A. D., Ord, J. K., 1973. *Spatial Autocorrelation*. Pion Limited.
- Coles, S., 2001. *An Introduction to Statistical Modeling of Extreme Values*. Springer Verlag.
- Dagan, G., Oct. 1982. Analysis of Flow Through Heterogeneous Random Aquifers – 2. Unsteady Flow in Confined Formations. *Water Resources Research* 18 (5), 1571–1585.
- Dagan, G., Aug. 1984. Solute Transport in Heterogeneous Porous Formations. *Journal of Fluid Mechanics* 145, 151–177. Available from: <http://journals.cambridge.org/action/displayAbstract?aid=390342>, doi:10.1017/S0022112084002858.
- Dagan, G., Sep. 1988. Time-Dependent Macrodispersion for Solute Transport in Anisotropic Heterogeneous Aquifers. *Water Resources Research* 24 (9), 1491–1500.
- Dagan, G., Jun. 1990. Transport in Heterogeneous Porous Formations: Spatial Moments, Ergodicity, and Effective Dispersion. *Water Resources Research* 26 (6), 1281–1290.

- Davis, A. D., Sep. 1986. Deterministic Modeling of Dispersion in Heterogeneous Permeable Media. *Ground Water* 24 (5), 609–615. doi:10.1111/j.1745-6584.1986.tb03709.x.
- de Marsily, G., 1986. *Quantitative Hydrogeology: Groundwater Hydrology for Engineers*, 440 pp. Academic Press, Inc.
- Delhomme, J.-P., 1978. Kriging in the Hydrosociences. *Advances in Water Resources* 1 (5), 251–266.
- Delhomme, J. P., 1979. Spatial Variability and Uncertainty in Groundwater Flow Parameters: A Geostatistical Approach. *Water Resources Research* 15 (2), 269–280.
- Dietrich, P., Butler Jr, J. J., Faiß, K., 2008. A Rapid Method for Hydraulic Profiling in Unconsolidated Formations. *Ground Water* 46 (2), 323–328. doi:10.1111/j.1745-6584.2007.00377.x.
- Dupuis, D. J., Jul. 2007. Using Copulas in Hydrology: Benefits, Cautions, and Issues. *Journal of Hydrologic Engineering* 12 (4), 381–393. doi:10.1061/ASCE1084-0699200712:4381.
- Durante, F., Sempi, C., May 2010. Copula Theory and Its Applications. Vol. 198 of *Lecture Notes in Statistics*. Springer Berlin Heidelberg, Berlin, Heidelberg. Available from: http://www.springerlink.com/index/10.1007/978-3-642-12465-5_1, doi:10.1007/978-3-642-12465-5.
- Emery, X., Jan. 2002. Conditional Simulation of Nongaussian Random Functions. *Mathematical Geology* 34 (1), 79–100.
- Evin, G., Favre, A.-C., 2008. A New Rainfall Model Based on the Neyman-Scott Process Using Cubic Copulas. *Water Resources Research* 44 (W03433). doi:10.1029/2007WR006054.
- Farlie, D., 1960. The Performance of Some Correlation Coefficients for a General Bivariate Distribution. *Biometrika* 47, 307–323. Available from: <http://www.jstor.org/stable/2333302>.
- Fiori, A., Janković, I., Dagan, G., Cvetković, V., Sep. 2007. Ergodic Transport Through Aquifers of Non-Gaussian Log Conductivity Distribution and Occurrence of Anomalous Behavior. *Water Resources Research* 43 (9), W09407. Available from: <http://www.agu.org/pubs/crossref/2007/2007WR005976.shtml>, doi:10.1029/2007WR005976.
- Frank, M. J., Feb. 1978. On the Simultaneous Associativity of $F(X,Y)$ and $X+Y-F(X,Y)$. *Aequationes Mathematicae* 18 (1-2), 266–267. Available from: <http://www.springerlink.com/index/10.1007/BF01844082>, doi:10.1007/BF01844082.
- Frees, E. W., Valdez, E. A., 1998. Understanding Relationships using Copulas. *North American Actuarial Journal* 2 (1), 1–25.

