

Technical University of Dresden

Faculty of Environmental Science

**Development of a GIS and model-based method for optimizing the
selection of locations for drinking water extraction by means of
riverbank filtration**

A dissertation in partial fulfillment of the requirements for the degree of
Doctor rerum naturalium (Dr. rer. nat.)

Presented by:

Yan Zhou

born on 22. November 1989

in Zhejiang, China

(Defense date: 22.09.2020)

Reviewer:

Prof. Dr. Rudolf Liedl

Technical University of Dresden, Institute of Groundwater Management

Prof. Dr. Olaf Kolditz

Helmholtz-Centre for Environmental Research, Department Environmental Informatics

Prof. Dr. Holger Weiß

Helmholtz-Centre for Environmental Research, Department Environmental Informatics

Übereinstimmungserklärung

Die Übereinstimmung dieses Exemplars mit dem Original der Dissertation zum Thema:

“Development of a GIS and model-based method for optimizing the selection of locations for drinking water extraction by means of riverbank filtration”

wird hiermit bestätigt.

Ort, Datum

Unterschrift (Vorname Name)

Abstract

The lack of safe drinking water worldwide has drawn the attention of decision makers to riverbank filtration (RBF) for its many advantages in purifying surface water. This study provides an overview of the hydrogeologic, fluvial, and environmental influences on the performance of RBF systems and aims to develop a model for RBF site selection. Using multi-attribute utility theory (MAUT), this study structured the RBF siting problem and assessed a multiplicative utility function for the decision maker. In a case study, geostatistical methods were used to acquire the necessary data and geographic information systems (GIS) were used to screen sites suitable for RBF implementation. Those suitable sites were then evaluated and ranked using the multi-attribute utility model. The result showed that sites can be identified as most preferred among the selected suitable sites based on their expected utility values. This study definitively answers the question regarding the capability of MAUT in RBF site selection. Further studies are needed to verify the influences of the attributes on the performance of RBF systems.

Zusammenfassung

Der weltweite Mangel an sauberem Trinkwasser hat die Entscheidungsträger auf die Uferfiltration (RBF) aufmerksam werden lassen, die vielfältige Vorteile bei der Reinigung von Oberflächenwasser bietet. Diese Studie bietet einen Überblick über die hydrogeologischen, fluvialen und umweltrelevanten Einflüsse auf die Leistung von RBF-Systemen und zielt darauf ab, ein Modell für die Auswahl von Standorten zur Wassergewinnung mittels Uferfiltration zu entwickeln. Unter Verwendung der multiattributiven Nutzwertanalyse (MAUT) strukturiert diese Studie das Problem der Standortauswahl zur Wassergewinnung mittels Uferfiltration und bewertet eine multiplikative Nutzenfunktion für den Entscheidungsträger. In einer Fallstudie werden geostatistische Methoden zur Erfassung der erforderlichen Daten und geografische Informationssysteme (GIS) zur Auswahl von Standorten verwendet, die für die Implementierung der Uferfiltration geeignet sind. Diese geeignete Standorte werden dann unter Verwendung des Nutzwertmodell bewertet und eingestuft. Das Ergebnis zeigt, dass die am besten geeigneten Standorte anhand ihrer erwarteten Nutzwerte unter die ausgewählte Standorten identifiziert werden können. Diese Studie beantwortet definitiv die Frage nach der Eignung von MAUT bei der Standortauswahl zur Wassergewinnung mittels Uferfiltration. Weitere Studien sind erforderlich, um den Einfluss der Attribute auf die Leistung von RBF-Systemen zu überprüfen.

Acknowledgments

I wish to express my sincere appreciation to my committee, Professor Rudolf Liedl, Professor Olaf Kolditz, and Professor Holger Weiß, who convincingly guided and encouraged me to be professional and do the right thing even when the road got tough. Without their persistent help, the goal of this project would not have been realized. I wish to show my deepest gratitude to my advisor, Professor Marc Walther, for his consistent support, guidance, constructive criticism and patience. The completion of my thesis would not have been possible without the support and nurturing by him. I am extremely grateful to Professor Thomas Grischek, Jürgen Schubert, Dr. Mick Wu, Professor Wibhu Kutanan, Dr. Yiwen Zhan, and Dr. Songcan Chen, who gave me valuable advice, invaluable experience, and profound belief in my work. Many thanks to Holger Weiß, Marc Walther, and Wibhu Kutanan for reviewing the thesis.

The China Scholarship Council (CSC), the Department of Environmental Informatics of the Helmholtz Centre for Environmental Research (UFZ), and the Helmholtz Interdisciplinary GRADuate School for Environmental Research (HIGRADE) are truly appreciated. Without their support and funding, this project could not have reached its goal.

I would like to extend my sincere thanks to colleagues at Jilin University and Tongji University, especially Professor Xiaosi Su, Professor Yalei Zhang, Dr. Xinqiang Du, Dr. Geng Cui, Hexuan Zhang, and Dr. Yiming Su, for their assistance with the collection of my data.

Also, thanks should go to the colleagues of the Department of Environmental Informatics and the entire team of former department Groundwater Remediation at UFZ, especially Ralf Trabitzsch, Dr. Erik Nixdorf, and Dr. Wenqing Wang, who never wavered in their support. Special thanks to Eva Joanna Koper and the whole staff of the Graduate School HIGRADE for their relentless support and helpful advice during this time.

I would like to recognize the help that I received from Dr. Zhiyong Song and Sarah Knechtel, who extended a great amount of assistance in preparing the maps. I gratefully acknowledge the assistance of the staff of the UFZ library, Jingzheng Zhang, Jürgen Schubert, Dr. Frank Thomas Lange, Thomas Grischek, Dr. Bin Hu, Professor Ralph L. Keeney, and numerous people, thank you for helping me in finding the sources for my thesis.

Finally, I would like to thank my family and friends for supporting me during the compilation of this thesis. Especially, I am grateful to Jiayan Gao for her patience and support in many ways.

Table of Contents

Abstract.....	iii
Zusammenfassung.....	iv
Acknowledgments.....	v
Table of Contents.....	vi
List of Tables.....	viii
List of Figures.....	x
Definition of terms.....	xiii
1. Abbreviations.....	xiii
2. Symbols.....	xiii
Part I Introduction.....	1
1. Introduction.....	2
2. Statement of purpose.....	2
3. Research questions.....	3
4. Overview of methodology.....	3
5. Organization of the dissertation.....	3
Part II Fundamentals and Literature Review.....	5
1. The definition of bank filtration.....	6
2. The Significance of RBF.....	7
2.1 RBF in drinking water supply.....	7
2.2 Benefits of RBF for China.....	14
3. RBF Site Selection.....	19
3.1 RBF site selection model.....	20
3.2 Definition of successful RBF sites.....	24
4. Factors Affecting RBF Site Selection.....	26
4.1 River hydrology/hydraulics.....	27
4.2 Geology.....	28
4.3 Land cover.....	36
4.4 Well field location.....	36
4.5 Water quality.....	37
4.6 Aquifer properties.....	38
4.7 Distance to river.....	41
4.8 Riverbed characteristics.....	43
5. Effect of Clogging on Yield.....	46
6. Summary.....	51
Part III Developing a Multi-attribute Utility Model for RBF Site Selection.....	53
1. Introduction.....	54
2. Objectives and Attributes.....	54

3.	Assessment of the Utility Function.....	57
3.1	Investigation of the qualitative preference structure.....	58
3.2	Assessment of component utility function.....	62
3.3	Assessment of the scaling constants	63
4.	Results.....	67
5.	Discussion.....	69
6.	Summary.....	74
Part IV Case Study.....		75
1.	Introduction.....	76
2.	Materials and Methods.....	78
2.1	GIS data collection.....	78
2.1.1	Geologic data.....	79
2.1.2	Land cover data	79
2.1.3	Groundwater quality data	80
2.1.4	Aquifer properties data.....	80
2.1.5	Surface water area data	80
2.1.6	Surface water quality data	81
2.1.7	Streambed material data	81
2.2	Kriging the saturated thickness.....	91
2.3	Aggregation of all constraint maps	103
3.	Results.....	105
3.1	Kriging.....	105
3.2	Suitable sites	105
4.	Discussion.....	109
4.1	A discussion of the kriging results.....	109
4.2	A discussion of the multi-attribute utility model results.....	117
5.	Summary.....	122
Part V Conclusions and Recommendations		123
1.	Conclusion and Recommendation	124
Appendix 1 Environmental quality standards for surface water (GB 3838-2002).....		125
Appendix 2 Quality standard for groundwater (GB14848-93).....		127
Appendix 3 Explanation to Germany's RBF site location data.....		130
Appendix 4 Layer information of drillings.....		133
Appendix 5 Streambed materials used by Schälchli (1993)		141
Appendix 6 Interview and questionnaires.....		143
Appendix 7 Surface water area of Jilin City.....		150
Bibliography		152

List of Tables

Table 1: Public water supply by State in Germany, 2016.....	9
Table 2: Public water supply by Basin in Germany, 2016	9
Table 3: Public water supply of Germany, 1991-2016	14
Table 4: Hydrogeologic condition of six-successful sites in Egypt.....	24
Table 5: Site conditions of 33 successful RBF sites in Central Europe.....	26
Table 6: RBF well/well field in Germany.....	30
Table 7: Processes of surface rock formations at RBF sites.....	31
Table 8: Petrography of surface rock formations at RBF sites	33
Table 9: Stratigraphy of surface rock formations at RBF sites.....	33
Table 10: Most frequent surface rock formations at RBF sites	34
Table 11: Size and hydraulic conductivity of sand filters.....	39
Table 12: Streambed materials used by Schälchli (1993).....	44
Table 13: Parameters of Beyer-Banscher’s quasi-stable final state of clogging	48
Table 14: List of attributes for the RBF siting problem.....	55
Table 15: Indifferent pairs and their implied relative scaling constants.....	65
Table 16: Component utility functions and scaling constants	68
Table 17: Assumed attributes distribution	70
Table 18: Partial rank correlation for the relationship between final utility and attributes.....	72
Table 19: Source water (surface water) quality conditions in Jilin City, 2018.....	76
Table 20: Grain sizes of streambed material sample	91
Table 21: Saturated thickness data.....	92
Table 22: Empirical semivariogram (omnidirectional) of saturated thickness data ($N=39$) with a lag spacing of 1000 m	95
Table 23: Parameters of the fitted semivariogram models	97
Table 24: Comparison of six suitable sites	107
Table 25: The utility for six suitable sites.....	108
Table 26: A comparison of the residuals ($N=39$).....	112
Table 27: A comparison of the predicted values to the true values ($N=39$)	113
Table 28: Residual distribution ($N=40$)	114
Table 29: Cross validation ($N=40$).....	115
Table 30: Scaling constants and p from 10 simulations	117
Table 31: Ranking results from the rank order weight simulation (six suitable sites)	119
Table 32: Improvement of each attribute needed for each site to be ranked as the most preferred.....	121
Table A1: Basic indicators and limitations of environmental quality standards for surface water.....	126

Table A2: Indicators of groundwater quality.....	128
Table A3: Rating of each indicator.....	129
Table A4 Category defined by comprehensive rating	129
Table A5: Properties of streambed materials the associated thickness of clogging layer in experiment flume	142

List of Figures

Figure 1: An illustration of bank filtration process.....	6
Figure 2: A notice posted by the Cranstonhill Water Works Co on 1 October 1808	8
Figure 3a: Public water supply of Germany by state, 2016.....	12
Figure 3b: Public water supply of Germany by drainage basin, 2016.....	13
Figure 4: Length of water supply pipelines in China, from 1978 to 2016.....	14
Figure 5: Drinking water production capacity and population, 2006-2016.....	15
Figure 6: Surface water & groundwater quality in China, 2016.....	17
Figure 7: Bottled water consumption in China, from 2006 to 2016	18
Figure 8: Lee’s decision model on RBF site selection	21
Figure 9: Archwichai’s decision model on RBF site selection.....	22
Figure 10: Wang’s decision model on RBF site selection.....	23
Figure 11: Simulated DOC concentration from RBF well using Lenk’s equation.....	25
Figure 12: Constraints on applying riverbank filtration	26
Figure 13: Distribution RBF well/well field and their dominant formations in Germany	29
Figure 14: The distribution of fluvial deposits and populated cities in China.....	35
Figure 15: Hydraulic conductivity of unconsolidated deposits	39
Figure 16: Begin of the breakup of the armor layer Θ_A/Θ_{cr} in depended on d_{mA}/d_m ratio .	44
Figure 17: Relationship between d_{mA}/d_m and d_{90}/d_m , Θ_A/Θ_{cr} and d_{90}/d_m	45
Figure 18: Change of flow rate over time depending on the change in hydraulic conductivity.....	47
Figure 19: Change of flow rate depending on the thickness of aquifer strata at quasi- stable final state of clogging	49
Figure 20: Change of flow rate (in unit) depending on the thickness of aquifer strata at quasi-stable final state of clogging	50
Figure 21: Hierarchical structure of attributes for the RBF siting problem.....	55
Figure 22: Validating the preferential independence.....	59
Figure 23: Validating the utility independence and certainty equivalents for X_I	60
Figure 24: Check the additive independence	62
Figure 25: An illustration of assessing a component utility function for X_I	63
Figure 26: Component utility functions.....	64
Figure 27: A lottery that generates the fifth equation	66
Figure 28: Distributions of attribute values at RBF sites.....	69
Figure 29: Empirical cumulative distributions	70
Figure 30: Influence of individual input variable/attribute on model output.....	71
Figure 31: Partial rank correlation coefficients of five attributes.....	73
Figure 32: Comparing the utility changes of mean DOC concentration for range 0-10 mg/L and 50-60 mg/L	73

Figure 33: Study area	77
Figure 34: A flowchart of methodology for site selection and evaluation	79
Figure 35: Alluvium map of Jilin City.....	83
Figure 36: Land cover map of Jilin City.....	84
Figure 37: Groundwater quality map of Jilin City.....	85
Figure 38: Hydraulic conductivity class map	86
Figure 39: Unsuitable distance to surface water (60 meters buffer).....	87
Figure 40: Suitable distance to surface water (310 meters buffer)	88
Figure 41: Locations of surface water quality observation in Jilin City.....	89
Figure 42: Locations of streambed material in Jilin City	90
Figure 43: Grain size distribution of streambed material samples	91
Figure 44: Borehole records of saturated thickness.....	93
Figure 45: Variogram (omnidirectional) of saturated thickness data	94
Figure 46: Semivariogram (omnidirectional) of saturated thickness data ($N=39$)	95
Figure 47: Empirical semivariogram (omnidirectional) with different lag spacing	96
Figure 48: Empirical semivariogram (omnidirectional) with a largest lag distance of 13000 m	97
Figure 49: Mathematical fitting of the semivariogram models	98
Figure 50: Interpolation grid and sample points	99
Figure 51: An illustration of neighbor search strategy	100
Figure 52: An illustration of predicting an unknown location.....	101
Figure 53: Neighbor-search strategy with radius = 3500m	102
Figure 54: Neighbor-search strategy with radius = 3500m (at least 4 neighbors).....	103
Figure 55: Overlay of constraint maps.....	104
Figure 56: Saturated thickness map	106
Figure 57: Suitable areas.....	107
Figure 58: Six suitable sites	108
Figure 59: Comparison of two search-strategies using ordinary kriging ($N=39$)	110
Figure 60: Comparison of two search-strategies using inverse distance weighting ($N=39$).....	111
Figure 61: Residual distribution ($N=39$).....	111
Figure 62: Cross validation ($N=39$)	113
Figure 63: Residual distribution ($N=40$).....	114
Figure 64: Cross validation ($N=40$)	116
Figure 65: Ranks of suitable sites after 10 simulations	118
Figure 66: A heat map of the rank frequency of six suitable sites.....	119
Figure A1: An illustration of RBF wells identification	132
Figure A2: Layer information of drilling 1, with sandy formation on the surface, near lake Müggelsee	134

Figure A3: Layer information of drilling 2, with sandy formation on the surface, near lake Müggelsee	135
Figure A4: Layer information of drilling 3, with mud (silt or clay) formation on the surface, near river Havel	136
Figure A5: Layer information of drilling 4, with mud (silt or clay) formation on the surface, near river Havel	137
Figure A6: Layer information of drilling 5, with mud (silt or clay) formation on the surface, near river Havel	138
Figure A7: Layer information of drilling 6, with lowland moor formation on the surface, near river Havel.....	139
Figure A8: Layer information of drilling 7, with lowland moor formation on the surface, near lake Müggelsee	140
Figure A9: Surface water area of Jilin City	151

Definition of terms

1. Abbreviations

AHP	Analytic Hierarchy Process
BGR	Federal Institute for Geosciences and Natural Resources of Germany
DOC	Dissolved Organic Carbon
GAC	Granular Activated Carbon
GIS	Geographic Information System
IDW	Inverse Distance Weighting
LBF	Lake Bank Filtration
LHS	Latin Hypercube Sampling
LK	Lognormal Kriging
MAR	Managed Aquifer Recharge
MAUT	Multi-attribute Utility Theory
MEE	Ministry of Ecology and Environment of the People's Republic of China
MOH	Ministry of Health of the People's Republic of China
MOHURD	Ministry of Housing and Urban-Rural Development of the People's Republic of China
OK	Ordinary Kriging
OMPs	Organic Micropollutants
PRCC	Partial Rank Correlation Coefficient
RBF	Riverbank Filtration
SAC	Standardization Administration of the People's Republic of China

2. Symbols

\succcurlyeq	Reads "is at least as desirable as"
μ	Mean (Part III); Lagrange parameter (Part IV)
b	Saturated aquifer thickness (m)
c	$h/h(t)$ ratio (Part II); partial sill (Part IV)
c_0	Nugget
C_U	Uniformity coefficient
CV	Coefficient of variation
d	Constant water depth (m)
d_i	Sediment diameter (grain size) for which i percent of the sediment sample is finer (mm)
d_m	Average diameter (grain size) of the streambed material (mm)

d_{mA}	Average diameter (grain size) of the armor layer (mm)
E	Expectation operator
F	Comprehensive rating value of groundwater quality
h	Thickness of the aquifer strata (m) (Part II); distance (m) (Part IV)
$h(t)$	Thickness of the clogging layer at time t (m)
h_f	Length of groundwater-flow paths (m)
h_{ij}	Euclidean distance between location i and j (m)
IQR	Interquartile range of a dataset
k	Scaling constant in the multiplicative utility function
K	Hydraulic conductivity (m/s)
$K(t)$	Hydraulic conductivity of the clogging layer at time t (m/s)
K_c	Hydraulic conductivity class
K_{gravel}	Hydraulic conductivity of the gravelly aquifer strata (m/s)
k_i	Scaling constant of attribute X_i in the multi-attribute utility function
K_m	Average hydraulic conductivity of the two-layer aquifer system (m/s)
K_{sand}	Hydraulic conductivity of the sandy aquifer strata (m/s)
L	Distance between wells and surface water (m) (Part II); lottery (Part III)
m	Mean of a dataset
M	Median (second quartile) of a dataset
MAE	Mean absolute errors
max	Maximum of a dataset
min	Minimum of a dataset
MSE	Mean squared errors
n	Number of neighbors
N	Sample size
p	Probability (Part III); power parameter (IDW) (Part IV)
q	Initial saturated infiltration flux (m/s)
Q	Initial flow rate (m ³ /s)
$q(t)$	Saturated infiltration flux at time t (m/s)
$Q(t)$	Flow rate at time t (m ³ /s)
$Q1$	Lower (first) quartile of a dataset
$Q3$	Upper (third) quartile of a dataset
q_{gravel}	Initial infiltration flux of gravelly materials (m/s)
Q_{gravel}	Initial flow rate of gravelly materials (m ³ /s)
q_{sand}	Initial infiltration flux of sandy materials (m/s)
Q_{sand}	Initial flow rate of sandy materials (m ³ /s)
r	Correlation coefficient; range (Part IV)
R	Radius (m)
S	Area of infiltration (m ²)
SSE	Error Sum of Squares

T	Transmissivity (m^2/s)
$u(\mathbf{x})$	$u(\mathbf{x}) = u(x_1, x_2, \dots, x_5)$, which is the value of utility u at \mathbf{x} (here five attributes), including each multi-attribute consequence, <i>i.e.</i> , each RBF site condition
u_i	A utility function over the single attribute X_i
v_i	Sample value at location i
w_i	Weight of neighbor value v_i
X_i	Attribute, $i=1, 2, \dots, 5$ (Part III)
$\gamma(h_{ij})$	Semivariogram of the pair of location i and j
Θ_A	Critical dimensionless shear stress of the armor layer to start breaking up
Θ_{cr}	Critical dimensionless shear stress of bed load begin
σ	Standard deviation of a dataset

Part I
Introduction

1. Introduction

Before the *Vibrio cholerae* was discovered as the cause of the disease cholera, John Snow (1813-1858), one of the founding fathers of epidemiology, argued that cholera is a specific disease spread by water contaminated by faeces hence he recommended in his 1849 essay *On the Mode of Communication of Cholera* that water be “filtered and well boiled before it is used” (Snow 1849, 30).

Bank filtration is a natural system through which waterborne pathogens (such as *Vibrio cholerae*) and many other contaminants from surface waters are inactivated or removed. It is a cost-effective treatment method for water utilities because if well operated, the system provides high-quality raw water that needs only a few additional treatments. Bank filtration has so many advantages that it has been applied in many countries throughout the world to improve the quality of raw water for the production of drinking water.

In many regions of the world, surface water is the main source of drinking water. For example, in 2016, lacking clean groundwater sources, 85% of the drinking water in China was directly withdrawn and treated from surface water, such as rivers, lakes, and reservoirs. Unlike groundwater and bank-filtrated surface water, surface water directly withdrawn for drinking water production has certain disadvantages, and it often contains a large number of pathogens, which are the main concern to public health. Using bank filtration as the first step of drinking water production will greatly improve drinking water safety and reduce its overall treatment cost. Increasing awareness of drinking water safety and the declining water quality caused by poor management have made people eager to apply and promote bank filtration technology, particularly in less-developed countries.

Every successful bank filtration system starts with siting the system at a suitable place. Presently, only a few models on bank filtration site selection have been developed and which factors most affect the performance of bank filtration systems in these models are still not clear. These models account for trade-offs among attributes but do not take uncertainty into account. Furthermore, they lack procedures to test the robustness of the model. This thesis aims to develop a new model to select and evaluate suitable sites for the implementation of bank filtration systems associated with rivers or streams, *i.e.*, riverbank filtration (RBF), with consideration of the encountered uncertainty during decision making.

2. Statement of purpose

The purpose of this qualitative and quantitative study is to better articulate the objectives and attributes of the RBF site selection problem. This project aims to develop a new model of site selection for drinking water extraction by means of RBF in cities with perennial rivers considering the encountered uncertainty during decision making.

3. Research questions

This research aims to answer the following questions: How do the influencing factors affect the performance of a RBF system used for drinking water supply? What are the most important factors for maximizing the yield of a drinking water extraction well and at the same time optimizing the quality of the well water? How can multi-attribute utility theory (MAUT) support the decision maker in choosing the most preferred RBF site for drinking water supply? How robust is the multi-attribute utility model under uncertainty of trade-offs considered by the decision maker?

4. Overview of methodology

To answer these research questions, first of all, empirical studies were reviewed and the most important factors affecting the performance of RBF system were identified, based on which the RBF siting problem was structured. Objectives and attributes for the RBF siting problem were then articulated and defined.

Multi-attribute utility theory (MAUT) was applied to develop an evaluating system for the site selection of RBF systems. Component utility functions for the attributes and multi-attribute utility function for the sites suitable for RBF implementation were built, the assessment of which were based on the decision maker's preferences. Several consistency checks were made during the assessment procedures. Data necessary for our model were collected from the authorities and empirical studies. Most of the data can be directly converted to thematic maps or layers used for further analysis. Maps that cannot be obtained directly were built based on the available data. For instance, geostatistical techniques were used as spatial interpolation tools for creating the saturated aquifer thickness map, and "Buffer" as a technique of geographic information system (GIS) in ArcMap 10.6 was used to create the buffer maps (distances to the river) important for RBF siting. Geostatistical techniques were conducted using the programming language R version 3.6.1.

Thematic maps acquired from GIS-technique and geostatistical techniques were compiled using the ArcMap platform in order to screen the suitable sites for RBF. The associated attribute values for those suitable sites were also assigned or estimated based on the thematic maps and other GIS data. Then, the identified suitable sites were evaluated using the multi-attribute utility model developed from the previous steps. Based on the final utilities of the suitable RBF sites, the most preferred can be chosen. A sensitivity analysis was performed to identify the most important attribute under the influence of the variable distribution and range. Additionally, a weight-based sensitivity analysis was performed to check the influence of the uncertainty during the decision making process.

5. Organization of the dissertation

The dissertation consists of five parts. Part I introduces the topic. Part II gives the general coverage of RBF and identifies the most important factors affecting RBF site selection for drinking water supply. As a background of the study, it gives insights for better structuring the RBF siting problem before assessing the multi-attribute utility function. Part III defines the objectives and associated attributes of the RBF siting problem and performs a new approach on RBF site selection. The new decision model takes uncertainty into account. As a major finding of the thesis, this part gives a multiplicative utility function that can be used to calculate the utility of all suitable sites, based on which they can be ranked; then, the most suitable can be recommended. Part IV contains a case study of RBF site selection, which is a rather straightforward application of the multi-attribute utility model. Additionally, this part presents a method of delineating suitable areas or sites for RBF implementation. This part also contains an application of using geostatistical techniques to interpolate spatial data. A conclusion of the thesis is given in Part V; limitations of the thesis and recommendations for further research are also given.

Part II

Fundamentals and Literature Review

1. The definition of bank filtration

Bank filtration systems are characterized by a series of abstraction wells by a stream or lake, which results in groundwater depletion that forces river water or lake water to infiltrate into the subsurface towards abstraction wells (Figure 1) (Heath 1983; Schmidt *et al.* 2003; Tufenkji, Ryan, and Elimelech 2002; C. Ray *et al.* 2002; Hiscock and Grischek 2002). For this reason, groundwater withdrawn by water facilities from those wells is actually a mixture of true groundwater from the adjacent aquifer and induced surface water, the latter also called *bank-filtered surface water* or *bank filtrate*. The riverbed acts as a filter medium, which retains the solid ingredients (fine particles) of the leachate during infiltration (Schälchli 1992).

If the associated surface water body is a river or stream, the process is called *riverbank filtration* (RBF); for lakes, it is called *lake bank filtration* (LBF) (Gupta *et al.* 2015; Dash *et al.* 2008; Rocha and Marques 2018; Dillon *et al.* 2019). Bank filtration, induced bank filtration, induced infiltration, or induced recharge essentially all have the same meaning.

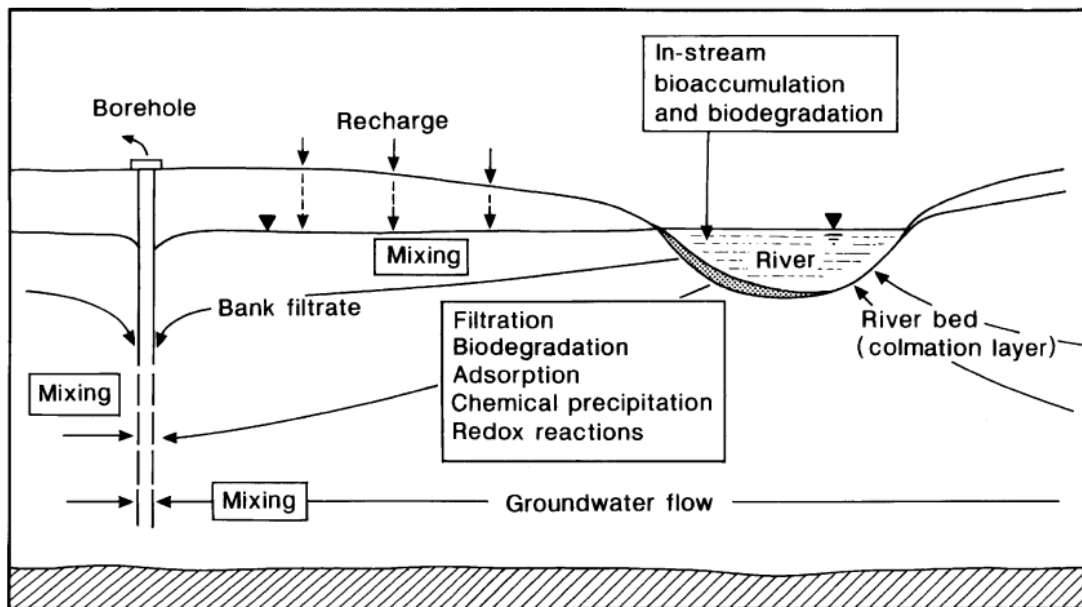


Figure 1: An illustration of bank filtration process

Source: Hiscock and Grischek (2002).

Waterworks applying bank filtration have multifold advantages including storage and recharge of local groundwater resources by diminishing the deficit of necessary water

in the aquifer, providing protection against shock loads to the source water, equilibration of temperature, and reliable attenuation of turbidity, dissolved organic carbon (DOC), pathogens, pharmaceutical contaminants, and other water constituents of concern (Kühn and Müller 2000; Schijven, Berger, and Miettinen 2002; Schubert 2002a, 2004a; Drewes and Summers 2002; Martín-Alonso 2004; Hoppe-Jones, Oldham, and Drewes 2010; Bradley *et al.* 2014). Frequently, bank filtration works together with *artificial groundwater recharge* or artificial infiltration to store water in the alluvial aquifer or to repel contaminated landside groundwater.

2. The Significance of RBF

2.1 RBF in drinking water supply

People have been using RBF in supplying drinking water for more than two hundred years (Glasgow City Archives 2020; Burnet 1869, 11; BMI 1985, 17; Ray *et al.* 2002, 2). Since 1808 in the United Kingdom, the Glasgow Waterworks Company started to supply filtered water from the Clyde River to customers in the City of Glasgow and suburbs (Figure 2) (Glasgow City Archives 2020; Burnet 1869, 5, 11; BMI 1985, 17). Water was filtered through the sandy riverbank and percolated into a brick tunnel from the Clyde River (Burnet 1869, 11). It was the first known utility to use RBF for purposes of water supply (Ray *et al.* 2002, 2). Since then, other cities in the United Kingdom have followed the example of Glasgow, such as Nottingham, Perth, Derby, and Newark. In the 1850s, it was officially adopted by other European countries or regions, especially along the Rhine, Elbe, and Danube Rivers (Ray *et al.* 2002). The purposeful extraction of bank-filtrated river water in Germany dates back to the 1870s (Kühn and Müller 2006). RBF has been used for drinking water supplies for more than 140 years in Germany. Well-known examples are Flehe Waterworks (1870) in Düsseldorf on the Rhine River and Saloppe Waterworks (1875) in Dresden on the Elbe River (Kühn and Müller 2006). In the Netherlands, the first bank-filtrated river water pumped for public drinking water supplies was probably in 1879 along the Rhine River at pumping station Hijmegen (Stuyfzand, Juhász-Holterman, and de Lange 2004).

Nowadays, water utilities still use RBF for drinking water supplies, and in some regions, particularly developed countries, it plays a critical role (Hiscock and Grischek 2002; Stuyfzand, Juhász-Holterman, and de Lange 2004; Lee and Lee 2010; Schubert 2002b). In 2014, 68 million m³ of bank-filtered river water was abstracted in the Netherlands, which equals 5.6% of the total abstracted amount for drinking water production (Geudens 2015). In 2016, the share of bank-filtered surface water and artificially enriched groundwater in public supply was about 17% in Germany (Destatis

2019b). A total of 97% of the population (about 80 million people) were connected to the public drinking

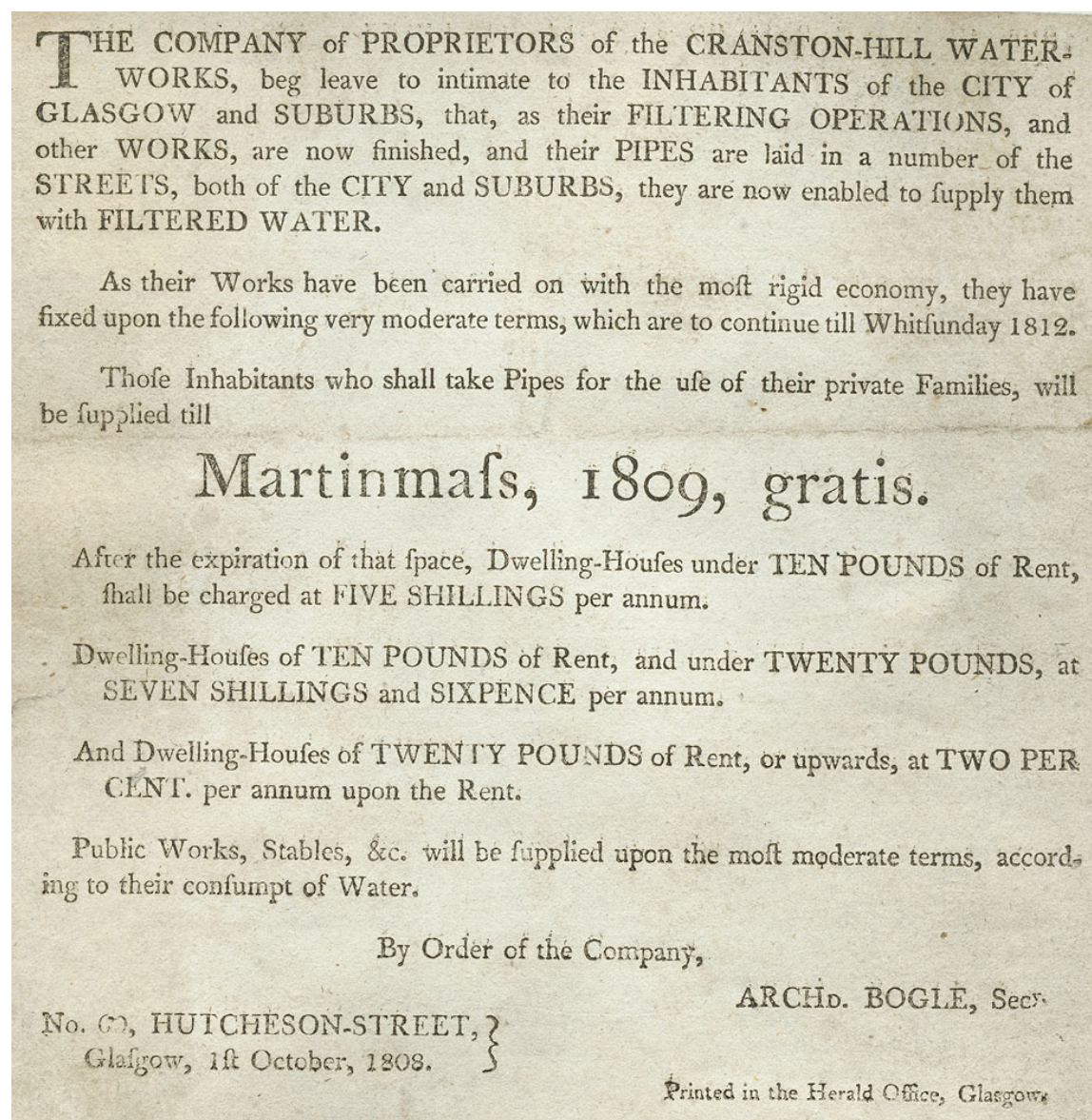


Figure 2: A notice posted by the Cranstonhill Water Works Co. on 1 October 1808

Source: Glasgow City Archives (<https://www.theglasgowstory.com/image/?inum=TGSA02064>).

water supply in Germany in 2016 (Destatis 2018). In total, the public water supply companies have gained around 5.2 billion cubic meters (inclusive total water supply to the public, losses or measurement differences, and waterworks' self-use) of drinking water (see Table 1). As reported by the Federal Office of Statistics (2019b), drinking water in

Germany was predominantly produced from true groundwater (61%). In addition to spring water (8%), groundwater (including true groundwater and spring water) is the most preferred drinking water source in Germany. Surface waters, in general, are less preferred. Among all surface water types, lake and reservoir water (12%) had the highest share of the

Table 1: Public water supply by State in Germany, 2016

State	Total water supply	True groundwater	Spring water	Bank-filtrated surface water	Artificially enriched groundwater	Lake and reservoir water	River water
Baden-Württemberg	678	52%	18%	1%	1%	23%	6%
Bavaria	868	72%	18%	7%	-	3%	-
Berlin	221	29%	-	57%	14%	-	-
Brandenburg	130	96%	-	2%	2%	-	-
Bremen	14	100%	-	-	-	-	-
Hamburg	116	100%	-	-	-	-	-
Hesse	353	76%	12%	-	12%	-	-
Mecklenburg-Vorpommern	96	84%	-	3%	-	-	13%
Lower Saxony	555	85%	2%	-	-	13%	-
North Rhine-Westphalia	1207	40%	2%	11%	31%	16%	1%
Rhineland-Palatinate	252	72%	13%	12%	-	4%	-
Saarland	63	97%	3%	-	-	-	-
Saxony	278	22%	4%	20%	3%	49%	1%
Saxony-Anhalt	74	73%	1%	5%	18%	4%	-
Schleswig-Holstein	179	100%	-	-	-	-	-
Thuringia	121	43%	13%	-	-	44%	-
National total	5204	61%	8%	8%	9%	12%	1%

Note: total water supply values are in million cubic meters, water source values are shown in percent; values in this table may not sum to totals because of independent rounding; source: Destatis (2019b).

total withdrawals; however, two specific types of surface water—bank-filtered surface water, although usually defined separately from surface water by German authorities, and artificially enriched groundwater, abstracted by 412 facilities from 124 waterworks in total in Germany—delivered as much as 17% of safe drinking water to residents in 2016 (Destatis 2003, 2019b). These two types of water sources contributed the most to surface water withdrawals. Direct intake of river water was reduced to a limited amount, probably because of pollution (Kühn and Müller 2000). Germany’s capital, Berlin, depended 57% on bank filtration and 14% on artificial groundwater recharge. This represents the highest share of bank-filtrated surface water among all states. North Rhine-Westphalia, the most populous state of Germany, had the largest amount of total public water withdrawals

among all states, and abstracted the most bank-filtrated surface water as well as artificially enriched groundwater. In this state, the share of bank-filtrated surface water and artificially enriched groundwater reached 42%. In Saxony, a state in eastern Germany, 20% of raw water production was based on bank filtration. The most notable use of bank filtration and artificial groundwater recharge for drinking water supply could be found within the Elbe basin (18% from bank-filtered surface water) and the Rhine basin (18% from artificially enriched groundwater) (Destatis 2019b)(see Table 2). See also Figure 3 for states and basins mentioned in the text.

Table 2: Public water supply by Basin in Germany, 2016

River basin	Total water supply	True groundwater	Spring water	Bank-filtrated surface water	Artificially enriched groundwater	Lake and reservoir water	River water
Danube	810	70%	17%	7%	-	1%	5%
Rhine	2290	51%	9%	7%	18%	14%	-
Ems	196	92%	-	-	7%	-	-
Weser	534	77%	7%	-	1%	15%	-
Elbe	1056	57%	2%	18%	5%	17%	-
Oder	33	90%	4%	3%	3%	-	-
Meuse	115	62%	-	-	-	37%	-
Eider	38	100%	-	-	-	-	-
Schlei/Trave	64	100%	-	-	-	-	-
Warnow/Peene	67	80%	-	1%	-	-	18%
National total	5204	61%	8%	8%	9%	12%	1%

Note: total water supply values are in million cubic meters; values in this table may not sum to totals because of independent rounding; source: Destatis (2019b).

Since the reunification (1989/1990), total drinking water supply/consumption in Germany continually decreased until 2013, especially the amount of water abstracted from underground sources (see Table 3). Population shows an inverted U-shaped trend: first an increase and then a decrease. The decreasing trend of population also stopped in 2013, yet the decreasing trend of bank-filtrated surface water had already stopped in 2001 and then increased significantly; its share to total public water supply increased from 5% to 8%. Apparently, raw water abstracted by RBF systems was still a favorable drinking water source in Germany, despite its old implementation history.

Apart from developed countries with further industrialization and increasing awareness of drinking water safety, some newly industrialized countries and developing countries are also paying attention to RBF (Archwichai *et al.* 2011; Ghodeif *et al.* 2016; Hu *et al.* 2016; Sandhu *et al.* 2011; Freitas *et al.* 2012; Dillon *et al.* 2019). Ray (2008) argued that there was a worldwide potential of RBF. Research on the potential of RBF has been conducted for India and Egypt (Sandhu *et al.* 2011; Ghodeif *et al.* 2016). The application of RBF in China started in the 1930s, and the first Chinese RBF facility was

constructed in the northeast (Hu *et al.* 2016). According to a recent review on RBF implementation in China (Hu *et al.* 2016), there were about 300 RBF facilities present in China in 2002; however, lacking management standards, operation technology guidelines, and waterworks construction standards for RBF systems, the promotion of this technique in China was slow.

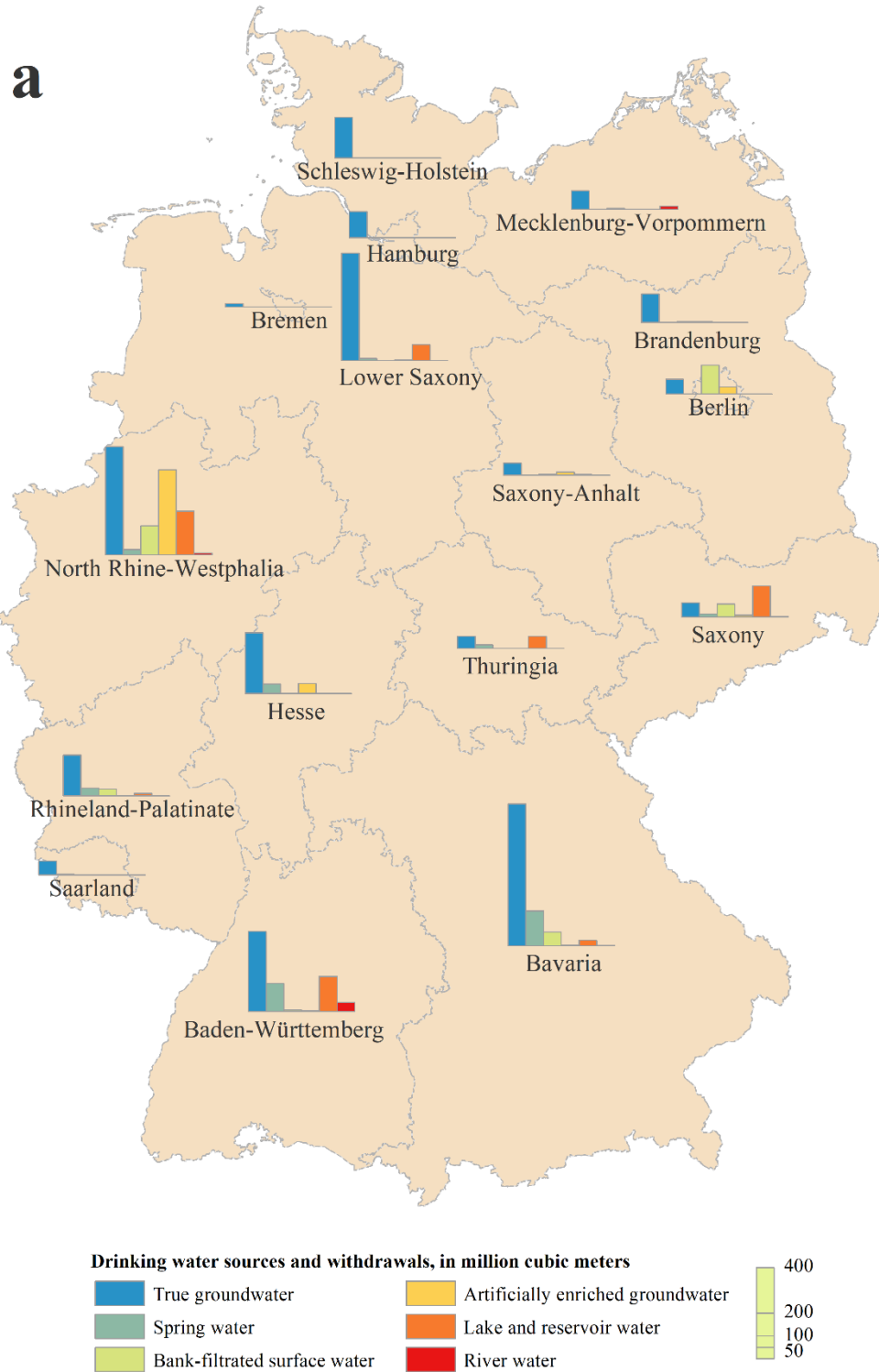


Figure 3a: Public water supply of Germany by state, 2016

Source: data from Destatis (2019b); Administrative Map of Germany from GADM (2019).

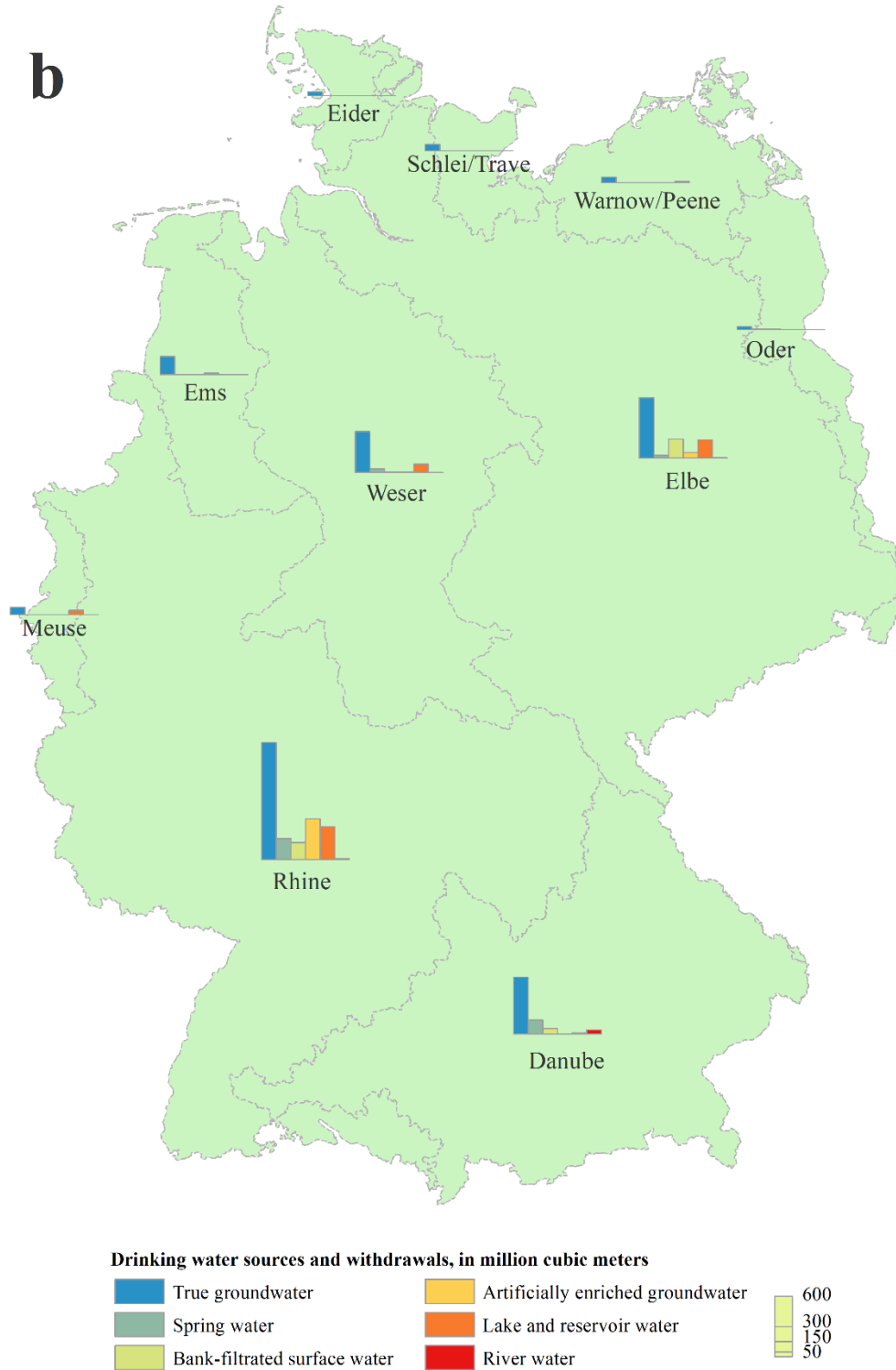


Figure 3b: Public water supply of Germany by drainage basin, 2016

Source: data from Destatis (2019b); River basin shapefile from EEA (2019).

Table 3: Public water supply of Germany, 1991-2016

Year	Total water supply	Groundwater	Bank-filtrated surface water	Artificially enriched groundwater	Other surface water sources	Population
1991	6516	4693	393	619	811	80.27
1995	5810	4224	304	563	719	81.82
1998	5557	4103	268	478	709	82.04
2001	5409	4011	280	427	691	82.44
2004	5372	3953	284	429	705	82.5
2007	5128	3581	410	464	673	82.22
2010	5081	3535	395	468	682	81.75
2013	5053	3499	436	444	675	80.77
2016	5204	3598	417	484	706	82.52

Note: water supply values are in million cubic meters, population values are in millions, groundwater including true groundwater and spring water, other surface water sources including lake and reservoir water and river water, values may not sum to totals because of independent rounding. Source: data of water supply from Destatis (2003, 2013, 2019b), data of population from Destatis (2019a).

2.2 Benefits of RBF for China

Taking China as an example, RBF has the potential to help ensuring drinking water safety, reduce bottled water use and carbon emissions (Hu *et al.* 2016, 924; Ray 2008, 224; Gadgil 1998; Schmidt *et al.* 2003).

Ensuring drinking water safety. The length of water supply pipelines increases by several tens of thousands of kilometers per year in China (Figure 4) (MOHURD 2018).

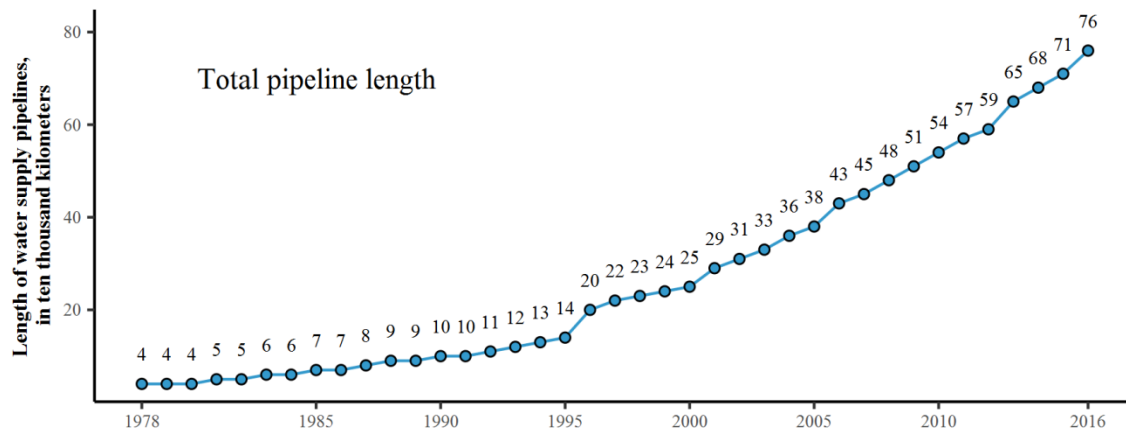


Figure 4: Length of water supply pipelines in China, from 1978 to 2016

Source: MOHURD (2018).

Tap water is the main source of drinking water for most of the people. The public access to piped water increased from 30% in 1985 to 77% in 2007 and, despite a massive growth in urban population, access for urban residents reached nearly 94% in 2007 (Zhang *et al.* 2010).

Along with the addition of total pipeline length, the amount of tap water supply also increases steadily, the most of which is from surface water sources (Figure 5). As the population grows in the urban and county seat area, surface water withdrawals (calculated by the *integrated production capacity* of public suppliers in urban and county seat area) increased significantly from 2006 to 2016, whereas groundwater withdrawals barely changed. Instead of groundwater, surface water is the main source of drinking water in China. In 2016, 85% of drinking water were abstracted from surface water sources, such as rivers, lakes, and reservoirs; the rest of them (15%) were from underground aquifers (MOHURD 2018).

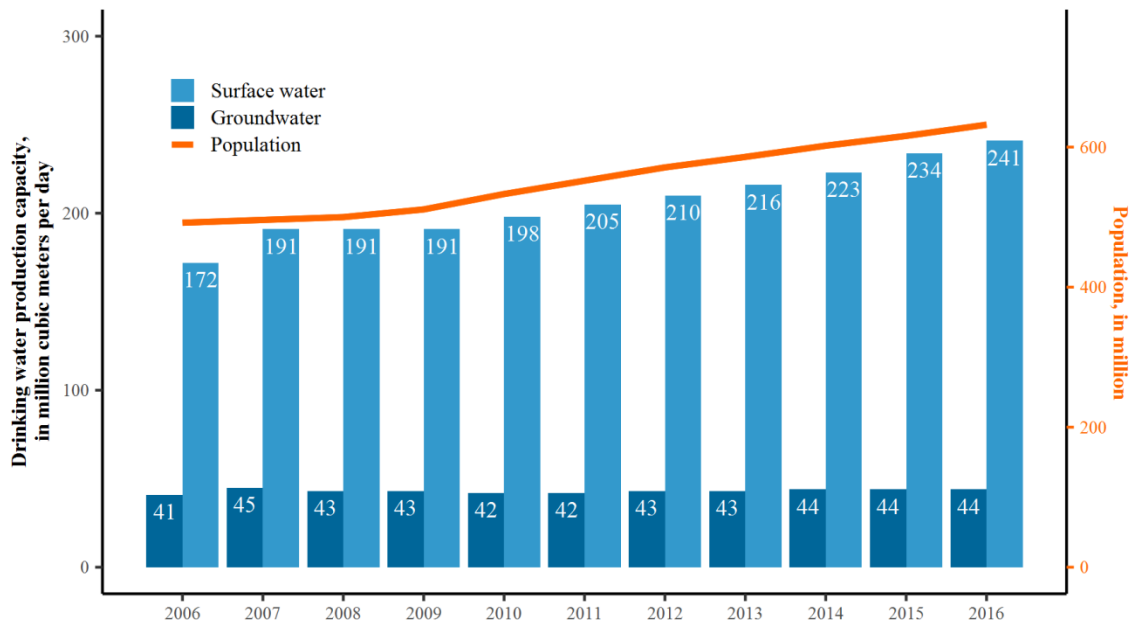


Figure 5: Drinking water production capacity and population, 2006-2016

Note: public suppliers in urban and county seat areas, capacity equals integrated production capacity. Source: MOHURD (2016-2018).

Groundwater constitutes 97% of global freshwater and is an important source of drinking water in many regions of the world, such as in Europe and the United States (BMI 1985). Groundwater is a preferred source for drinking water because it is typically of more stable quality and better microbial quality than surface water. It often requires little or no

treatment to serve as drinking water whereas surface water bodies are particularly vulnerable to contamination hence often need to be treated extensively (WHO 2006, 2016); however, groundwater seldom meets the demand for large cities, due to increasing demand for water by growing population and industry. To meet this gap, supplies will have to increasingly be sourced from surface water. When surface waters are used as raw water in drinking water supplies, appropriate filtration serves as a barrier for microbial pathogens and should be considered wherever feasible (Snow 1849, 30; Gadgil 1998, 268).

Apart from the increasing demand of drinking water, the pollution of tap water source is a constant cause of concern in China. At the end of 2006, the Ministry of Health of the People's Republic of China (MOH) and the Standardization Administration of the People's Republic of China (SAC) established new standards (GB 5749-2006) for drinking water quality to replace the old standards (GB 5749-1985), which have been used since 1985 (MOH and SAC 2006). The new issued standards are more stringent, adding 71 more indices on drinking water quality. As a consequence, people found that in 2009, over 40% of the drinking water in the country was not safe to drink (China Construction Newspaper 2012). An investigation made by the Ministry of Housing and Urban-Rural Development of the People's Republic of China (MOHURD), responsible for providing safe drinking water, showed that among all purified water samples from 4457 waterworks in urban area in 2008 and 2009, only 58.2% of them met the new quality standards (China Construction Newspaper 2012). After upgrading those waterworks and pipelines, improving operation management, and replacing raw water sources, MOHURD conducted another national investigation on drinking water quality two years later. The proportion of safe drinking water in the cities has increased from 58.2% to 83.0% by 2011; the investigated waterworks or the sample size, however, was not available (China Construction Newspaper 2012).

Not all water bodies are suitable as sources for drinking water. In 2016, 67.8% of all surface water bodies in China were safe for drinking water supply, while only 39.9% of all groundwater bodies were safe (Figure 6). This contradicts with the WHO statement on the previous page. Poor awareness of groundwater protection has caused severe groundwater pollution and may make it even less preferred in China.

To provide safe tap water to the large populations in urban and suburban areas, a reliable and robust water purification process is essential as well as a carefully designed system of collection, storage, and distribution. RBF systems can at least greatly improve the safety of drinking water because first of all, most raw water comes from bank-filtered surface water (especially river water), and only a small portion comes from true groundwater (Lenk *et al.* 2006). This is particularly true for the source water situation in China. Parallel processes such as infiltration or artificial groundwater recharge may also

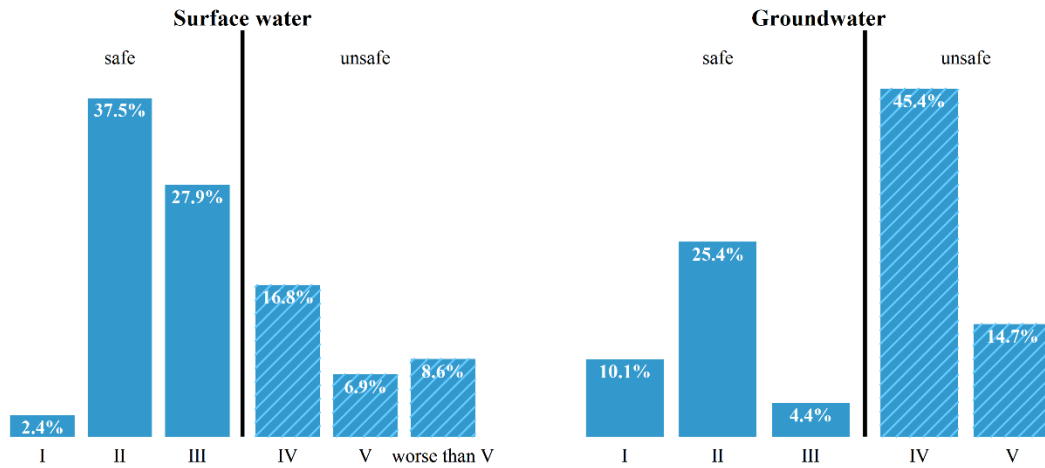


Figure 6: Surface water & groundwater quality in China, 2016

Note: “safe” or “unsafe” is classified according to water quality standard GB 3838-2002 (2002) (surface water) and GB 14848-93 (1993) (groundwater) for water bodies to be used as the drinking water source, with I being the highest water quality category; for detailed water quality indicators please see Appendix 1 and 2. Percentages of water quality category in Chinese surface water bodies (including stream, lake, and reservoir) and groundwater bodies (national groundwater monitoring wells) may not sum to total because of independent rounding. Source: data from MEE (2016).

be used during drinking water treatment if true groundwater is heavily polluted (Kühn and Müller 2000). Secondly, RBF is safe because it enhances the effectiveness of all disinfection methods by reducing turbidity and the concentration of chemicals and pathogens (Gadgil 1998, 268; WHO 2016, 55).

Reduce bottled water use and carbon emissions. Out of safety concerns, many people turn to bottled water instead of drinking tap water directly. Bottled water’s consumption in China has been steadily growing over the past years (Figure 7), and the Chinese spend hundreds of billions of yuan each year on it (Daxue Consulting 2019). China contributes to more than one-quarter of the world’s total bottled water consumption in 2017, ranked No.1 among other nations since 2013 (Rodwan 2018). Bottled water that originates from protected groundwater sources is less likely to contain contaminants that pose a health risk, but not all groundwater is well protected, and no water is guaranteed to be completely free of contaminants (U.S. EPA 2005). Different kinds of contaminations at unsafe levels have been found in bottled water. Cases of bottled water contaminated by disease-causing microbes, organic materials, and inorganic compounds were not rare and reported both in China and overseas (NMPA 2019; Lewis 2019; Farsaci 2019; Bao 2019). In 2008, Guiyang reported an outbreak of Hepatitis A (a virus) caused by drinking contaminated bottled water and led to over 300 hospitalizations (Xinhuanet 2008). Bottled water consumption creates a huge amount of plastic waste. Plastic pollution that accumulated in the Earth’s environment due to mismanagement will adversely affect wildlife, wildlife habitat, and humans (Parker 2018). Initiatives have been taken around the world to reduce plastic waste. For instance, the San Francisco International Airport (SFO)

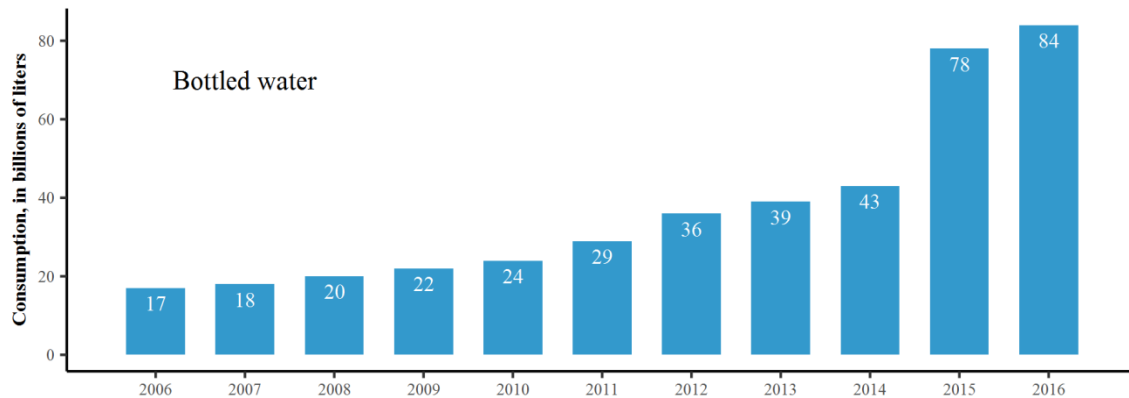


Figure 7: Bottled water consumption in China, from 2006 to 2016

Source: Beverage Marketing Corporation (2006-2016).

banned bottled water (starting August 20, 2019) to reduce net carbon emissions and energy (Alony 2019).

In China, people prefer drinking boiled tap water or bottled water from producers because they think it is safer than their tap water. The most common cause of waterborne illness is bacteria, such as *E. coli*, *Vibrio cholerae*, *Salmonella*, and *Legionella*, but illness can also be caused by protozoa (including *Giardia* and *Cryptosporidium*), viruses (like Hepatitis A and Rotavirus) and various chemical pollutants (Mayo Clinic 2018). Boiling, or more accurately, pasteurization can kill most of the pathogens; in fact, it is the most effective way of killing *Cryptosporidium*, a microscopic parasite that causes illness in healthy adults, and severe illness and even death for people with weakened immune systems (U.S. EPA 2005). For safe drinking water supply, attention has been paid to *Cryptosporidium* (oocyst) due to its resistance to chlorine-based disinfectants (U.S. EPA 2005). People may, however, be exposed to waterborne pathogens and become infected through water used for brushing teeth, making ice cubes, and washing fruits and vegetables – not just through the water they drink (U.S. EPA 2005). Despite the habit of drinking boiled water, waterborne epidemics hit China occasionally. In 2010, a Paratyphoid A outbreak occurred in Guangxi, leading to 80 individuals hospitalized, including 50 students (The Disease Daily 2010). In 2012, Guizhou reported an outbreak of typhoid, the total number of suspected people affected was 141 people (120 cases at a school) (Xiao *et al.* 2015). A cholera outbreak occurred in Hubei in 2012, and nine people were infected (Edmundson 2012). Contaminated water or food was the main cause of these outbreaks. Although pasteurization is an effective method to get rid of pathogens, it is not appropriate when chemical contamination is present (NYSDOH 2018). This may increase exposure to chemicals that pose a health risk such as nitrates and solvents by concentration in the boiled water or by volatilization into the breathing zone (NYSDOH 2018). Furthermore, boiling

is a very energy-consuming method to acquire safe drinking water; thus, it is economically unrealistic and environmentally unsustainable to the developing world (Gadgil 1998). Advanced drinking water treatment processes can be used to remove high levels of contaminants from source water; however, at the same time, they increase the cost for water utilities and may not be environmentally friendly (Tao and Xin 2014). The new standards certainly give a challenge for the waterworks in China and yet the improvement of drinking water quality should not be a burden on the environment. Besides the necessity of upgrading distribution systems (*e.g.*, replace old pipes to prevent re-contamination), a cost-effective and more reliable drinking water treatment process is also needed for people to choose tap water over bottled water in China (Cosier and Shen 2009; Tao and Xin 2014). Because RBF systems represent an energy-efficient, viable method of water purification that minimizes the use of chemical and active carbon and may require less treatment, its use and the resulting reduction in consumption of bottled water would reduce carbon emissions (Kühn and Müller 2000, 61; 2006, 52; Ray *et al.* 2002, 1).

3. RBF Site Selection

However, before implementing an RBF system and benefiting from the advantages of RBF for drinking water supply, it is very important to choose a suitable location to install it. The consideration of RBF site selection began probably in the 1810s (Glasgow City Archives 2020). In 1818, the Cranstonhill Water Works Company relocated its facilities up-river at Dalmarnock due to the increasing pollution of the Clyde River (Glasgow City Archives 2020; Burnet 1869, 7). Since then, RBF systems have been built around Europe for drinking water supply, but records of RBF systems have been scarce until a few decades ago.

Suggestions on site selection first appeared in some German literature (Beyer and Banschler 1975, 569-570; BMI 1975, 32-33; Fokken 1996a, 491-492). Factors to consider are fluvial processes (deposition, bed load, and erosion), the distance between wells and river, and land use.

In the English literature, the latest books on RBF are mainly the results of relevant conferences held between 2000 and 2010 (Jülich 2000; Ray 2002; Hubbs 2004; Linsky, Ray, and Melin 2004; Ray and Shamrukh 2011). These texts cover various aspects of RBF including but not limited to site selection.

The most detailed descriptions of RBF site selection among the German and English literature are probably the works of Grischek, Schoenheinz, and Ray (2002), Caldwell (2004), Lenk *et al.* (2006), and U.S. EPA (2010). Some factors, such as well field location, aquifer properties, and surface water quality, have appeared in more than one publication and are critical for RBF selection.

3.1 RBF site selection model

Existing RBF site selection models were mainly established in the last decade (Lee and Lee 2010; Archwichai *et al.* 2011; Wang, Ye, and Du 2016).

Lee and Lee (2010) applied geographic information system (GIS) and analytic hierarchy process (AHP) to select suitable locations for RBF in the highly urbanized area of Seoul, along the Han River, South Korea. Suitable RBF sites or candidate sites have been selected from the preliminary analysis; therefore, the detailed process of how these suitable sites were delineated was not documented in their work. Lee and Lee's RBF evaluation model is rather complete, and includes various hydrogeologic, water quality, and socioeconomic factors. Based on the existing RBF projects in South Korea, 21 (or 20) most influential elements were selected by the authors (Figure 8). These influential elements can be categorized into three Level 2 criteria, six Level 3 criteria, 17 Level 4 criteria, and four Level 5 criteria (Level 1 is the suitability index). The assessment of the scaling constants were based on 15 experts' knowledge, considered to be specialists in RBF and site suitability analysis. The authors then compared the experts' scaling constants with those derived from the previous projects; the results turned out to be similar. A pairwise comparison was conducted by the authors to determine the relative importance of those criteria. The result shows that, among the Level 2 criteria, *possibility* is considered to be more important than *urgency* and *efficiency*. Finally, four candidate sites in the Seoul area of the Han River Basin were evaluated. The recommended most-preferred site has the characteristics of "good water quality" and "a close connection to an existing water purification facility". Lee and Lee's 2010 work conducted opinions from multiple experts, which enhanced the credibility of the AHP analysis. Their work failed, however, to record how these criteria were chosen. The method lacks an effective validation or uncertainty analysis and the information of candidate RBF sites is not complete, the latter of which makes it impossible to compare the results with other models.

Archwichai *et al.* (2011) identified key parameters affecting the performance of RBF systems and developed a three-step procedure on RBF site selection. The first step is to identify RBF potential areas at river basin level in Thailand; the second step is at the locality level within potential river basins; while the third step is at site level within the potential local areas. The authors used weighted overlays to evaluate and select potential areas at each level. The first two steps screened suitable areas using criteria such as Quaternary geology, hydrogeology, hydrology, and water quality. In the final step, they used nine criteria from a total of five components to evaluate and rank the final suitable site (Figure 9). The scores and weights of those criteria are assigned based on the authors' judgement. The main advantage of this site selection method is that it can delineate suitable areas at the river basin scale as well as at the local site scale, which enhances the applicability. The disadvantage is the absence of model validation or a sensitivity analysis of variables and weights.

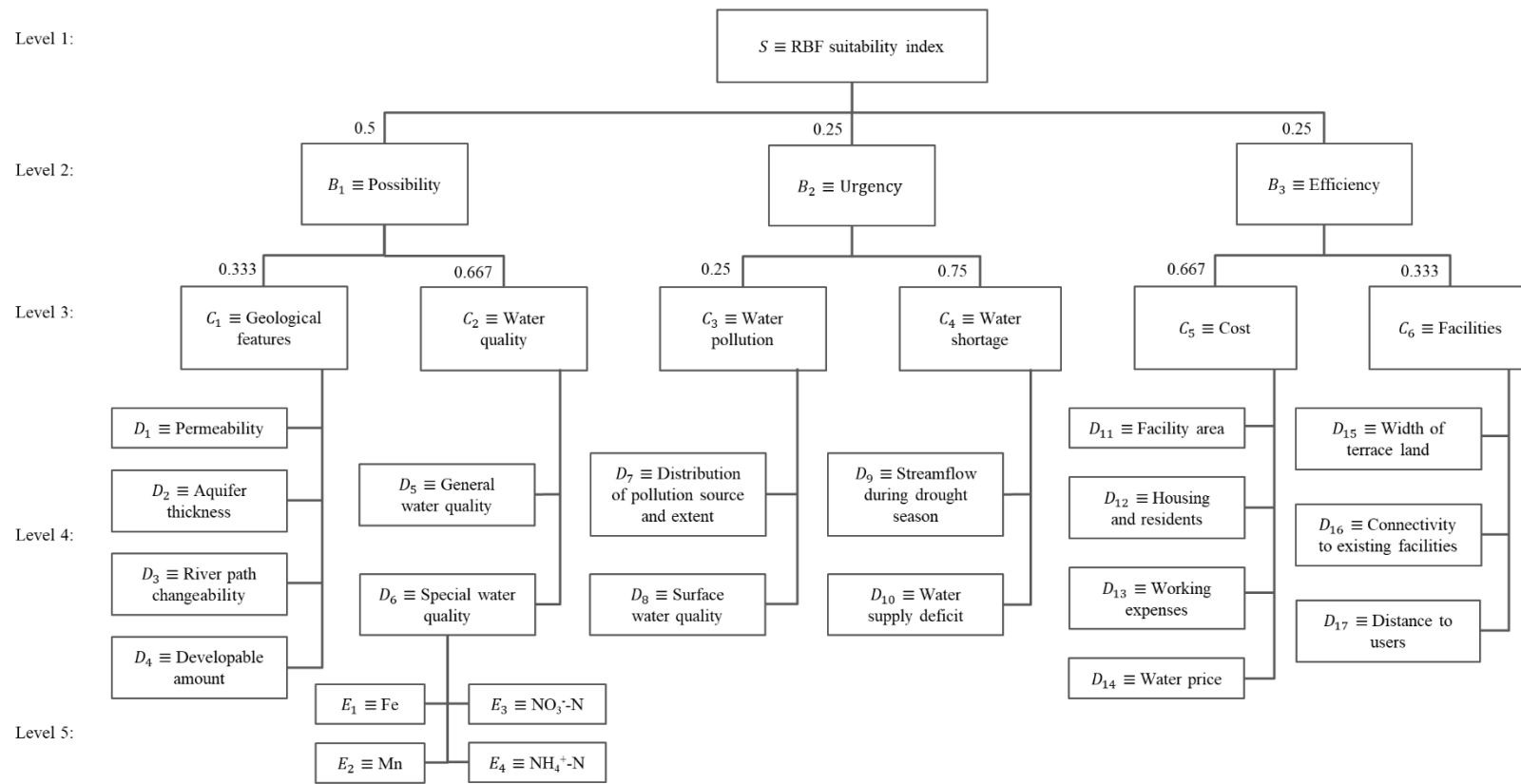


Figure 8: Lee and Lee's decision model on RBF site selection

Note: \equiv means "is defined as". Source: modified after Lee and Lee (2010).

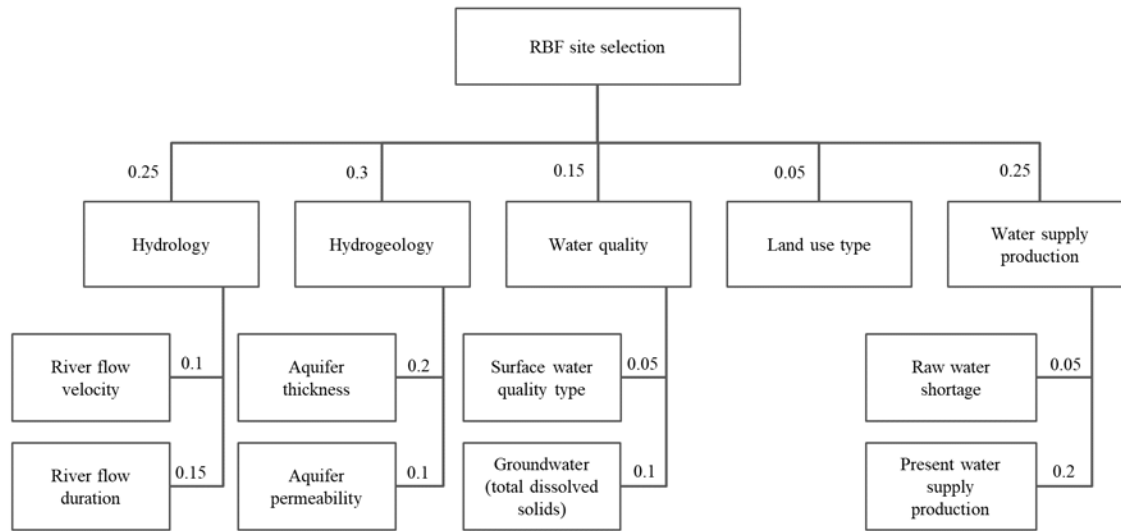


Figure 9: Archwichai's decision model on RBF site selection

Source: modified after Archwichai et al. (2011).

Wang *et al.* (2016) developed an index system and integrated it into GIS to investigate the suitability of RBF along the Second Songhua River and its major tributaries in China. Yin *et al.* (2018) used the same index system and applied it to the RBF site selection problem in the Songhua River and its tributaries, the study area of which is larger than that of Wang *et al.* (2016). Wang's evaluation system consists of suitability indexes such as water quality, water quantity, interaction intensity between surface water and groundwater, and exploitation conditions of the aquifer. The scores and scaling constants of those indexes were assessed based on experts' opinions (Figure 10). Weighted summation of each criterion was used by the authors to calculate the final suitability scores, based on which the authors established five suitability grades, from "highly suitable areas" to "unsuitable areas". According to the authors, the top three grades were considered suitable for RBF implementation, whereas the worst two grades were unsuitable. Those unsuitable areas have characteristics such as "bad water quality" and "insufficient recharge". Limitations exist for the further application of this index system because of the specific hydrogeologic site conditions, uncertainties associated with limited field data, and the absence of model validation. The authors pointed out the necessity of developing other decision making methods for the RBF selection problem (Wang, Ye, and Du 2016; Yin *et al.* 2018).

Grischek (2019) argued that, probably due to site specifics in different regions, a general site selection procedure cannot be made, and there is still no guideline or proven model yet for RBF site selection. In fact, the recent works on selection of suitable RBF

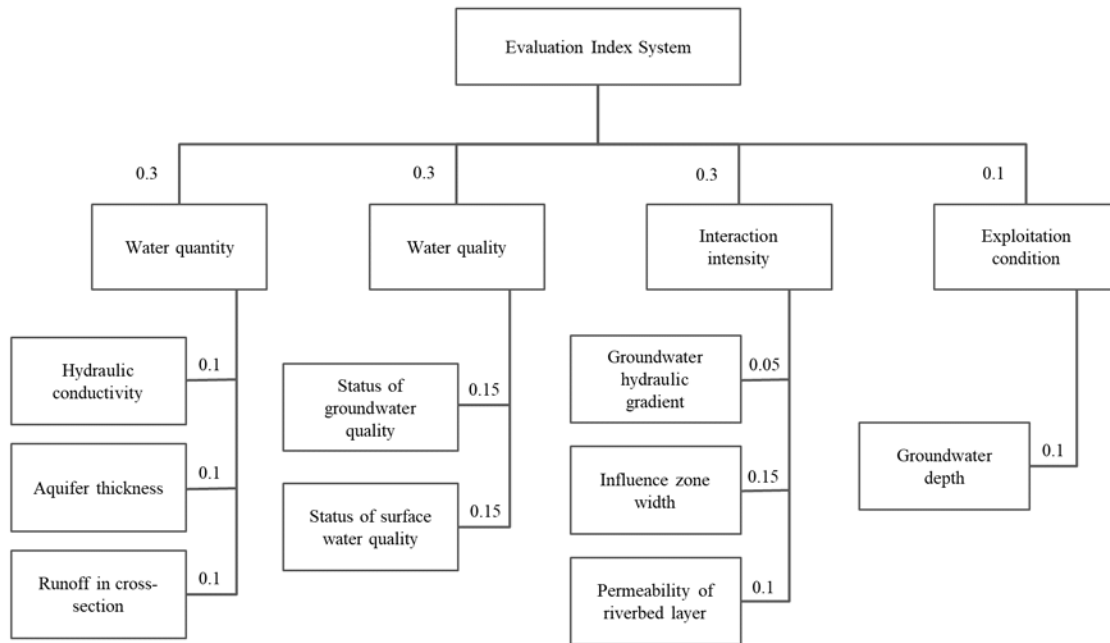


Figure 10: Wang's decision model on RBF site selection

Source: modified after Wang *et al.* (2016).

sites are still based on just a few factors. Sandhu *et al.* (2011) summarized some existing RBF sites in India and discussed the potential of implementing RBF systems for other cities in the country. The authors identified four potential RBF sites or cities along the large rivers in India based on the distribution of alluvial deposits. Ghodeif *et al.* (2016) studied existing RBF sites in Egypt and recommended that, for site selection of new RBF systems, detailed investigation of hydrogeologic conditions, water quality (both river water and groundwater), and river hydrology are essential. Coarse materials such as sand and gravel from the riverbed and aquifer provide suitable conditions for RBF; however, the preference for these potential sites was not given.

These existing site selection methods relies on a weighted linear function, from which the trade-offs between criteria are constant; however, this is rarely the case and uncertainties do exist when considering the trade-offs among attributes (Keeney and Wood 1977). For example, how much more important is water quality than water quantity for an RBF system (if it is so for the decision maker)? This determines how much better one alternative is over another.

New decision analysis is needed for the site selection of RBF systems, from which uncertainties are taken into account.

3.2 Definition of successful RBF sites

The purpose of the RBF site selection procedure is to maximize the performance of the RBF system after its completion, which might be measured by its long-term yield and purification capability. A successful RBF system should have a high yield of bank-filtered river water and a high purification capability as well. The share of bank-filtered river water from RBF wells has often been considered as a criterion for successful RBF sites (Ghodeif *et al.* 2016; Lenk *et al.* 2006).

Ghodeif *et al.* (2016) evaluated eight existing RBF sites in Egypt, with respect to the share of bank-filtrated river water and well water quality. Six successful sites were considered to be those with high share (>50%) of bank-filtrated river water, whereas two sites with <10% share of bank-filtrated river water were considered not successful, the problem of which was caused by unfavorable hydrogeologic conditions (presence of a low hydraulic conductivity layer with a high clay content). Insufficient hydrogeologic investigations is therefore the main reason for the failure of these two sites. The site condition of the successful sites can be summarized as below (Table 4). The authors did not measure the purification capability of those RBF sites in Egypt, but the well water quality evaluated from those sites were rather good, most of which met the drinking water standard.

Table 4: Hydrogeologic condition of six-successful sites in Egypt

Parameter	Scale of measure	Range
K	m/s	2.3×10^{-4} - 1.39×10^{-3}
b	m	16-500
T	m^2/s	0.017-0.347

Note: hydraulic conductivity (K), aquifer thickness (b), transmissivity (T). Source: Ghodeif *et al.* (2016).

Lenk *et al.* (2006) summarized 33 RBF sites in Central Europe; they defined the sites as RBF sites (including actually both river- and lake bank filtration, the term RBF was used for consistency) if the share of bank-filtered surface water in the production well was higher than 50%. We might consider those RBF sites as successful sites, too, in terms of a high share of bank filtrate. The authors then analyzed the parameters affecting the removal efficiency of dissolved organic carbon (DOC). The result shows that DOC concentration from surface water, residence time (or distance between wells and surface water), and transmissivity have a positive correlation with DOC removal in RBF wells, the empirical models of which is shown as equation (1):

$$Y = -0.503 + 0.811 \ln(X_1) + 0.236X_2^{0.437} + 7.428X_3 \quad (1)$$

where Y is DOC removal (mg/l), X_1 is DOC concentration from surface water ranging between 1.5-9.0 mg/l; X_2 is residence time (d) with a range between 0.01-210 d; X_3 is transmissivity with a range between 0.003-0.230 m^2/s . The equation is based on 43 measurement points ($N=43$) with a coefficient of determination $R^2=0.74$.

An RBF system with higher DOC concentration from surface water, longer residence time, and higher transmissivity tends to have a higher purification capability (measured as DOC removal). According to this regression equation, the final DOC concentration measured in the RBF well can be estimated (Figure 11).

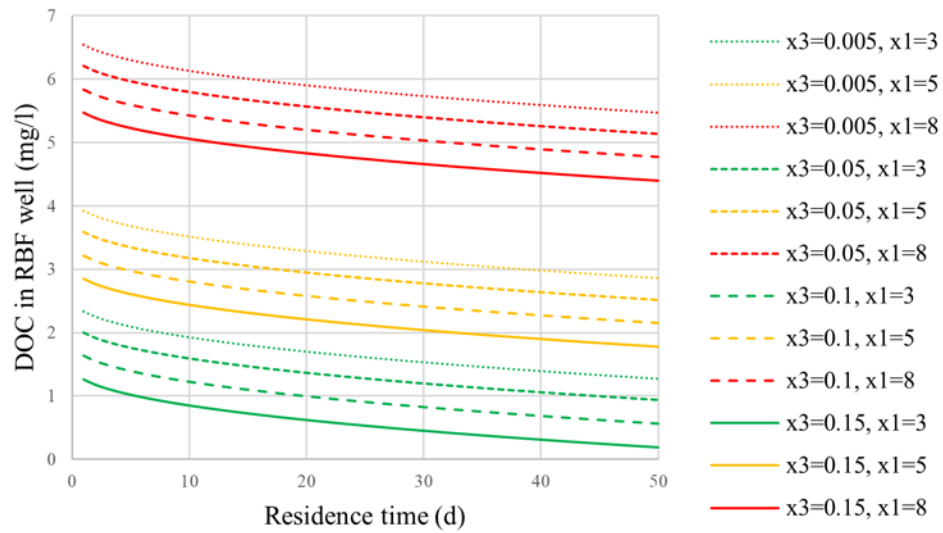


Figure 11: Simulated DOC concentration from RBF well using Lenk's equation

Note: x_1 is DOC concentration from surface water, in mg/l; x_3 is transmissivity, in m^2/s . Source: modified after Lenk et al. (2006).

As can be seen from Figure 8, higher DOC concentrations from surface water yield higher DOC removal, yet the final DOC concentration in RBF wells are much higher than those with lower DOC concentrations from surface water (if residence time is fixed).

Using purification capability as a criterion for RBF site selection might be inappropriate because it may not truly reflect the objective of the decision maker (for instance, to yield bank filtrate with good quality). The purification capability of the RBF system is, however, highly dynamic and no general consensus has merged yet.

The site information of those successful RBF sites is summarized below (Table 5).

Table 5: Site conditions of 33 successful RBF sites in Central Europe

Parameter	Mean	Median	Minimum	Maximum	Standard deviation
K (m/s)	4.2×10^{-3}	1.7×10^{-3}	1×10^{-4}	2×10^{-2}	5.3×10^{-3}
b (m)	21.5	16.0	4.0	70.0	16.3
T (m ² /s)	0.056	0.033	0.003	0.230	0.057
DOC (mg/l)	4.4	3.6	1.4	9.0	2.0
Residence time (d)	31.2	16.5	0.01	210	42.4
L (m)	55.8	23.6	0.2	310	67.7

Note: hydraulic conductivity (K), aquifer thickness (b), transmissivity (T), DOC concentrations are measured in river water, statistics of residence time are obtained from 58 measurement points, distance between wells and surface water (L), statistics of the distance are obtained from 66 measurement points. Source: Lenk *et al.* (2006).

4. Factors Affecting RBF Site Selection

It is necessary to review the major factors affecting RBF systems (Figure 12), some of which are rather important yet not taken into account from the existing site selection models, for instance, well field location (Grischek, Schoenheinz, and Ray 2002; Caldwell 2004), riverbed characteristics (with respect to riverbed scour) (Schälchli 1992; Hubbs 2004), and distance between wells and surface water or residence time (Lenk *et al.* 2006).

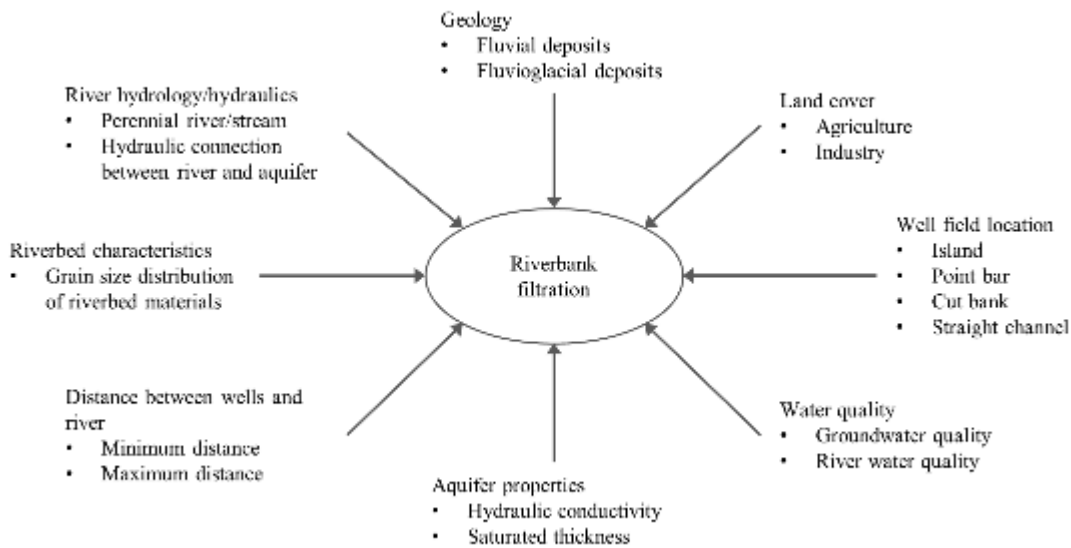


Figure 12: Constraints on applying riverbank filtration

4.1 River hydrology/hydraulics

River hydrology determines, firstly, how much water can be infiltrated from the river into the aquifer. Compared to seasonal or intermittent streams that flow for weeks and months of the year, firstly, *perennial streams* have continuous flow all year round and are therefore preferred in RBF (Kühn and Müller 2006; Lenk *et al.* 2006; Caldwell 2004). In Europe, RBF is well adapted along perennial streams such as the Rhine, the Elbe, the Danube, and the Ruhr. River hydrology has been considered as a criterion in RBF site selection in previous works. Lee and Lee (2010) used “streamflow during drought season” in their work for RBF site selection, from which higher flows were preferred. “Average minimum yearly discharge near the RBF site” was suggested by Caldwell (2004) for factors affecting yield. “River flow duration” from 9 to 12 months was used by Archwichai *et al.* (2011) in their RBF evaluation system, which means intermittent streams are also suitable for RBF in their model. Despite this, perennial streams that have a flow duration for 12 months was considered suitable in this study because the impact of a no-flow period on RBF systems was still unknown.

Secondly, river hydrology has an impact on sediment transport and renewal capability of the river, which indirectly affects the water yield from the abstraction well due to the process of clogging (Caldwell 2004). The removal of the clogging layer during high flow period increases the hydraulic conductivity of the riverbed, causing an increase of the yield of bank-filtrated river water. According to Schubert (2004a), however, some dam-regulated rivers may miss this renewal advantage of RBF due to the cutting of peak flow of flood waters. Dams can strongly affect the seasonal fluctuation of flows by cutting the peak flow during a flood, but they cannot completely remove the impact of flooding on cities and channels downstream. So, although a river is regulated by dams, the clogging layer can still be somewhat removed by the temporally high flow caused by a flood.