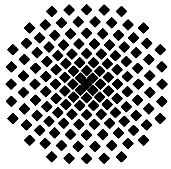
- Freeze, R. A., 1975. A Stochastic-Conceptual Analysis of One-Dimensional Groundwater Flow in Nonuniform Homogeneous Media. *Water Resources Research* 11 (5), 725–741.
- Freeze, R. A., Cherry, J. A., 1979. *Groundwater*. Prentice Hall.
- Freyberg, D. L., 1986. A Natural Gradient Experiment on Solute Transport in a Sand Aquifer – 2. Spatial Moments and the Advection and Dispersion of Nonreactive Tracers. *Water Resources Research* 22 (13), 2031–2046.
- Frind, E. O., Muhammad, D., Molson, J. W., 2002. Delineation of Three-Dimensional Well Capture Zones for Complex Multi-Aquifer Systems. *Ground Water* 40 (6), 585–598.
- Frind, E. O., Sudicky, E. A., Schellenberg, S. I., 1987. Micro-Scale Modelling in the Study of Plume Evolution in Heterogeneous Media. *Stochastic Hydrology and Hydraulics*, 263–279.
- Gambolati, G., Volpi, G., 1979. Groundwater Contour Mapping in Venice by Stochastic Interpolators – 1. Theory. *Water Resources Research* 15 (2), 281–290.
- Gelhar, L. W., 1992. *Stochastic Subsurface Hydrology*. Prentice Hall College Div.
- Gelhar, L. W., Axness, C. L., 1983. Three-Dimensional Analysis of Macrodispersion in Aquifers. *Water Resources Research* 19, 161–180.
- Genest, C., Sep. 1987. Frank's Family of Bivariate Distributions. *Biometrika* 74 (3), 549–555. Available from: <http://biomet.oxfordjournals.org/cgi/content/abstract/74/3/549>, doi:10.1093/biomet/74.3.549.
- Genest, C., Favre, A.-C., Béliveau, J., Jacques, C., 2007. Metaelliptical Copulas and Their Use in Frequency Analysis of Multivariate Hydrological Data. *Water Resources Research* 43 (W09401). doi:10.1029/2006WR005275.
- Genest, C., Gendron, M., Bourdeau-Brien, M., 2009. The Advent of Copulas in Finance. *The European Journal of Finance* 15 (7/8), 609–618. doi:10.1080/13518470802604457.
- Genest, C., Rivest, L.-P., 1993. Statistical-Inference Procedures for Bivariate Archimedean Copulas. *Journal of the American Statistical Association* 88 (423), 1034–1043. Available from: <http://www.jstor.org/stable/2290796>.
- Goltz, I. K., 1991. Spatial Variability of Hydraulic Conductivity in a Sand Aquifer at North Bay, Ontario. Master's thesis, University of Waterloo, Waterloo, Ontario.
- Gómez-Hernández, J. J., Wen, X.-H., 1998. To be or not to be Multi-Gaussian? A Reflection on Stochastic Hydrogeology. *Advances in Water Resources* 21 (1), 47–61. doi:10.1016/S0309-1708(96)00031-0.
- Goovaerts, P., 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press.
- Goovaerts, P., 1999. Geostatistics in Soil Science: State-of-the-Art and Perspectives. *Geoderma* 89, 1–45.