Stream water interacts with groundwater basically through flux exchanges (Vandas, Winter, and Battaglin 2002). If the flux exchanges exist, we say the stream and the ambient aquifer are *hydraulically connected*, which is fundamental to a successful RBF system (Doussan *et al.* 1997; Grischek, Schoenheinz, and Ray 2002; Lenk *et al.* 2006). Surface waters and unconfined fluvial/fluvioglacial sand and gravel aquifers are often very closely connected (Brunke and Gonser 1997; Province of British Columbia 2016). Many streams gain water from groundwater inflow in some reaches and lose flow in other reaches (Alley, Reilly, and O. Lehn 1999). The losing stream (a stream that loses water as it flows downstream) condition is usually caused by the rising of stream water table during storms; it is also influenced by anthropogenic activity such as groundwater pumping. Continuous infiltration of stream water by groundwater pumping causes accumulation of particles on the porous media of the streambed that can alter fluid flow properties and cause the reduction of the streambed permeability (McDowell-Boyer, Hunt, and Sitar 1986). The *unsaturated zone* (or *zone of aeration*, *vadose zone*) occurs with this relatively low-permeable *clogging layer* on the streambed may cause hydraulic disconnection between

the stream and the ambient groundwater (Sophocleous 2002). The hydraulic connection or the existence of an unsaturated zone below the streambed has to be checked on site before planning a drinking water supply system using RBF (Beyer and Banscher 1975). Numerous methods can be applied to check the hydraulic connection between stream and groundwater, such as direct measurement of water flux, heat tracer methods, Darcy's Law based methods, and mass balance approaches (Kalbus, Reinstorf, and Schirmer 2006).

4.2 Geology

Geologic formations that yield enough water for public waterworks and their ability to filter out various kinds of contaminants from the surface water are of great importance for an RBF facility. Sand seems to be the ideal kind of formation for RBF systems because it filters and yields water well. Prior studies show that material with a grain size between 0.1 mm and 0.5 mm (or between 0.15 mm and 0.35 mm) and a uniformity coefficient (C_U) smaller than 5 has a rather good filtering ability (Kühn and Müller 2006; Sánchez *et al.* 2006). The uniformity coefficient is a numerical expression of the variety in particle sizes in mixed natural soils, defined as the ratio of the diameter of a grain that has 60 percent (by weight) of the sample finer than itself to the diameter of a grain that has 10 percent finer than itself (Meinzer 1923). This preferred material is within the grain size range of fine sand and medium sand. Clay is the least suitable material as a source of water supply; it contains a large amount of water but it mostly cannot be yielded under ordinary hydrostatic pressure (Meinzer 1923). Gravel formations yield water very well, but they may have low efficiency in removing contaminants (*e.g.*, bacteria) during RBF if fine-grained riverbed sediments are absent (Schijven, Berger, and Miettinen 2002). According to Sprenger *et al.* (2017), most of the RBF facilities are located on the fluvial and glacial deposits.

To investigate suitable geological formations, the RBF systems in Germany were studied. A total of 1269 RBF (including both river- and lakebank filtration; the term RBF was used in this text for consistency) drinking water production wells or well fields in Germany were analyzed (see Figure 13 and Table 6). All well or well field were located in drinking water protection zones. Appendix 3 shows an illustration of how those wells or well fields were identified (Figure A1).

In order to acquire the geologic formation at each site, data from the Federal Institute for Geosciences and Natural Resources of Germany (BGR) was used. Surface rock formations at RBF sites were extracted from the General Geologic Map of the Federal Republic of Germany (GÜK250) (BGR 2019). This step was followed by using the "Overlay" tool in ArcMap 10.6, a geospatial processing program.

Following the RBF statistics of Germany from the previous chapter, only the RBF systems that located in the federal state of Berlin, North Rhine-Westphalia, Saxony, and Rhineland-Palatinate were analyzed. RBF wells or well fields in other federal states were not taken into consideration in this study, either due to their limited contribution to drinking



Figure 13: Distribution RBF well/well field and their dominant formations in Germany

Note: map was generated using ArcGIS Pro (2018). Sources: shapefile of countries and river network are from Natural Earth(2019b); shapefile of geology is from the Geologic overview map (GÜK250) (BGR 2019). Well or well field locations are from the following sources: Berlin Senate Department for Urban Development and Housing (SenSW 2009), Saxon State Office for Environment, Agriculture and Geology (LfULG 2018), North Rhine-Westphalian State Agency for Nature, Environment and Consumer Protection (LANUV 2015), and Ministry of Environment, Energy, Food and Forests of Rhineland-Palatinate (MUEEF 2019).

Table 6: RBF well/well field in Germany

State	Water body	Number of wells/well fields
Berlin	Spree, Müggelsee (lake)	539
	Harvel, Lake Tegel, Wannsee (lake)	332
North Rhine-Westphalia	Rhine	170
	Ruhr	36
	Lippe	49
Saxony	Elbe	41
	Mulde	15
Rhineland-Palatinate	Rhine	84
	Moselle	3

Source: well or well field locations are from Berlin Senate Department for Urban Development and Housing (SenSW 2009), Saxon State Office for Environment, Agriculture and Geology (LfULG 2018), North Rhine-Westphalian State Agency for Nature, Environment and Consumer Protection (LANUV 2015), and Ministry of Environment, Energy, Food and Forests of Rhineland-Palatinate (MUEEF 2019).

water supply or because of the limited time frame for this project.

Among the available RBF wells or well fields, 871 of them are extensively operated by the water utilities in Berlin. In Berlin, lake bank filtration (LBF) wells are densely placed around the lakes Tegel and Müggelsee. In North Rhine-Westphalia, RBF systems are widespread in the Lower Rhine region and along the Ruhr River, *e.g.*, in the cities of Cologne, Düsseldorf, and Schwerte. In Saxony, large RBF plants are located in the cities of Dresden, Meissen, and Torgau. Waterworks alone on the Elbe River (*e.g.*, Hosterwitz Waterworks in Dresden, Torgau-Ost Waterworks in Torgau) provide safe drinking water for 1.5 million people and are of great importance to the big cities in Saxony (Griseck *et al.* 1998).

The following part of the analysis, *i.e.*, the formative process, petrography, and stratigraphy of the rock formations, are based on the General Geologic Map of the Federal Republic of Germany (GÜK250) (BGR 2019).

The formative process of the surface rock formations at RBF sites are shown in Table 7. Two kinds of major formative processes can be found: clastic sedimentary rocks and extrusive igneous rocks (volcanic rocks). Clastic sedimentary rocks are the predominant rock formations, made by fluvial, limnic, fluvio-glacial, aeolian, terrestrial, and phytogenic processes, while extrusive igneous rocks were only found at three single well locations at the bank of Rhine River. *Fluvial deposits* and *fluvio-glacial deposits* are the most frequent rock formations in RBF sites. This finding is consistent with other studies on geologic or hydrogeologic properties of RBF sites in Europe (Sprenger *et al.* 2017).

Fluvial deposits are made in conjunction with running water. They consist largely of gravel and sand, which are excellent water-bearing formations, especially in large river valleys; however, not all deposits made by running water are well sorted. The bulk of the fluvial deposits are poorly sorted, which are a mixture of coarse sediments (such as pebbles and boulders)

Table 7: Processes of surface rock formations at RBF sites

Formative process	Main component	Main location of BF well	Number of wells/well fields
Fluvial	Stream deposits; overbank deposits; terrace deposits; stream deposits to permafrost feature	At the banks of Rhine, Elbe, Ruhr, and Mulde River	451
Limnic	Limnic deposits	Berlin	36
Fluvioglacial	Outwash plain (valley) deposits, foreset deposits	Berlin	654
Aeolian	Aeolian sand	Berlin	44
Terrestrial	Redistributed deposits	At the bank of Rhine River near Koblenz	2
Phytogenic	Lowland moor peat	Berlin	79
Volcanic	Volcanic formations	At the bank of Rhine River near Koblenz	3

Sources: geologic information is from the geologic overview map (GÜK250) (BGR 2019); well or well field locations are from the same sources as Table 6.

as well as fine sediments (such as clay and silt) (Meinzer 1923). The poorly-sorted deposits are probably a result of erratic stream flow conditions, for instance, during a flood event. A particular place may suddenly receive coarse sediments by a raging flood and soon after receive only clayey sediments when the runoff dropped rapidly, which filled the spaces between the coarse deposits (Meinzer 1923).

Glacial deposits are made by moving glaciers. They consist of all sizes of materials from clay to huge boulders. A good explanation of the glacial deposits has been given by a well-known hydrogeologist, Oscar Edward Meinzer (1923, 126), in his monumental treatise on *The occurrence of ground water in the United States with a discussion of principles*:

The deposits of this debris carried by glaciers are called “glacial drift.” The drift is in part deposited directly by the ice and in part carried farther by the waters resulting from the melting of the ice or by the wind. That which is deposited directly by the ice forms heterogeneous unassorted mixtures called “till” or “boulder clay”; that deposited by escaping streams forms chiefly water-bearing “outwash gravel”; that deposited by lakes impounded by ice or till forms chiefly impervious clay beds; and that deposited by the wind forms accumulations of loess or dune sand.

Fluvioglacial deposits are the part of drift that results from the actions of glacial meltwater, which is also highlighted on Figure 13. For RBF systems in Germany, outwash valley deposits and foreset deposits (in German, *Vorschüttungsablagerungen*) are the most important fluvioglacial deposits. They contribute even more than the fluvial deposits to public water supply with respect to the number of abstraction wells or well fields. Ice sheets or continental glaciers covered large areas of Europe during the last ice age or Pleistocene glaciation, including the northern part of Germany. These continental ice sheets carried and laid down all kinds of glacial drift. Among the 871 drinking water production wells in

Germany's capital city Berlin, most of them were placed on the glacial drift, more specifically, that deposited by the meltwater streams flowing in front of or beyond a glacier, and formed by coalescing "outwash valley", "outwash fans" or "outwash plains". In German, the outwash valley near Berlin is called the "*Berliner Urstromtal*", which means Berlin's "ancient stream valley", "meltwater valley", or "ice-marginal valley". Unlike unassorted till, outwash gravel and sand sorted by running water may be highly porous (Meinzer 1923). Near Lake Tegel in Berlin, the hydraulic conductivity (K) of the outwash sand aquifer is about 6.0×10^{-4} to 1.1×10^{-3} m/s with a mean value of 3.0×10^{-4} m/s, which is lower than some well-known RBF sites (e.g., 4×10^{-3} to 2×10^{-2} m/s at site Flehe; 1×10^{-3} to 2×10^{-3} m/s at site Torgau-Ost) that are based on fluvial deposits but serve well in public supply (Henzler, Greskowiak, and Massmann 2014; Schubert 2002b; Grischek and Bartak 2016).

Both formative processes (fluvial and fluvio-glacial) are related to running water, which is a great sorting agent and perhaps the most important agent for the production of water-bearing formations (Meinzer 1923).

Soft rocks are found more frequently than hard rocks at RBF sites, mainly because of their porous property, which is favorable as a filter in water treatment. In Germany, psammitic soft rocks predominantly consisting of sand is the most frequent rock formation at all RBF sites, with respect to petrography (see Table 8). The psammitic soft rocks formation is a common component of fluvial and fluvio-glacial deposits, the accumulated sandy sediments can reach almost 40 meters thick (see Appendix 4, Figure A2 and Figure A3). Pelitic soft rock formations consisting of silt and clay are the second most frequent rock formation occurring at RBF sites. The majority of them formed during the Holocene fluvial and flooding process, while fewer wells or well fields located on the limnic fine sediments that transported and deposited in freshwater lakes existed before this latest geologic age. Among all deposits, silt and clay are the least likely formation to yield water (Meinzer 1923). It contains water, has a high porosity, but as "the constituent particles of clay are impalpably small the interstices between the particles are so minute that they hold tenaciously to all their water, rendering the clay impervious under ordinary hydrostatic pressure" (Meinzer 1923). In practice, RBF wells penetrate this silt and clay layer on the surface and drill deep into the underlying, better water-bearing formations—sand or gravel—in order to yield water freely. The pelitic soft rock formation spreads over the surface of the floodplain of constantly flowing waters (rivers and streams), which is not very likely to be thicker than 2 to 3 meters for a large area (Fuhrmann 1999). This thin layer helps further to reduce the potential of groundwater pollution (Aller *et al.* 1987); thus, it is favorable in implementation of RBF systems. However, the top clay layer reported at the lower reach of some large river systems may reach a thickness of about 50 m (Ghodeif *et al.* 2016); therefore, careful hydrogeologic investigation at each potential RBF site is needed.

Table 8: Petrography of surface rock formations at RBF sites

Petrographic name	Main component	Main location of BF well	Number of wells/well fields
Pelitic soft rocks	Clay/silt	Berlin; at the banks of Rhine, Elbe, Ruhr, Moselle, and Mulde River	269
Psammitic soft rocks	Sand	Berlin; at the banks of Rhine, Elbe, and Mulde River	894
Psephitic soft rocks	Gravel	Lower Rhine region	16
Phytogenic soft material	Peat	Berlin	79
Pyroclastic soft rocks	Tephra	At the bank of Rhine River near Koblenz	3
Psephitic hard rocks	Breccia	At the bank of Rhine River near Mainz	6
Not specified	Compound material	At the bank of Rhine River near Koblenz	2

Sources: geologic information is from the geologic overview map (GÜK250) (BGR 2019); well or well field locations are from the same sources as Table 6.

Layer information from some drillings at a RBF site in Berlin is shown in Appendix 4, Figure A4, Figure A5, and Figure A6.

Stratigraphic information of the rock formations was also analyzed, which indicated that they were mostly formed in late geologic time, Holocene and Pleistocene, which together formed the Quaternary period (see Table 9). Clastic sedimentary rocks that formed during this latest geologic period are mostly unconsolidated; therefore, these formations may serve as a good filter for drinking water supply.

Table 9: Stratigraphy of surface rock formations at RBF sites

Stratigraphic name	Main location of BF well	Number of wells/well fields
Holocene	Berlin; at the banks of Rhine, Elbe, Ruhr, Mulde, and Moselle River	394
Pleistocene to Holocene	Berlin	103
Pleistocene	Berlin; at the banks of Rhine and Mulde River	766
Permian	At the bank of Rhine River near Mainz	6

Sources: geologic information is from the geologic overview map (GÜK250) (BGR 2019); well or well field locations are from the same sources as Table 6.

RBF systems are also found in the deposits or formations that resulted from other processes, for instance, lowland moor peat as a product of phytogenic process and sandy

deposits as a product of the aeolian process (wind). Wells drilled into lowland moor peat are mostly found in Berlin in this study, more specifically, near the Havel River. The bog peat (German: *Moortorf*) at this RBF site is usually less than 1 meter thick, while at some wells located on bog peat, it may be over 1 meter thick, *e.g.*, in Rahnsdorf, near the lake Müggelsee (see Appendix 4, Figure A7 and Figure A8). The underlying formation is mostly sand. Wells with aeolian formations can be found mostly in Berlin. Aeolian deposits are those carried, transported, and deposited by wind during one or more periglacial periods; thus, they are very fine-grained, such as loess. Loess is a silt-sized porous material that makes an excellent soil but a poor aquifer (Meinzer 1923). Aeolian sand deposits in this study are, however, a better material to yield water. Near the city of Mainz, six wells are located on the Permian fluvial breccia formation, which is a type of clastic sedimentary rock. Due to its early origins, this kind of sediment is cemented into hard rock that is not likely to serve as a filter for RBF system and, to the author’s knowledge, has not been reported as a dominated material in RBF sites. The layer information at this RBF site is not available, but this kind of formation may have a very limited importance in this case due to its low occurrences.

A summary of some most frequent surface rock formations on which RBF sites in Germany were located are shown in Table 10, with respect to formative process, petrography, and stratigraphy. Although only the surface formations were analyzed, it is rather difficult to examine the underlying formations at all sites within the time frame of this project. Instead, layer information near some wells located on those formations is given in Appendix 4. If certain areas are considered to have the potential for an RBF system, a more detailed investigation on-site is needed to see if any clay or other confining units are present.

Table 10: Most frequent surface rock formations at RBF sites

Rock formation	Number of wells/well fields
Pleistocene glacial sand	654
Holocene fluvial clay/silt	229
Pleistocene fluvial sand	103
Holocene lowland moor peat	79
Pleistocene to Holocene fluvial sand	59
Pleistocene to Holocene aeolian sand	44
Holocene limnic clay/silt	36
Holocene fluvial sand	34

Sources: geologic information is from the geologic overview map (GÜK250) (BGR 2019); well or well field locations are from the same sources as Table 6.

After reviewing the existing RBF sites in Germany, a short conclusion of how the German experience may be used in the implementation of a new RBF system were given as follows:

- Fluvial and fluvio-glacial deposits are the most important material for an RBF system.
- Sand is the most frequent material occurring at RBF sites.
- Although gravel is the best formation to yield water, it does not occur as frequently as sand in this study.
- Fluvial clayey or silty deposits that cover and spread over the surface of sandy material is a common case in RBF implementation and has a favorable result.
- Aeolian sand and lowland moor peat spread over sandy material can also be used by RBF.
- Other formations are less promising; they either lack the ability to yield water or to act as an effective filter.

Fluvial aquifers are important sources of drinking water in many places, for example, China (Figure 14). Many large cities are located in the eastern part of the nation,



Figure 14: The distribution of fluvial deposits and populated cities in China

Sources: shapefile of China is from NFGIS, cited in Liao (2012); shapefile of other countries from Natural Earth (2019a); river network from Natural Earth (2019b); formation of fluvial deposits was digitalized after Zhang et al. (1990).

which is topographically lower than the west. As some of the large rivers (*e.g.*, Yellow River, Yangtze River) flow to the east, a great amount of debris is carried and deposited by them and their tributaries, which provides favorable geologic formations for RBF.

4.3 Land cover

Whether a specific land type should be selected and used for RBF system or not depends on the local authorities. Theoretically, all types of land cover or land use could be used or reconstructed for the purpose of drinking water supply using RBF, as long as the underlying geologic formation is suitable and the site is close to the river. Land cover, however, is closely related to the vulnerability of groundwater (Aller *et al.* 1987); therefore, it has an impact on the raw water quality of RBF system. Human activities change the original land cover and may increase the risk of groundwater being contaminated. Archwichai *et al.* (2011) developed a scoring and weighting system for RBF site selection, and set an ascending score for community/industry, agriculture, and open/public land use.

Nitrate accumulation in the shallow groundwater from the source of agricultural activities and landfills in urban areas is a problem that cannot be ignored anymore (Gu *et al.* 2013). Wellhead protection zones have to be delineated carefully to ensure a certain residence time of potential contaminants before they reach RBF wells; in addition, slow sand filtration combined with infiltration basins can be used to increase the share of infiltrated surface water in the production well or repel and dilute the contaminated groundwater bodies. Although this method is frequently applied together with RBF by waterworks, only the process of RBF was considered in this study.

In this study, impervious surfaces such as roads, airports, and other artificial structures were considered not feasible for RBF system implementation because of the high reconstruction costs.

4.4 Well field location

Yield is also affected by site geometry or well field location. RBF wells are often placed along the straight channel, at the cut bank or point bar, and on an island. Wells placed at the point bar or on an island are considered to be the best because of the high proportion of bank-filtrated surface water (Grischek, Schoenheinz, and Ray 2002). It is due to the movable ground of the streambed which reduces the risk of severe clogging as well as the gradient of the stream level which causes the natural cross-stream flow (Schubert 2004b). Cut bank is considered to be the worst due to the lower renewal capability of the streambed protected by large stones (Caldwell 2004).

4.5 Water quality

The quality of water discussed here includes that from the underground as well as from the rivers or streams on the land surface.

Groundwater quality. Groundwater is a component of raw water from RBF systems. Groundwater bodies to be used as drinking water sources have to meet the safety standard adopted by the local authorities. For example, groundwater used as a drinking water source should meet water quality category III (indicators and associated levels given in Appendix 2), as adopted by the Chinese authorities (AQSIQ 1993).

Surface water quality. Surface water quality has to be taken into consideration because bank-filtrated surface water is another component of the raw water from RBF systems. Surface water quality affects RBF systems in two ways. Firstly, although RBF systems have the ability to improve water quality, they cannot completely eliminate some critical contaminants from the surface water, for example, dissolved organic carbon (DOC), the amount of which in the raw water is rather important for water facilities. DOC is a food source of some organisms. A high concentration of DOC has a negative effect on the biological stability of drinking water in the distribution system (Grünheid, Amy, and Jekel 2005). Chlorine is one of the most widely-used disinfectants to kill pathogens in drinking water (U.S. EPA 2000). Organic matter in water reacting with chlorine will form harmful chemical by-products and therefore, is a primary concern (U.S. EPA 2001). Granular activated carbon (GAC) is often used in water treatment as an extra filter for absorbable species, such as organic micropollutants (OMPs) (Kühn and Müller 2000), which are generally present in water bodies in low concentrations but can have harmful effects on humans (UBA 2018). Although RBF systems have the capability of reducing OMP concentrations (Storck *et al.* 2012; Nagy-Kovács *et al.* 2018), a low concentration of DOC helps to enhance the removal of OMPs from GAC; thus, the run time of GAC can be extended and the cost can be saved (Kühn and Müller 2000). Furthermore, the increasing concentration of DOC significantly increases the mobility of heavy metals and arsenic from the soil (Antoniadis and Alloway 2002; Kalbitz and Wennrich 1998). According to Lenk *et al.* (2006), the DOC concentration at 33 RBF sites in Central Europe varies between 1.4 and 9 mg/L. In rivers, the concentration of DOC varies typically from 1 to 15 mg/L, in extreme cases, from 30 to 60 mg/L (BMI 1975). Wang *et al.* (2016) reported that the concentration of DOC in the Songhua River (a river in China) varies between 10.2 and 58.4 mg/L. Sites with better surface water quality, especially lower DOC concentration, should be considered first although sometimes there is no surface water quality requirement for RBF systems.

Secondly, severely contaminated river water will cause complete clogging on the riverbed, which can lead to reduction of yield and may even cut off the hydraulic connection between the surface water and groundwater (Grischek and Bartak 2016; Hubbs 2004). The process of clogging and the influencing factors were reviewed here due to its important role in RBF systems. Two types of clogging can be identified: external clogging

(German: *äußere Kolmation*), which is the compaction of the riverbed material through the deposition of suspended load and precipitate of river on the surface; and internal clogging (*innere Kolmation*), which is caused by infiltration of fine sediments into the pores or interstices of riverbed material (Beyer and Banscher 1975; Lisle 1989; Blaschke *et al.* 2003).

Clogging of the riverbed is well described and summarized in Beyer and Banscher (1975), Schälchli (1992), and Baveye *et al.* (1998), the process of which is influenced by physical, chemical, and biological variables. The water-quality-associated variables that cause riverbed clogging can be summarized as suspended load, grain size distribution, and shape of suspended particles (physical); type and concentration of dissolved matter (chemical); the variety of invertebrates, algae and their commonness, and the extent of eutrophy of the water (biological) (Schälchli 1992).

Every river contains a variable amount of suspended fine particles. An increasing concentration of the suspended load in the surface water can accelerate the clogging process (Schälchli 1992; Baveye *et al.* 1998), more precisely, the development of internal clogging (Schälchli 1993). According to Schälchli (1993), the grain size distribution of the suspended load influence mainly the external clogging process, and a higher proportion of the finest particles can accelerate the reduction of riverbed hydraulic conductivity. Coarser, less graded suspended materials tends to cause less clogging (at low velocities) than finer, better graded materials (Cunningham *et al.* 1987). Under certain chemical conditions (pH value, redox potential), the precipitation of mineralogical components such as $\text{Fe}(\text{OH})_3$ and FeCO_3 from Fe^{2+} rich surface water at riverbed surface or interstices of riverbed materials through infiltration is the main cause of riverbed clogging (BMI 1975). Biological (or microbiological) clogging might be caused by the accumulation of cells in the pore space of riverbed materials (Baveye *et al.* 1998).

Baveye *et al.* (1998) reviewed the few existing mathematical models in predicting the process of biological clogging. Schälchi (1993) found two equations that describe the process of physical clogging, with consideration of the concentration of suspended load; however, there is still no proven model to simulate and forecast the complete process of riverbed clogging; recent studies based on parameters that are rather incomplete (Griseck and Bartak 2016; Pholkern *et al.* 2015).

4.6 Aquifer properties

Hydraulic conductivity. The hydraulic conductivity (K) refers to the water-transmitting characteristic of aquifer material in quantitative terms (Heath 1983). Based on the review of RBF sites in Germany, sandy material is the most common aquifer medium for RBF systems. According to Domenico and Schwartz (1997), unconsolidated sandy materials have a K value that varies from 2×10^{-7} to 6×10^{-3} m/s (see Table 11). Based on the

Table 11: Size and hydraulic conductivity of sand filters

Characteristics	Fine sand	Medium sand	Coarse sand
Particle size (mm)	>0.063 to ≤0.20	>0.20 to ≤0.63	>0.63 to ≤2.0
K (m/s)	2×10^{-7} to 2×10^{-4}	9×10^{-7} to 5×10^{-4}	9×10^{-7} to 6×10^{-3}

Source: particle size from ISO (2017), hydraulic conductivity from Domenico & Schwartz (1997).

work of Freeze and Cheery (1979), it varies over a range of 10^{-7} to 10^{-2} m/s (Figure 15).

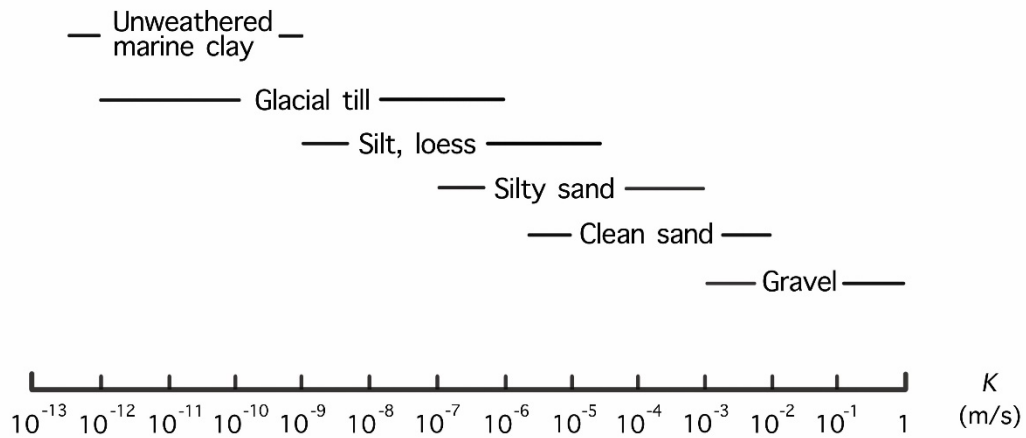


Figure 15: Hydraulic conductivity of unconsolidated deposits

Source: modified after Freeze and Cheery (1979).

Not all sandy materials within these ranges are suitable for RBF. Previous work suggests at least 1×10^{-4} m/s is required for RBF systems (Grischek, Schoenheinz, and Ray 2002; Lenk *et al.* 2006); however, if other site conditions are favorable, a lower value, *e.g.*, 5×10^{-5} m/s could also be sufficient (Grischek 2019). K value reported at well fields of Cedar Rapids (U.S., with 7.5×10^{-5} to 10^{-3} m/s) and Lake Tegel (Germany, with 5×10^{-5} to 10^{-4} m/s) supported this statement (Grischek, Schoenheinz, and Ray 2002; Hoffmann and Gunkel 2011). Archwichai *et al.* (2011) set hydraulic conductivity values higher than 1 m/d (ca. 1.2×10^{-5} m/s) for selecting potential areas used for RBF; however, no highest threshold value was set. The highest K value in the field was likely to be about 10^{-2} m/s, for instance, at well field Auf dem Grind (10^{-3} to 10^{-2} m/s), Böckingen (10^{-2} m/s), at one site at the middle reach of the Ruhr River (2×10^{-2} m/s), and on an island in the Danube River in Regensburg (2×10^{-2} m/s) (Grischek, Schoenheinz, and Ray 2002; Lenk *et al.* 2006). All

these sites are located in Germany. In practice, RBF systems located at gravel aquifers with a K value greater than 10^{-2} m/s works well if the quality of river water is favorable. To conclude, suitable K values were set to range between 10^{-5} and 10^{-1} m/s in order to include all possible conditions, and suitable aquifer mediums should be dominated by sand and gravel.

Hydraulic conductivity is one of the very few physical parameters that varies on such a wide range of magnitude (see Figure 15); expressed in units of hydraulic conductivity class K_c may be more meaningful (Bear 1972):

$$K_c = -\log_{10}K(m/s). \quad (2)$$

Saturated thickness. In an unconfined aquifer, the saturated thickness (b) is defined as the vertical distance between the water table surface and the aquifer base. In Lenk *et al.*'s 2006 work, the aquifer thickness of 33 RBF sites ranges from 4 m to 70 m, with a mean value of 21.5 m (Lenk *et al.* 2006). At the lower reach of some continental rivers, where the alluvial deposits may reach a thickness of hundreds of meters, a thicker aquifer can be expected (Ghodeif *et al.* 2016; Zhang *et al.* 2009). The saturated thickness determines the aquifer's ability to yield water in low streamflow conditions as the amount of induced river water is limited during this period. While in high streamflow conditions, a relatively large share of induced river water will be extracted from the pumping wells. As the aquifer becomes thinner, the amount of water in the underground may not sufficient for water facilities with a high extraction rate. Experience shows that at least 5 m of saturated thickness should exist for a feasible bank filtration site (Griseck, Schoenheinz, and Ray 2002; Kühn and Müller 2006). This more or less matches with the lowest saturated thickness (4 m) of the aquifer at one site near the Neckar, a river in Germany (Lenk *et al.* 2006).

Transmissivity. For an aquifer of hydraulic conductivity K , and saturated thickness b , the transmissivity is defined as:

$$T = Kb. \quad (3)$$

Griseck (2019) argued that transmissivity could be used as a parameter for RBF site selection instead of using hydraulic conductivity and saturated thickness. According to Freeze and Cheery (1979), aquifers that are good for groundwater exploitation should have transmissivities greater than 0.015 m²/s. In practice, transmissivity reported at RBF sites ranges from 0.003 to 0.347 m²/s (Ghodeif *et al.* 2016; Lenk *et al.* 2006).

If transmissivity is used as a single attribute for RBF site selection, hydraulic conductivity and saturated thickness have to be checked, particularly in areas with low transmissivity caused by unsuitable low hydraulic conductivity.

4.7 Distance to river

The distance between well and river (L) has an impact on two factors for an RBF system: the share of bank-filtrated river water and its quality.

The share of bank-filtrated river water in the well generally decreases with increasing distance from the river. Riverbed clogging will reduce the rate of river water infiltration, so if severe clogging is expected, production wells should be placed close to the riverbank to ensure a planned share of bank-filtrated river water during operation (Grischek, Schoenheinz, and Ray 2002).

River water or bank-filtrated river water is usually softer and has lower nitrate concentration than true groundwater. If for certain reasons a high fraction of bank filtrate is desired, to promote a softer or raw water with lower nitrate concentration, for example, less distance from the surface water is required (Fokken 1996b).

A greater distance (over ca. 8 m) may significantly improve water quality before bank-filtrated river water reaches the well because it is favorable for the removal of pathogens (such as *Cryptosporidium* oocyst) from the surface water (U.S. EPA 2010). Bartak *et al.* (2015) stresses a distance between 20 and 70 m is sufficient to remove coliform bacteria and algae. As the distance from the riverbed or bank increases, a reduced zone is formed which means the aquifer becomes less re-aerated, which leads to precipitation of iron and manganese onto the sediments in the subsurface (Bourg and Bertin 1993). Oxygen is usually significantly depleted within ca. 1.5 m - 4.6 m of the riverbed, due to microbial activity in this zone (U.S. EPA 2010). As the distance increases, dissolved oxygen increases and manganese decreases, which means these chemical changes occurring in the reduced zones are reversible (Bourg and Bertin 1993); therefore, Bourg and Bertin (1993) suggested boreholes to be located outside of this zone.

The horizontal aquifer passage, the distance between production wells and water bodies (more specifically, river banks), should be optimized based on expected infiltration rates, preferred flow path length, and preferred residence time (Grischek, Schoenheinz, and Ray 2002; Kühn and Müller 2006). The flow path of the infiltrated surface water is likely longer than the horizontal aquifer passage because it does not travel in a straight line.

The existing RBF sites have a wide range of horizontal aquifer passage from a few meters (*e.g.*, 10 m at Matan in Lanzhou, China) to few kilometers (*e.g.*, 3.5 km at Velddriel and Sellik in Aalst, the Netherlands) (Hu *et al.* 2016; Stuyfzand, Juhász-Holterman, and de Lange 2004).

From the existing RBF systems, the distance varies not only from site to site but also within a site. In China, the distance ranges from 10 m to 3000 m (Hu *et al.* 2016). In Europe, the distance at most sites is not less than 50 m (Grischek, Schoenheinz, and Ray 2002). Similar findings given by Ray *et al.* (2002), who noted that the distance between well (both vertical filter well galleries and horizontal collector wells) and the riverbank along the Rhine River in Europe varies from 50 m to approximately 250 m. The minimum and maximum distance given by Sontheimer (1991) is 20 m (Düsseldorf-Flehe) and 860 m

(Krefeld), respectively, both located in sites along the Rhine River. Experience by Cologne waterworks on the Rhine River concluded that a distance of 150 m to 400 m from wells to the edge of the main channel is favorable, so long as the wells are located flood-free (Fokken 1996b). Sprenger *et al.* (2017) generated a database for managed aquifer recharge (MAR) sites in Europe, relating that MAR methodologies were induced bank filtration, surface spreading, and well injection. The database showed that for induced bank filtration in Europe, the horizontal aquifer passage (*i.e.*, distance from well to surface water) varies between 50 m to 1270 m ($N=78$). Another study on Central European sites with a share of bank-filtrated surface water higher than 50% indicated that the largest distance from well to water body is 310 m (Lenk *et al.* 2006). Stuyfzand *et al.* (2004) investigated 21 RBF well fields that have records of distance to the edge of the surface body in the Netherlands. They vary between 130 m and 3500 m (Stuyfzand, Juhász-Holterman, and de Lange 2004). The U.S. EPA has suggested a minimum of ca. 8 m between the edge of the surface water body at its 100-year flood extent and the pumping well; such a distance is necessary in order to remove pathogens from the surface water (U.S. EPA 2010). Based on the aquifer medium, the Ministry of Ecology and Environment of the People's Republic of China (MEE) has suggested a minimum distance of 30 m between the edge of primary protected area of a drinking water source and the pumping well located in the unconfined aquifer is required in order to prevent direct contamination from anthropogenic activities (MEE 2018b).

Rather long distances were used in modeling suitable areas for RBF in some previous works. Yin *et al.* (2018) used a buffer area within 15 km from the Second Songhua River and 10 km from the tributary as the main study areas for RBF feasibility study. Wang *et al.* (2016) used a buffer area within 20 km from the riverbank as the core study range of RBF (*i.e.*, the feasible region of riverbank filtration), with respect to hydraulic connection between river water and groundwater; however, wells located at such a distance may abstract only limited share of bank-filtrated water.

High water level or flooding water level, however, is a critical factor for the distance between the edge of the surface water body and the well. Area under the risk of floods should also be delineated and may further be defined as unsuitable for RBF. According to the EPA, RBF wells should not be placed in the 100-year flood inundation area (U.S. EPA 2010). Fokken (1996b) argued that the moderate flood inundation area (rather than the 100-year flood) is not suggested for the siting of RBF wells. Although large floods, *e.g.*, 100-year flood, may cause devastating damage to the water facilities along the riverbank, considering the relatively low chance of encountering a 100-year flood, the long operation of RBF system during the flood-free period may have greater importance to the people that rely on this system. If the information of historical floods is available for the area, more specifically moderate floods that happen more frequently and threaten the RBF system, it should be used to delineate the unsuitable area along the riverbank.

4.8 Riverbed characteristics

River hydrology has an impact on sediment transport and renewal capability of the river, while the riverbed characteristics determine the ease at which the sediment is resuspended by flows as well as the thickness of the clogging layer.

For many streams, an armor layer with coarse materials at the streambed surface will protect trapped sediment from resuspension during high flow events and prevent the restoration of capacity in RBF systems (Caldwell 2004). The breakup of the armor layer increases the riverbed hydraulic conductivity greatly. For this flushing process, not only the high flow condition is required, but also the grain size distribution of riverbed materials is of importance. The ease at which this armor layer is flushed can be measured by shear stress.

Günter (1971) and Schälchli (1993) studied the relationship between the critical dimensionless shear stress of the armor layer (German: *Deckschicht*) to start breaking up and the d_{mA}/d_m ratio. A well-proven formula in Switzerland was given by Günter (1971), as cited in Schälchli (1993, 1995):

$$\theta_A/\theta_{cr} = (d_{mA}/d_m)^{2/3}, \quad (4)$$

where θ_A is the critical dimensionless shear stress of the armor layer to start breaking up, θ_{cr} is the critical dimensionless shear stress of bed load begin, which is usually a constant value of 0.05. d_{mA} is the average grain size of the armor layer and d_m is average grain size of the streambed material.

Schälchli (1993) studied the process of streambed clogging and the required flow-conditions for flushing the clogging layer. In the author's work, Günter's formula was verified by using a natural mixture of streambed materials, as can be seen in Figure 16, the sample points (see Table 12) gathered close to the empirical line. The two authors' works show a phenomenon that in order to break up the armor layer, streambeds with greater d_{mA}/d_m ratio need higher shear stress, which means they are more difficult to be removed during flood conditions.

Gessler (1965) introduced a method to sample and determine the average grain size of the armor layer d_{mA} . In practice, however, streambed materials sampled by previous investigations may not be the purpose of studying the armor layer; thus, it may not be analyzed separately. The d_{mA} size of the streambed materials may be missing. The d_{90}/d_m ratio is, however, much easier to obtain. The d_{90} diameter indicates the sediment

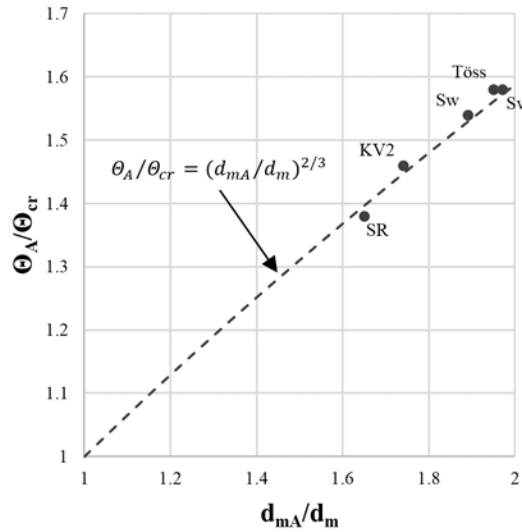


Figure 16: Begin of the breakup of the armor layer θ_A/θ_{cr} in depended on d_{mA}/d_m ratio

Source: formula is after Günter (1971), as cited in Schälchli (1993); points are streambed samples used by Schälchli (1993).

Table 12: Streambed materials used by Schälchli (1993)

Material name	θ_A	θ_A/θ_{cr}	d_{mA}	d_m	d_{mA}/d_m	d_{90}/d_m
Sw	0.077	1.54	51.2	27	1.89	2.37
Sv	0.079	1.58	63.2	32	1.97	2.44
SR	0.069	1.38	32.8	20	1.65	2
KV2	0.073	1.46	11.9	6.9	1.74	2.17
Töss	0.079	1.58	43.2	22	1.95	2.36

Note: $\theta_{cr} = 0.05$; grain size d_i in millimeters; d_{mA} was calculated using d_{mA}/d_{10} and d_{10} given by the author. For more detail information of the materials please see Appendix 5.

diameter/size for which 90 percent (by weight) of the sediment sample is finer, which is also the 90th percentile of the cumulative grain size distribution curve (also applies to other diameters). Schälchli (1995) claimed that d_{mA} size can be approximately taken as d_{90} size of the streambed material. The relationship of d_{90}/d_m ratio and d_{mA}/d_m ratio is shown in Figure 17 (left).

Based on the streambed sample collected by Schälchli (1993), the d_{mA}/d_m ratio showed a positive correlation with d_{90}/d_m ratio and a regression function was given here. In fact, it is not hard to find that the d_{90} size is a bit larger than d_{mA} in this case. Nevertheless, based on this empirical function we can see that θ_A/θ_{cr} and d_{90}/d_m ratio also fulfill Günter's formula (Figure 17, right). This proved that the d_{90}/d_m ratio can appropriately describe the

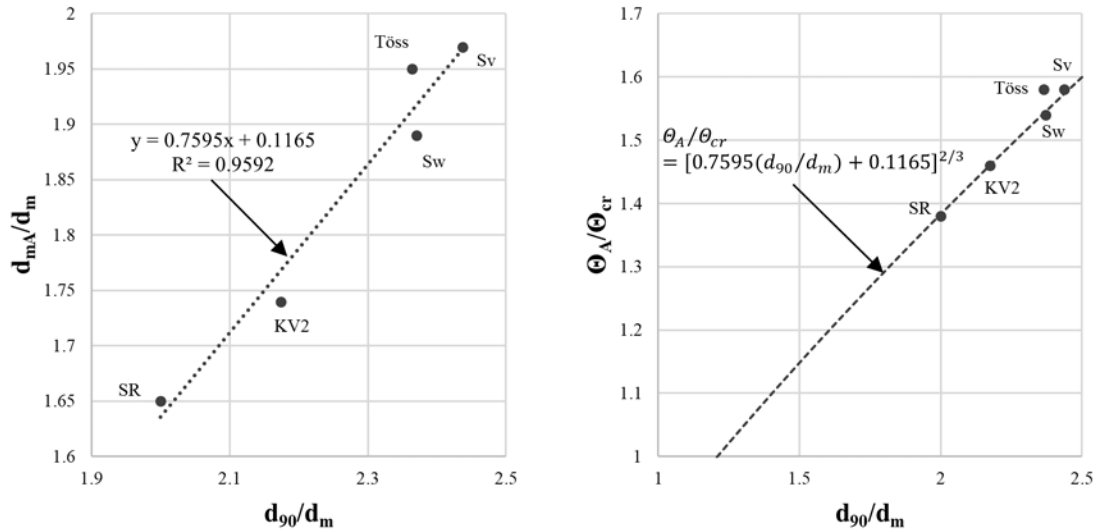


Figure 17: Relationship between d_{mA}/d_m and d_{90}/d_m , θ_A/θ_{cr} and d_{90}/d_m

Source: data from Schälchli (1993).

property of the armor layer during a flood. To remove the armor layer, a higher shear stress is needed for the material of a streambed with a greater d_{90}/d_m ratio.

According to Schälchli (1993), the d_{mA}/d_m ratio of natural streambed materials usually varies within the range of 1.5 and 1.9, while for very coarse streambed materials in the Alps, they may reach a value of 2.6. Based on this and assuming that the relationship of d_{mA}/d_m and d_{90}/d_m showed in Figure 17 (left) fulfills all streambed materials in nature, a proper range of d_{90}/d_m ratio may be between 1.8 and 3.2. The ratio of other streambed samples used by Schälchli (1993) was also found within this range (see Appendix 5, Table A5).

Caldwell (Caldwell 2004) suggested d_{90}/d_{10} ratio can be used to predict the renewal capacity of RBF system during a flood event. A large d_{90}/d_{10} ratio, for instance, indicates the presence of an armor layer at the streambed surface, which protects trapped sediment from resuspension during high flow events and prevents the restoration of capacity in RBF systems (Caldwell 2004); however, studies on the effect of this ratio are rare, and a cross-check is needed.

Another parameter describing the riverbed characteristics is the average grain size of the riverbed materials (d_m), which determines the thickness of the clogging layer on the riverbed (Schälchli 1993). The thickness of the clogging layer is an important factor affecting the infiltration rate; thus, it also has an impact on the yield of bank filtrate. This parameter or criterion was reviewed in more detail in the following section—“Effect of clogging on yield”.

5. Effect of Clogging on Yield

The degree to which clogging affecting the yield of bank filtrate can be described using two parameters: the hydraulic conductivity of the clogged layer and its thickness (BMI 1975). The continuous reduction of hydraulic conductivity and the addition of thickness of the clogged layer both affect the yield negatively.

Let $K(t)$ and $h(t)$ be the hydraulic conductivity and the thickness of the clogging layer at time t , for fully saturated conditions, the average hydraulic conductivity (K_m) of the two-layer system (illustrated in Figure 18) is given by Bear (1979), as cited in Brunner *et al.* (2009):

$$K_m = \left(\frac{1}{h(t)+h} \left(\frac{h(t)}{K(t)} + \frac{h}{K} \right) \right)^{-1}, \quad (5)$$

where h and K are the thickness and hydraulic conductivity of the aquifer strata.

Under conditions of steady state flow or constant water table, the saturated infiltration flux ($q(t)$) at time t is the average hydraulic conductivity multiplied by the hydraulic gradient, according to Brunner *et al.* (2009); it reduces to

$$q(t) = (h(t) + h + d) \left(\frac{h(t)}{K(t)} + \frac{h}{K} \right)^{-1}, \quad (6)$$

where d is the constant water depth.

Compared to the beginning of the infiltration, *i.e.*, when $K(t)=K$, the change of flow rate ($Q(t)/Q$) or infiltration flux ($q(t)/q$) over an area S over time t can be expressed as

$$\frac{Q(t)}{Q} = \frac{q(t)S}{qS} = \left(\frac{c}{1+c} + \frac{1}{1+c} \left(\frac{K(t)}{K} \right)^{-1} \right)^{-1}, \quad (7)$$

where $c = h/h(t)$.

The change of the flow rate over time depending on the change in hydraulic conductivity as well as the influence of $h/h(t)$ ratio is shown in Figure 18.

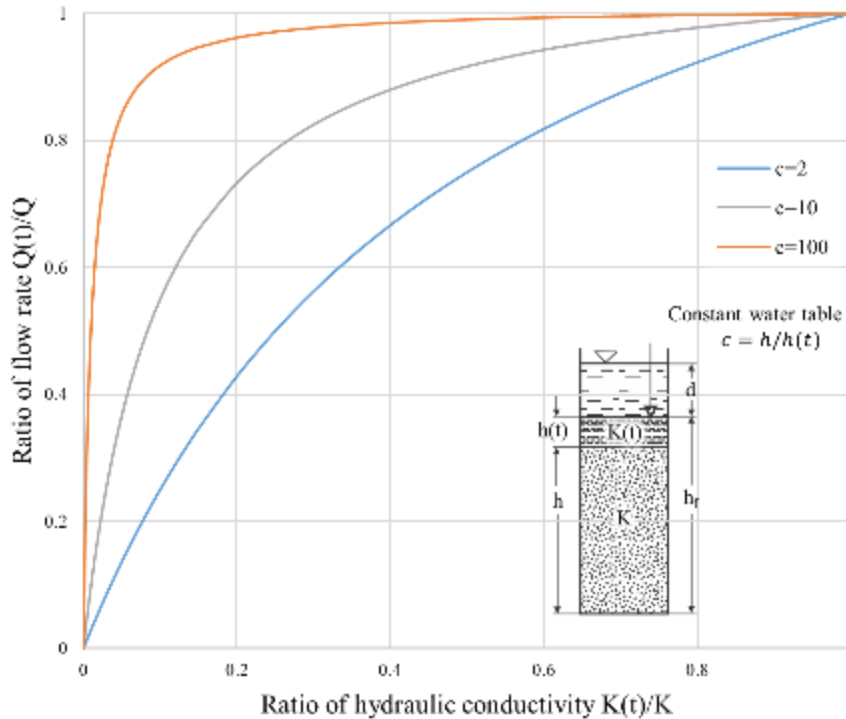


Figure 18: Change of flow rate over time depending on the change in hydraulic conductivity

Source: modified from van Riesen (1975, 109) and BMI (1975, 5).

The length of groundwater-flow paths (h_f) is defined as

$$h_f = h(t) + h, \quad (8)$$

which is just slightly larger than h and can therefore be approximately considered as the length of bank-filtrated river-water-flow paths during RBF. For large $h/h(t)$ ratios, a significant reduction of flow rate only happens at an extremely small $K(t)/K$ ratio. For instance, given a realistic scenario $h/h(t)=100$, a 90% reduction of hydraulic conductivity due to clogging will lead to only about 8% reduction of the flow rate (BMI 1975), which is not much for RBF systems. Significant reduction of the flow rate happens when the clogging layer is remarkably thick and its hydraulic conductivity is extremely low (BMI 1975). A large $h/h(t)$ ratio is common at RBF sites because the distance between well and river water (L) varies from a few meters to a few thousand meters (see section “Distance to river”), while the thickness of the clogging layer is usually just some centimeters; moreover, h or h_f is usually much larger than L (Schälchli 1993). Extremely low hydraulic conductivity of the clogging layer, however, can be expected in RBF. Beyer and Banscher (1975) reported a *quasi-stable final state* (German: *quasi-stabiler Endzustand*) of the

internal clogging layer, after which the hydraulic conductivity changes (reduces) only very slightly. The quasi-stable final state of $K(t)$ is about 10^{-7} m/s for both sandy and gravelly riverbed materials. The thickness of the clogged layer $h(t)$ is about 0.03 m and 0.1 m for sandy and gravelly materials, respectively (Table 13).

Table 13: Parameters of Beyer-Banscher's quasi-stable final state of clogging

Parameter	Scale of measure	Materials	
		Sand	Gravel
K	m/s	4×10^{-4}	1.2×10^{-3}
$K(t)$	m/s	10^{-7}	10^{-7}
$h(t)$	m	0.03	0.1

Source: Beyer and Banscher (1975).

We can assume here that the character of riverbed material also represents the aquifer medium. At Beyer-Banscher's final state of clogging, the $K(t)/K$ ratio is as small as 2.5×10^{-4} and 8.3×10^{-5} for sandy and gravelly materials, respectively. On the contrary, the clogging layer at gravel riverbed is thicker than that of a sandy bed, which is consistent with Schälchli (1993, 124). An equation of the thickness of the clogged layer ($h(t)$) depending on average grain size (d_m) of the coarse riverbed materials was given by Schälchli (1993, 124):

$$h(t) = 3d_m + 0.01, \quad (9)$$

while for fine riverbed materials, the clogged layer is about $5d_m$ thick; thus, for aquifer strata of the same thicknesses (h) or RBF system of the same length of flow path (h_f), a much stronger effect of clogging on flow rate can be expected for gravelly riverbed materials than sandy materials (Figure 19).

For gravel materials, more than 50% reduction of flow rate is expected, even though the flow path is as long as 1000 m; whereas for sandy materials, a 500 m length can still maintain about 81% of the original flow rate. The curves also show that a longer h causes less reduction of the flow rate.

The aforementioned reduction of the flow rate is measured in percent; however, the initial flow rate ($t=0$) of gravelly materials (Q_{gravel}) is much higher than sandy materials (Q_{sand}) because

$$\frac{Q_{sand}}{Q_{gravel}} = \frac{q_{sand}S}{q_{gravel}S} = \frac{K_{sand}}{K_{gravel}}, \quad (10)$$

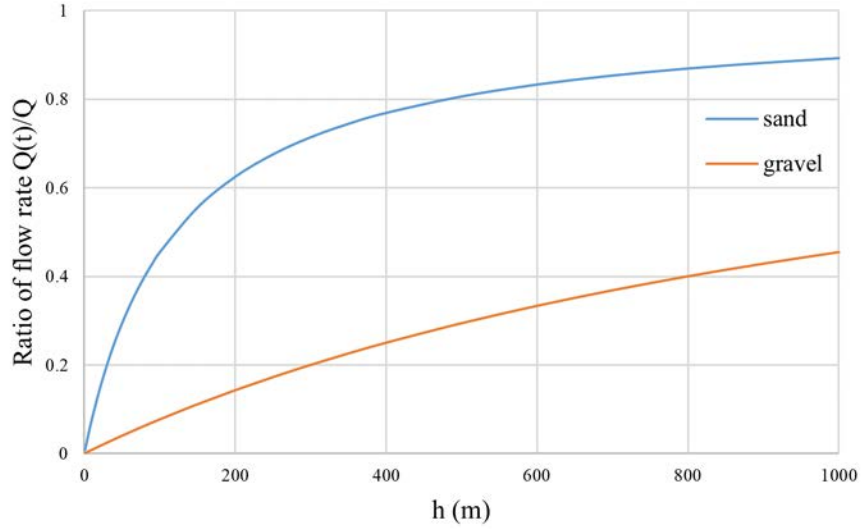


Figure 19: Change of flow rate depending on the thickness of aquifer strata at quasi-stable final state of clogging

Note: curves are acquired using equation (10) and parameters from Beyer and Banscher (1975), see Table 13.

where q_{sand} and q_{gravel} are the initial infiltration flux of sandy and gravelly aquifer strata; K_{sand} and K_{gravel} are the hydraulic conductivity of the sandy and gravelly aquifer strata.

For Beyer-Banscher's case, the Q_{sand}/Q_{gravel} ratio is about 0.33. If we assign the initial flow rate Q_{sand} for 1 unit (m^3/s), the Q_{gravel} should have 3 units. Then, the unit-reduction (instead of percent-reduction) of Q due to clogging can be illustrated as the curves in Figure 20.

Although gravelly materials have more percent reduction than sandy materials (for the same h), the final flow rate of gravelly materials exceeds that of sandy materials after the thickness of aquifer strata (h) increased to about 420 m. This gives us some insight in RBF implementation: if we know the estimated average flow path (h_f) is much shorter than 420 m, sandy materials can yield more bank filtrate than gravelly materials; whereas for longer h_f , gravelly materials can yield more bank filtrate than sandy materials. In order to yield a high proportion of bank filtrate, the distance between well and river water (L) in RBF is usually less than 310 m (Lenk *et al.* 2006). As has been explained in the previous text, the flow path h_f is usually much longer than L and is not constant (Schälchli 1993). This increases the uncertainty in choosing an appropriate distance between wells and river water. More study on the influence of the average grain size (d_m) of riverbed materials to the final yield of bank filtrate after full clogging is therefore needed.

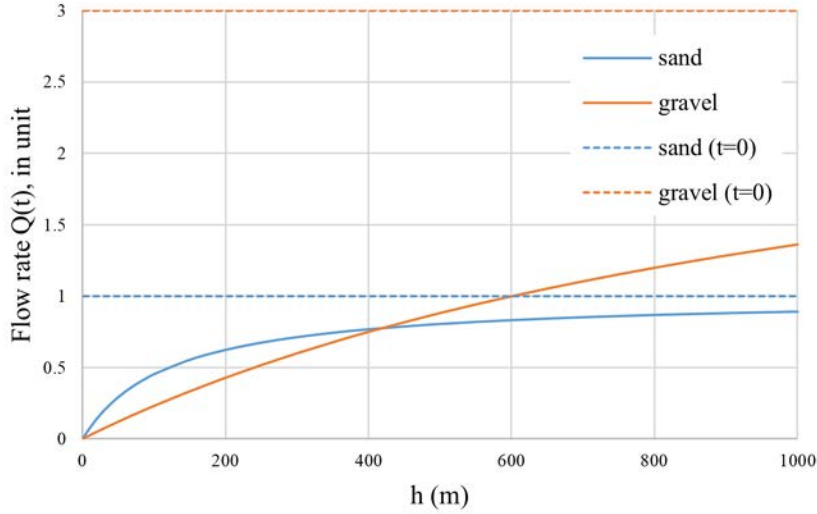


Figure 20: Change of flow rate (in unit) depending on the thickness of aquifer strata at quasi-stable final state of clogging

Note: curves are acquired using equation (10) and parameters from Beyer and Banscher (1975), see Table 13. The dotted lines refer to the initial state before clogging.

These results are based on theoretical simulations that only hold when the two-layer system remains saturated or hydraulically connected. Brunner *et al.* (2009) claimed that the soil beneath the clogging layer has the potential to become desaturated under the condition

$$\frac{K(t)}{K} \leq \frac{h(t)}{d+h(t)}. \quad (11)$$

This condition is not fulfilled when the water depth (d) is greater than 119.97 m for sandy materials, and 1199.9 m for gravel materials (at Beyer-Banscher's final state of clogging). This means, in reality, the two-layer system has probably long become disconnected since the river water seldom has these depths.

However, this model might be too simple to simulate the complete RBF process, since, in reality, the clogged area will spread out the riverbed, and the intensity of clogging is somewhat different in this area. According to Schubert (2002b), this almost impermeable (10^{-8} m/s) and fully clogged layer covers a region of the riverbed nearest to the wells, which stretches from the well-side bank about 80 m to the middle of the river (*e.g.*, well field Flehe); the surrounding infiltration zone (partly clogged) has, however, a rather high hydraulic conductivity (3×10^{-3} m/s), which only causes a slight reduction of flow rate. Furthermore, clogging is a dynamic process, which is also influenced by river flow conditions. Erosion or bed load transport condition will lead to a self-cleaning of the

riverbed, which can slow down the formation of clogging layer (Schubert 2002b). It took about 20 years in order to reach Beyer-Banscher's quasi-stable final state of clogging at RBF site along the Elbe River and Mulde River (Germany) (Beyer and Banscher 1975). This indicated that, in practice, the overall effect of clogging on flow rate or yield is somewhat smaller than the cases illustrated in Figure 19.

In conclusion, clogging might have only a limited impact on the yield of bank filtrate. Under the condition of Beyer-Banscher's quasi-stable final state of clogging, the effect of clogging on yield can be summarized as follows:

- given the same length of flow path, coarse riverbed materials (or aquifer mediums) tend to cause more reduction (in percent) on yield than a riverbed that consists of fine materials, and,
- for short distances, fine riverbed materials can provide more yield than coarse riverbed materials, however,
- for long distances, coarse riverbed materials can provide more yield than fine riverbed materials.

6. Summary

In this part, the fundamentals of RBF and its positive influence on safe drinking water supplies for the less developed world were described. Models on RBF site selection were reviewed and their pros and cons were briefly discussed. The result indicates the need for developing a new decision making approach, from which the uncertainties of the trade-offs between the attributes can be taken into account. From those models, aquifer properties (such as hydraulic conductivity, saturated aquifer thickness, and transmissivity), surface water quality, and groundwater quality were used as influencing factors consistently. Higher hydraulic conductivity, saturated thickness, transmissivity, and water quality are preferred. A comprehensive review of major factors affecting RBF site selection were made. Some factors, such as well field location, the renewable ability of streambed hydraulic conductivity, and DOC concentration in the surface water, were not considered as criteria in existing models yet have a rather high impact on the yield or quality of the bank-filtered river water. The d_{90}/d_m ratio of streambed materials was for the first time introduced as a criterion for RBF site selection. Due to the special meaning of the clogging layer to RBF systems, the effect of clogging on the yield of bank filtrate was also reviewed, the result of which showed that clogging of streambed might only have a limited impact on the yield of bank filtrate. Nevertheless, the grain sizes of the aquifer medium (particularly the streambed materials) should be taken into consideration not only because it determines the aquifer hydraulic conductivity but also the intensity of the clogging effect. Furthermore, under the condition of severe clogging, the grain sizes of the aquifer medium and the distance between wells and river water should be considered simultaneously in order to maximize yield.

Part III

Developing a Multi-attribute Utility Model for RBF Site Selection

1. Introduction

Ever since Daniel Bernoulli (1700-1782), a Swiss mathematician and physicist, introduced *utility* to describe the price that a person is willing to pay for a gamble (St. Petersburg Paradox) in 1738 (Bernoulli 1954), numerous mathematicians have contributed to the development of the *expected utility theory* (mainly after World War II) and the foundation of decision theory (von Neumann and Morgenstern 1953; Fishburn 1989).

Multi-attribute utility theory (MAUT) is a decision theory that developed for dealing with problems with multiple objectives (Keeney and Raiffa 1993). It allows the decision maker to choose among a set of alternatives being evaluated on the basis of two or more attributes (Dyer *et al.* 1992). For about half a century, MAUT has been successfully adopted in fields of operations research, management science, computer science, artificial intelligence, management, business, applied mathematics, civil/environmental/industrial/manufacturing engineering, economics, information systems, energy, and water resources (Wallenius *et al.* 2008).

As introduced in Part II, several approaches have been already conducted by scholars in order to solve the problem of RBF site selection, for example, the analytic hierarchy process (AHP) (Saaty 1987). People prefer to use AHP than MAUT in some decision problems because it is more user-friendly than MAUT (Chen and Lee 2000). MAUT, however, despite this limitation, has a well-established axiomatization and a complete theoretical foundation that differs from those methods already applied. There is a need to develop a complete model for RBF site selection based on MAUT.

2. Objectives and Attributes

In this section, objectives and attributes for the MAUT model used for RBF site selection problem are set. Any utility assessment starts with a discussion of the concepts of the problem and structuring the problem. The basic or overall objective of selecting and ranking these RBF sites is to develop long-term operating RBF systems used for drinking water supply, with respect only to factors such as the quantity and quality of the raw water. How the wells are distributed on the field, which type of wells are used, and how the drinking water is distributed are out of the scope of this study; thus, they were not introduced here. The major objectives are two: maximize well yield and optimize well water quality. The problem of selecting the most suitable site for RBF system might be established as a simple model in Figure 21.

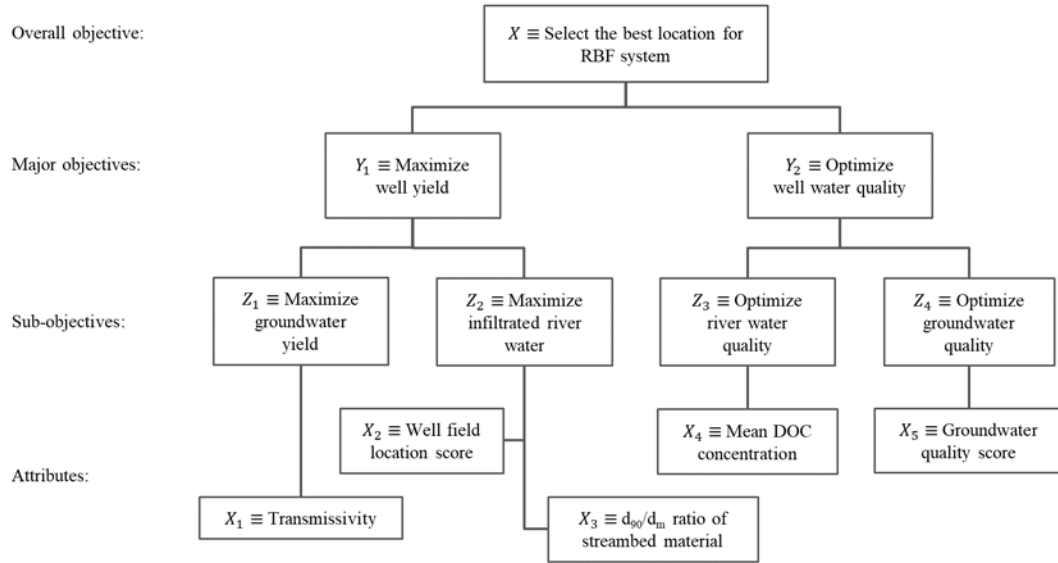


Figure 21: Hierarchical structure of attributes for the RBF siting problem

Note: \equiv means “is defined as”.

Water abstracted from RBF wells is a mixture of true groundwater and bank-filtrated river water, which is why the quality and quantity of both are of concern. The objectives may conflict. For example, a larger distance between wells and rivers is favorable for the quality of bank-filtrated river water but its amount or share in abstracted well water will be reduced. This is particularly true if the amount of stored groundwater is limited and a large share of bank-filtrated river water is required.

Factors affecting RBF system performance were addressed in the previous section. Five attributes were set and used later in assessing the utility functions to evaluate the suitable RBF sites (Table 14). Threshold values are either based on global RBF site conditions or historical field investigations of the study area.

Table 14: List of attributes for the RBF siting problem

Attribute	Scale of measure	Worst	Best
$X_1 \equiv$ Transmissivity	Square meter per second (m^2/s)	0.001	0.35
$X_2 \equiv$ Well field location score	Subjective	0	100
$X_3 \equiv d_{90}/d_m$ ratio of streambed material	Ratio scale	3.5	1.5
$X_4 \equiv$ Mean DOC concentration	Milligram per liter (mg/L)	60	0
$X_5 \equiv$ Groundwater quality score	Subjective	0	100

Transmissivity. Transmissivity (T) is the product of hydraulic conductivity (K) and aquifer thickness (b). For the objective “maximize the groundwater yield”, greater transmissivity means a higher capability of the aquifer to transmit water; thus, an aquifer

with a higher transmissivity is better for well exploitation. The minimum and maximum T values are based on Lenk *et al.* (2006) and Ghodeif *et al.* (2016). A higher hydraulic conductivity may lead to higher transmissivity but shorter residence time of the bank filtrate, which may cause a lower pollution removal efficiency; however, Lenk *et al.*'s 2006 work reported that the removal efficiency of surface water DOC concentration (range: 1.5-9.0 mg/L) was positively correlated with aquifer transmissivity (range: 0.003-0.230 m²/s). This indicated that to a certain degree greater transmissivity is favorable for the raw water quality. Furthermore, research at well field Torgau-Ost (Germany) showed that, already about half of the total DOC removal achieved at the colmation/clogging layer on the surface of the riverbed (Nestler *et al.* 1997). One of the key factors for RBF as an effective water pre-treatment technique is the clogging layer, the formation of which during continuing infiltration of surface water is unavoidable. That research indicated that the influence of transmissivity on pollution removal is probably not as important as the clogging layer; therefore, one assumption is made here: for transmissivity within the model range (0.001-0.350 m²/s), the higher the transmissivity of the aquifer, the better the yield and DOC removal efficiency.

Well field location score. Subjective values are given for this attribute: 100 for wells located “on an island”, 66 for “at the point bar”, 33 for “along the straight channel”, and 0 for “at the cut bank”.

d_{90}/d_m ratio of streambed material. According to the work of Schälchli (1993), the d_{90}/d_m ratio of streambed material from natural rivers varies between 1.8 and 3.2. See Part II for more detail. For the multi-attribute utility model, the range was therefore set from 1.5-3.5.

Mean DOC concentration. The range of mean DOC concentration of surface water was set based on the Federal Ministry of the Interior (Germany) report (BMI 1975). See Part II for more detail.

Groundwater quality score. Groundwater quality score was set subjectively within a range of 0-100, which was based on the regulations made by the local authority. The Chinese authorities, for instance, have set groundwater quality into five categories: category I-V (AQSIQ 1993). Category I is the best, and V is the worst. According to the Chinese regulation (AQSIQ 1993), groundwater used as a source of drinking water should have a quality not worse than category III. Classification method of these categories and the values of the associated indicators was shown in Appendix 2. The categories are assessed based on the comprehensive rating value F (see Appendix 2), which has a range of 0-4.25 for category I-III. A lower F value indicates better groundwater quality. If the F values are available, it can be rescaled using the equation

$$new\ score = \frac{new_{max} - new_{min}}{old_{max} - old_{min}} * (old\ score - old_{min}) + new_{min} \quad (12)$$

or

$$\text{Groundwater quality score} = \frac{100-0}{0-4.25} * (F - 4.25) + 0. \quad (13)$$

However, if the F values are not available, 100 could be set for the best water quality—“category I”, 50 for “category II”, and 0 for “category III”.

3. Assessment of the Utility Function

The RBF site evaluation problem was already structured in the previous section; based on those objectives and attributes, a model could be build using multi-attribute utility theory (MAUT). This section will be a straightforward application of MAUT; more details on the theory can be found in one of the standard texts written by Keeney and Raiffa (1993).

The developing of MAUT in evaluating the RBF sites consisted of five steps: investigation of the qualitative preference structure, assessment of component utility function, assessment of the scaling constants, evaluation of suitable sites using utility function, and identification of the most suitable site. An illustration of the assessment has been made by Keeney and Wood (1977). In this part, the author focused on the first three steps, while the latter two steps were considered as results and presented in Part IV. Consistency checks were made throughout the assessment process. It is a rather difficult task to evaluate utility functions, which requires a considerable amount of interaction between the analyst and the decision maker (Keeney and Wood 1977). This process was simplified in this study; the author played as the role of decision maker as well as the analyst, and his preferences were given throughout the study. For consistency, the term—“decision maker” was used henceforth, but the reader should be aware that he/she referred to the author.

In order to appropriately structure the RBF siting problem, a simple model has already been established in the previous section (Figure 21) and attributes X_1, X_2, \dots, X_5 were identified. Then, the remaining task is to assess an object function or utility function $u(\mathbf{x}) = u(x_1, x_2, \dots, x_5)$ over the 5 attributes ($u(\mathbf{x})$ is the value of utility u at \mathbf{x} , which includes each multi-attribute consequence or in our case, each RBF site condition) (Keeney 1974). According to Keeney and Raiffa (Keeney and Raiffa 1993), one possible way is to find an appropriate function, call it f , with a simple form such that

$$u(x_1, x_2, \dots, x_5) = f[u_1(x_1), u_2(x_2), \dots, u_5(x_5)], \quad (14)$$

where u_i designates a utility function over the single attribute X_i .

According to Keeney and Raiffa (1993), if there are two probability distributions A and B over multi attribute consequence $\tilde{\mathbf{x}}$ or in our case, two RBF sites with condition A and B , *site A is at least as desirable as B if and only if*

$$E_A[u(\tilde{\mathbf{x}})] \geq E_B[u(\tilde{\mathbf{x}})], \quad (15)$$

where E_A and E_B are the expectation operators taken with respect to distribution measures A and B , respectively. In other words, we just assert that expected utility is the appropriate criterion to use in choosing among alternative RBF sites. According to the von Neumann-Morgenstern expected utility theory (von Neumann and Morgenstern 1953) (as cited in Fishburn (1967), Keeney and Raiffa (1993)), for *alternative site A and B*, \mathbf{x}^A is at least as desirable as \mathbf{x}^B if and only if

$$u(\mathbf{x}^A) \geq u(\mathbf{x}^B), \quad (16)$$

which can also be written as the property:

$$\mathbf{x}^A \succcurlyeq \mathbf{x}^B \Leftrightarrow u(\mathbf{x}^A) \geq u(\mathbf{x}^B), \quad (17)$$

where \succcurlyeq reads “is at least as desirable as”; therefore, the performance of a RBF site can be ranked based on their expected utilities.

3.1 Investigation of the qualitative preference structure

The qualitative structure is important for multi-attribute (more than two) problems, and it should be investigated before assessing a utility function. It indicates the appropriateness of the actual utility function being assessed (Keeney and Wood 1977). By doing this, assumptions such as *preferential* and *utility independence* are investigated (Keeney 1974).

By definition (Keeney and Raiffa 1993), if the value trade-offs between the pair of attributes X and Y do not depend on the attribute Z , we say the pair of attribute X and Y is preferentially independent of Z .

To check whether the attributes pair $\{X_1, X_2\}$ was preferentially independent of the other attributes, the trade-offs between X_1 (transmissivity) and X_2 (well field location score) were considered. An interactive way to investigate preferential independence is by conducting a questionnaire between the analyst and the decision maker. In Appendix 6, an example that adapted after Keeney and Raiffa (1993) can be found (see Interview A and Questionnaire A). The result of the questionnaire indicated preferential independence. Other techniques can also be used to check the preferential independence, one of which is well explained in Keeney and Raiffa (1993, 96-100). The author adapted this by asking the decision maker, “If I fix $\{X_3, X_4, X_5\}$ at their best levels, which (x_1, x_2) pair would you prefer, $(0.001, 100)$ or $(0.35, 0)$? In other words, if you were at $(0.001, 0)$ would you rather push X_1 up to its limit of 0.35 or X_2 up to its limit of 100?” The answer was $(0.35, 0)$, which means for the decision maker, the X_1 attribute is more critical. Then, the author followed

by asking the decision maker, “OK then. Let $\{X_3, X_4, X_5\}$ still fixed at their best levels, give me a value x_1 such that you are *indifferent* between $(x_1, 0)$ and $(0.001, 100)$. In other words, I’m asking you to consider the following. Imagine that you are at $(0, 0)$. How much would you have to push X_1 up to be equivalent to X_2 going from 0 to 100?” The decision maker then gave a rough number of 0.25, which means he/she is indifferent between the pair $(0.001, 100)$ and $(0.25, 0)$, in other words, they yield the same utility for the decision maker. Then, the author asked the decision maker, “If I change $\{X_3, X_4, X_5\}$ to their worst levels or any other levels, do you still feel indifferent between the pair $(0.001, 100)$ and $(0.25, 0)$?” The answer was “Yes.” For other specific (x_1, x_2) pairs the same condition was verified. Now it seemed appropriate to assume that $\{X_1, X_2\}$ was preferentially independent to other attributes, *i.e.*, $\{X_3, X_4, X_5\}$. Similar validation was made for any other pair of attributes, it turned out they were also preferentially independent of other three attributes. As an illustration, in Figure 22, the decision maker was indifferent between $x' = (x_1 = 0.001, x_2 = 100)$ and $x'' = (x_1 = 0.25, x_2 = 0)$ no matter at which levels other attributes were fixed.

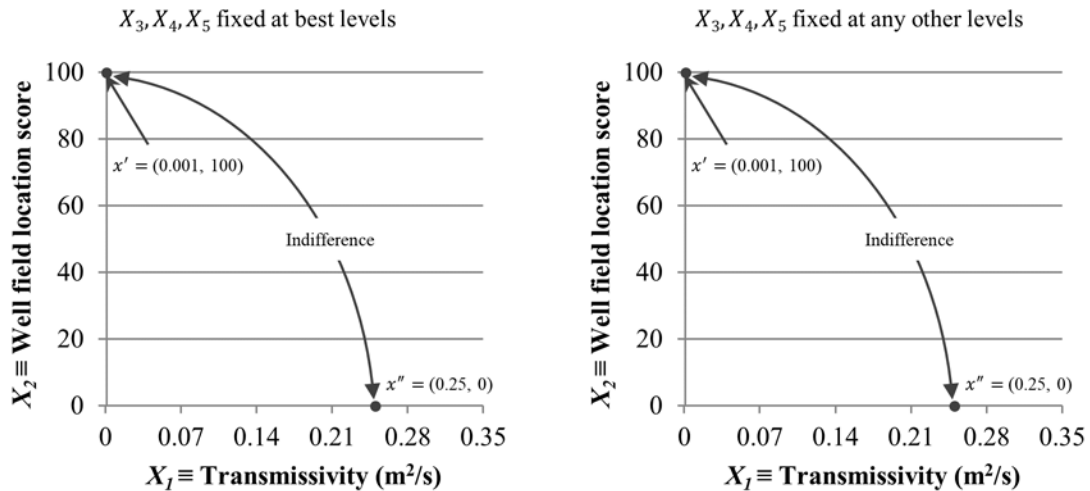


Figure 22: Validating the preferential independence

Utility independence was investigated using von Neumann-Morgenstern lotteries (von Neumann and Morgenstern 1953). According to Keeney and Wood (1977) it can be assessed by asking the decision maker to assign a specific value \hat{x}_i (for attribute X_i) such that it was indifferent as to whether it received the value \hat{x}_i for certain, or the lottery L that yielded the best consequence x_i^* with probability 0.5 or the worst consequence x_i^0 with probability 0.5.

The utility function for this lottery L can be written as

$$u(\hat{x}_i) = E[u(\tilde{x}_i)] = 0.5u(x_i^*) + 0.5u(x_i^0), \quad (18)$$

where \tilde{x} are the uncertain consequences of the lottery L . If we set $u(x_i^*) = 1$ and $u(x_i^0) = 0$, $u(\hat{x}_i) = 0.5$.

Alternatively, as Dyer and Miles (1974) assessed in their study of the trajectory selection for the Mariner (Voyager) Jupiter-Saturn 1977 project, the decision maker can also be required to assign a specific probability number p_i (for attribute X_i) such that it was indifferent as to whether it received the value \hat{x}_i' for certain, or the lottery L that yielded the best (most preferred) consequence x_i^* with probability p_i or the worst (least preferred) consequence x_i^0 with probability $1 - p_i$, where \hat{x}_i' can be any value between x_i^0 and x_i^* .

In this study, the assessment used was derived from Keeney and Wood (1977). Taking X_1 as an example, the procedure to check the utility independence between X_1 and other attributes was performed by asking the decision maker to consider a lottery or some lotteries. Firstly, the author fixed the attributes other than X_1 and asked the decision maker for the value of \hat{x}_1 such that \hat{x}_1 is, for sure, indifferent to a 50-50 chance that yielded $x_1 = 0.001$ or $x_1 = 0.35$. The response was $\hat{x}_1 = 0.15$. This is referred to as the *certainty equivalent* for the lottery yielding either $x_1 = 0.001$, with probability 0.5, or $x_1 = 0.35$, with probability 0.5 (Figure 23, left) (see “certainty equivalent” (Keeney and Raiffa 1993, 142)).

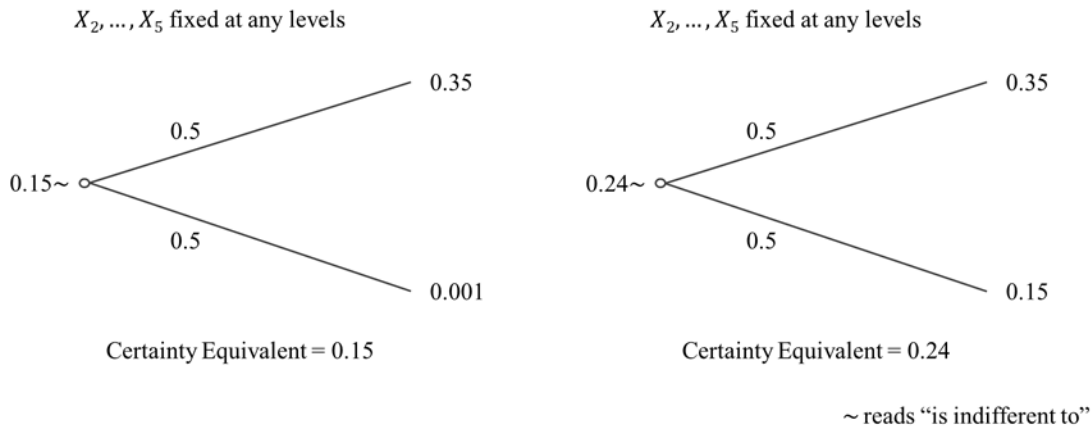


Figure 23: Validating the utility independence and certainty equivalents for X_1

Secondly, the author found that, for the decision maker, the certainty equivalent $\hat{x}_1 = 0.15$ did not change when only the levels of the other attributes X_2, \dots, X_5 were varied. The author can also continue to ask the decision maker to consider more lotteries for consistency, for instance, the certainty equivalent for the 50-50 lottery yielding either 0.15 or 0.35 was assessed to be 0.24 (Figure 23, right), and this also did not depend on the levels of attributes other than X_1 . Now, the author felt more certain to assume that X_1 was

utility independent of $\{X_2, \dots, X_5\}$. Similar assessment was made to check the utility independence of other attributes, which confirmed their utility independences.

After validating the preferential independence and utility independence, the *additive independence* was required to be checked in order to choose an appropriate utility function f over all attributes. Two utility functions—*additive* utility function and *multiplicative* utility function—were developed during the mid-1960s and early 1970s, and they were widely accepted in solving multi attribute problems (Fishburn 1967; Pollak 1967; Keeney 1974). Their adapted forms for this study were illustrated in equation (19) and (20) (Keeney and Raiffa 1993).

Additive utility function:

$$u(x_1, x_2, \dots, x_5) = \sum_{i=1}^5 k_i u_i(x_i) \quad (19)$$

Multiplicative utility function:

$$1 + ku(x_1, x_2, \dots, x_5) = \prod_{i=1}^5 [1 + k k_i u_i(x_i)] \quad (20)$$

where:

1. u is normalized by $u(x_1^0, x_2^0, \dots, x_5^0) = 0$ and $u(x_1^*, x_2^*, \dots, x_5^*) = 1$, with $(x_1^0, x_2^0, \dots, x_5^0)$ being the worst consequence and $(x_1^*, x_2^*, \dots, x_5^*)$ the best.
2. $u_i(x_i)$ is a conditional/component utility function on X_i normalized by $u_i(x_i^0) = 0$ and $u_i(x_i^*) = 1$, with x_i^0 being the worst consequence and x_i^* the best, $i = 1, 2, \dots, 5$.
3. $k_i = u(x_i^*, \bar{x}_i^0)$, with \bar{x}_i^0 being the complement of x_i^0 . For example, $k_1 = u(x_1^*, x_2^0, \dots, x_5^0)$.
4. $k > -1$ is a nonzero scaling constant that is a solution to $1 + k = \prod_{i=1}^5 (1 + k k_i)$.

According to Keeney and Wood (1977), the appropriate utility function can be chosen by asking the decision maker on his/her preference or indifference on the flowing lotteries A and B :

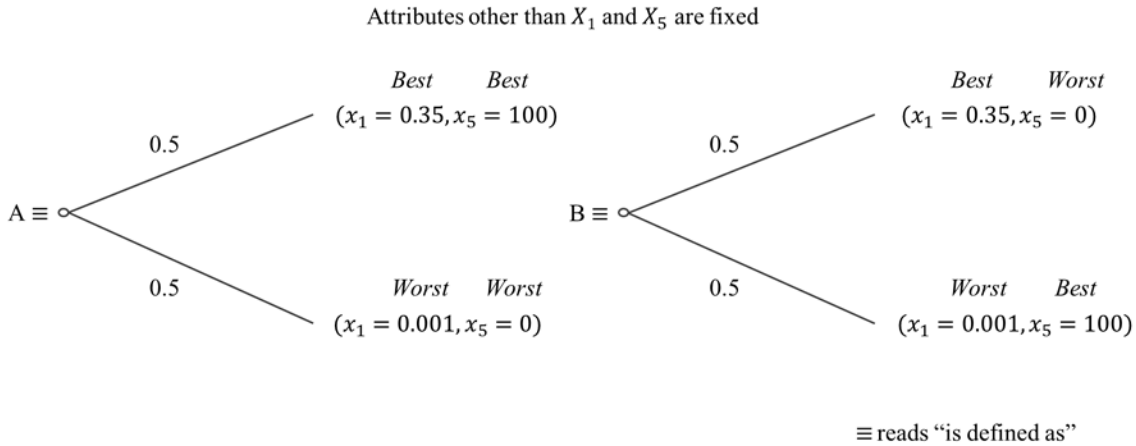


Figure 24: Check the additive independence

Note: $X_1 \equiv$ Transmissivity (m^2/s), $X_5 \equiv$ Groundwater quality score.

Together with the preferential independence and utility independence conditions, if lottery A and B is indifferent to the decision maker, the additive utility function is appropriate for the case (theoretically, all pairs of attributes should be checked); if either lottery A or B is preferred for the decision maker, the multiplicative utility function is appropriate (Keeney and Wood 1977). The response was that the decision maker had a preference for lottery B ; therefore, the multiplicative utility function was chosen in this study.

3.2 Assessment of component utility function

A component utility function is a one-attribute utility function. In our case, there were five component utility functions. For assessing these utility functions, the author simplified the process by assessing just one certainty equivalent for each attribute (see previous step), the same assessment has been used in the work on water resource planning by Keeney and Wood (1977). An exponential utility function for attribute X_1 was fit to the assessed points, *i.e.*, $(x_1 = 0.001, u_1 = 0)$, $(x_1 = 0.15, u_1 = 0.5)$, and $(x_1 = 0.35, u_1 = 1)$ (Figure 25, right). Exponential utility functions were assessed for other attributes as well by using the same technique (Figure 26).

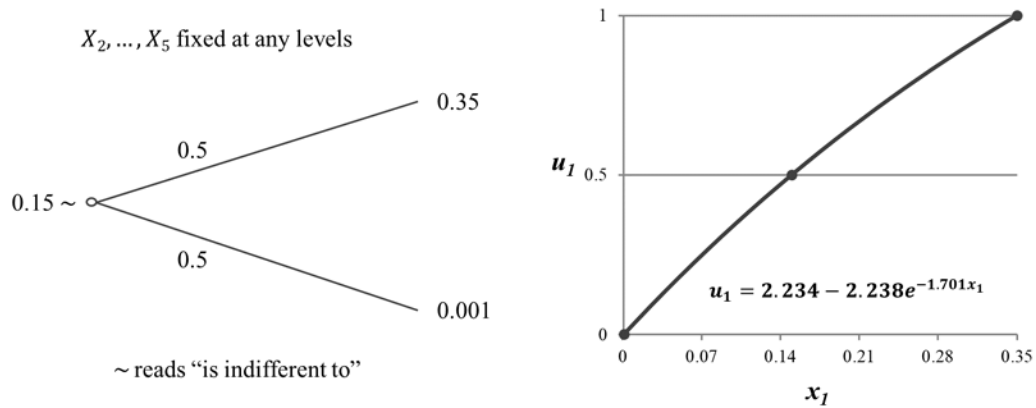


Figure 25: An illustration of assessing a component utility function for X_1

3.3 Assessment of the scaling constants

After all component utility functions were specified, the next step was to assess the scaling constants k_i and k in equation (20) in order to solve the multiplicative utility function. k_i were first ordered based on the preference of the decision maker. As in the process applied in the work of Keeney and Raiffa (1993) and Keeney and Wood (1977), all five attributes in Table 14 were set at their worst (least preferred) consequences or levels, the decision maker was then asked, “If only one attribute among those five could be raised to its best level, which one would be preferred?” The response was attribute X_4 (mean DOC concentration). According to Keeney and Wood (Keeney and Wood 1977), this indicated that k_4 must be the largest of the k_i , and in our case, it had the most influence on the final choice of a RBF site; it was not correct, however, to conclude that X_4 is more important than other attributes. This “common misinterpretation” was discussed in detail by Keeney and Raiffa in their text (Keeney and Raiffa 1993, 271-273). If it was indifferent for the decision maker to move either X_i or X_j to its best level, then $k_i = k_j$. After several confirmations were made, the result was that

$$k_4 > k_1 > k_5 > k_2 = k_3. \quad (21)$$

To evaluate the relative scaling constants, the multiplicative utility function was used several times. The method used here was derived from Keeney and Wood (1977) by considering the trade-offs between two attributes at a time. This is, however, a rather difficult task for the decision maker. One possible and less painful process can be used by thinking about the trade-offs between X_4 (the attribute with the highest scaling constant) and X_i ; this is particularly useful when encountering more attributes. In practice, we needed

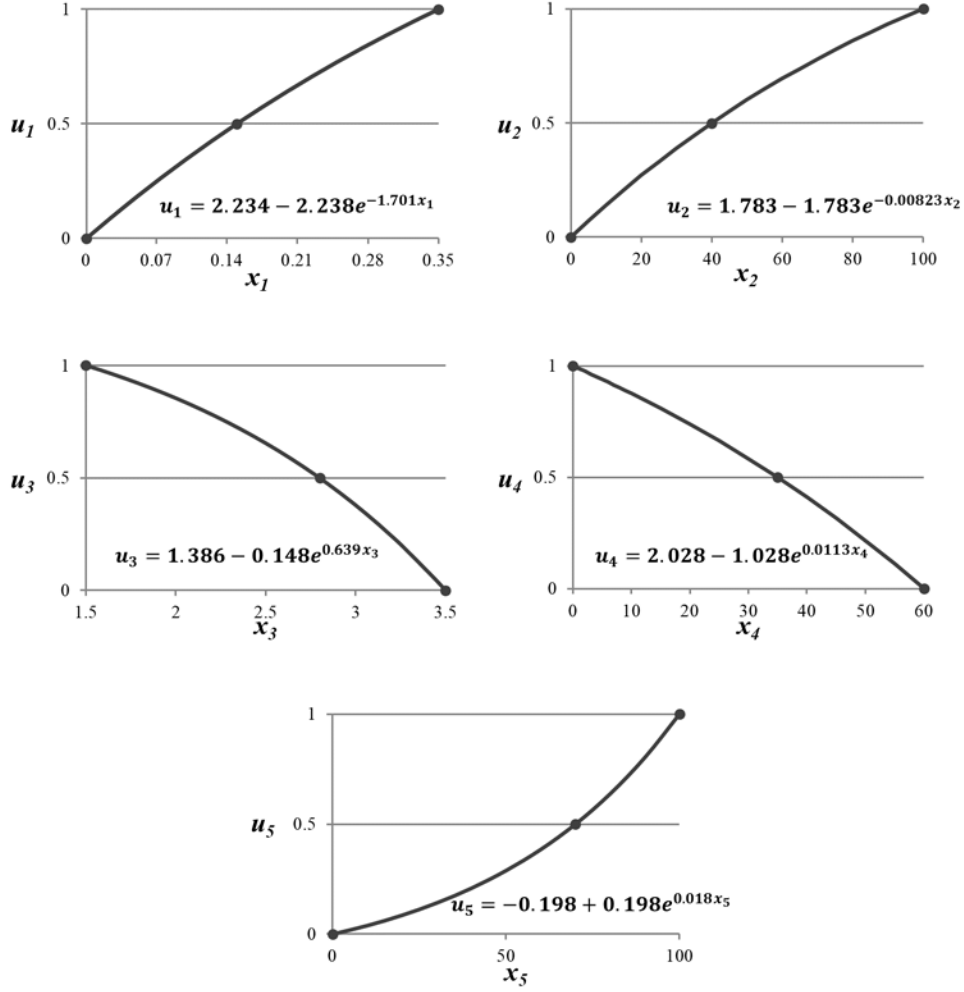


Figure 26: Component utility functions

to acquire some indifferent pairs of (x_4, x_i) , while attributes other than X_4 and X_i are fixed at any levels, $i = 1, 2, 3, 5$. For example, on the one hand, the author fixed X_4 at its worst level and X_1 at its best level; on the other hand, the author fixed X_1 at its worst level and left X_4 empty. Then, the author asked the decision maker, “At which level of X_4 are the two pairs on each hand indifferent”. The response was $x_4 = 10$, so the indifferent pair was $(x_4 = 10, x_1 = 0.001)$ and $(x_4 = 60, x_1 = 0.35)$, which indicated that

$$u(x_1 = 0.001, x_2^0, x_3^0, x_4 = 10, x_5^0) = u(x_1 = 0.35, x_2^0, x_3^0, x_4 = 60, x_5^0). \quad (22)$$

From equation (20), we know that $k_i = u(x_i^*, \bar{x}_i^0)$. We also have set $u_i(x_i^0) = 0$, and $u_i(x_i^*) = 1$, $i = 1, 2, \dots, 5$, so by fixing X_2, X_3, X_5 at their worst levels, and using the

multiplicative utility function (20) and equation (22), the following relationship between k_1 and k_4 was obtained

$$[kk_4u_4(10) - 1]/k = (k_1 - 1)/k \text{ or } k_1 = k_4u_4(10). \quad (23)$$

$u_4(x_4)$ was solved before (see Figure 26); thus,

$$k_1 = 0.877k_4. \quad (24)$$

We can acquire another three indifferent pairs by doing this through all attributes; however, remember that we have already considered the trade-offs between X_1 and X_2 , and got one indifferent pair— $(x_1 = 0.001, x_2 = 100)$ and $(x_1 = 0.25, x_2 = 0)$ —from the previous section. Fixing X_3, X_4, X_5 at their worst levels, the relative importance between X_1 and X_2 can be solved as

$$k_2 = k_1u_1(0.25) = 0.771k_1. \quad (25)$$

Furthermore, for the decision maker, X_2 is equally important to X_3 , since $k_2 = k_3$ (see equation 21); thus, to acquire all relative scaling constants, one more indifferent pair between X_4 and X_5 is needed. All indifferent pairs and the relative scaling constants are shown in Table 15. For consistency, the scaling constants k_i should have the same order as equation 21.