- Goovaerts, P., 2001. Geostatistical Modelling of Uncertainty in Soil Science. *Geoderma* 103, 3–26.
- Goovaerts, P., Avruskin, G., Meliker, J. R., Slotnick, M. J., Jacquez, G. M., Nriagu, J. O., 2005. Geostatistical Modeling of the Spatial Variability of Arsenic in Groundwater of Southeast Michigan. *Water Resources Research* 41 (W07013). doi:10.1029/2004WR003705.
- Grimaldi, S., Serinaldi, F., Apr. 2006. Design Hyetograph Analysis with 3-Copula Function. *Hydrological Sciences-Journal* 51 (2), 223–238.
- Guthke, P., Mar. 2010. Effects of Non-Gaussian Spatial Dependence of Hydraulic Conductivity Fields on Hydrodynamic Macrodispersion. Master's thesis, University of Stuttgart.
- Haslauer, C. P., Bárdossy, A., Sudicky, E. A., 2008. Geostatistical Analysis of Hydraulic Conductivity Fields Using Copulas. *Geostats 2008 – Proceedings of The Eighth International Geostatistical Congress 02*, 889–898.
- Hinterding, A., Streit, U., 2002. Automatic Model Selection for Spatial Interpolation. In: *Proceedings of the 8th annual conference of the international association for mathematical geology IAMG2002*. pp. 87–92.
- Isaaks, E. H., Srivastava, R. M., 1989. *An Introduction to Applied Geostatistics*. Oxford University Press.
- Joe, H., 1997. *Multivariate Models and Dependence Concepts*. Monographs on Statistics and Applied Probability. Chapman & Hall / CRC.
- Journal, A., 1989. *Fundamentals of Geostatistics in Five Lessons*. Tech. rep., American Geophysical Union, Washington, D. C.
- Journal, A., Alabert, F., 1989. Non-Gaussian Data Expansion in the Earth Sciences. *TERRA Nova* 1, 123–134.
- Kazianka, H., Pilz, J., 2009. Copula-Based Geostatistical Modeling of Continuous and Discrete Data Including Covariates. *Stochastic Environmental Research and Risk Assessment* 24 (5), 661–673. doi:10.1007/s00477-009-0353-8.
- Kitanidis, P. K., May 1997. *Introduction to Geostatistics: Applications in Hydrogeology*. Cambridge University Press.
- Krige, D. G., 1951. A Statistical Approach to Some Basic Mine Valuation Problems on the Witwatersrand. *Journal of the Chemical Metallurgical & Mining Society of South Africa* 52 (6), 119–139.
- Li, J., 2010. *Application of Copulas as a New Geostatistical Tool*. Ph.D. thesis, Universität Stuttgart.
- Li, J., Bárdossy, A., Guenni, L., Liu, M., 2011. A Copula Based Observation Network Design Approach. *Environmental Modelling & Software* 26 (11), 1349–1357. Available from: <http://www.sciencedirect.com/science/article/pii/S136481521100106X>, doi:10.1016/j.envsoft.2011.05.001.