Table 15: Indifferent pairs and their implied relative scaling constants

Indifferent pair	Relative scaling constants
$(x_4 = 10, x_1 = 0.001) \sim (x_4 = 60, x_1 = 0.35)$	$k_1 = 0.877k_4$
$(x_1 = 0.001, x_2 = 100) \sim (x_1 = 0.25, x_2 = 0)$	$k_2 = 0.771k_1$
-	$k_3 = k_2$
$(x_4 = 20, x_5 = 0) \sim (x_4 = 60, x_5 = 100)$	$k_5 = 0.739k_4$

The above four equations have five unknowns. One more equation is needed to acquire the specific values of these scaling constants. Keeney and Raiffa (1993) used “probabilistic scaling” to acquire another function, the decision maker was required to give a probability of p so that he/she was indifferent to the following lottery, for example (Figure 27).

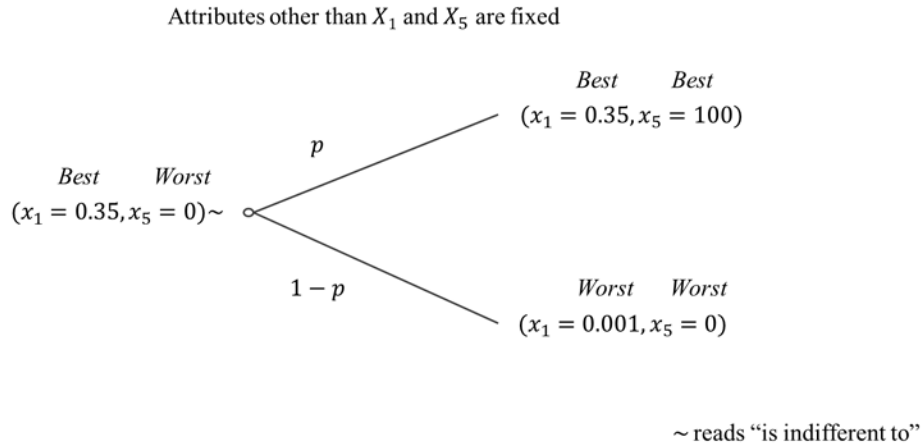


Figure 27: A lottery that generates the fifth equation

For the decision maker, the consequence $(x_2 = 0.35, x_5 = 0)$ was indifferent to a lottery yielding either $(x_2 = 0.35, x_5 = 100)$ with a probability of $p = 0.3$ or $(x_2 = 0.001, x_5 = 0)$ with a probability of $1 - p = 0.7$. Fixing other attributes at their worst levels and using the multiplicative utility function (20) again, an extra (the fifth) function was obtained

$$k_1 = 0.3(k_1 + k_5 + k k_1 k_5). \quad (26)$$

The probability p given by the decision maker should be consistent with the relative importance of scaling constants in Table 15, which were given by him-/herself. According to Keeney and Raiffa (1993, 347), the scaling constants should fulfill

$$1 + k k_i > 0. \quad (27)$$

Combining the equations in Table 15 and the equation obtained from the lottery in Figure 27:

$$k_1 = p(k_1 + k_5 + k k_1 k_5), \quad (28)$$

we could obtain

$$0.877k_4 = p(0.877k_4 + 0.739k_4 + k * 0.877 * 0.739k_4^2), \quad (29)$$

or

$$k k_4 = \frac{(1 - p) * 0.877 - p * 0.739}{p * 0.877 * 0.739}, \quad (30)$$

which should be greater than -1 (according to equation 27). Then, p should fulfill

$$p < 0.3873. \quad (31)$$

This was consistent with the probability of the lottery (Figure 27) given by the decision maker.

As listed directly after additive utility function (19) and multiplicative utility function (20), scaling constant k is a solution to

$$1 + k = \prod_{i=1}^5 (1 + k k_i), \quad (32)$$

which can be obtained if all attributes were set at their best levels and then using the multiplicative utility function (20). Now all scaling constants can be solved using equations (26) and (32) together with the previous four equations; they yielded

$$\begin{array}{lll} k_1 = 0.0153 & k_2 = 0.0118 & k_3 = 0.0118 \\ k_4 = 0.0175 & k_5 = 0.0129 & \end{array} \quad (33)$$

and

$$k = 115.26. \quad (34)$$

4. Results

A summary of the scaling constants is shown in Table 16. Scaling constants (33) and (34), together with the component utility functions in Figure 26, the multiplicative utility function (20) can now be specified. Finally, the utilities of each suitable RBF site were evaluated under their specific site condition, in other words, based on their attribute values, an example was given in Part IV.

Table 16: Component utility functions and scaling constants

Attribute	Scale of measure	Range	Utility function	Ranking of scaling constant	Indifference equivalent	Relative scaling constant	Scaling constant
$X_1 \equiv$ Transmissivity	m ² /s	0.001-0.35	$u_1 = 2.234 - 2.238e^{-1.701x_1}$	2	$(x_4 = 10, x_1 = 0.001)$ $\sim(x_4 = 60, x_1 = 0.35)$	$k_1 = 0.877k_4$	0.0153
$X_2 \equiv$ Well field location score	Subjective	0-100	$u_2 = 1.783 - 1.783e^{-0.00823x_2}$	4	$(x_1 = 0.001, x_2 = 100)$ $\sim(x_1 = 0.25, x_2 = 0)$	$k_2 = 0.771k_1$	0.0118
$X_3 \equiv d_{90}/d_m$ ratio of streambed material	Ratio scale	1.5-3.5	$u_3 = 1.386 - 0.148e^{0.639x_3}$	4	-	$k_3 = k_2$	0.0118
$X_4 \equiv$ Mean DOC concentration	mg/L	0-60	$u_4 = 2.028 - 1.028e^{0.0113x_4}$	1	-	k_4	0.0175
$X_5 \equiv$ Groundwater quality score	Subjective	0-100	$u_5 = -0.198 + 0.198e^{0.018x_5}$	3	$(x_4 = 20, x_5 = 0)$ $\sim(x_4 = 60, x_5 = 100)$	$k_5 = 0.739k_4$	0.0129

5. Discussion

To fully analyze the effects of the attribute uncertainties on our model output, the Latin Hypercube Sampling (LHS) method was performed to provide stratified values of input attributes, which was first introduced by McKay *et al.* (1979). When using LHS, the range of each attribute is divided into N intervals of equal marginal probability $1/N$, and each interval was sampled once (McKay, Beckman, and Conover 1979; Iman, Helton, and Campbell 1981). LHS has the advantage of using fewer samples than simple random sampling to achieve the same accuracy and ensures each of the attribute is presented in a fully stratified manner, no matter which attributes might turn out to be important (McKay, Beckman, and Conover 1979).

The samples of transmissivity and mean DOC concentration of the surface water at RBF sites distribute rather lognormal (Figure 28) (Lenk *et al.* 2006). According to Lenk *et al.* (2006), the log-transformed transmissivity values at 30 existing RBF sites from Central Europe have a mean of -3.38 and a standard deviation of 1.07. The log-transformed mean DOC concentration values (from 31 existing RBF sites) have a mean of 1.37, and a standard deviation of 0.48. The means of the two attributes were kept unchanged, whereas the standard deviation was adjusted here in order to sample those values over their entire ranges from our MAUT model and to prevent oversampling in the outer ranges of the interval as well. The distribution of other attributes are not well known (particularly in the case of RBF); they were simply assumed to be normally distributed with a known mean (using median of the range) and standard deviation (20% of the range) (Table 17).

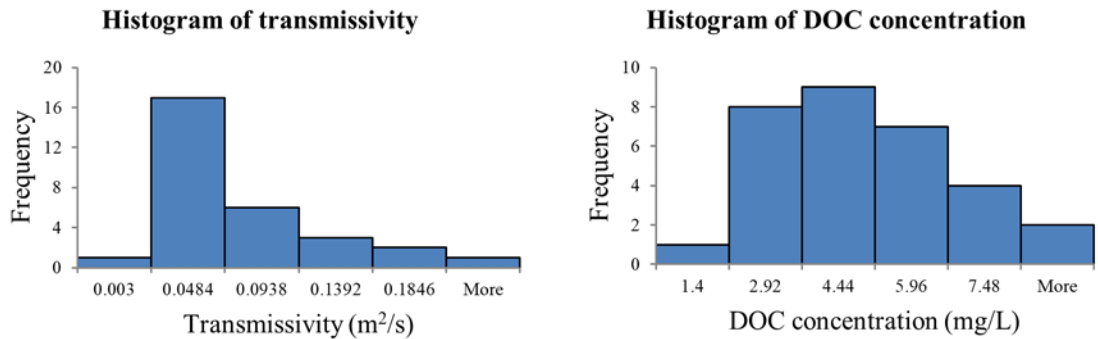


Figure 28: Distributions of attribute values at RBF sites

Note: most of the surface water DOC concentrations are mean values. Source: data from Lenk *et al.* (2006).

The LHS was performed using R-package “pse” (Chalom and Prado 2017). Two hundred sets of attributes/variables were generated and selected using the LHS technique

($N = 200$). The distribution of the results showed that the utility of simulated RBF sites was most likely between 0.05 and 0.2 (Figure 29 (f)).

Table 17: Assumed attribute distributions

Attribute	Scale of measure	Distribution	Arguments
$X_1 \equiv$ Transmissivity	m ² /s	Lognormal	$\mu = -3.38, \sigma = 0.9$
$X_2 \equiv$ Well field location score	Subjective	Normal	$\mu = 50, \sigma = 20$
$X_3 \equiv d_{90}/d_m$ ratio of streambed material	Ratio scale	Normal	$\mu = 2.5, \sigma = 0.4$
$X_4 \equiv$ Mean DOC concentration	mg/L	Lognormal	$\mu = 1.37, \sigma = 1.0$
$X_5 \equiv$ Groundwater quality score	Subjective	Normal	$\mu = 50, \sigma = 20$

Note: μ is mean value, σ is standard deviation.

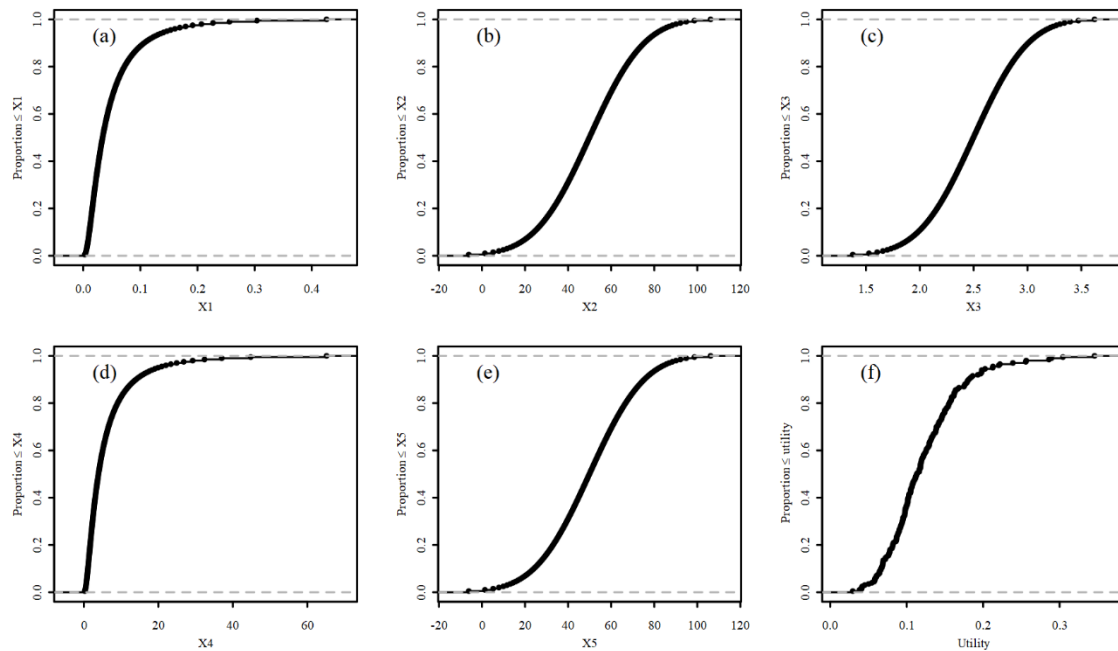


Figure 29: Empirical cumulative distributions

Note: $N = 200$. (a), (b), (c), (d), (e): attributes; (f): utility of simulated RBF sites.

The change in utility of simulated RBF sites (simply called “final utility” henceforth) over the change in individual attribute value is shown as five scatterplots (Figure 30), with respect to five individual attributes. As can be seen, most of the attribute values are within their individual ranges. For attribute values beyond the individual ranges, the component utilities were either set to 0 (the same as the lowest utility) or 1 (the highest utility). A linear model for each attribute was added to the scatterplot (black line), and the

final utility was found to either increase or decrease with attribute values. They are consistent

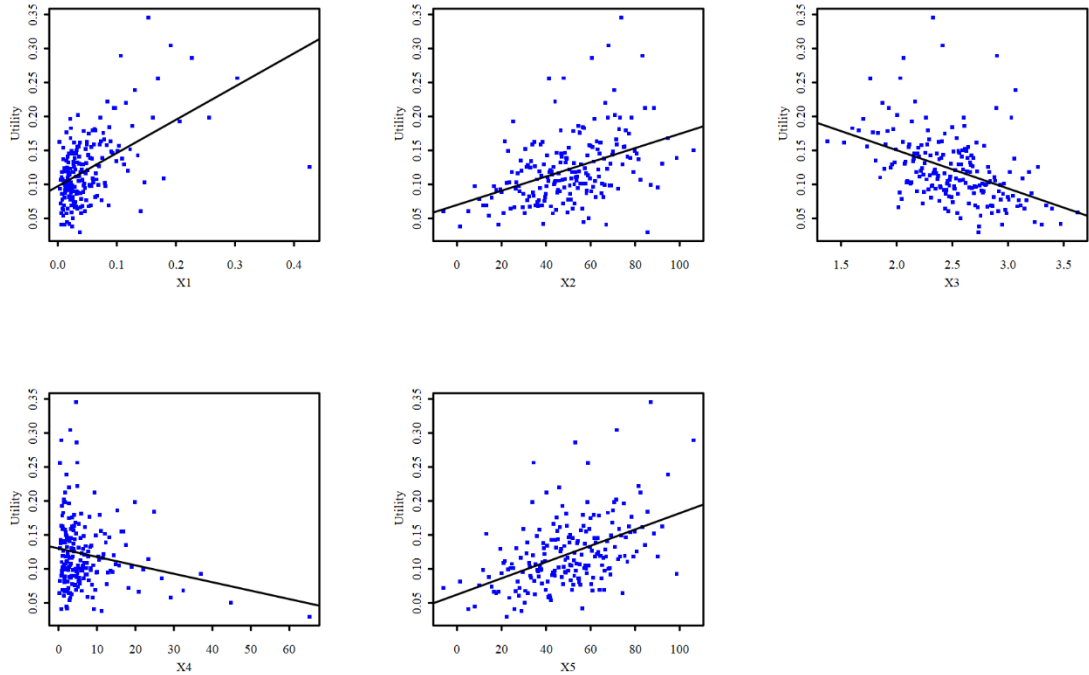


Figure 30: Influence of individual input variable/attribute on model output

with the decision maker’s preference because, for instance, higher values of transmissivity, well field location score, and groundwater quality score are preferred (higher utility), whereas lower values of d_{90}/d_m ratio and mean DOC concentration are preferred as well. Notice that transmissivity and mean DOC concentration values were more frequently sampled at their bottom ranges (ca. 0-0.15 m²/s and 0-20 mg/L, separately). To assess the sensitivity of our outcome variable (the final utility) to the variation of each attribute, an uncertainty analysis using partial rank correlation coefficient (PRCC) was proven to be useful (Marino *et al.* 2008). Partial rank correlation characterizes the linear relationship between an input and output (both are rank-transformed) after the linear effects on output of the remaining inputs are discounted (Marino *et al.* 2008).

A partial rank correlation was therefore used to evaluate the null hypothesis that there is no significant relation between final utility and individual input attribute after controlling for the effects of the four additional attributes ($N = 200$). For instance, there was strong, positive, fourth-order partial rank correlation between final utility ($\mu = 0.122$, $\sigma = 0.050$) and transmissivity (μ and σ are shown in Table 17), controlling for well field location score, d_{90}/d_m ratio, mean DOC concentration, and groundwater quality score, $r =$

0.79, $p = 0.000$. Results of the zero-order correlation yielded that there was strong, positive correlation between final utility and transmissivity, $r = 0.47$, $p = 0.000$, indicating that controlling for the other four attributes had a rather high effect on the strength of the relationship of between the two variables. Correlation and partial correlation analysis between final utility and the other three attributes (well field location score, d_{90}/d_m ratio, and groundwater quality score) yielded similar results, among which d_{90}/d_m ratio was negatively correlated with final utility. Results of the fourth-order partial rank correlation between final utility and mean DOC concentration yielded a moderately strong, negative, partial correlation, $r = -0.38$, $p = 0.000$. Results of the zero-order correlation showed that there was a weak, negative correlation between final utility and mean DOC concentration, $r = -0.12$, $p = 0.082$, indicating that controlling for the attributes other than mean DOC concentration had a rather high effect on the strength of the relationship between the two variables. The results of correlation and partial correlation for all attributes are shown in Table 18.

Table 18: Partial rank correlation for the relationship between final utility and attributes

		Transmissivity	Well field location score	d_{90}/d_m ratio	Mean DOC concentration	Groundwater quality score
Final utility	Zero-order correlation	0.47	0.44	-0.50	-0.12	0.46
	p-value	0.000	0.000	0.000	0.082	0.000
Final utility	Fourth-order correlation	0.79	0.79	-0.81	-0.38	0.82
	p-value	0.000	0.000	0.000	0.000	0.000

Note: fourth-order partial correlation was conducted using R-package “ppcor” (Kim 2015); zero-order correlation was conducted using R-package “stats v.3.6.2” (<https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/cor.test>).

The results provided evidence to reject the partial correlation null hypothesis because the significance levels for the fourth-order partial correlation were less than 0.05.

Replicating (or bootstrapping) the LHS 50 times, the original PRCCs (shown in Table 18) and the associated confidence intervals (95%) are shown in Figure 31. Based on these results, we can more confidently say that the final utility was rather less sensitive to mean DOC concentration than other attributes although mean DOC concentration was the most important attribute according to the decision maker’s preference. The mean DOC concentration and transmissivity were both more frequently sampled at their bottom ranges due to lognormal distribution; however, transmissivity turned out to be more important than mean DOC concentration, with respect to higher partial rank correlation. One possible reason is that the range of their component utilities varied differently within these particular attribute ranges. The decision maker was risk averse because the utility functions of

transmissivity and mean DOC concentration are concave (Figure 26). This indicated that the average slope of the transmissivity utility function within its bottom range (0-0.15 m²/s) is greater than that of mean DOC concentration (0-20 mg/L), which means a greater variability of the component utility (0-0.5 for transmissivity, 0.74-1 for mean DOC concentration). This further indicated that the influence of an individual attribute on final utility depends on the range of attribute values from the suitable/candidate sites. For instance, a mean DOC concentration range of 0-10 mg/L will have less influence on the component utility than a range of 50-60mg/L (Figure 32). If all attributes except mean DOC concentration are held constant, the model will yield fewer different final utilities for mean DOC concentration with ranges of 0-10 mg/L than 50-60 mg/L.

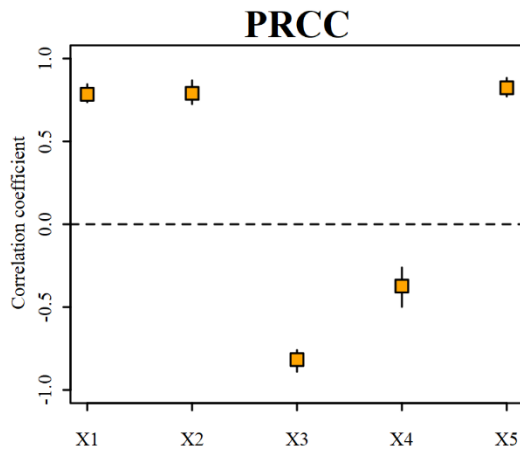


Figure 31: Partial rank correlation coefficients of five attributes

Note: orange box shows the original PRCC; bar shows 95% confidence interval, generated by bootstrapping 50 times.

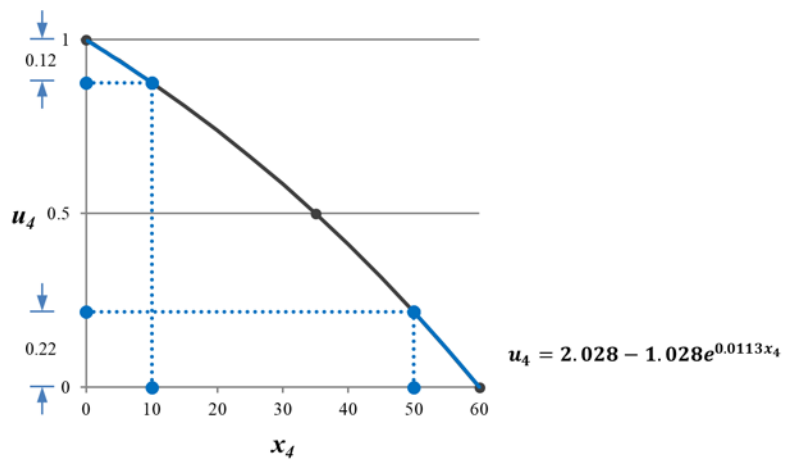


Figure 32: Comparing the utility changes of mean DOC concentration for range 0-10 mg/L and 50-60 mg/L

6. Summary

In this part, two major objectives and five attributes were set for the RBF site selection problem. A multiplicative utility function was assessed through investigation of the qualitative preference structure, assessment of component utility function, and assessment of the scaling constants. The decision maker's preferences were reflected throughout the analysis. To check the performance of the model over the entire range of the attribute values, a sensitivity analysis based on Latin Hypercube Sampling (LHS) and partial rank correlation coefficient (PRCC) was conducted. The result showed that the distribution of attribute values and the range of attributes from suitable RBF sites had an impact on the relative importance of the attribute.

Part IV

Case Study

1. Introduction

To illustrate how the model works as well as test its performance, a study area was selected in Jilin City along the Second Songhua River in China (Figure 33).

The Second Songhua River originates from the mountains on the border between China and North Korea, and meets the Nen River (Nenjiang) near Songyuan to form the Songhua River, a tributary of the Amur River (Heilongjiang). The Second Songhua River is a dam-regulated river; several dams such as Fengman Dam, Baishan Dam, and Hongshi Dam have altered the natural flow conditions of the river. About 20 km downstream of the Fengman Dam is located one of the largest cities in the region: Jilin. Figure 33 shows the most populated regions in Jilin City. Because of the dam regulation, seasonal fluctuation of flows is not obvious. The gauging station at Jilin City showed that the average monthly discharge varies between 300 m³/s and 700 m³/s, while the water level varies between 186.4 m above sea level and 186.8 m above sea level during the period 1979-2006 (Liu 2017). Dams have failed to protect the city from flooding: a 2010 flood caused over 100 dead or missing in Jilin Province and spilled chemicals into the Second Songhua River in Jilin City (Xinhua 2010).

Public water supply in Jilin City relies mostly on surface water intake from the Second Songhua River by four waterworks. Groundwater abstraction (capacity) equals 2.8% from all sources, which played a limited role in drinking water supply (Liu 2017). Due to heavy industry and its mismanaged effluent flowing into the river, water quality of drinking sources from four waterworks seldom meets the standards (see Table 19). To provide safe drinking water with low pollution risks, the city's decision makers are now searching for new alternative drinking water sources.

Table 19: Source water (surface water) quality conditions in Jilin City, 2018

Month	Waterworks I	Waterworks II	Waterworks III	Waterworks IV
January	Unsafe	Unsafe	Unsafe	Unsafe
February	Unsafe	Unsafe	Unsafe	Unsafe
March	Unsafe	Unsafe	Unsafe	Unsafe
April	Safe	Unsafe	Unsafe	Unsafe
May	Safe	Safe	Safe	Safe
June	Safe	Safe	Safe	Safe
July	Safe	Unsafe	Unsafe	Unsafe
August	Safe	Safe	Safe	Safe
September	Safe	Unsafe	Unsafe	Unsafe
October	Safe	Safe	Safe	Safe
November	Safe	Safe	Safe	Safe
December	Safe	Safe	Safe	Safe

Note: water quality conditions were evaluated according to Environmental Quality Standards for Surface Water (GB 3838-2002) (MEE and AQSIQ 2002). Source: MEE (2018a)

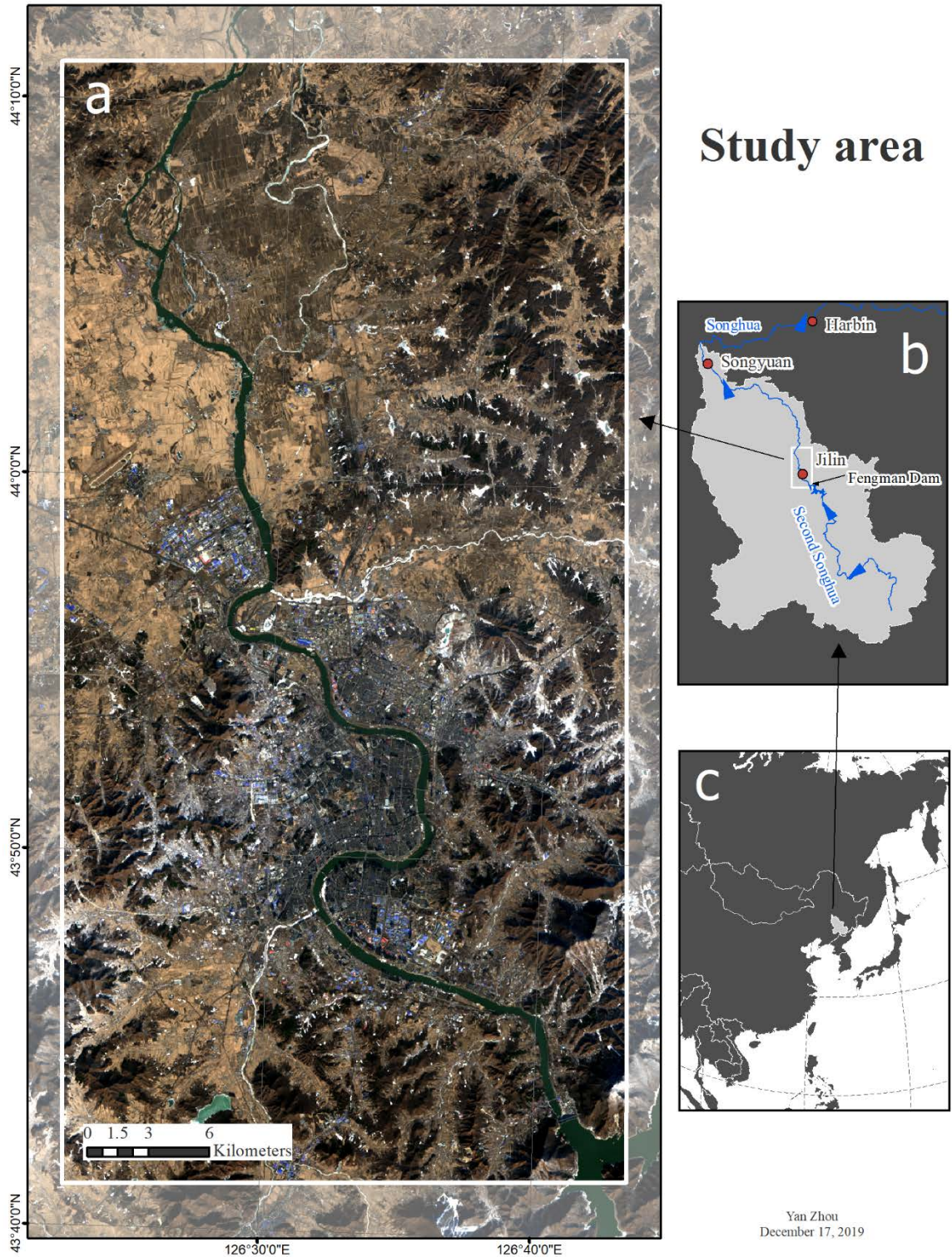


Figure 33: Study area

Note: (a) Natural color image of Jilin City area, projection: WGS_1984_UTM_Zone_52N. (b) The Second Songhua River and its drainage basin. Blue arrows show the flow direction. (c) Location of the drainage basin. Source: image was captured by Landsat 8 satellite on December 17, 2019 (<https://usgs.gov>); shapes are from Natural Earth (2019a, 2019b, 2019c); drainage basin was delineated from ~30 meter DEM data from NASA JPL (2000).

Drinking water sources along the Second Songhua River and the Songhua River were affected by the Jilin chemical plant explosion in Jilin City in November 13, 2005. The incident caused great public concern because a large population used surface water directly as their drinking water source. For example, in Songyuan and Harbin, drinking water from river water sources served populations of 1.2 million and 1.9 million in 2004, respectively (Peng, Su, and Liu 2006). Approximately 100 tons of benzene, nitrobenzene, and aniline were released into the Second Songhua River in this incident (MEE 2005). Water utilities had to be shut down after the incident; however, water from groundwater abstraction wells along the riverbank were considered to be safe based on a simulation by Peng *et al.* (2006). Contaminants (such as nitrobenzene) had been reduced to a safe concentration before reaching the pumping wells. A similar finding was reported after the Sandoz chemical spill in 1986 in Switzerland, from which the Rhine was heavily polluted; however, RBF systems along this river showed the capability of protection against the shock loads (Kühn and Müller 2000). In order to provide a safe drinking water source, search for better water management, and at the same time control the pollution discharge into the surface water, the Ministry of Ecology and Environment (MEE) has set up a research program on spreading the application of RBF in areas along the Songhua River (MEE 2013).

2. Materials and Methods

Unfortunately, the suitable areas/sites for RBF systems implementation in Jilin City have not been delineated; therefore, a GIS technique was used here to determine the suitable RBF sites. Using the multiplicative utility function developed in Part III, suitable RBF sites can be evaluated and ranked. A flowchart of methodology is shown in Figure 34.

2.1 GIS data collection

Some GIS data mentioned here was only used for developing constraint maps, such as geologic data, land cover data, and surface water area data, while groundwater quality data and aquifer properties data were used for both developing the constraint maps and characterizing the suitable sites. Surface water quality data and streambed material data offered insufficient but necessary information of each suitable site. Thematic maps were made using the GIS technique (found at end of this section), which consisted of constraint maps for screening the suitable RBF sites and also maps for assigning the attribute values at each site.

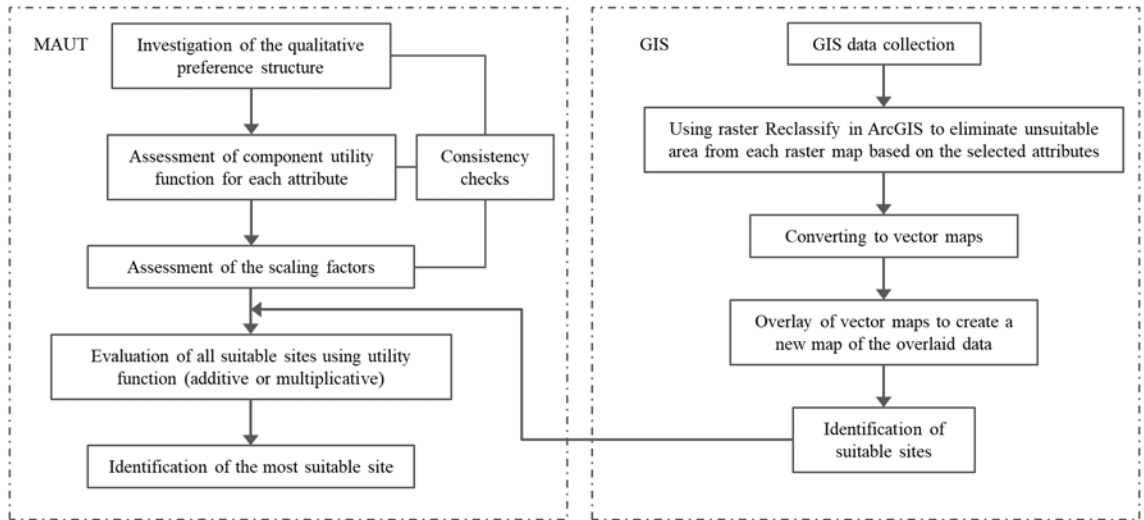


Figure 34: A flowchart of methodology for site selection and evaluation

2.1.1 Geologic data

Glacial drift formations played a limited role in the study area which is why fluvial deposits or alluvium of Quaternary age were considered suitable geologic formations for RBF systems. Ancient fluvial deposits older than those from the Quaternary age were not considered suitable—they are likely already or partly consolidated due to the lithification process (Boggs 2006).

The alluvium map (Figure 35) was obtained and modified from the Hydrogeologic Map of Jilin City (unpublished), at a scale of 1:200 000.

No fluvio-glacial deposits are found on the land surface. Downstream of the Fengman Dam, the width of the alluvium aquifer is rather small, varying between 1 km and 3 km. The landform is mainly mountain valley plains. As the streamflow of the Second Songhua River reaches the urban area or the first meander, the aquifer becomes wider. Two tributaries brought numerous debris from the southwest and east, which makes the aquifer more favorable for RBF. The northern part of the study area is known as the Yishu syncline (southwest-northeast direction) filled with hundreds of meters of Tertiary sediments, while on the surface, it is mostly covered by Quaternary alluvial and diluvial soft rock with a thickness of 10-50 m (Liao and Su 2007). Geologically, this part of the study area seems to be most suitable for RBF.

2.1.2 Land cover data

Land cover data was generated by Gong *et al.* (2019), who obtained the 10-m resolution global land cover map based on Sentinel-2 data. Land cover classifications of the study area are cropland, forest, grassland, shrubland, wetland, water, impervious surface, and bareland [*sic*].

Most of the urban area or artificial structures were classified as impervious surface (red color) in this study, which was considered unsuitable for RBF implementation (Figure

36). In the highly urbanized area, little space at the riverbank can be used for other purposes. Downstream of the urban area is dominated by agricultural use, and it seems that the undeveloped riverbank areas are suitable for RBF.

2.1.3 Groundwater quality data

A groundwater quality map (Figure 37) of the Second Songhua River basin was modified from the Groundwater environment map of Jilin Province that was provided by the National Geological Archives of China, at a scale of 1:1 800 000 (NGA 2013). The map classified groundwater quality into two categories in our study area: category IV and category lower than IV. More detailed classification is not available from the data source. These groundwater quality categories were based on the quality standard for groundwater (GB/T 14848-1993) (AQSIQ 1993). The map was first made available in 2013 from the archives; this is why this old standard was used as reference instead of the newly released standard in 2018 (GB/T 14848-2017). According to the old standard, groundwater of category I, II, and III are suitable as drinking water source, category IV can also be served as drinking water source after proper treatment. For safety concern, category V and worse than V shall not be used as drinking water source (AQSIQ 1993). Detailed information of quality category for groundwater is shown in Appendix 2.

Groundwater was, however, highly polluted in the northern part of the study area, probably due to the intensive agricultural activity and the pollution from the urban area.

2.1.4 Aquifer properties data

Hydraulic conductivity class map (Figure 38, see page 40-41 for how hydraulic conductivity class is defined) of the Jilin City area was modified from Qiu *et al.* (2015). In general, the fluvial deposits in the upper reaches are coarser than the lower reaches, and the hydraulic conductivity of which from the upper reaches is also higher. As expected, the hydraulic conductivity of the aquifer in the upper reaches was higher, presented as higher hydraulic conductivity class (smaller numbers). The saturated aquifer thickness map was acquired using kriging, the procedure of which will be explained in more detail in the next section.

The saturated thickness of the study area was based on 40 borehole records of groundwater table and bedrock elevation, which were collected from the Major Science and Technology Program for Water Pollution Control and Treatment during the 11th Five-Year Plan of China (2006-2010). The saturated aquifer thickness was defined as the difference of elevation between the groundwater table and bedrock. More detail source information is not available, for example, the measurement date of the water table.

2.1.5 Surface water area data

In order to perform the RBF site suitability analysis, information about the seasonal surface water area/boundary is necessary. A surface water data/map was acquired from the Global Surface Water database produced under the Copernicus Programme (Pekel *et al.* 2016). Surface water seasonality data is based on a single year observation value—from January 2018 to December 2018—and has a spatial resolution of 20 to 30 meters (see

Appendix 7, Figure A9). The data was used in this study for delineating surface water area at different water levels as well as generating the constraint maps of distance to the water body.

The Buffer tool from the Analysis toolset in ArcMap 10.6 (ESRI, 1999-2017) was used to determine two buffer areas of surface water body. Before generating the buffers, two definitions needed to be set. Based on the Global Surface Water database (Pekel *et al.* 2016), “moderate water level” water surface was defined in this work as the area that is underwater for more than 6 months, while the area that is underwater for at least 1 month was defined as “high water level” water surface. For waterworks, the share and quality of bank-filtrated surface water are influenced by “moderate water level” water surface for most of the time, whereas RBF wells should be placed outside of the flooded area when surface waters are at a “high water level”.

The historical flood data was not available in the study area; therefore, the unsuitable area was defined as the area between 60 m from the water body at a “high water level” and the water body itself. This buffer distance of 60 meters was delineated to prevent direct contamination from the surface water bodies. The exact distance was set, firstly, based on the regulation of Ministry of Ecology and Environment (MEE) (2018b). Drinking water abstraction wells should be placed at least 30 m away from source of pollution. Secondly, the distance was extended to 60 m because of the 20 m to 30 m spatial resolution of water bodies (Pekel *et al.* 2016). A buffer distance of 310 m away from the edge of the surface water body at “moderate water level” was defined as the area suitable for placing RBF wells. Wells placed in this area are likely to provide a high proportion of surface water. This distance was based on a database of RBF sites located in Central Europe, where the shares of bank-filtrated surface water in all RBF wells or monitoring wells are above 50% (Lenk *et al.* 2006). However, there is no guarantee that wells placed within this distance would provide a 50 percent (or higher) share of bank-filtrated surface water since this is determined by the specific conditions at each site.

The distance-to-surface-water-body analysis delineated some unsuitable and suitable areas for RBF systems close to the Second Songhua River (Figure 39 and Figure 40). The unsuitable areas used a 60-meter buffer based on water area at the high water level, while the suitable areas used a 310-meter buffer from the water area at the moderate water level.

2.1.6 Surface water quality data

DOC concentration was used here as the critical indicator of surface water quality; however, it has rather low quality. DOC concentration of surface water from the study area has been investigated by Wang *et al.* (2016). A total of four observations or samples were collected from a field campaign at Jilin City launched during September and November, 2014 (Figure 41). The lowest concentration was 10.2 mg/L, whereas the highest concentration was 29.6 mg/L.

2.1.7 Streambed material data

Locations of streambed material samples were published in You (2016); three sample locations are found in our study area (Figure 42). The grain size information of those samples were obtained from Dr. Du of Jilin University (2019) (Table 20).

Usually, fluvial transported sediments tend to distribute approximately lognormal if particles are expressed in millimeters, which means the sieve curve increases slowly with fine and large particles but rapidly with medium particle sizes (Bunte and Steven R. 2001). The result of sieve analysis (Figure 43) shows that, however, our samples distributed rather less lognormally. Their median grain sizes varied from 1.2 millimeters and 5.2 millimeters, which represented sediment from a sand-bed and gravel-bed, respectively. The average grain size of streambed material d_m was calculated using the formula given by Folk and Ward (1957):

$$d_m = \frac{d_{16} + d_{50} + d_{84}}{3}, \quad (35)$$

where d_{16} diameter indicates the sediment diameter/size for which 16 percent of the sediment sample is finer, and likewise for d_{50} , d_{84} .

The highest d_{90}/d_m ratio was 2.9, whereas the lowest ratio was 1.9. The sorting coefficient (S) of the samples indicated that they were very poorly sorted (coefficient between 2 and 4) (Folk and Ward 1957).

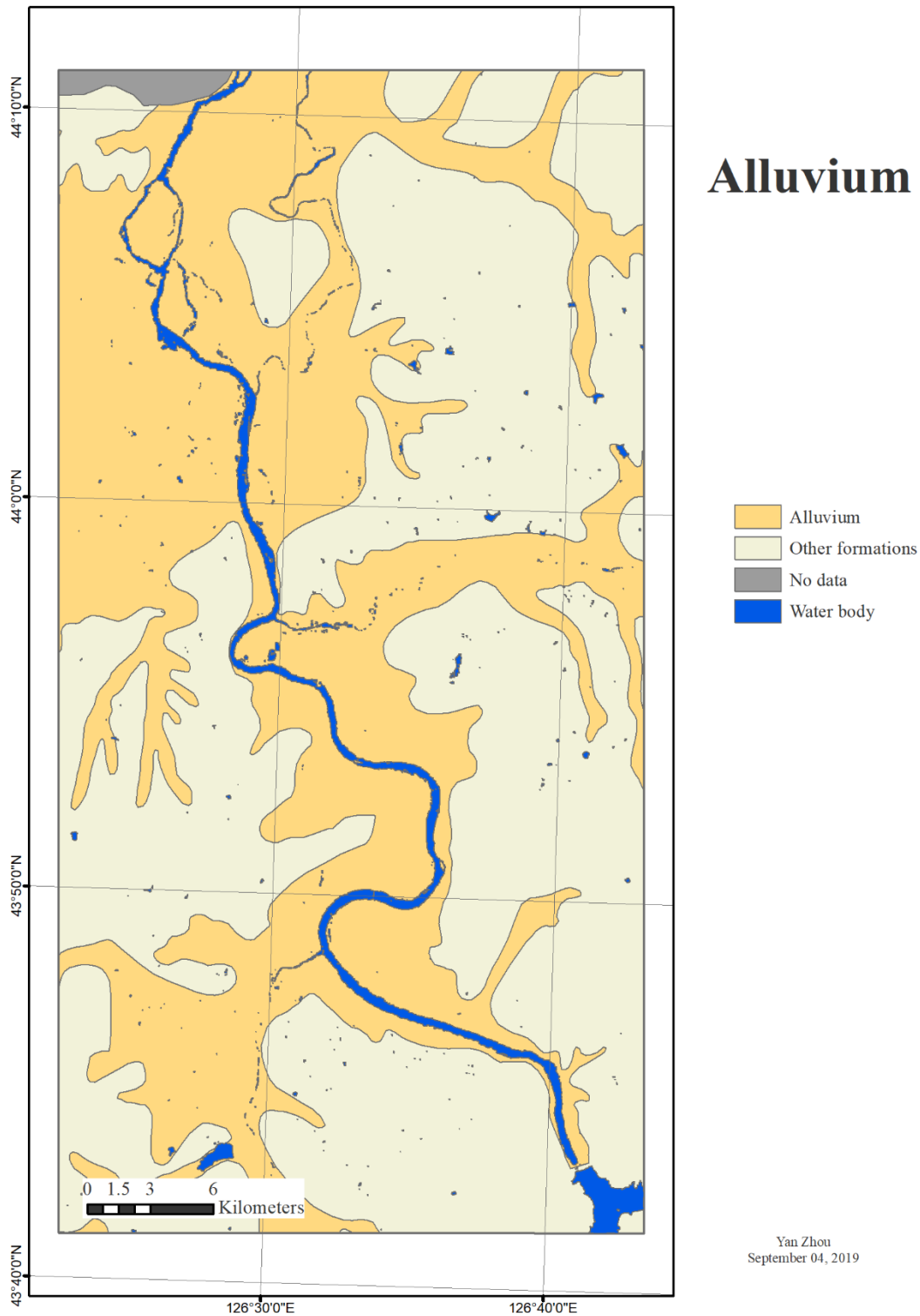


Figure 35: Alluvium map of Jilin City

Source: map modified from 1:200000 Hydrogeologic Map of Jilin City (unpublished); water body was from EC JRC/Google (2016). Projection: WGS_1984_UTM_Zone_52N.

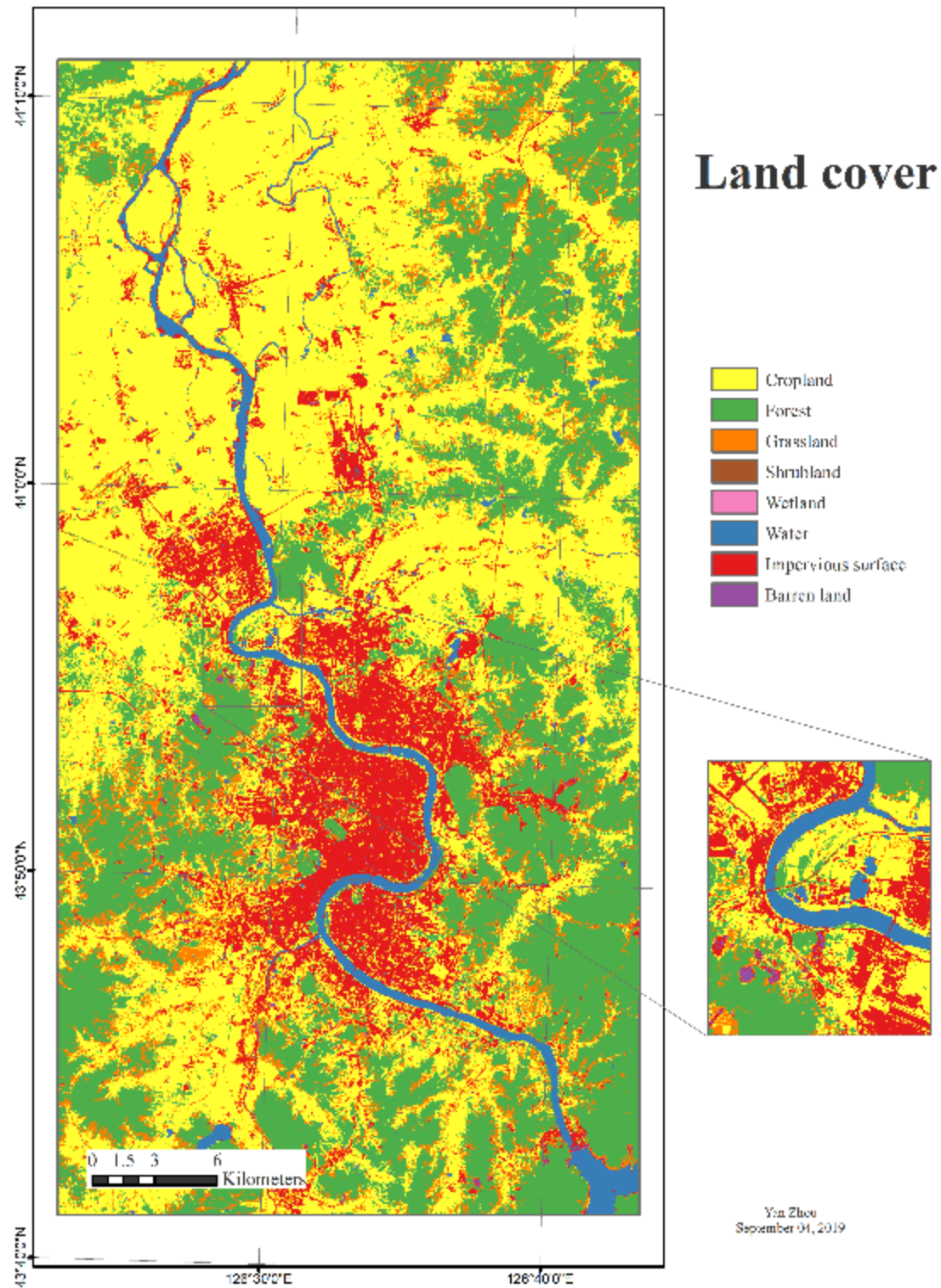


Figure 36: Land cover map of Jilin City

Source: map modified from Gong et al. (2019) (<http://data.ess.isinghua.edu.cn/>). Projection: WGS_1984_UTM_Zone_52N.

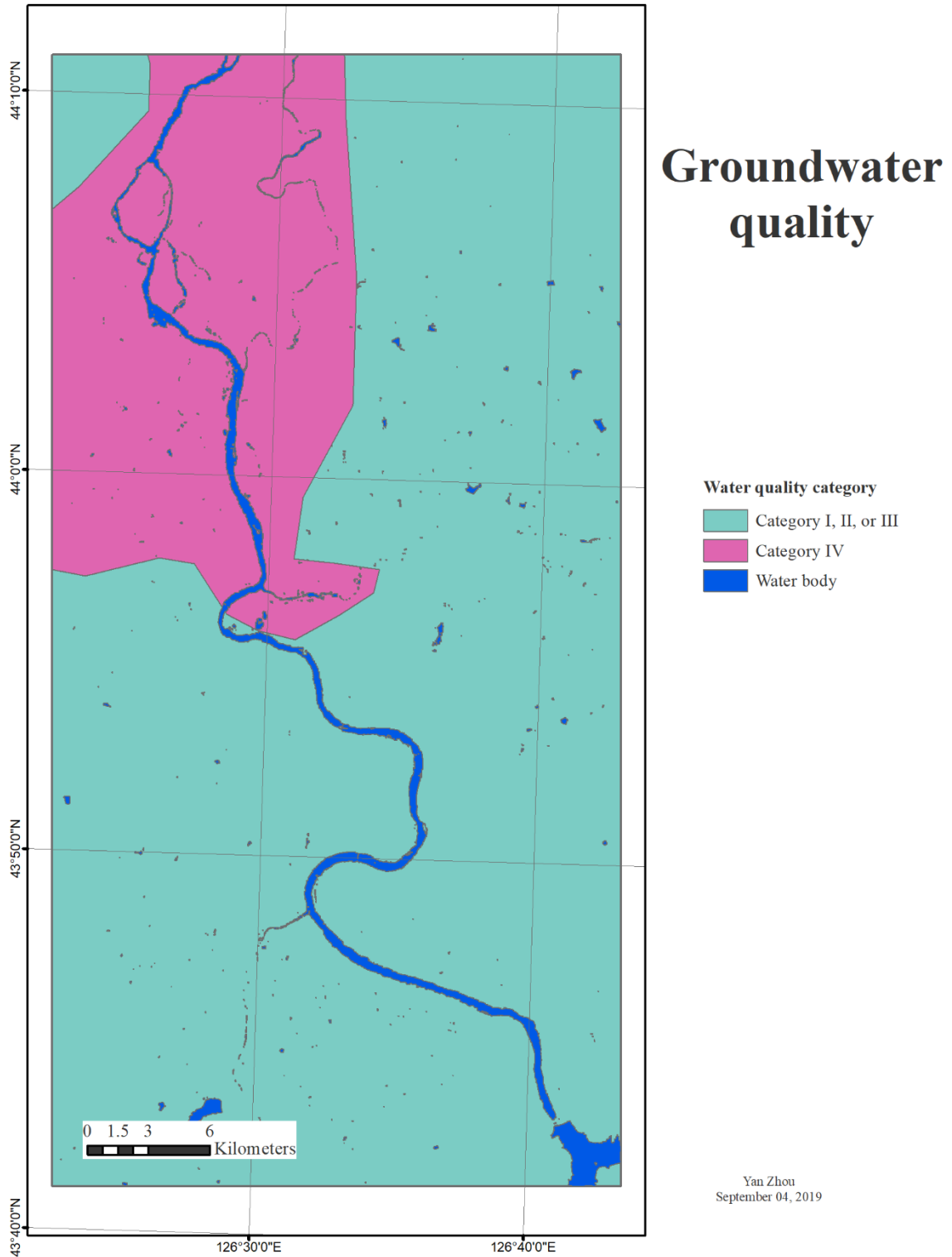


Figure 37: Groundwater quality map of Jilin City

Note: category I, II, and III are safe to be used as drinking water sources, category IV is unsafe. Source: map modified from 1:1800000 Groundwater Environment Map of Jilin Province (available from NGA (2013)); water body was from EC JRC/Google (2016). Projection: WGS_1984_UTM_Zone_52N.

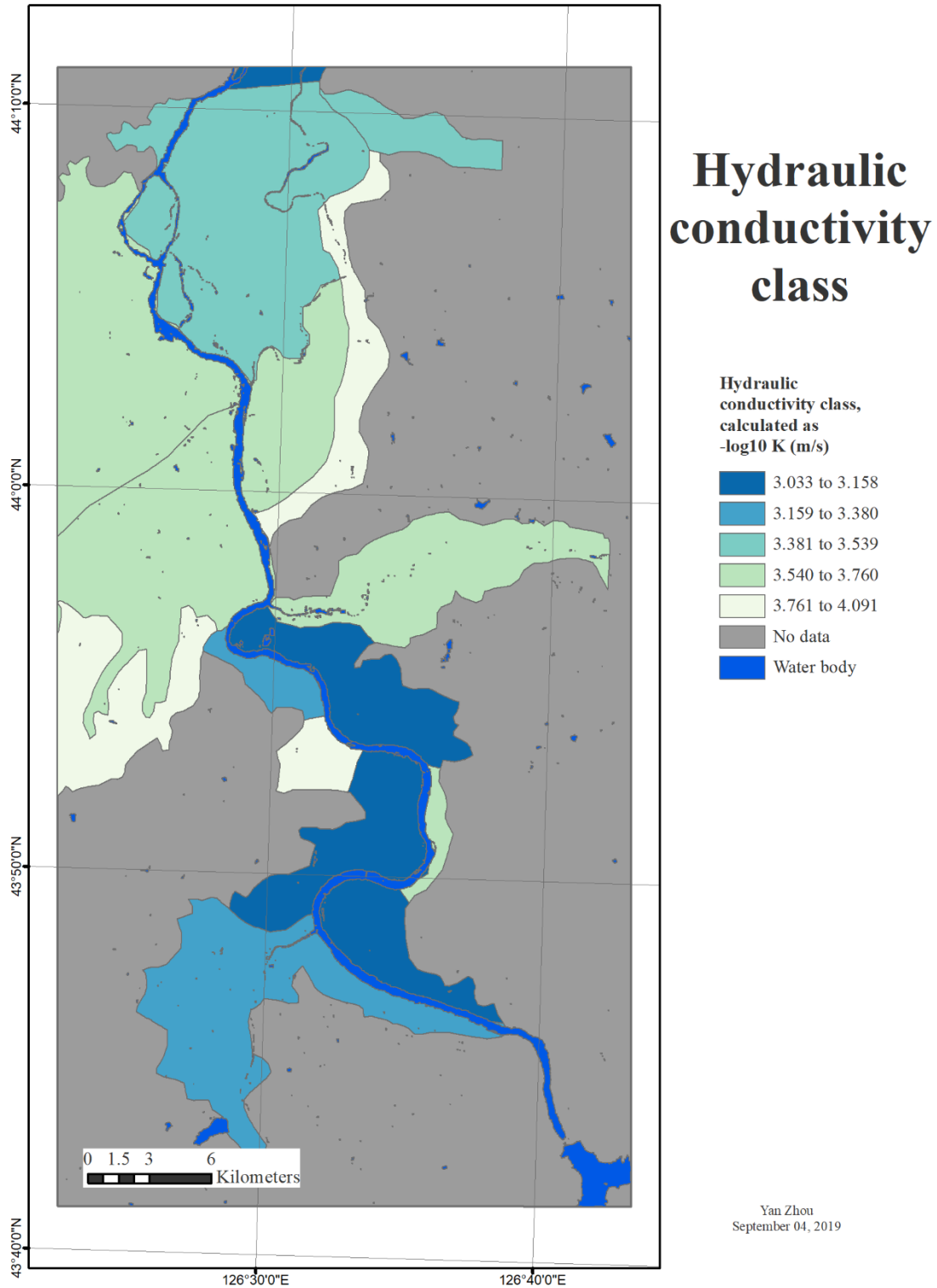


Figure 38: Hydraulic conductivity class map

Source: map modified from Qiu et al. (2015); water body was from EC JRC/Google (2016). Projection: WGS_1984_UTM_Zone_52N.

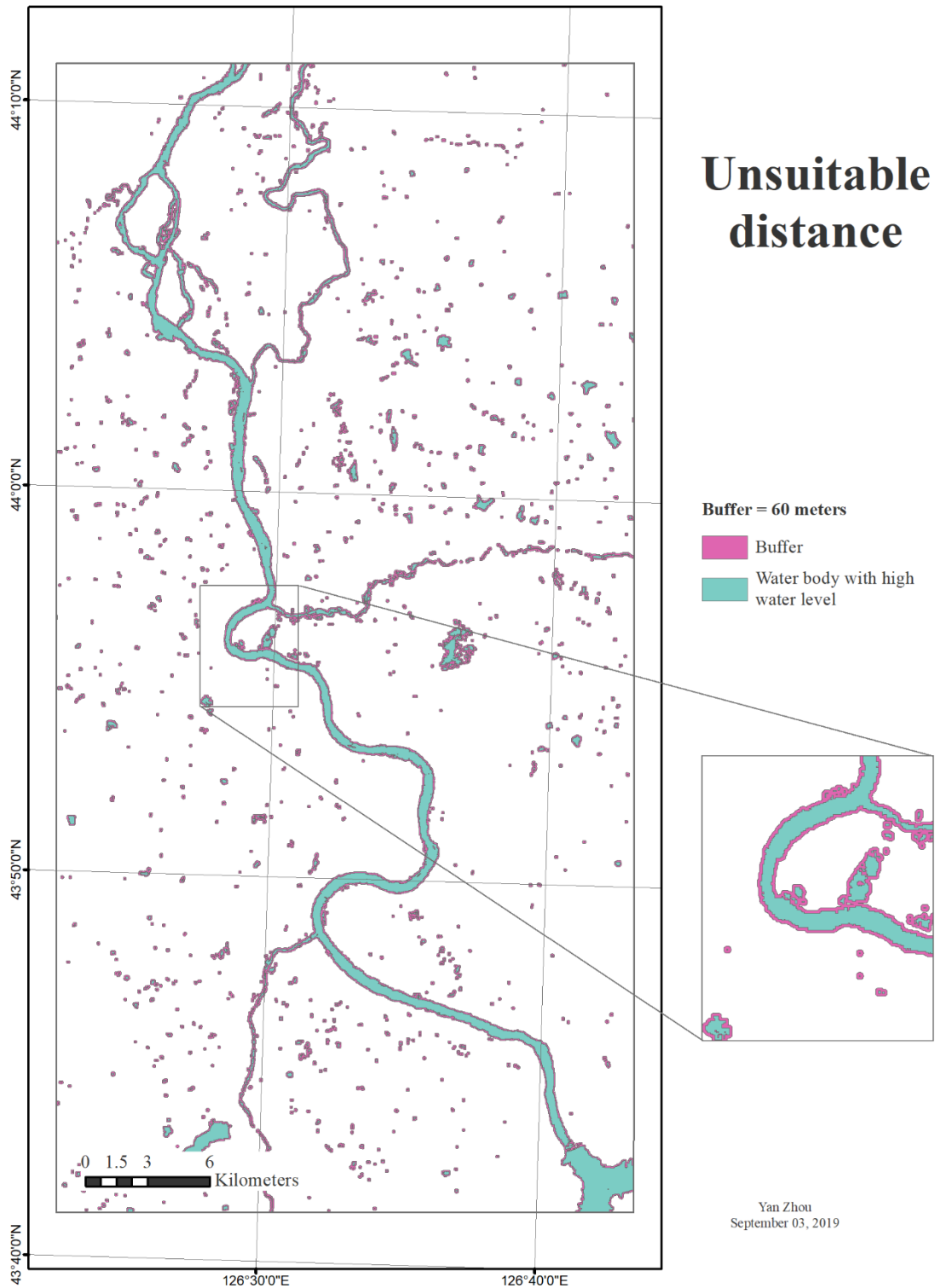


Figure 39: Unsuitable distance to surface water (60 meters buffer)

Sources: water body is from EC JRC/Google (2016). Projection: WGS_1984_UTM_Zone_52N.

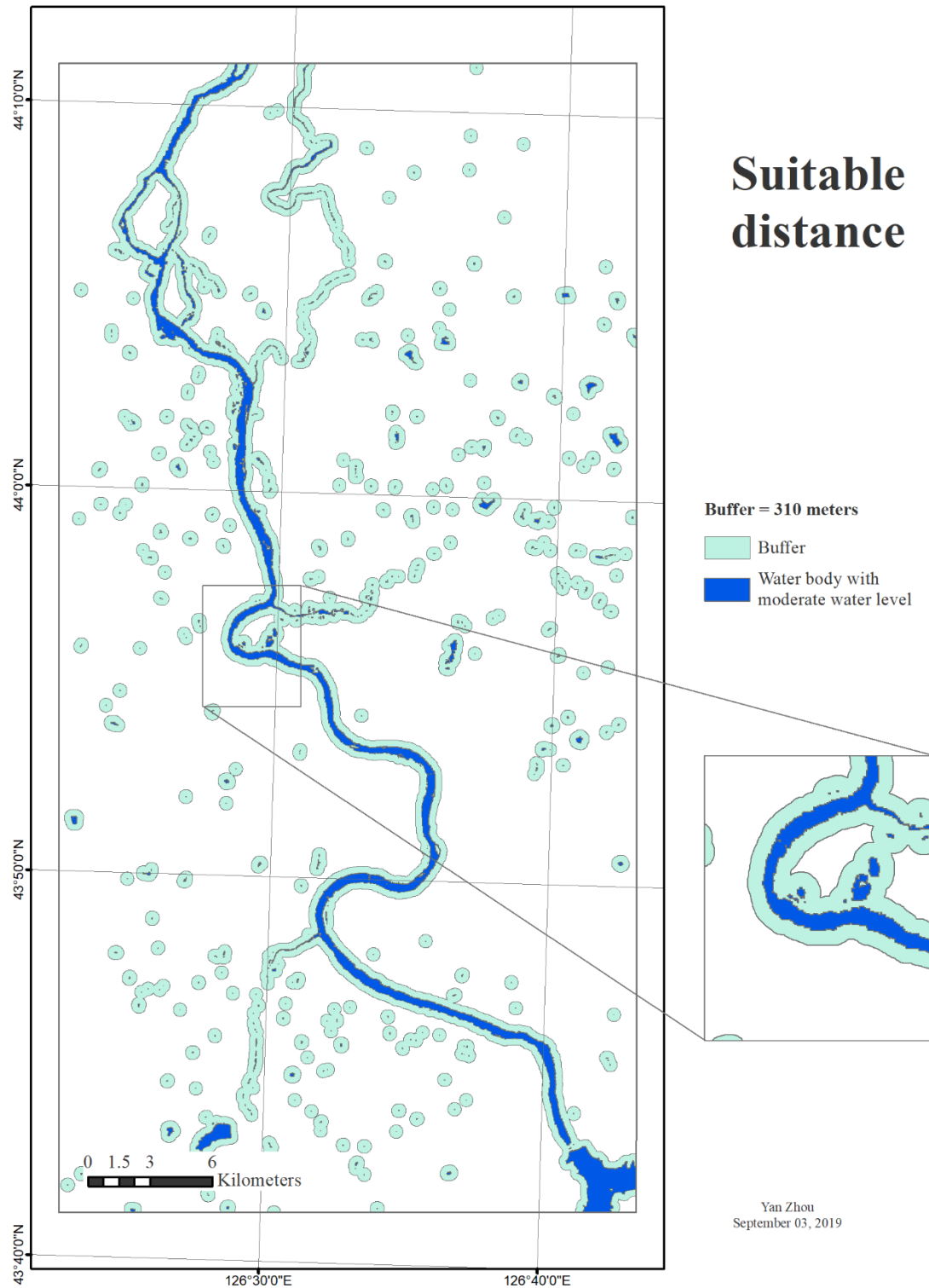


Figure 40: Suitable distance to surface water (310 meters buffer)

Sources: water body is from EC JRC/Google (2016). Projection: WGS_1984_UTM_Zone_52N.

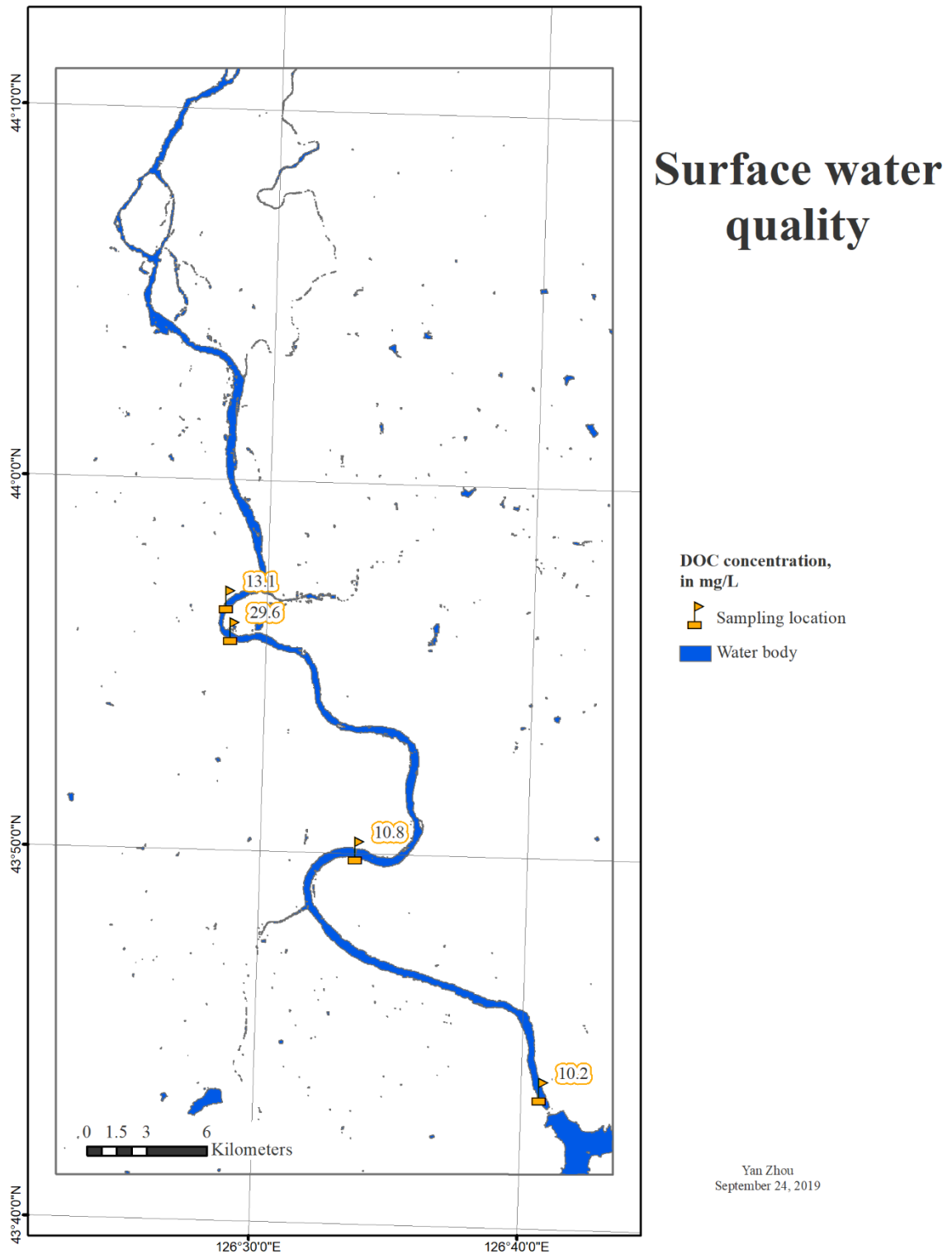


Figure 41: Locations of surface water quality observation in Jilin City

Source: DOC data is from Wang et al. (2016), water body is from EC JRC/Google (2016). Projection: WGS_1984_UTM_Zone_52N.

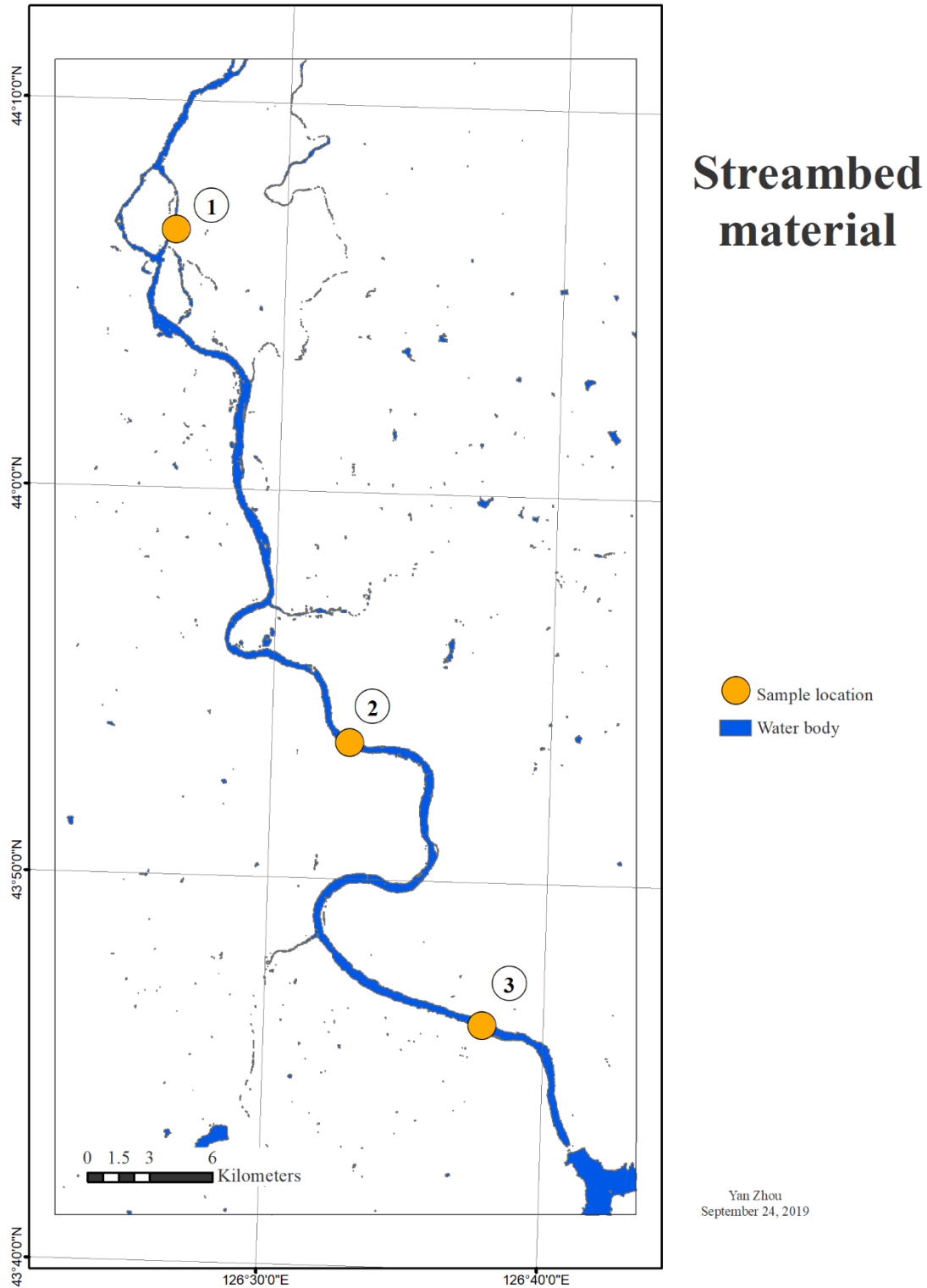


Figure 42: Locations of streambed material in Jilin City

Source: sample locations are from You (2016), water body is from EC JRC/Google (2016). Projection: WGS_1984_UTM_Zone_52N.

Sieve results



Figure 43: Grain size distribution of streambed material samples

Source: Du (2019).

Table 20: Grain sizes of streambed material sample

Sample	d_5	d_{10}	d_{16}	d_{30}	d_{50}	d_{60}	d_{84}	d_{90}	d_{95}	d_m	C_U	S	d_{90}/d_m
1	0.18	0.26	0.36	1.0	5.2	5.8	8.1	8.8	9.3	4.6	22	3.3	1.9
2	0.14	0.2	0.36	0.56	1.2	2.2	6.5	7.7	8.6	2.7	11	2.8	2.9
3	0.14	0.23	0.36	1.2	3.3	5.3	7.7	8.5	9.1	3.8	23	3.2	2.2

Note: diameter d_i in millimeters, indicates the sediment diameter/size for which i percent of the sediment sample is finer; d_m is the average grain size of the streambed material. C_U is the uniformity coefficient, equals to d_{60}/d_{10} , S is sorting coefficient after Folk and Ward (1957), which equals to $(d_{84}-d_{16})/4+(d_{95}-d_5)/6.6$. Source: Du (2019).

2.2 Kriging the saturated thickness

Kriging is a geostatistical interpolation method, which has been elaborated to tackle increasingly complex problems in mining, petroleum engineering, pollution control and abatement, and public health (Matheron 1963; Oliver and Webster 2014). In hydrogeologic investigation, kriging has also been successfully applied in estimating water-table altitudes, hydraulic head, and transmissivity or hydraulic conductivity (Kholghi and Hosseini 2006;

Motaghian and Mohammadi 2011; Dunlap and Spinazola 1981; Rivest, Marcotte, and Pasquier 2008).

To build the saturated aquifer thickness map for our study area, the following data of aquifer saturated thickness was used (Table 21, visualized in Figure 44)

Table 21: Saturated thickness data

Record	Longitude	Latitude	Thickness	Record	Longitude	Latitude	Thickness
1	126.4772	43.96209	30.2	21	126.5056	43.91359	19.33
2	126.4863	43.95218	32.4	22	126.5144	43.91476	16.93
3	126.4499	43.97592	28.6	23	126.5411	43.91566	13.16
4	126.4721	43.95252	14.6	24	126.5014	43.9066	18.09
5	126.4282	43.96489	27.9	25	126.5354	43.91088	15.56
6	126.4665	43.93432	13	26	126.5888	43.9122	9.6
7	126.5658	43.88934	14.3	27	126.5155	43.88613	6.62
8	126.5523	43.87747	9.5	28	126.3474	44.08751	13.52
9	126.4044	44.03831	51.25	29	126.4922	44.06025	45.59
10	126.4496	44.03662	49.02	30	126.3468	44.06928	73.98
11	126.5189	44.04526	20.69	31	126.4367	43.98707	39.38
12	126.3615	43.99143	48.65	32	126.4984	43.9955	18.64
13	126.4027	43.97958	44.43	33	126.5246	44.00891	21.9
14	126.4925	43.92367	18.9	34	126.4891	44.02138	24.76
15	126.4908	43.92028	17.35	35	126.5222	44.02384	15.99
16	126.4897	43.9145	18.23	36	126.5113	44.03185	17.35
17	126.5475	43.93902	18.17	37	126.5083	44.03841	25.1
18	126.5273	43.93187	39.41	38	126.5045	44.0509	17.8
19	126.5083	43.92176	18.55	39	126.4729	44.06585	26.7
20	126.5188	43.91858	17.64	40	126.4958	44.06988	32.2

Note: unit of saturated thickness is meter. Source: the 11th Five-Year Plan of China (2006-2010).

The Second Songhua River is also shown on the map (Figure 44), which flows from south to north, and the alluvium in this area is also shown. The saturated thickness samples seemed rather not equally distributed over the alluvium, with significantly fewer samples in the northwestern area, where the saturated thickness might be rather thick. Luckily, most of the samples were collected in areas relatively close to the river, which are more important than any other areas for our study.

Geostatistics can be supported by the first law of geography—“everything is related to everything else, but near things are more related than distant things” (Tobler 1970). It assumes, for example, all values in our study area are the result of a random process and the values are somewhat dependent on each other, or in other words, there is a spatial dependence or autocorrelation between the sample locations. In this case, the saturated aquifer thickness was assumed to vary continuously throughout the region. According to Matheron (1963), this spatial autocorrelation can be acquired by examining the squared difference of the sample values at all pairs of distances, which should increase with distance

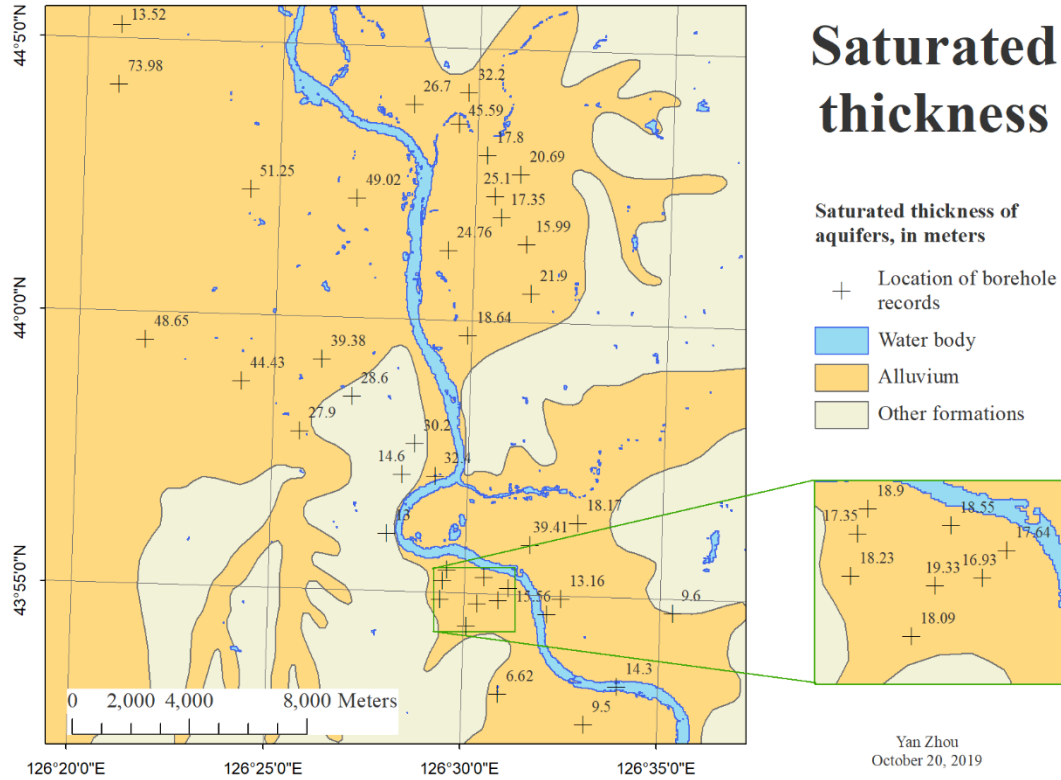


Figure 44: Borehole records of saturated thickness

Note: map shows a total of 40 borehole records of saturated thickness. Source: alluvium map was modified from 1:200000 Hydrogeologic Map of Jilin City (unpublished), borehole records are from the 11th Five-Year Plan of China (2006-2010), water body is from EC JRC/Google (2016). Projection: WGS_1984_UTM_Zone_52N.

within an influence zone. This squared difference is called a *variogram* by Matheron in his 1962 study (as cited in Cressie 1993, 58) and the *semivariogram* has been defined as half of the variogram:

$$\text{Variogram} = 2\gamma(h_{ij}) = (v_i - v_j)^2, \quad (36)$$

where

$$h_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, \quad (37)$$

and $\gamma(h_{ij})$ is the semivariogram of the pair of location i and j , v_i is the sample value at location i , and v_j is the value at location j . Locations or spatial coordinates are given as (x, y) , so h_{ij} is the Euclidean distance between (x_i, y_i) and (x_j, y_j) .

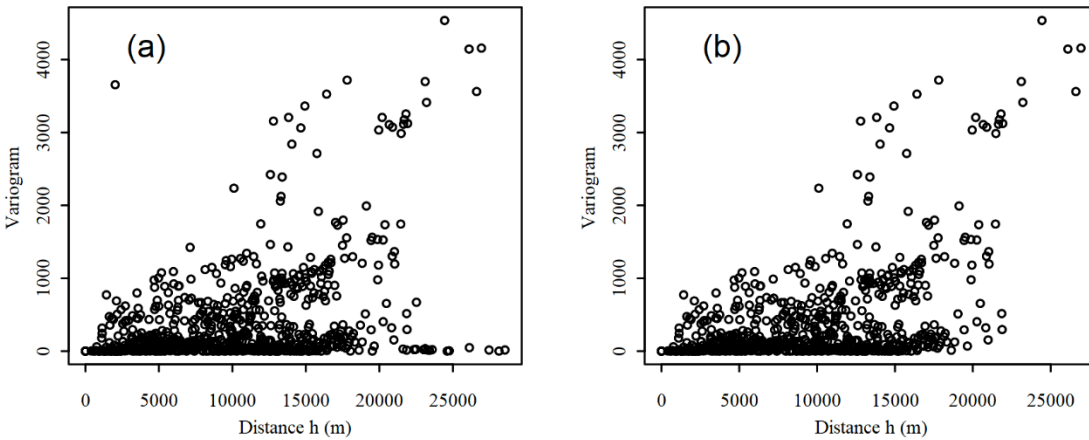


Figure 45: Variogram (omnidirectional) of saturated thickness data

Note: (a) $N=40$ (N is the sample size), with 780 pairs of distances and squared differences between saturated thickness values; (b) $N=39$, with 741 pairs.

In Figure 45 (a), most variogram values first increase with distance and then decrease, except some pairs which always yield high variograms regardless of the distance. Notice that some extreme values may cause this problem, for instance, one pair that yielded a rather high variogram at distance about 2000 meters, significantly differed from other pairs at the same distance. This pair was found to be exactly the result of two borehole records located on the top-left in Figure 44 (value=13.52 and 73.98). The extreme low value of 13.52 seemed very unusual compared to its neighbors, which all had rather high values. To see whether this extreme value or outlier would cause unfavorable conditions for our model, the prediction process of the unknowns was made using the original dataset ($N=40$, Figure 45 (a)) as well as the dataset without this outlier ($N=39$, Figure 45 (b)). The latter was used to illustrate the method of ordinary kriging henceforth.

To build the geostatistical model, the semivariogram is of concern, rather than the variogram. As already shown before, their relationship is

$$\text{Semivariogram} = \frac{1}{2} * \text{Variogram} = \gamma(h_{ij}). \quad (38)$$

Yet, for our model, there are still too many pairs of variogram/semivariogram and distance, which makes the semivariogram structure unclear and difficult to fit a model to. To deal with this problem, pairs of locations must be grouped by a certain distance, and then the semivariograms of each group were averaged, and a model was fitted to each group (Figure 46). This distance used for grouping is usually referred to as *lag spacing*, *lag increment* (Isaaks and Srivastava 1989). The tolerance allowed on this distance was taken as half of the lag spacing.

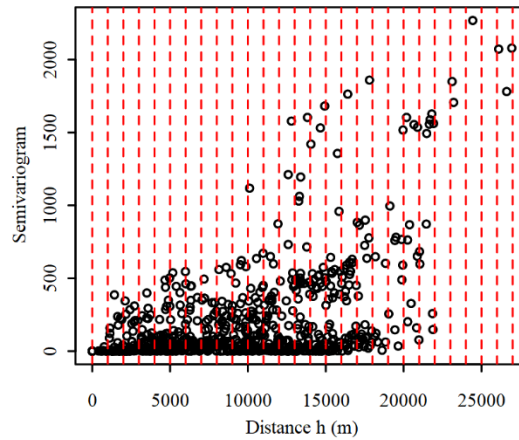


Figure 46: Semivariogram (omnidirectional) of saturated thickness data ($N=39$)

Note: the distance or space between the red dotted lines referred to a lag size of 1000 m.

By doing this, the average value of the semivariograms at each lag distance was obtained, which is called the *experimental* or *empirical semivariogram*. On the one hand, the lag spacing directly determines the number of empirical semivariograms, which should be neither too small nor too large. On the other hand, as Journel and Huijbregts suggested in their work (1978), the number of sample pairs at each lag distance should be more than 30 (for most of them); thus, the appropriateness of the chosen lag tolerance needs also to be checked. An illustration of different lag spacing was shown in Figure 47. A lag spacing of 1000 meters was considered as appropriate for our model, which contained an appropriate number of sample pairs at each lag distance and empirical semivariograms as well, without them being smoothed out (Table 22).

Table 22: Empirical semivariogram (omnidirectional) of saturated thickness data ($N=39$) with a lag spacing of 1000 m

No. of pairs	Lag	$\gamma(h)$	No. of pairs	Lag	$\gamma(h)$
20	740.876	4.0431	72	14526.5	320.5906
74	1535.683	49.40439	66	15447.86	269.2557
78	2482.33	66.98744	62	16481.91	333.8909
102	3544.123	60.0226	44	17482.32	368.8255
98	4502.225	92.76745	14	18454.33	228.9764
86	5462.623	97.5247	20	19586.62	633.4952
86	6458.331	101.2982	20	20581.92	774.0665
84	7523.572	108.4269	20	21581.56	1038.427
78	8504.016	143.0235	0	-	-
92	9562.207	134.7584	4	23162.65	1777.992
94	10511.04	183.0846	2	24439.24	2268.685
84	11486.36	170.9701	0	-	-
78	12504.61	183.7457	6	26556.03	1977.36
98	13460.62	272.594	-	-	-

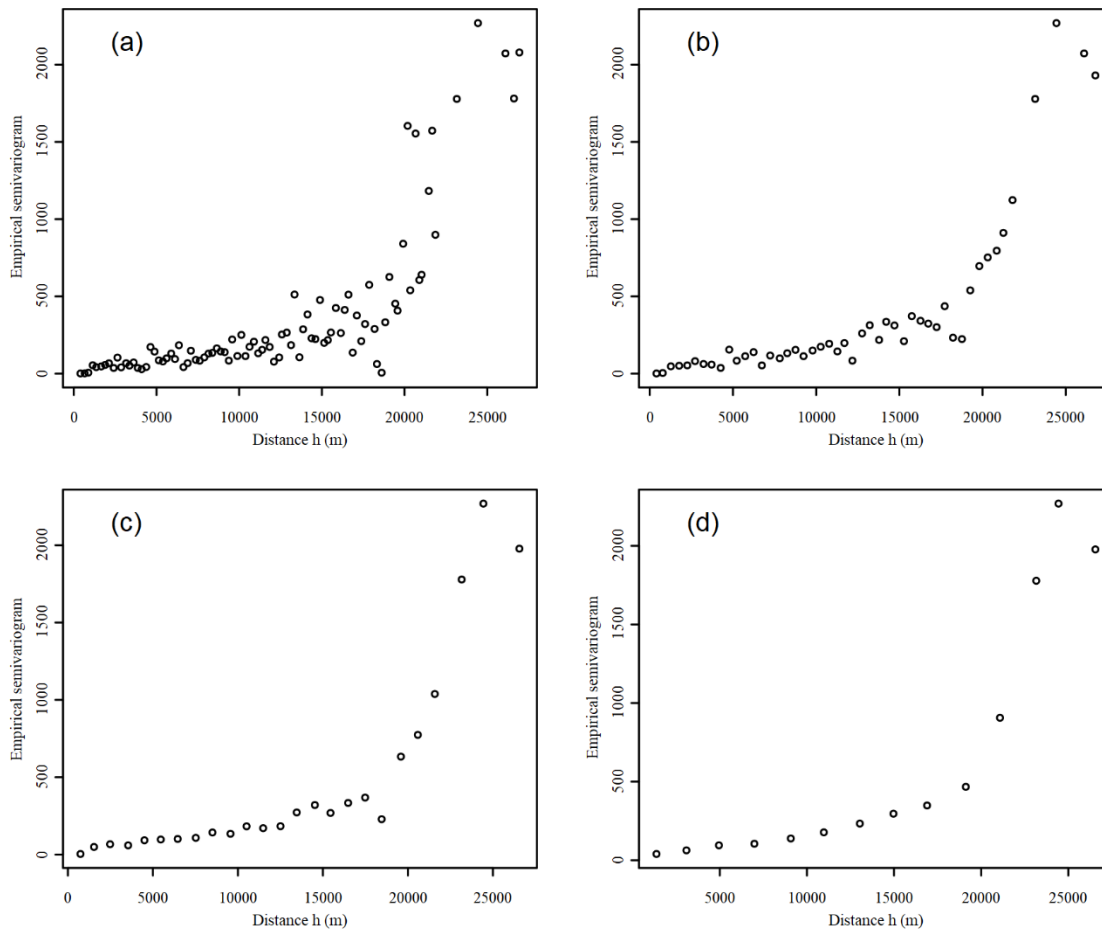


Figure 47: Empirical semivariogram (omnidirectional) with different lag spacing

Note: (a) bin size of 250 m, (b) 500 m, (c) 1000 m, and (d) 2000 m.