- Liu, G., Butler Jr, J. J., Bohling, G. C., Reboulet, E., Knobbe, S., Hyndman, D. W., 2009. A New Method for High-Resolution Characterization of Hydraulic Conductivity. *Water Resources Research* 45 (W08202). doi:10.1029/2009WR008319.
- Louis, H., Fischer, K., 1960. *Allgemeine Geomorphologie*. de Gruyter.
- Marcotte, D., Gloaguen, E., 2008. A Class of Spatial Multivariate Models Based on Copulas. In: Ortiz, J. M., Emery, X. (Eds.), *Geostats 2008 – Proceedings of The Eighth International Geostatistical Congress*. GECAMIN, pp. 177–186.
- Matheron, G., 1963. Principles of Geostatistics. *Economic Geology* 58, 1246–1266.
- Matheron, G., 1971. *The Theory of Regionalized Random Variables and its Applications*. Tech. Rep. 5.
- Morgenstern, D., 1956. Einfache Beispiele zweidimensionaler Verteilungen. *Mitt. Math. Statist* 8, 234–235.
- Naff, R. L., Haley, D. F., Sudicky, E. A., Apr. 1998. High-Resolution Monte Carlo Simulation of Flow and Conservative Transport in Heterogeneous Porous Media – 1. Methodology and Flow Results. *Water Resources Research* 34 (4), 663–677. doi:10.1029/97WR02712.
- Nelsen, R. B., 1999. *An Introduction to Copulas*. Vol. 139 of *Lecture Notes in Statistics*. Springer Verlag.
- Neuman, S. P., Di Federico, V., 2003. Multifaceted Nature of Hydrogeologic Scaling and Its Interpretation. *Reviews of Geophysics* 41 (3), 4–1–4–31.
- Neuman, S. P., Zhang, Y.-K., May 1990. A Quasi-Linear Theory of Non-Fickian and Fickian Subsurface Dispersion – 1. Theoretical Analysis With Application to Isotropic Media. *Water Resources Research* 26 (5), 887–902.
- Rajaram, H., Gelhar, L. W., Jun. 1991. Three-Dimensional Spatial Moments Analysis of the Borden Tracer Test. *Water Resources Research* 27 (6), 1239–1251.
- Rajaram, H., Gelhar, L. W., Sep. 1993. Plume Scale-Dependent Dispersion in Heterogeneous Aquifers – 1. Lagrangian Analysis in a Stratified Aquifer. *Water Resources Research* 29 (9), 3249–3260.
- Rau, M., 2011. *Geostatistical Analysis of Three-Dimensional Hydraulic Conductivity Fields by Means of Maximum Gauss Copula*. Master's thesis, Universität Stuttgart.
- Ritzi, R. W. J., Allen-King, R. M., 2007. Why Did Sudicky [1986] Find an Exponential-Like Spatial Correlation Structure for Hydraulic Conductivity at the Borden Research Site? *Water Resources Research* 43 (W01406). doi:10.1029/2006WR004935.
- Rubin, Y., 1990. *Stochastic Modeling of Macrodispersion in Heterogeneous Porous Media*. *Water Resour. Res.* Available from: <http://www.agu.org/journals/wr/wr9001/89WR01522/89WR01522.pdf>.

- Salmon, F. Recipe for Disaster: The Formula That Killed Wall Street [online], Feb. 2009. Available from: http://www.wired.com/techbiz/it/magazine/17-03/wp_quant.
- Salvadori, G., de Michele, C., Jul. 2007. The Use of Copulas in Hydrology: Theory and Practice. *Journal of Hydrologic Engineering* 12 (4), 369–380. doi:10.1061/ASCE1084-0699200712:4369.
- Schweizer, B., Wolff, E. F., Jul. 1981. On Nonparametric Measures of Dependence for Random Variables. *The Annals of Statistics* 9 (4), 879–885. Available from: <http://www.jstor.org/stable/2240856>.
- Serinaldi, F., Grimaldi, S., Jul. 2007. Fully Nested 3-Copula: Procedure and Application Hydrological Data. *Journal of Hydrologic Engineering* 12 (4), 420–430. doi:10.1061/ASCE1084-0699200712:4420.
- Sklar, A., 1959. Fonctions de répartition a n dimensions et leur marges. *Publ. Inst. Stat. Paris* 8, 229–131.
- Smith, L., Schwartz, F. W., 1980. Mass Transport – 1. A Stochastic Analysis of Macroscopic Dispersion. *Water Resources Research* 16 (2), 303–313.
- Strebelle, S., 2002. Conditional Simulation of Complex Geological Structures Using Multiple-Point Statistics. *Mathematical Geology* 34 (1), 1–21.
- Sudicky, E. A., 1983. Migration of Contaminants in Groundwater at a Landfill: A Case Study – 4. A Natural Gradient Dispersion Test. *Journal of Hydrology* 63, 81–108.
- Sudicky, E. A., 1986. A Natural Gradient Experiment on Solute Transport in a Sand Aquifer: Spatial Variability of Hydraulic Conductivity and Its Role in the Dispersion Process. *Water Resources Research* 22 (13), 2069–2082.
- Sudicky, E. A., Illman, W. A., Goltz, I. K., Adams, J. J., McLaren, R. G., 2010. Heterogeneity in Hydraulic Conductivity and its Role on the Macroscale Transport of a Solute Plume: From Measurements to a Practical Application of Stochastic Flow and Transport Theory. *Water Resources Research* 46. doi:10.1029/2008WR007558.
- Sun, A. Y., Ritzi, R. W. J., Sims, D. W., 2008. Characterization and Modeling of Spatial Variability in a Complex Alluvial Aquifer: Implications on Solute Transport. *Water Resources Research* 44. doi:10.1029/2007WR006119.
- Therrien, R., Sudicky, E. A., 1996. Three-Dimensional Analysis of Variably-Saturated Flow and Solute Transport in Discretely-Fractured Porous Media. *Journal of Contaminant Hydrology* 23, 1–44.
- Todini, E., Pellegrini, F., 1999. A Maximum Likelihood Estimator for Semi-Variogram Parameters in Kriging. In: *GeoENV II: Geostatistics for Environmental Applications: Proceedings of the Second European Conference on Geostatistics for Environmental Applications Held in Valencia, Spain, November 18-20, 1998*. Kluwer Academic Pub, p. 187.