A common choice for the largest lag distance that allowed for our semivariogram model is, as a rule of thumb, approximately half of the largest distance of the dataset (Johnston *et al.* 2001, 66; Journel and Huijbregts 1978, 194). As can be seen from Table 22 and Figure 47, there were fewer pairs of samples at large distances, and the empirical semivariograms became erratic; thus, the upper limit of 13000 m for the lag distance was considered to be appropriate.

A normally distributed data or a Gaussian process is paradigmatically employed in ordinary kriging, a transformation of the original data is sometimes necessary (*e.g.*, gold grades) (Cressie 1993). Journel and Huijbregts (1978) claimed that if the sample distribution is clearly lognormal, a log-transformation of the data is usually preferred (in the sense of a lesser estimation variance). The distribution of our original dataset looked rather lognormal (Figure 48 (a)), a log-transformation (Figure 48 (b)) of the original data

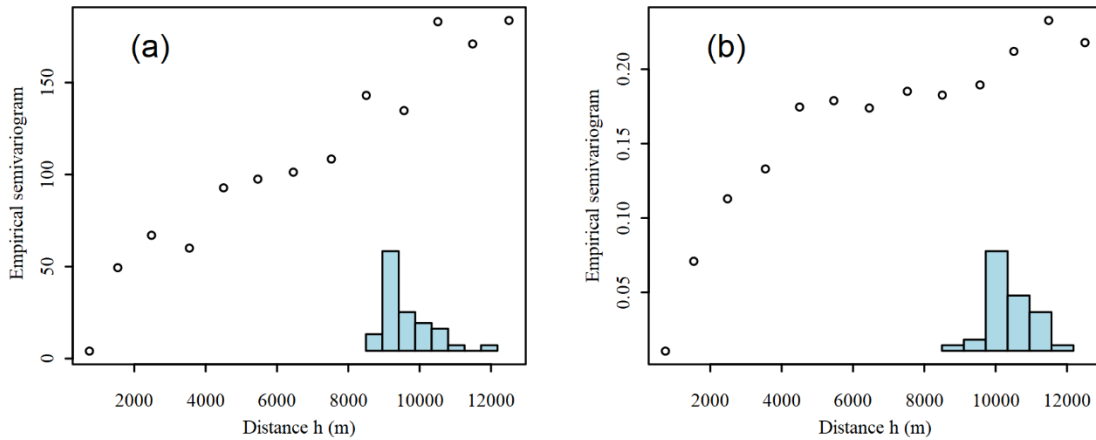


Figure 48: Empirical semivariogram (omnidirectional) with an upper limit of 13000 m for the lag distance

Note: lag spacing of 1000 m; (a) original saturated thickness dataset, (b) log-transformed saturated thickness dataset. Histograms of both dataset are shown on the bottom-right of the plots.

seemed advisable. *Shapiro-Wilk Test* was applied to test the null hypothesis of normality for both datasets, and the R-package “stats v3.6.1” was used. According to Royston (1995), a p-value > 0.05 implied the distribution of the data was not significantly different from normal distribution. A p-value of 0.0002 was obtained from the original dataset, while 0.421 from the log-transformed dataset, indicating that the null hypothesis can be rejected for the original dataset but cannot in the case of the log-transformed dataset.

The log-transformed (natural logarithm) saturated thickness dataset was used for prediction henceforth, which was back-transformed after kriging; therefore, the technique is called “lognormal kriging” (LK), which is a specific form of ordinary kriging (OK). In order to choose an appropriate model, a mathematical fitting of different semivariogram models was made. The method used here was *Least Squares*, with a purpose to reduce the residuals or the error sum of squares (*SSE*) to the minimum (Cressie 1993, 94). The fitting results were shown in Table 23 and Figure 49.

Table 23: Parameters of the fitted semivariogram models

Model	c_0	r	c	ρ	<i>SSE</i>
Spherical	-0.00984	6752	0.209	0.963	0.00347
Exponential	-0.03902	2911	0.253	0.980	0.00183
Gaussian	0.0199	3288	0.179	0.961	0.00364

Note: c_0 is Nugget, r is range, c is partial sill, ρ is correlation coefficient, *SSE* is error sum of squares.

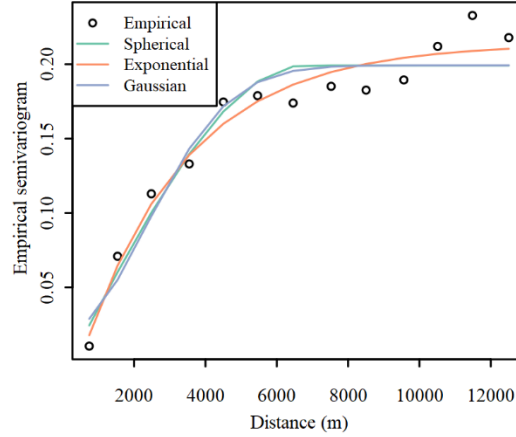


Figure 49: Mathematical fitting of the semivariogram models

All three semivariogram models yielded a rather good fit, among which the exponential model yielded the best fit, with a correlation coefficient of 0.98 and a sum of the squared errors of 0.00183. The spherical model was a little bit better than the Gaussian model. The rather small *nugget effect* (c_0 close to 0) of three models indicated a continuity at the origin (Cressie 1993). A negative value of c_0 has little practical meaning, this was resulted by the optimization process using *Least Squares*, and the value c_0 has not been forced to be non-negative. Notice that from Table 23 and Figure 49, the empirical semivariogram value at the first lag (lag=741) for the original and the log-transformed saturated thickness dataset was 4 and 0.01, separately. They were both non-negative and close to 0, indicating small nugget effect. For this reason, a better technique would be fitting these semivariogram models without nugget effect (c_0). This has, however, not been corrected here due to the limited time of this project. Thus, for interpolation of the unknown locations using ordinary kriging, the exponential model was chosen, which was written as a form

$$\tilde{\gamma}(h_{ij}) = -0.03902 + 0.253 \left(1 - \exp\left(\frac{-h_{ij}}{2911}\right) \right), \quad (39)$$

where $\tilde{\gamma}(h_{ij})$ or $\tilde{\gamma}_{ij}$ denotes the modeled semivariogram value based on the distance between location i and j , *i.e.*, h_{ij} .

A 105×105 gridded area was generated, with a grid size of 250×250 meters (Figure 50). Based on Hengl (2006), the grid resolution should be between $0.05\sqrt{A/N}$ and $0.1\sqrt{A/N}$, where A is surface of study area, N is sample size. So, in our case, a resolution of 250 meters is considered appropriate. The geometric center of these grids was the location to be predicted, which represented the value of the grid or pixel area.

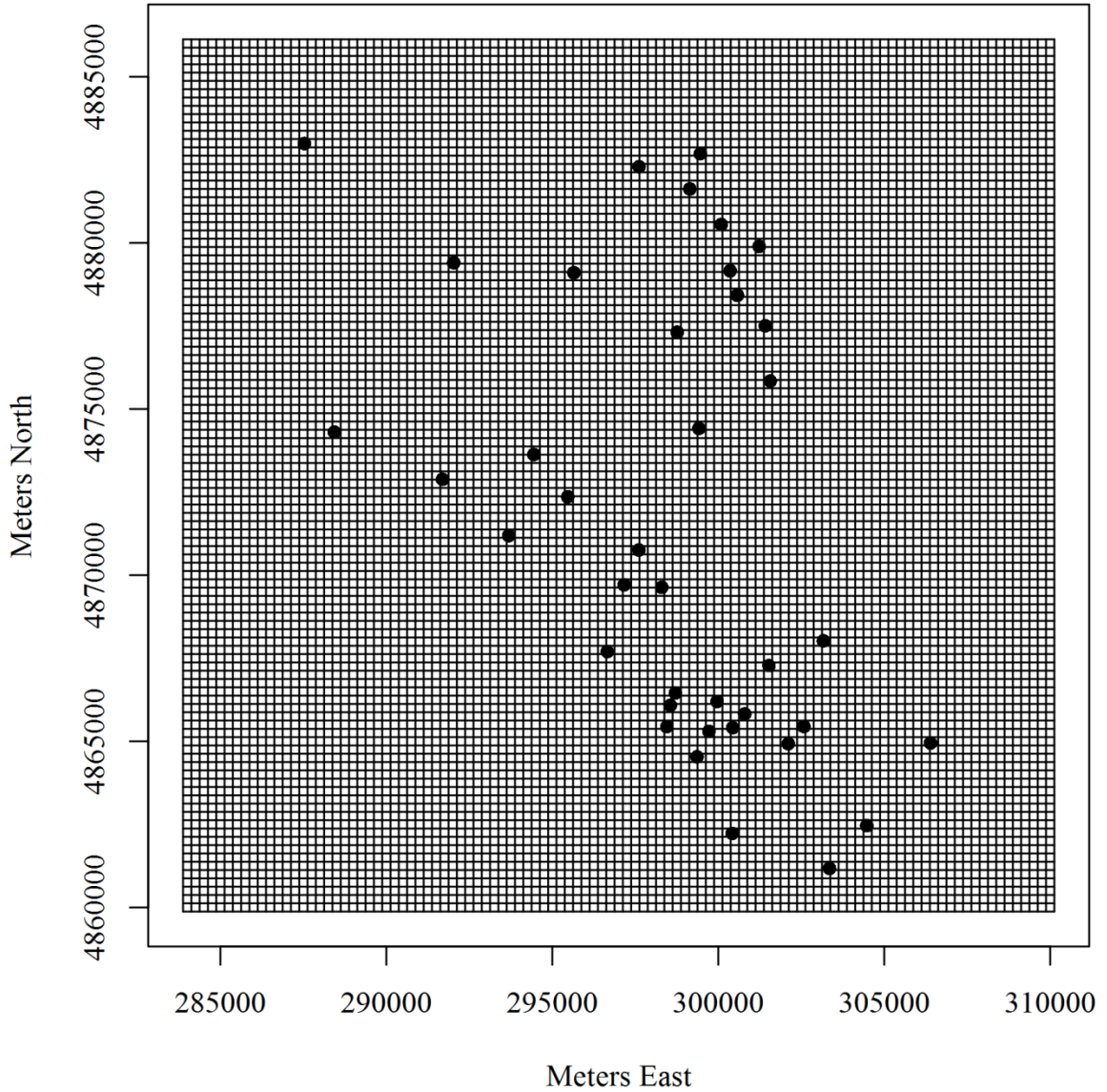


Figure 50: Interpolation grid and sample points

Kriging is a linear prediction; to predict the value \hat{v}_0 at an unknown location, the values of its neighboring samples and their weights have to be specified (see equation (40)). These weights can be solved exactly using the semivariogram model specified in the previous step.

$$\hat{v}_0 = \sum_{i=1}^n w_i v_i, \quad (40)$$

where v_i is the sample value at location i , w_i is the weight of neighbor value v_i , n is the number of neighbors.

But before solving the kriging weights, it is necessary to develop a strategy of how the neighbors are to be defined or selected. An illustration of how neighbors were selected was shown in Figure 51.

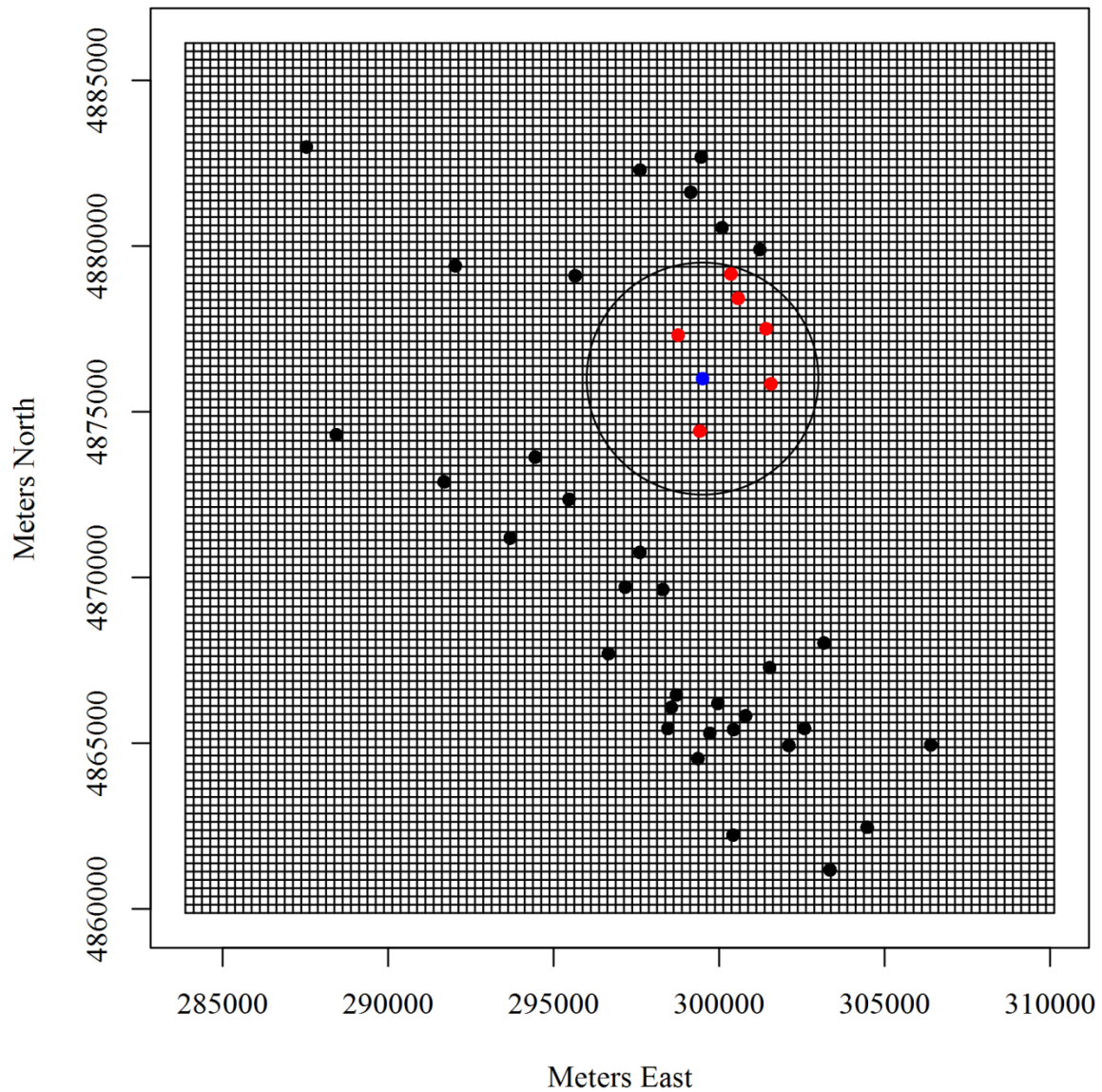


Figure 51: An illustration of neighbor search strategy

In Figure 51, the blue point was the unknown location to be predicted, while the red points were the neighbors considered. According to Isaaks and Srivastava (1989), the search neighborhood should be slightly larger than the average spacing between the sample data, which can be calculated with a simple formula

$$\text{Average spacing between data} \approx \sqrt{\frac{\text{Total area covered by samples}}{\text{Number of samples}}}. \quad (41)$$

So for our case, the average spacing between samples was about 3248 m; thus, a simple search strategy was used—a circle with a radius (R) of 3500 m (Figure 52, black circle). Figure 52 illustrates how value \hat{v}_0 at an unknown location was predicted using kriging.

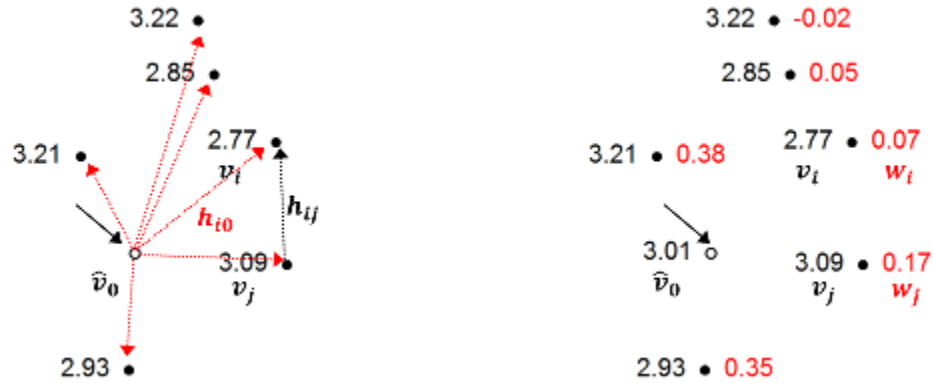


Figure 52: An illustration of predicting an unknown location

Note: the circle is the unknown location to be predicted (\hat{v}_0), black points are the neighbors. Log-transformed saturated thickness values (e.g., v_i, v_j) are placed on the left-side of the neighbor points, the solved kriging weights (e.g., w_i, w_j) are shown on the right-side of the neighbor points.

To solve the weights of the neighbors, equation (39) and (42) were required

$$\begin{pmatrix} \tilde{\gamma}_{11} & \dots & \tilde{\gamma}_{1n} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \tilde{\gamma}_{n1} & \dots & \tilde{\gamma}_{nn} & 1 \\ 1 & \dots & 1 & 0 \end{pmatrix} * \begin{Bmatrix} w_1 \\ \vdots \\ w_n \\ \mu \end{Bmatrix} = \begin{Bmatrix} \tilde{\gamma}_{10} \\ \vdots \\ \tilde{\gamma}_{n0} \\ 1 \end{Bmatrix}, \quad (42)$$

where w_i is the kriging weight for the location or neighbor i , $\tilde{\gamma}_{i0}$ is the modeled semivariogram value based on the distance between the unknown location to be predicted and the neighbor i , $i = 1, 2, \dots, n$, n is the number of neighbors, μ is the Lagrange parameter.

Unlike simple kriging, ordinary kriging used the Lagrange parameter in the equation (42) to ensure the *unbiasedness condition* (Isaaks and Srivastava 1989). The solved kriging weights combined with the log-transformed values of neighbors, the solution of equation (40) was $\hat{v}_0 = 3.01$ (Figure 52, right), which was also a log-transformed saturated thickness value at this unknown location.

Locations of borehole records for this study were, however, not evenly separated, as shown in Figure 44. Using the neighbor-search strategy defined earlier, the number of nearby samples for each unknown location varied between as much as 15 to only one (Figure 53). Unknown locations with 0 neighbors would not be predicted.

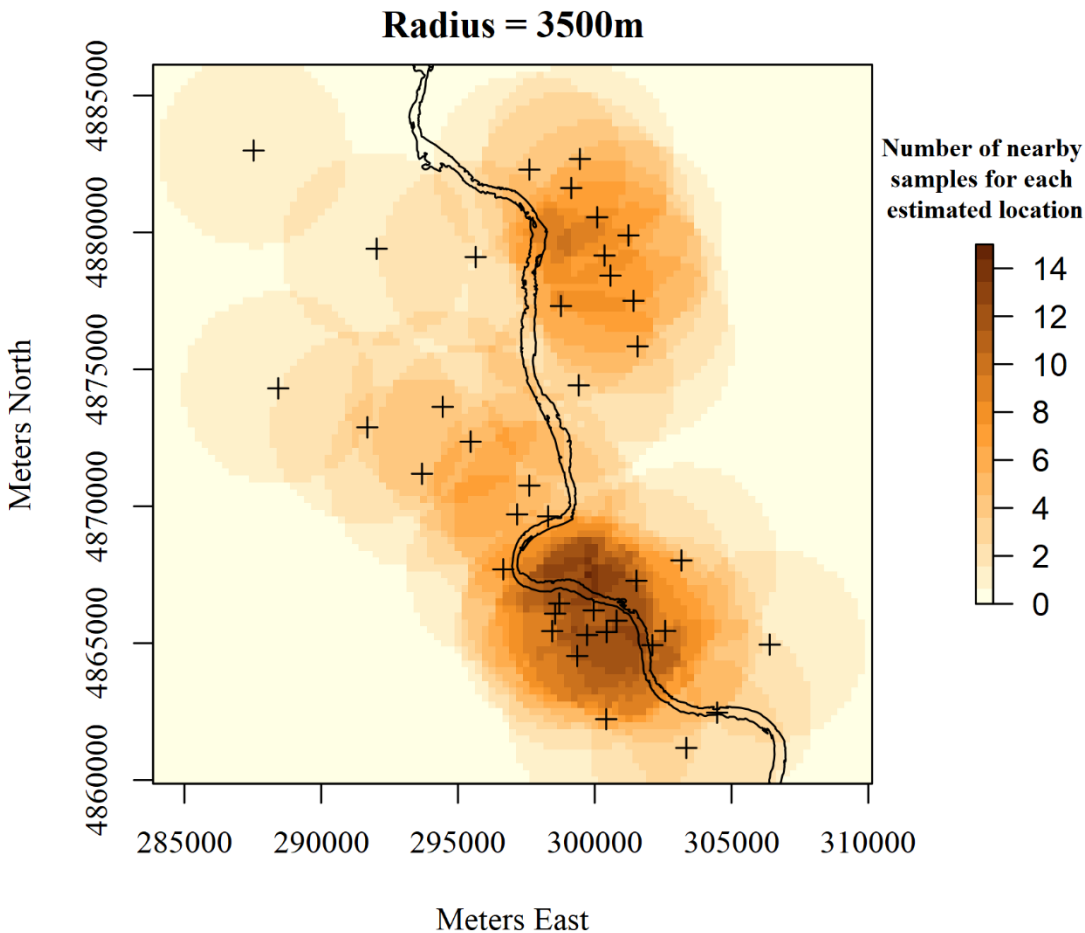


Figure 53: Neighbor-search strategy with radius = 3500 m

Note: samples were marked as crosses.

For robustness of the model, the number of nearby samples should be at least four, as Issaks and Srivastava (1989, 341) argued in their text. An extra condition was added to ensure that at least four of the closest nearby samples were considered for each predicted unknown location (Figure 54). By doing this, unknown locations for the whole area were predicted, the results were shown in the “Results” section of this part. A discussion of the kriging results were made in the “Discussion” section of this part.

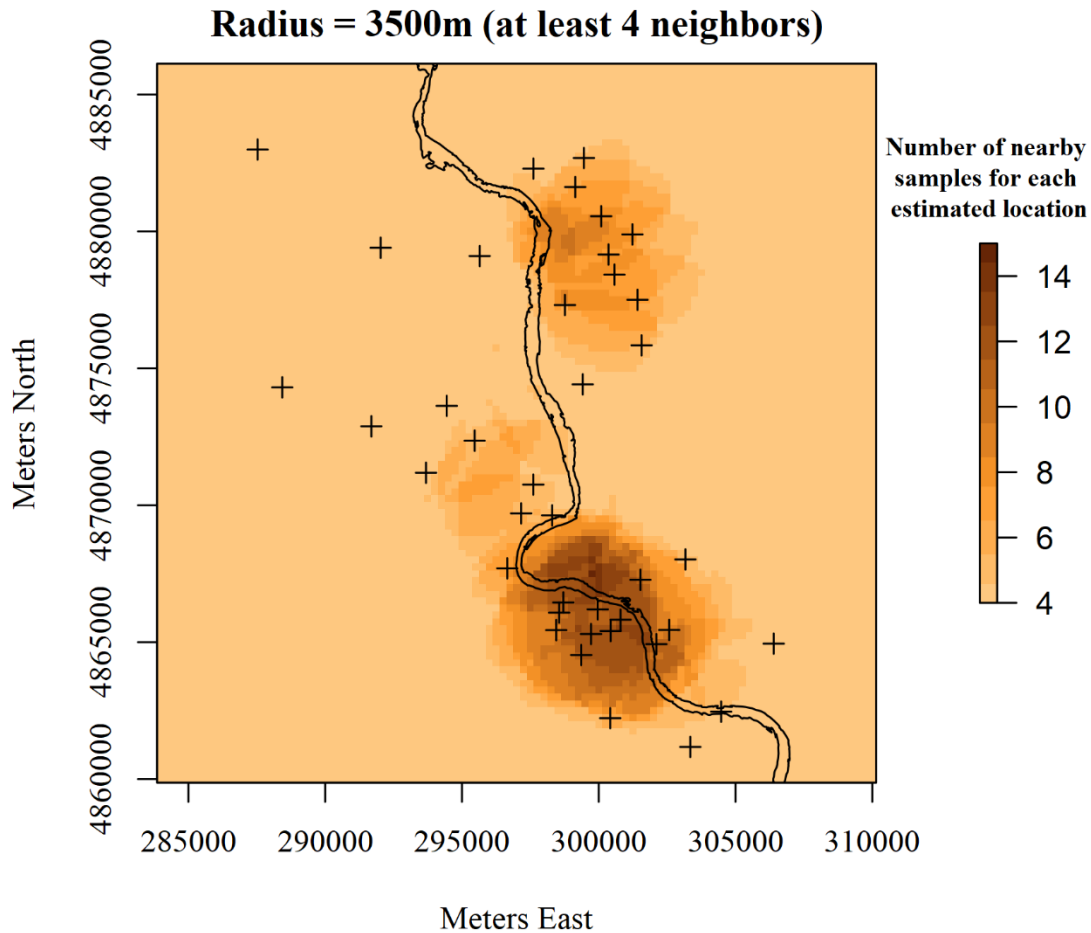


Figure 54: Neighbor-search strategy with radius = 3500 m (at least 4 neighbors)

2.3 Aggregation of all constraint maps

The screening of suitable areas or sites for RBF systems took place after all GIS data were collected and the associated constraint maps of Jilin City were built (Figure 55). The constraint maps include an alluvium map, a land cover map, a groundwater quality map, a hydraulic conductivity (class) map, a saturated thickness map (obtained from kriging), a buffer map of surface water area at high water level, a buffer map of surface water area at moderate water level, and a map of water area at high water level.

Constraint maps were all projected into the “WGS_1984_UTM_Zone_52N” (EPSG: 32652) coordinate system in ArcMap 10.6 (ESRI, 1999-2017). If necessary, raster datasets were firstly reclassified to “1” (suitable) or “0” (unsuitable) in ArcMap to distinguish suitable and unsuitable areas based on the selection attributes. Then, all maps were converted to vector datasets (polygon); similarly, suitable and unsuitable attribute values were set to “1” and “0”, respectively. All maps were aggregated into a single map using “Union” tool from Analysis toolset, which computed a geometric union of the input

features that contained all attributes. After that, “Field Calculator” was performed and the AND Boolean

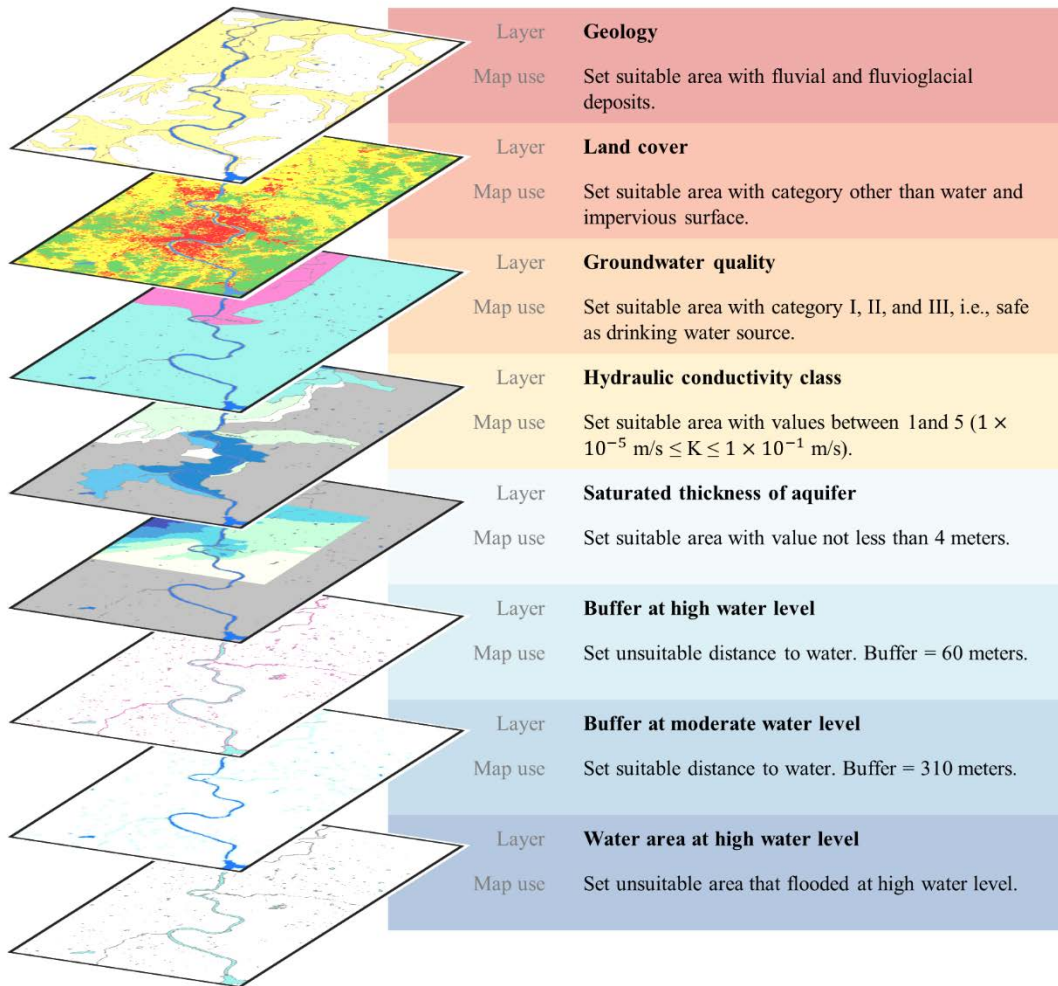


Figure 55: Overlay of constraint maps

operator was used to yield the final suitable areas that fulfilled all constraints. Among the suitable areas, those connected and having a relatively large combined area were considered to be a candidate site or suitable site for RBF. Note that areas far from the Second Songhua River could still be classified as suitable areas since the buffers were based on all surface waters. However, they need to be eliminated because only areas close to the main channel were considered suitable for RBF.

Constraints of the RBF site selection problem were defined as follows:

- Only perennial streams are considered,
- Fluvial and fluvioglacial deposits are the suitable geologic formations,
- Water and impervious surface (or artificial surfaces such as urban and associated areas) are the unsuitable land cover types,
- Groundwater quality should meet the regulation from the local authority (*e.g.*, according to Chinese regulations, not worse than category III),
- Hydraulic conductivity should be between 10^{-5} m/s and 10^{-1} m/s (or hydraulic conductivity class between 1 and 5),
- Saturated thickness of aquifer should not be less than 4 meters,
- Distance from wells to water should be between 60 and 310 meters,
- Areas flooded by seasonally high water are unsuitable.

3. Results

3.1 Kriging

The result of ordinary kriging is shown in Figure 56. As the Second Songhua River flows from the southeast to the northwest, the aquifer becomes thicker. In the study area, the estimated saturated thickness varied between 6.8 meters and 66.1 meters. For the areas close to the river (≤ 310 m), the minimum aquifer thickness was 10 meters, the maximum was about 47 meters, and the mean thickness was about 25 meters.

3.2 Suitable sites

Based on the constraint maps, the suitable areas for RBF system were delineated (Figure 57). Due to groundwater contamination (category IV), no suitable areas were identified from the northern part of the study area. Although the southern part was highly urbanized, there were still some fragmental areas identified as suitable. The relatively small area with saturated thickness information has gathered those suitable areas in the middle of the study area. From those suitable areas, six suitable RBF sites can be delineated, the lands of which were less incomplete (Figure 58). As explained in the previous section, areas delineated based on the main channel were considered suitable, while those based on oxbow or other surface water bodies were considered unsuitable and removed. Referring to the satellite map or land cover map of the study area, all sites were located downstream of the city center, close to the Second Songhua River. Three of them were found at the point bar or inner bend of meander (site 1, 4, 5), two were at the cut bank or outer bend of meander (site 2, 6), and one site was located near the straight channel of the river (site 3).

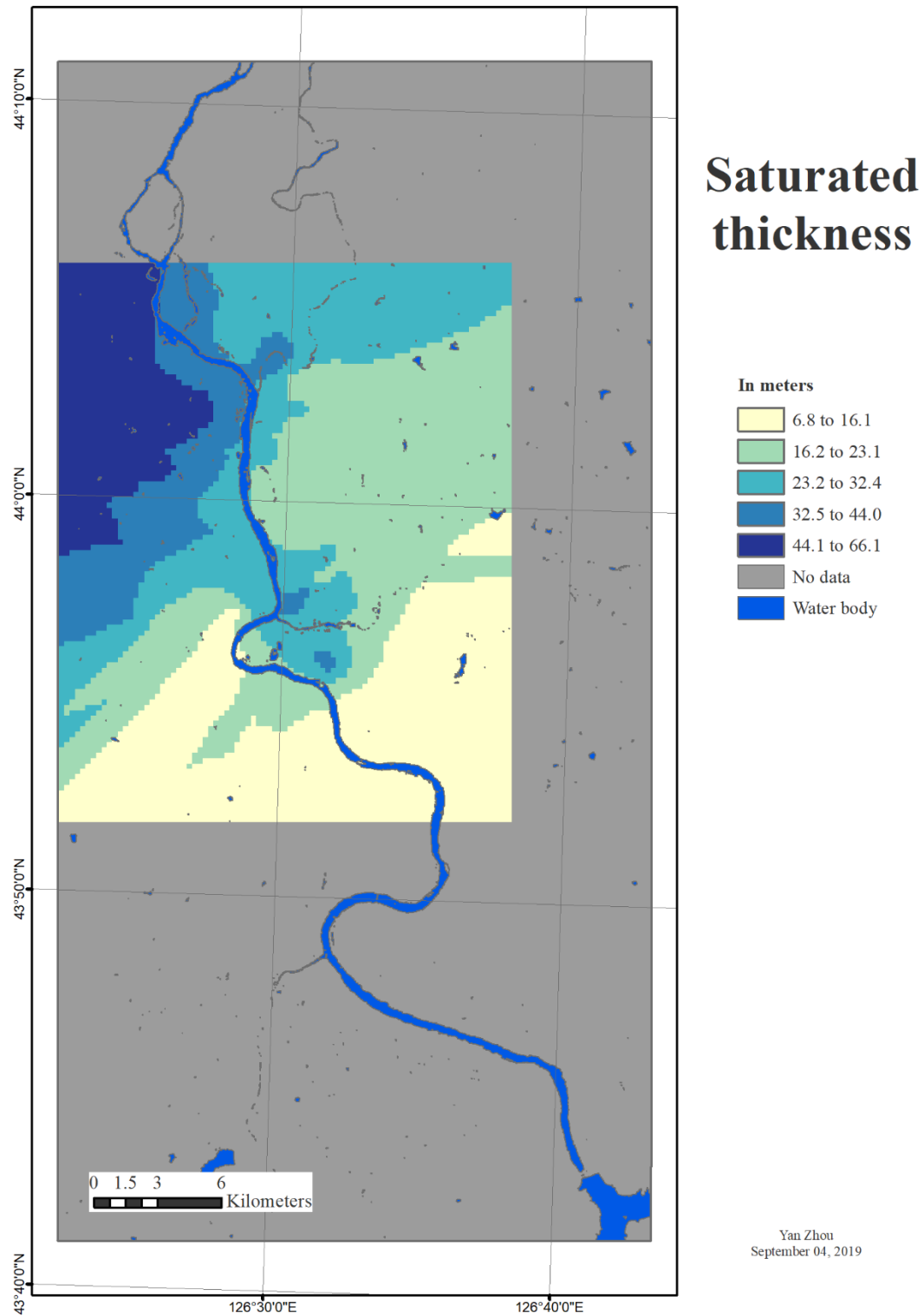


Figure 56: Saturated thickness map

Source: map was generated using kriging (see previous section); water body is from EC JRC/Google (2016).
Projection: WGS_1984_UTM_Zone_52N.

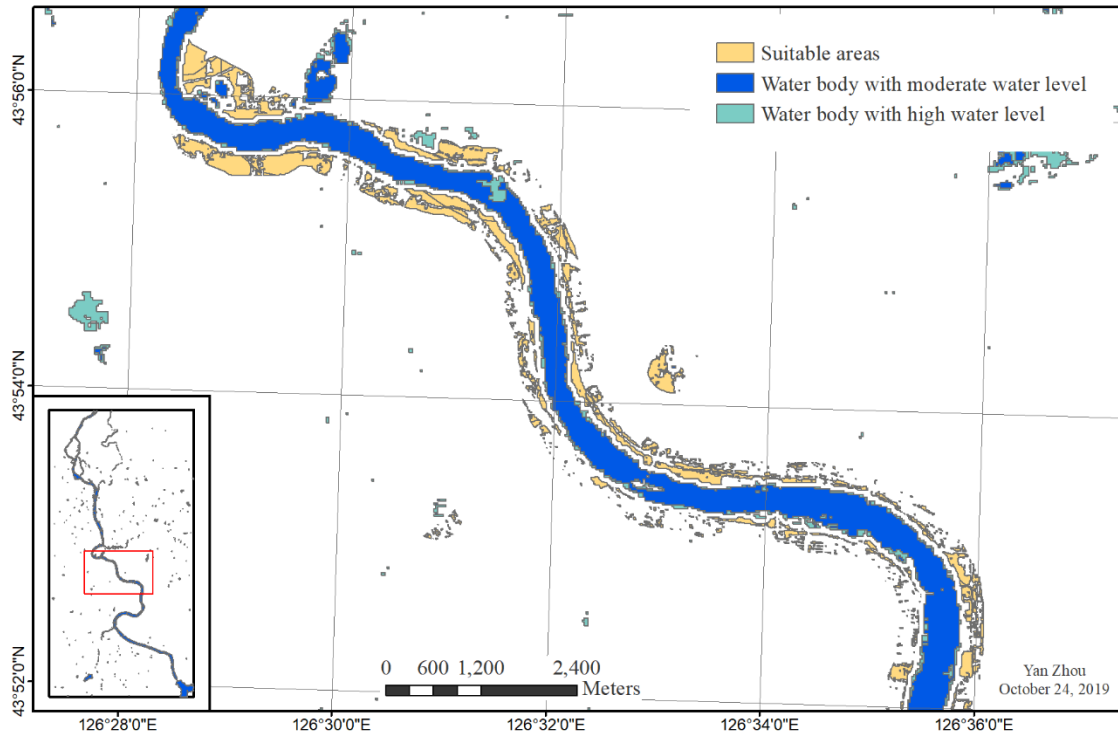


Figure 57: Suitable areas

Table 24: Comparison of six suitable sites

Attribute	Scale of measure	Alternatives					
		Site 1	Site 2	Site 3	Site 4	Site 5	Site 6
$X_1 \equiv$ Transmissivity	m ² /s	0.010	0.012	0.023	0.010	0.010	0.011
$X_2 \equiv$ Well field location score	Subjective	66	0	33	66	66	0
$X_3 \equiv d_{90}/d_m$ ratio of streambed material	Ratio scale	2.2 ^a	2.5 ^a	2.8 ^a	2.4 ^a	2.9	2.6 ^a
$X_4 \equiv$ Mean DOC concentration	mg/L	21.4 ^b	29.6	18.6 ^c	29.5 ^c	11.8 ^c	12.2 ^c
$X_5 \equiv$ Groundwater quality score	Subjective	68 ^d	41 ^d	19 ^d	36 ^d	45 ^d	66 ^d

Note: ^arandomly assigned from range 1.9-2.9.

^busing the average of 13.1 and 29.6 mg/L.

^crandomly assigned from range 10.2-29.6 mg/L.

^drandomly assigned from range 0-100.

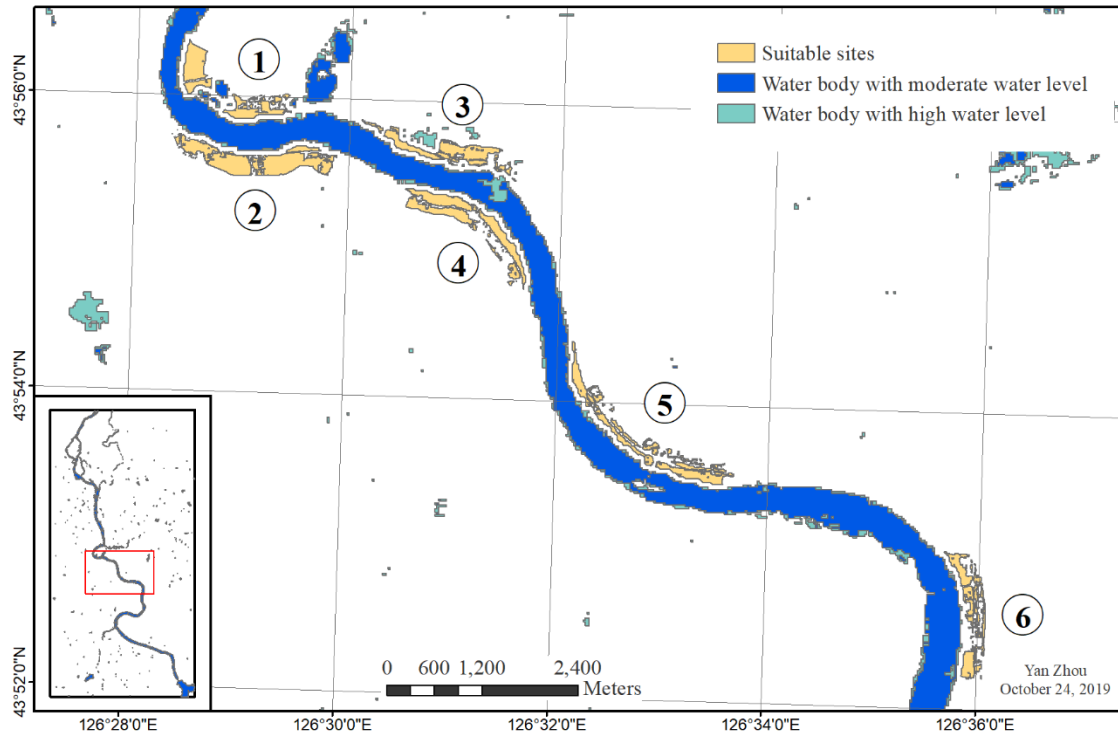


Figure 58: Six suitable sites

The associated attributes were assigned for the suitable sites (Table 24). Among them, transmissivity was calculated using equation (3), where the average saturated thickness and hydraulic conductivity of suitable sites was estimated using “Zonal Statistics” tool from the “Spatial Analyst” toolbox in ArcMAP 10.6.

Using the multiplicative utility function assessed in Part II, the final utility value of each suitable site is shown in Table 25.

Table 25: The utility for six suitable sites

Site	Utility Value	Rank
1	0.1519	1
2	0.0424	6
3	0.0661	5
4	0.0925	3
5	0.1021	2
6	0.0680	4

Based on this result, site 1 was the most preferred with a utility of 0.1519, which was better than sites 4 and 5, and much better than sites 2, 3, and 6. Site 2 was the least preferred site with a utility of 0.0424. Sites 4 and 5 were less preferred, and they were close to each other, but site 5 was slightly better. Similarly, sites 3 and 6 were very close to each other, but site 6 was slightly better. Overall, the six suitable sites yielded rather low utilities but succeeded in ranking the alternatives and selecting the most preferred. Site 1,

characterized by the most preferred well field location score, d_{90}/d_m ratio of streambed material, and groundwater quality score, yielded the highest utility among all suitable sites. The final decision or recommendation was to choose site 1 for implementing a new RBF system.

4. Discussion

4.1 A discussion of the kriging results

A comparison of spatial interpolation with 39 borehole records ($N=39$) has been shown in Figures 59 and 60. Inverse distance weighting (IDW) with a power parameter (p) of 2 was applied here for comparison with ordinary kriging (OK). The results were rather similar. As can be seen from the result maps above, the simple neighbor-search strategy of radius equal to 3500 meters failed to estimate all locations on the map, whereas searching with at least four neighbors was able to fix this limitation; however, the estimation errors or uncertainties increased with distance between the estimated location and its neighbors, and it was true for both of the search strategies. Searching with at least four neighbors smoothed out the estimated values and the variances of OK. The goal of the spatial interpolation was to estimate the saturated thickness values that are not far away from the Second Songhua River, as was illustrated in the maps. Luckily, most of these borehole records are located relatively dense close to the Second Songhua River, which means a good estimation for our target areas.

The results of spatial interpolation with the whole dataset ($N=40$, including outlier) were not shown here; a discussion instead was made by conducting cross-validation.

Cross-validation was used as a qualitative tool to examine the accuracy of the geostatistical model. The discussion consisted of a comparison of the residuals and a comparison of the estimated values and the true values. The influence of the outlier mentioned earlier was checked, meaning the interpolations were made between two datasets—the log-transformed borehole records of saturated thickness, excluding the outlier ($N=39$) and the original log-transformed borehole records ($N=40$). All samples were cross-validated due to the inadequate sample size.