- Volpi, G., Gambolati, G., 1978. On the Use of a Main Trend for the Kriging Technique in Hydrology. *Advances in Hydrosience* 1 (12), 345–349.
- Vomvoris, E. G., Gelhar, L. W., Jan. 1990. Stochastic Analysis of the Concentration Variability in a Three-Dimensional Heterogeneous Aquifer. *Water Resources Research* 26 (10), 2591–2602. Available from: <http://www.agu.org/pubs/crossref/1990/WR026i010p02591.shtml>, doi:10.1029/WR026i010p02591.
- Wang, J., Kitanidis, P. K., 1999. Analysis of Macrodispersion Through Volume Averaging: Comparison With Stochastic Theory. *Stochastic Environmental Research and Risk Assessment* 13, 66–84.
- Woodbury, A. D., Sudicky, E. A., 1991. The Geostatistical Characteristics of the Borden Aquifer. *Water Resources Research* 27 (4), 533–546.
- Woodbury, A. D., Sudicky, E. A., 1992. Inversion of the Borden Tracer Experiment Data – Investigation of Stochastic Moment Models. *Water Resour. Res.* 28 (9), 2387–2398.
- Zhang, L., Singh, V. P., Mar. 2006. Bivariate Flood Frequency Analysis Using the Copula Method. *Journal of Hydrologic Engineering* 11 (2), 150–164.
- Zhang, Y.-K., Neuman, S. P., May 1990. A Quasi-Linear Theory of Non-Fickian and Fickian Subsurface Dispersion – 2. Application to Anisotropic Media and the Borden Site. *Water Resources Research* 26 (5), 903–913.
- Zinn, B., Harvey, C. F., 2003. When Good Statistical Models of Aquifer Heterogeneity Go Bad: a Comparison of Flow, Dispersion, and Mass Transfer in Connected and Multivariate Gaussian Hydraulic Conductivity Fields. *Water Resources Research* 39 (3). doi: 10.1029/2001WR001146.



Institut für Wasserbau 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>

Direktoren

Prof. Dr. rer. nat. Dr.-Ing. András Bárdossy
Prof. Dr.-Ing. Rainer Helmig
Prof. Dr.-Ing. Silke Wieprecht

Vorstand (Stand 01.04.2009)

Prof. Dr. rer. nat. Dr.-Ing. A. Bárdossy
Prof. Dr.-Ing. R. Helmig
Prof. Dr.-Ing. S. Wieprecht
Jürgen Braun, PhD
Dr.-Ing. H. Class
Dr.-Ing. S. Hartmann
Dr.-Ing. H.-P. Koschitzky
PD Dr.-Ing. W. Marx
Dr. rer. nat. J. Seidel

Emeriti

Prof. Dr.-Ing. habil. Dr.-Ing. E.h. Jürgen Giesecke
Prof. Dr.h.c. Dr.-Ing. E.h. Helmut Kobus, PhD

Lehrstuhl für Wasserbau und Wassermengenwirtschaft

Leiter: Prof. Dr.-Ing. Silke Wieprecht
Stellv.: PD Dr.-Ing. Walter Marx, AOR

Versuchsanstalt für Wasserbau

Leiter: Dr.-Ing. Sven Hartmann, AOR

Lehrstuhl für Hydromechanik und Hydrosystemmodellierung

Leiter: Prof. Dr.-Ing. Rainer Helmig
Stellv.: Dr.-Ing. Holger Class, AOR

Lehrstuhl für Hydrologie und Geohydrologie

Leiter: Prof. Dr. rer. nat. Dr.-Ing. András Bárdossy
Stellv.: Dr. rer. nat. Jochen Seidel

VEGAS, Versuchseinrichtung zur Grundwasser- und Altlastensanierung

Leitung: Jürgen Braun, PhD
Dr.-Ing. Hans-Peter Koschitzky, AD

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 Flächen-tragwerke 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; Gesteungskosten 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 Binnenschifffahrt 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 Binnenschifffahrt 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 Gál, 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; Gál, 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 Binnenschifffahrt durch Schwimmerhebwerke*, 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; Gál, 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 rotations-symmetrisch 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 seegangbelasteter 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 horizontalebene 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 Schußbrinnen 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 Klufdaten zur realitätsnahen Generierung von Klufnetzen mit anschließender laminar-turbulenter Strömungsberechnung*, 1993, ISBN 3-92 1694-78-7
- 79 Marschall, Paul: *Die Ermittlung lokaler Stofffrachten im Grundwasser mit Hilfe von Einbohrloch-Meßverfahren*, 1993, ISBN 3-921694-79-5, vergriffen
- 80 Ptak, Thomas: *Stofftransport in heterogenen Porenaquiferen: Felduntersuchungen und stochastische Modellierung*, 1993, ISBN 3-921694-80-9, vergriffen
- 81 Haakh, Frieder: *Transientes Strömungsverhalten in Wirbelkammern*, 1993, ISBN 3-921694-81-7
- 82 Kobus, Helmut; Cirpka, Olaf; Barczewski, Baldur; Koschitzky, Hans-Peter: *Versuchseinrichtung zur Grundwasser und Altlastensanierung VEGAS, Konzeption und Programmrahmen*, 1993, ISBN 3-921694-82-5
- 83 Zang, Weidong: *Optimaler Echtzeit-Betrieb eines Speichers mit aktueller Abflußregenerierung*, 1994, ISBN 3-921694-83-3, vergriffen
- 84 Franke, Hans-Jörg: *Stochastische Modellierung eines flächenhaften Stoffeintrages und Transports in Grundwasser am Beispiel der Pflanzenschutzmittelproblematik*, 1995, ISBN 3-921694-84-1
- 85 Lang, Ulrich: *Simulation regionaler Strömungs- und Transportvorgänge in Karst-aquiferen 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 instationärer 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
- 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 Porous 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 Flusstauhaltungen*, 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 Sub-surface - 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 Yesheatesfa: *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; Stuhmann, 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 eingeeerdeter 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 CO₂ 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; Stuhmann, 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; Stuhmann, 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 CO₂ 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 CO₂ 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 modeling 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 Hochwasserabflussergebnissen*, 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 Bliedernicht, 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

Die Mitteilungshefte ab der Nr. 134 (Jg. 2005) stehen als pdf-Datei über die Homepage des Instituts: www.iws.uni-stuttgart.de zur Verfügung.