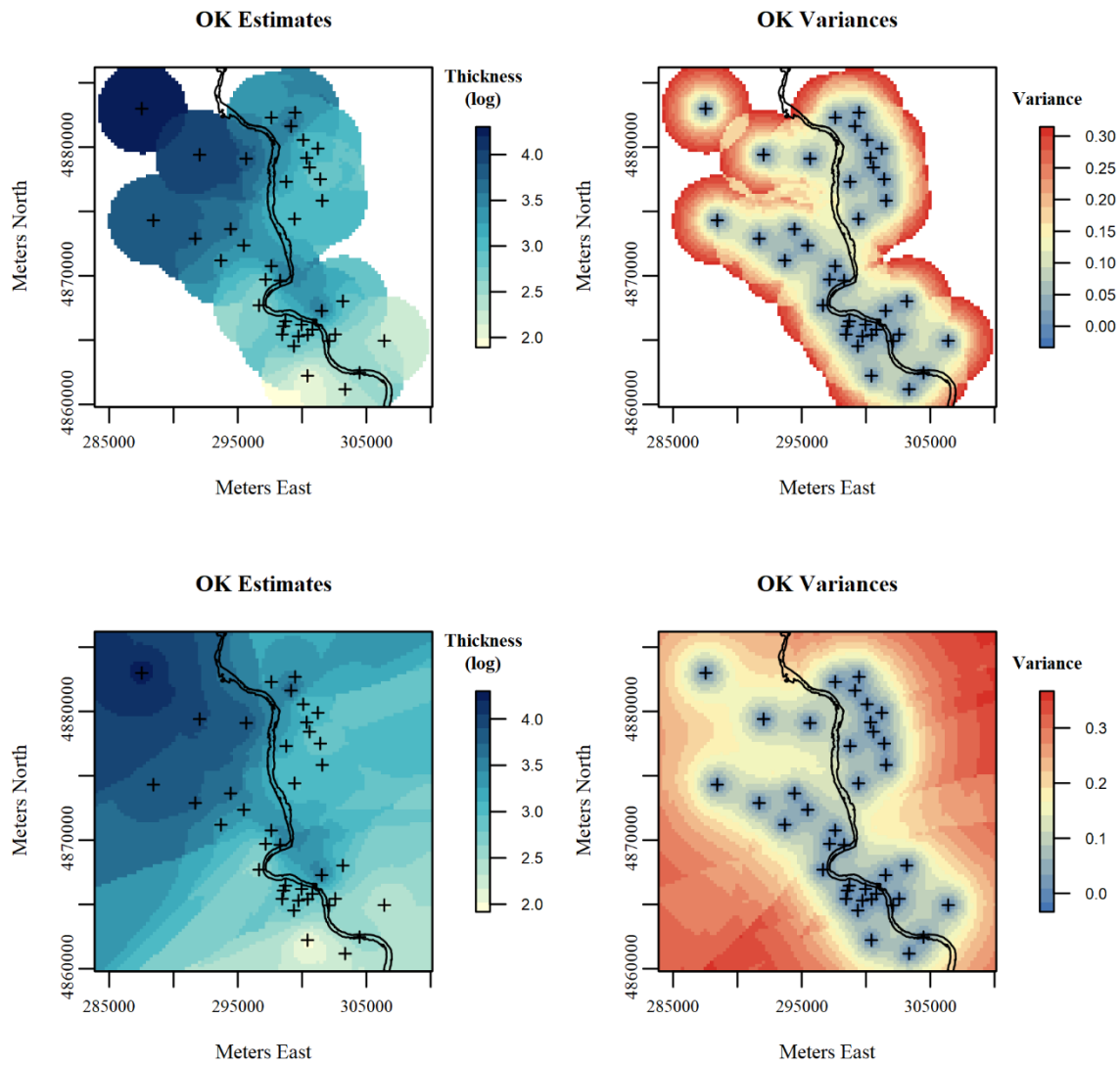


Figure 59: Comparison of two search-strategies using ordinary kriging ($N=39$)

Note: top two graphs are results of search-strategy with $R = 3500$ m; bottom two graphs with $R = 3500$ m but at least 4 neighbors. The Second Songhua River and borehole records (cross) are also shown on the maps.

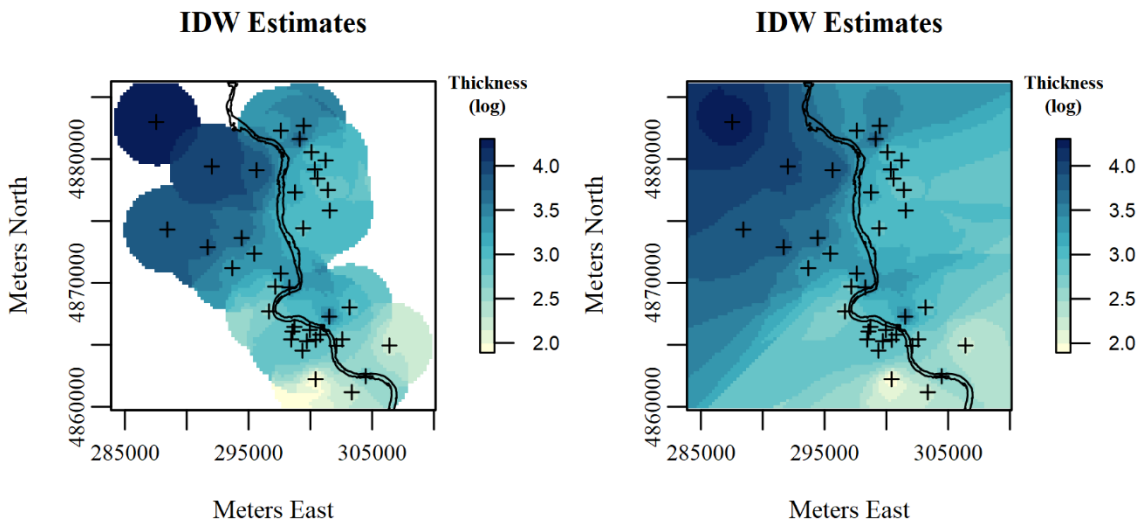


Figure 60: Comparison of two search-strategies using inverse distance weighting ($N=39$)

Note: $R = 3500m$ (left), and $R = 3500m$ but at least 4 neighbors (right).

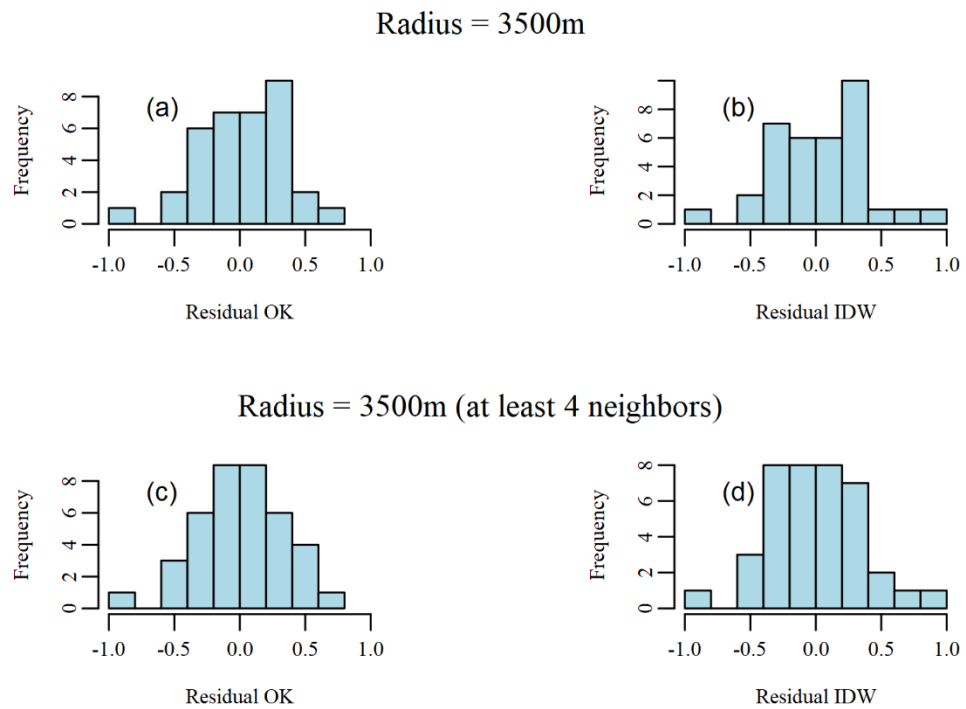


Figure 61: Residual distribution ($N=39$)

Note: (a), (c) ordinary kriging; (b), (d) inverse distance weighting, power parameter (p) = 2.

Table 26: A comparison of the residuals ($N=39$)

	Radius = 3500 m		Radius = 3500m (at least 4 neighbors)	
	OK	IDW	OK	IDW
N	35	35	39	39
m	0.018	0.025	-0.003	-0.002
σ	0.334	0.346	0.337	0.342
IQR	0.442	0.467	0.419	0.479
M	0.053	0.045	0.018	-0.002
MAE	0.262	0.263	0.263	0.259
MSE	0.109	0.117	0.111	0.114

Note: ordinary kriging (OK), inverse distance weighting (IDW); for IDW, power parameter (p) = 2, which is the power parameter, greater values of p indicates greater influence to values close to the predicted location; N is the number of cross-validated samples; m is the mean value of the dataset; σ is the standard deviation; IQR is the interquartile range, $IQR = Q3 - Q1$, where $Q1$ is the lower quartile, $Q3$ is the upper quartile; M is the median; MAE is the mean absolute error; MSE is the mean squared error, henceforth.

For the case of $N=39$, four samples in the northwest failed to be cross-validated if the radius was only set to 3500 m (referring to Figure 53). This was because of the large distances (> 3500 m) between those samples and their nearest neighbors. In Figure 61, the error distribution of four different approaches were more or less symmetrical.

In fact, the means (m) and medians (M) were both close to 0, indicating a balance of overestimates and underestimates (Table 26). It was easy to notice that searching with at least four neighbors yielded preferable results whose means and medians were closer to 0; however, there was no evidence to indicate which search-strategy was better based on other estimation criteria.

The variances (σ) of the errors were fairly close as well, from which ordinary kriging yielded smaller variances indicating a slightly better estimation. The same went for the interquartile ranges (IQR). As commonly-used estimation criteria, the mean absolute errors (MAE) and mean squared errors (MSE) incorporate both the bias (the mean) and the spread of the error distribution (the variance) (Isaaks and Srivastava 1989). Based on the values of MSE , ordinary kriging was slightly better than inverse distance weighting. Based on MAE , however, we could not conclude which method was better.

A comparison was made of predicted to true values for an additional check (Table 27). Searching with at least four neighbors yielded preferable results, based on the correlation. The same estimation criterion also indicated that ordinary kriging was also slightly better than inverse distance weighting. Scatterplots showed bivariate distributions of estimated and true values (Figure 62), which offered a way of checking *conditional bias* (Isaaks and Srivastava 1989). For all conditions, the estimated values matched the true values in a

Table 27: A comparison of the predicted values to the true values ($N=39$)

	Radius = 3500m			Radius = 3500m (at least 4 neighbors)		
	True	OK	IDW	True	OK	IDW
N	35	35	35	39	39	39
m	3.00	3.02	3.02	3.10	3.10	3.10
σ	0.44	0.33	0.29	0.52	0.38	0.35
CV	0.15	0.10	0.10	0.17	0.12	0.11
min	1.89	2.35	2.33	1.89	2.34	2.37
$Q1$	2.80	2.86	2.88	2.84	2.87	2.88
M	2.92	2.97	2.95	2.94	2.98	2.99
$Q3$	3.31	3.23	3.21	3.44	3.42	3.37
max	3.82	3.58	3.57	4.30	3.96	3.94
r		0.65	0.61		0.76	0.75

Note: CV is the coefficient of variation; min is the minimum of the dataset; max is the maximum; r is the correlation coefficient between the true values and predicted values.

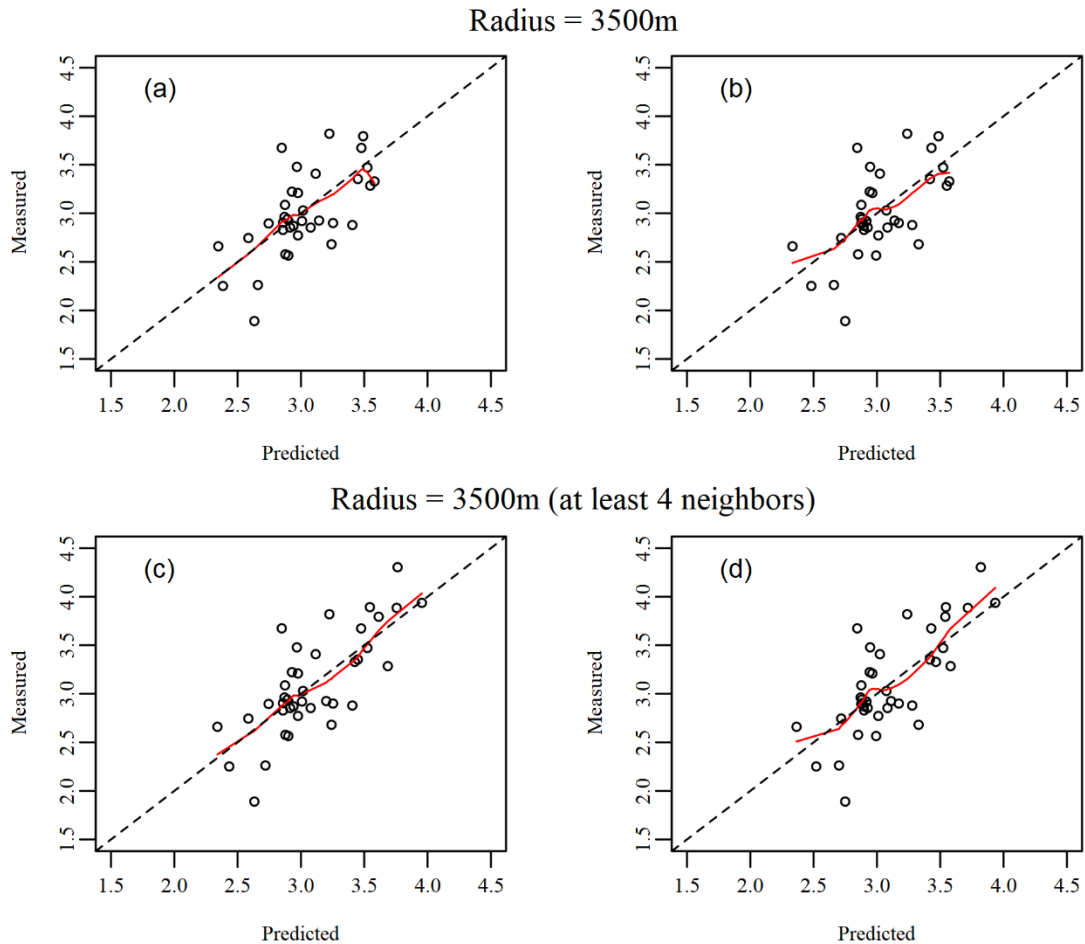


Figure 62: Cross validation ($N=39$)

Note: (a) and (c) ordinary kriging; (b) and (d) inverse distance weighting; conditional expectation curves are shown as red lines.

similar manner—they all gathered close to the 45-degree line. It was hard to tell which one was better. Adding a conditional expectation curve (red line) gave us additional information, which was often more useful than just looking at a scatterplot. The conditional expectation curve is a kind of *smoother*. Friedman’s “super smoother” was adopted here, which was achieved using package “graphics” in code R (Friedman 1984). Estimates using OK with at least four neighbors looked preferable than others since the conditional expectation curve fell closer to the 45-degree line.

Estimations with the whole dataset ($N=40$, including the outlier) yielded greater variances, *MAE*, and *MSE*; thus, they were less accurate than estimations without the outlier (see Table 28). Histograms of the residuals (Figure 63) showed also a wider error distribution than the previous results (Figure 61). Lower correlation coefficients of predicted to true values were reported (see Table 29). Furthermore, the outlier created some erratic phenomena when looking at the conditional expectation curves, especially in the highest and lowest classes (both ends of the curve) (see Figure 64).

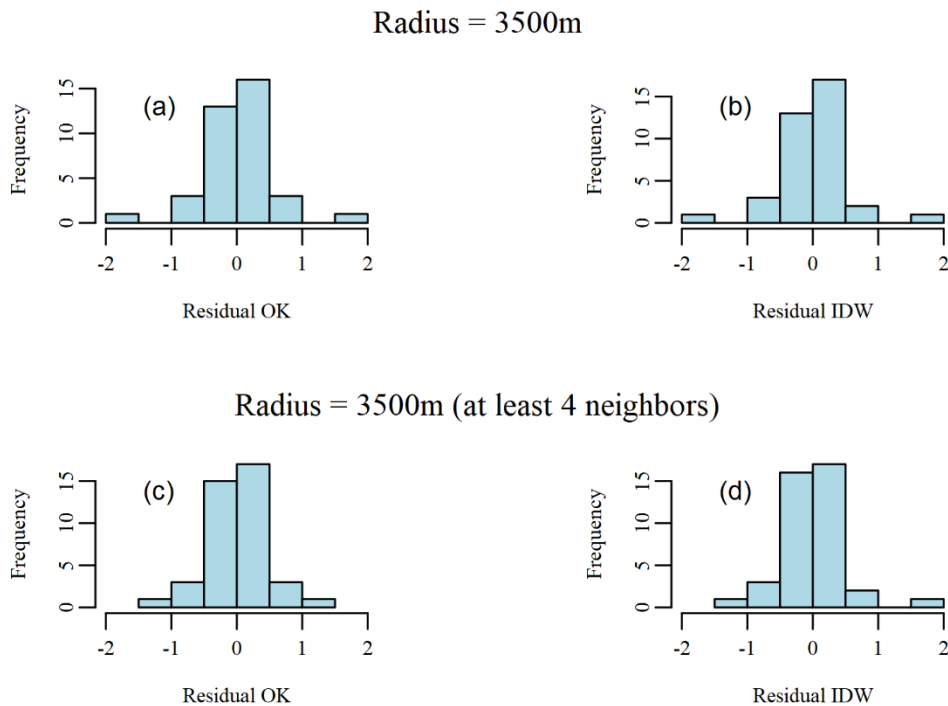


Figure 63: Residual distribution ($N=40$)

A sample size of 40 borehole records is rather small, which may affect the reliability, or accuracy, of the model. For this reason, caution should be taken when attempting to remove some outliers for prediction. A better way might be to explicitly remove the outliers first for the modeling; then, to use the whole dataset (outliers included) in the prediction

Table 28: Residual distribution ($N=40$)

	Radius = 3500m		Radius = 3500m (at least 4 neighbors)	
	OK	IDW	OK	IDW
N	37	37	40	40
m	0.017	0.024	0.018	0.015
σ	0.516	0.523	0.442	0.479
IQR	0.466	0.485	0.438	0.491
M	0.056	0.045	0.024	0.011
MAE	0.341	0.340	0.310	0.317
MSE	0.259	0.266	0.191	0.224

(Krause 2012); however, the removed outlier was of minor importance to our estimation, since it was located rather far from the Second Songhua River as well as other samples. Furthermore, areas that are close to the Second Songhua River are considered more important for RBF. Another way to check the kriging results is to only cross-validate the samples close to the Second Songhua River, for example, no farther than 2 km. One must also be aware that locations far from existing borehole records may be filled with great uncertainty; therefore, they are less reliable.

The original saturated thickness dataset looked lognormal, but the fact is, the samples were rather unevenly collected during previous investigations. It was not difficult to notice that a bunch of samples with relatively lower values were densely gathered in the southeastern area, while in the northwestern area, the samples with relatively higher values were fewer and sparse. This may result in the lognormal distribution of the dataset, rather than the nature of the saturated thickness dataset.

The omnidirectional variogram model was applied in this study; however, a check of directional influence was necessary, which may increase the estimation accuracy. The aquifer seems to become thicker as the Second Songhua River flows from the southeast to the northwest.

Table 29: Cross validation ($N=40$)

	Radius = 3500m			Radius = 3500m (at least 4 neighbors)		
	True	OK	IDW	True	OK	IDW
N	37	37	37	40	40	40
m	3.02	3.04	3.05	3.09	3.11	3.10
σ	0.48	0.38	0.36	0.52	0.39	0.38
CV	0.16	0.13	0.12	0.17	0.13	0.12
min	1.89	2.35	2.33	1.89	2.34	2.37
$Q1$	2.77	2.86	2.88	2.82	2.87	2.88
M	3.02	2.97	2.95	2.93	2.99	2.98
$Q3$	3.33	3.24	3.24	3.42	3.41	3.35
max	4.30	4.30	4.30	4.30	4.09	4.23
r		0.30	0.25		0.56	0.46

Nevertheless, these results were still considered acceptable. The purpose of kriging was to provide a saturated thickness map necessary for RBF site selection and evaluation; a more detailed investigation of the estimation approach was out of the scope of this study.

If important information is not available at the places to be evaluated, kriging can be an effective tool to fill in the blanks. For instance, the created saturated thickness map is fundamental for RBF system implementation, which also influences the performance of a RBF system as well.

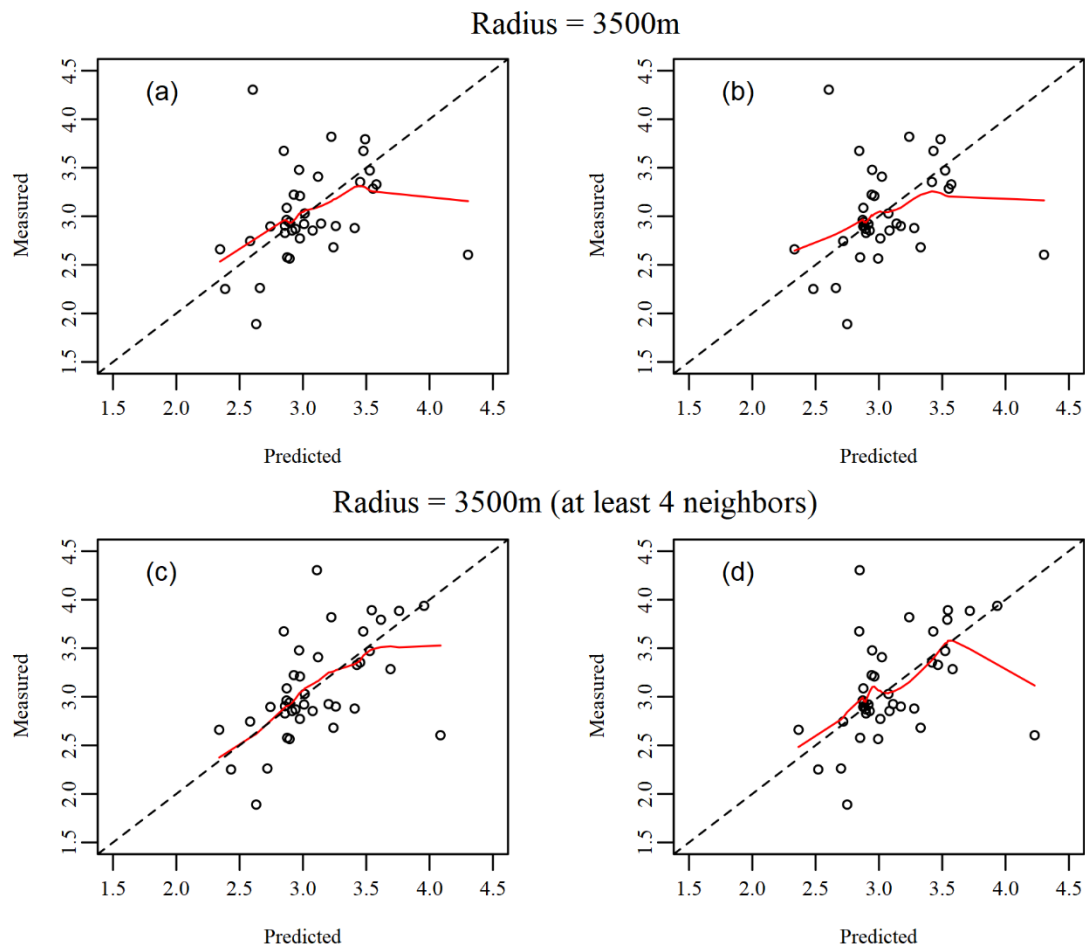


Figure 64: Cross validation ($N=40$)

Note: (a) and (c) ordinary kriging; (b) and (d) inverse distance weighting.

4.2 A discussion of the multi-attribute utility model results

The goal was to select the best (most preferred) site for RBF; thus, a sensitivity analysis was performed, which focused mainly on the conditions under which the best site changes. It was conducted by examining the influence of assigned weights and uncertainty of the attribute values.

Rank order weights. Considering tradeoffs between two attributes is a rather difficult task, thus, the underlying uncertainty may rise significantly, or in other words, the procedure may tend to be rather unstable. For this reason, a sensitivity analysis was made to check whether the relative importance between the attributes may influence our final choice. Butler *et al.* (1997) and Jiménez *et al.* (2003) argued that a sensitivity analysis on rank ordered weights is more meaningful than weights that were generated completely randomly because it is consistent with the decision maker's judgement of attribute importance. To do this, the order of scaling constants or weights in equation (17) was kept unchanged, and then the relative weights of k_i to k_4 , $i = 1, 2, 3, 5$, were randomly generated. This changed the values of x_4 at left-hand side of the indifferent pairs in Table 15 (regardless of the trade-offs made between X_1 and X_2). By doing this, the degree to which a single attribute is more important than others is random. Additionally, the probability p in Figure 27 was adapted for consistency. The probability p was important, because it was used to determine k_4 . The model was tested to see if the rank of suitable sites would change, more importantly, to see if the most preferred one will change. To do this, the simulations were run 10 times, meaning 10 sets of the scaling constants were generated. Scaling constants and probability p of these simulations are shown in Table 30.

Table 30: Scaling constants and p from 10 simulations

Run	k_1	k_2	k_3	k_4	k_5	k	p
1	8.47E-05	4.55E-05	4.55E-05	8.78E-05	8.26E-05	1.47E+05	0.071
2	3.04E-01	2.47E-01	2.47E-01	4.12E-01	2.92E-01	-6.88E-01	0.568
3	2.68E-02	1.11E-02	1.11E-02	3.10E-02	1.87E-02	7.27E+01	0.328
4	6.70E-02	2.33E-02	2.33E-02	7.32E-02	4.64E-02	1.81E+01	0.395
5	1.62E-02	3.20E-03	3.20E-03	4.55E-02	9.04E-03	1.57E+02	0.336
6	8.09E-01	2.01E-01	2.01E-01	8.39E-01	6.20E-01	-9.92E-01	0.868
7	1.90E-03	2.85E-04	2.85E-04	3.08E-03	1.16E-03	4.46E+03	0.148
8	1.61E-02	2.54E-03	2.54E-03	2.22E-02	5.33E-03	2.75E+02	0.357
9	6.84E-03	2.89E-03	2.89E-03	7.00E-03	4.21E-03	5.66E+02	0.250
10	8.28E-01	6.88E-01	6.88E-01	9.30E-01	8.16E-01	-1.00E+00	0.855

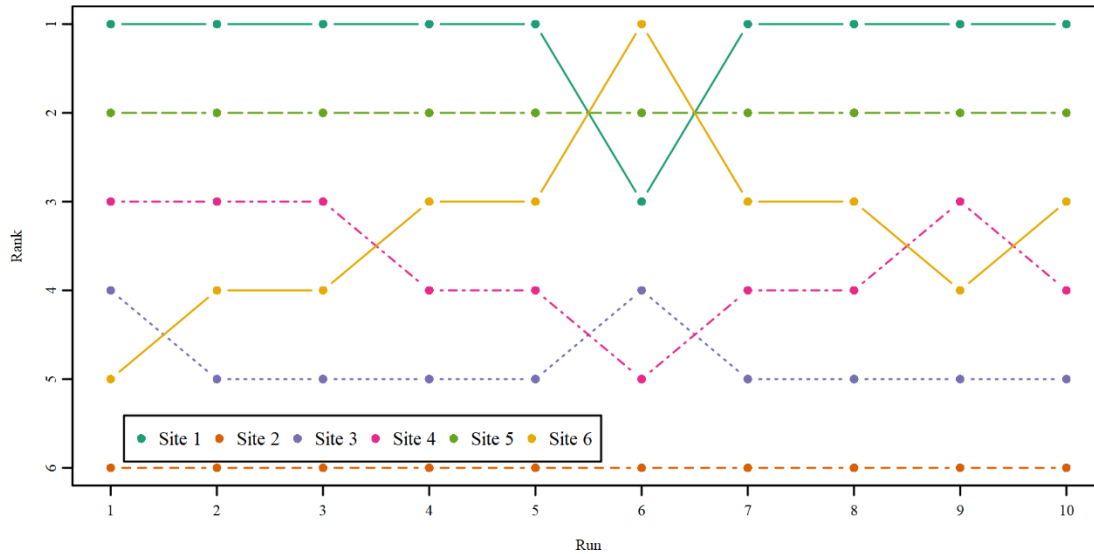


Figure 65: Ranks of suitable sites after 10 simulations

The ranks of the suitable sites of 10 simulations are shown in Figure 65. Only three simulations (run 2, 3, and 9) agreed with the previous results. As the scaling constants changed, the most preferred also varied between different sites; however, only one simulation yielded a different, most preferred site other than site 1. It turned out that site 6 could also rank as the most preferred. During these 10 simulations, the ranks of four sites did not remain constant, whereas site 2 ranked the least preferred and site 5 ranked the second position consistently. Site 3 was rather stable, and moved only one position. Site 1 and site 4 moved two positions, whereas site 6 moved as much as four positions. The worst rank of site 6 was at position 5, while its best rank was at position 1. This could be interpreted as site 6 was more sensitive to the changing scaling constants than any other suitable sites. However, site 1 ranked the best site more frequently than site 6.

Then, 10 000 simulations were performed to analyze the distributions of ranks for all suitable sites; the result is shown in Figure 66 and Table 31.

The results showed that as the scaling constants altered (their orders kept unchanged), there was about a 78% chance that site 1 was still the most preferred, which was consistent to the original recommendation. Site 5 and 6 could also be our final choice; however, they were much less frequently ranked as the most preferred. Site 6 was rather more sensitive to the scaling constants assigned than the other two candidates. Nevertheless, the most preferred site should be chosen from site 1, 5, and 6; other sites were not competitive.

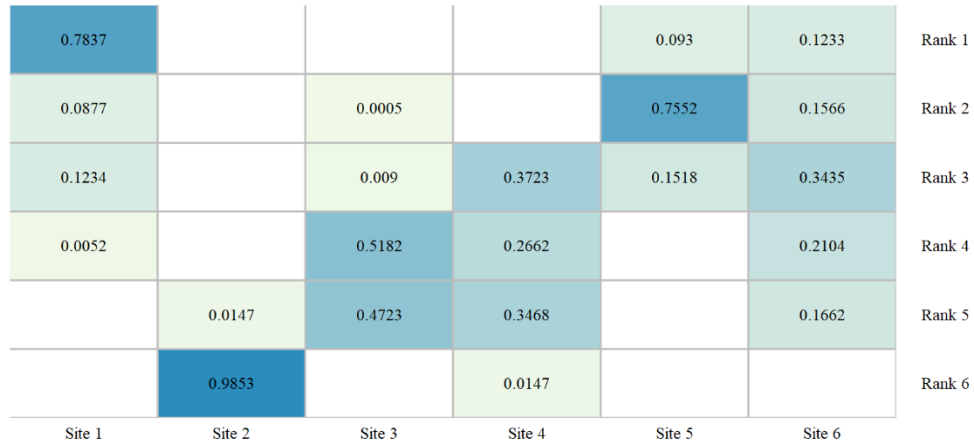


Figure 66: A heat map of the rank frequency of six suitable sites

Note: number in each cell means frequency, empty cells indicate no occurrence.

Table 31: Ranking results from the rank order weight simulation (six suitable sites)

	<i>min</i>	<i>Q1</i>	<i>M</i>	<i>Q3</i>	<i>max</i>	<i>m</i>	σ
Site 1	1	1	1	1	4	1.358	0.722
Site 2	5	6	6	6	6	5.986	0.120
Site 3	2	4	4	5	5	4.466	0.517
Site 4	3	3	4	5	6	3.997	0.880
Site 5	1	2	2	2	3	2.068	0.490
Site 6	1	2	3	4	5	3.145	1.224

Note: *min* is minimum, *Q1* is the lower quartile, *M* (*Q2*) is the median, *Q3* is the upper quartile, *max* is maximum, *m* is mean, σ is standard deviation.

Uncertainty of attribute values. To see if an improvement of a single attribute will lead to a change our final decision, the *relative change* and the *absolute change* of attributes can be checked (Triantaphyllou and Sánchez 1997). It is more intuitive for us to know the degree to which site 1 is better than others by providing both changes than just providing one of them. The sensitivity analysis was conducted through improving a single attribute while at the same time keeping other attributes unchanged, to see if at a certain level sites 2, 3, 4, 5, 6 will be the most preferred.

Table 32 showed the improvement of each attribute level needed for all sites other than site 1 to yield a final utility greater than the current utility of 0.1256. Among those attributes, X_3 (d_{90}/d_m ratio of streambed material) needed the lowest improvement, *i.e.*, a relative reduction of 45% for site 5. For attribute X_5 (groundwater quality score) an

additional 120%, 130%, and 412% for site 5, site 4, and site 3, respectively, will make them the most preferred. For attribute X_1 (transmissivity), an additional 769% for site 5, 1025% for site 4, 1081% for site 3, and 1910% for site 6 was needed to rank No.1. As for attribute X_4 (mean DOC concentration), any improvement within the ranges would not make a change in the most preferred site.

If measured by absolute change, site 1 was better than site 5 by 0.077 m²/s transmissivity, 1.3 d_{90}/d_m ratio of streambed material, and a groundwater quality score of 54. Site 1 was better than site 4 by 0.103 m²/s transmissivity and a groundwater quality score of 46.8. Site 1 was much better than site 3 by 0.249 m²/s transmissivity and a groundwater quality score of 78.3. Similarly, site 1 was much better than site 6 by a degree of 0.210 m²/s transmissivity and a well field location score of 73. Since no improvement of any attributes can make site 2 the most preferred, site 1 was clearly better than site 2.

For transmissivity, a measurement error of 0.077 m²/s or above for the local alluvium aquifer was not likely because the transmissivity of the study area has a range of about 0.001-0.036 m²/s.

In practice, improving the well field location score of site 3 means replacing the well field from the “straight channel” to the “point bar” or “island” because the score was set subjectively and discretely in this study as 0 for wells placed at the “cut bank”, 33 for along the “straight channel”, 66 for at the “point bar”, and 100 for on the “island”. The measurement error of this attribute was therefore the uncertainty of classifying the geometry of the river at the well field location. A more careful assessment of this attribute as well as its impact on an RBF system is needed. According to Grischek (2019), well field location score should be cross-checked because the utility of which is hard to balance against others. Schubert (2019) also mentioned that although the inner section for RBF sites is preferred, the outer section is also useable if the runoff dynamics are favorable.

Similar to transmissivity, a difference between the d_{90}/d_m ratio of 2.9 and 1.6 was relatively large for the measurement error of this attribute. Groundwater quality, however, can be optimized through well head protection or artificial groundwater recharge. This gives some insights for the decision maker, in case the final site has to be chosen from the less-preferred suitable sites.

This part of the sensitivity analysis on attributes can be considered as an initial evaluation of the impacts on the value uncertainties (Keeney and Wood 1977). The decision maker can now think whether it is worth choosing site 3, 4, 5, or site 6 by improving those attributes to the optimal levels.

Table 32: Improvement of each attribute needed for each site to be ranked as the most preferred

Attribute	Scale of measure	Site 2		Site 3		Site 4		Site 5		Site 6	
		Change	Optimal level	Change	Optimal level	Change	Optimal level	Change	Optimal level	Change	Optimal level
$X_1 \equiv$ Transmissivity	m ² /s	-	-	+1081%	0.272	+1025%	0.113	+769%	0.087	+1910%	0.221
$X_2 \equiv$ Well field location score	Subjective	-	-	-	-	-	-	-	-	-	73
$X_3 \equiv d_{90}/d_m$ ratio of streambed material	Ratio scale	-	-	-	-	-	-	-45%	1.60	-	-
$X_4 \equiv$ Mean DOC concentration	mg/L	-	-	-	-	-	-	-	-	-	-
$X_5 \equiv$ Groundwater quality score	Subjective	-	-	+412%	97.3	+130%	82.8	+120%	99.0	-	-

Note: “-” denoted that an improvement of the attribute is not possible or the attribute level must be higher than its best level in order to yield a better utility than site 1. A percent change of well field location score for site 6 is not possible, because the original score was 0.

5. Summary

In this part, the geographic information system (GIS) and multi-attribute utility theory (MAUT) was used for RBF site selection from a case study in Jilin City, China, where drinking water sources were rather vulnerable to pollution. Six suitable sites for RBF implementation along the Second Songhua River were delineated using GIS technique and evaluated using the multi-attribute utility model developed in Part III. The utilities of those suitable sites were calculated, followed by a successful ranking based on the calculations. The results showed that one site among those suitable sites yielded a significantly higher utility than others; therefore, it was recommended as the final choice for the decision maker. A sensitivity analysis was performed based on randomly generated scaling constants and uncertainties of attribute values. The result was consistent with the previous recommendation.

Geostatistical techniques were used to build the saturated aquifer thickness map for delineating the suitable RBF sites and assigning attributes for them. Ordinary kriging (OK) was used to estimate the values at unknown locations, which yielded better results than inverse distance weighting (IDW).

Part V
Conclusions and Recommendations

1. Conclusion and Recommendation

The research aimed to develop a new method for site selection of RBF system used for drinking water supply. Selecting a site for an RBF system is a decision with multiple objectives, and some physical parameters must be obtained in order to clearly reflect those objectives. Geostatistical methods and GIS are very useful tools in the acquisition of data necessary for the decision problem. Based on the GIS data obtained from various sources and methods such as kriging and buffering, suitable sites for RBF implementation in an urban area were identified.

Despite restricted data quality and quantity, further evaluation using the multi-attribute utility model developed in this study successfully ranked suitable sites. The model proved to be useful in evaluating the performance of RBF sites. The decision maker's preference was reflected throughout the analysis. The multi-attribute utility model took the uncertainty during the assessment of the relative importance between attributes into consideration. The result of a weight-based (scaling constant) sensitivity analysis was consistent with our final recommendation.

Due to the absence of site information of the empirical studies on RBF site selection (*e.g.*, Lee *et al.* (2010)), a comparison of the methods cannot be made. The individual-based decision analysis was a limitation. Decisions made based on a consensus of a group or *collective choice* by multiple groups of interest (*e.g.*, geologists, engineers, residents) could improve the trustworthiness of the model (Sen 1970; Dyer and Miles 1974). No matter if the final decision was individual-based or group-based, the study showed the possibility of using a new technique on RBF site selection.

Appendix 1

Environmental quality standards for surface water (GB 3838-2002)

Table A1: Basic indicators and limitations of environmental quality standards for surface water

No.	Indicator	Category and values				
		I	II	III	IV	V
1	Water temperature (°C)	The human induced changes in ambient water temperature should be limited to: Weekly average maximum temperature rise ≤ 1 Weekly average maximum temperature drop ≤ 2				
2	pH (Dimensionless)	6~9	6~9	6~9	6~9	6~9
3	Dissolved oxygen	\geq Saturation 90% (or 7.5)	≥ 6	≥ 5	≥ 3	≥ 2
4	Permanganate index	≤ 2	≤ 4	≤ 6	≤ 10	≤ 15
5	Chemical oxygen demand (COD)	≤ 15	≤ 15	≤ 20	≤ 30	≤ 40
6	Biological oxygen demand (BOD ₅)	≤ 3	≤ 3	≤ 4	≤ 6	≤ 10
7	Ammonia nitrogen (NH ₃ -N)	≤ 0.15	≤ 0.5	≤ 1.0	≤ 1.5	≤ 2.0
8	Total phosphorus (calculated in P)	≤ 0.02 (lake and reservoir 0.01)	≤ 0.1 (lake and reservoir 0.025)	≤ 0.2 (lake and reservoir 0.05)	≤ 0.3 (lake and reservoir 0.1)	≤ 0.4 (lake and reservoir 0.2)
9	Total nitrogen (lake and reservoir, calculated in N)	≤ 0.2	≤ 0.5	≤ 1.0	≤ 105	≤ 2.0
10	Cu	≤ 0.01	≤ 1.0	≤ 1.0	≤ 1.0	≤ 1.0
11	Zn	≤ 0.05	≤ 1.0	≤ 1.0	≤ 2.0	≤ 2.0
12	Fluoride (calculated in F ⁻)	≤ 1.0	≤ 1.0	≤ 1.0	≤ 1.5	≤ 1.5
13	Se	≤ 0.01	≤ 0.01	≤ 0.01	≤ 0.02	≤ 0.02
14	As	≤ 0.05	≤ 0.05	≤ 0.05	≤ 0.1	≤ 0.1
15	Hg	≤ 0.00005	≤ 0.00005	≤ 0.0001	≤ 0.001	≤ 0.001
16	Cd	≤ 0.001	≤ 0.005	≤ 0.005	≤ 0.005	≤ 0.01
17	Cr (hexavalent)	≤ 0.01	≤ 0.05	≤ 0.05	≤ 0.05	≤ 0.1
18	Pb	≤ 0.01	≤ 0.01	≤ 0.05	≤ 0.05	≤ 0.1
19	Cyanide	≤ 0.005	≤ 0.05	≤ 0.2	≤ 0.2	≤ 0.2
20	Volatile phenol	≤ 0.002	≤ 0.002	≤ 0.005	≤ 0.01	≤ 0.1
21	Petroleum	≤ 0.05	≤ 0.05	≤ 0.05	≤ 0.5	≤ 1.0
22	An-ionic surfactant	≤ 0.2	≤ 0.2	≤ 0.2	≤ 0.3	≤ 0.3
23	Sulfide	≤ 0.05	≤ 0.1	≤ 0.2	≤ 0.5	≤ 1.0
24	Fecal coliform bacteria (microbiota per liter)	≤ 200	≤ 2000	≤ 10000	≤ 20000	≤ 40000

Note: units in milligram per liter (if not specified), water quality category I is better than category V. Source: MEE & AQSIQ(2002).

Surface water quality category was evaluated by MEE and was based on the 21 indicators (except water temperature, total nitrogen, and fecal coliform bacteria) in Table A1 according to various standard limits (MEE (Ministry of Ecology and Environment of People's Republic of China) 2016). Then, single-factor evaluation method was used, and the worst water quality category is evaluated as the section water quality category.

Appendix 2

Quality standard for groundwater (GB14848-93)

Table A2: Indicators of groundwater quality

No.	Indicator	Category and values				
		I	II	III	IV	V
1	Colority (Hazen unit)	≤5	≤5	≤15	≤25	>25
2	Smell and taste	No	No	No	No	Yes
3	Turbidity (NTU)	≤3	≤3	≤3	≤10	>10
4	Visible matters to the naked eye	No	No	No	No	Yes
5	pH	6.5~8.5	6.5~8.5	6.5~8.5	5.5~6.5, 8.5~9	<5.5, >9
6	Total hardness (CaCO ₃)	≤150	≤300	≤450	≤550	>550
7	Total soluble solid	≤300	≤500	≤1000	≤2000	>2000
8	Sulfate	≤50	≤150	≤250	≤350	>350
9	Chloride	≤50	≤150	≤250	≤350	>350
10	Fe	≤0.1	≤0.2	≤0.3	≤1.5	>1.5
11	Mn	≤0.05	≤0.05	≤0.1	≤1.0	>1.0
12	Cu	≤0.01	≤0.05	≤1.0	≤1.5	>1.5
13	Zn	≤0.05	≤0.5	≤1.0	≤5.0	>5.0
14	Mo	≤0.001	≤0.01	≤0.1	≤0.5	>0.5
15	Co	≤0.005	≤0.05	≤0.05	≤1.0	>1.0
16	Volatile phenol (Benzene)	≤0.001	≤0.001	≤0.002	≤0.01	>0.01
17	An-ionic surfactant	Not Detected	≤0.1	≤0.3	≤0.3	>0.3
18	Permanganate index	≤1.0	≤2.0	≤3.0	≤10	>10
19	Nitrate (calculated in N)	≤2.0	≤5.0	≤20	≤30	>30
20	Nitrite (calculated in N)	≤0.001	≤0.01	≤0.02	≤0.1	>0.1
21	Ammonia nitrogen (NH ₄)	≤0.02	≤0.02	≤0.2	≤0.5	>0.5
22	Fluoride	≤1.0	≤1.0	≤1.0	≤2.0	>2.0
23	Iodide	≤0.1	≤0.1	≤0.2	≤1.0	>1.0
24	Cyanide	≤0.001	≤0.01	≤0.05	≤0.1	>0.1
25	Hg	≤0.00005	≤0.0005	≤0.001	≤0.001	>0.001
26	As	≤0.005	≤0.01	≤0.05	≤0.05	>0.05
27	Se	≤0.01	≤0.01	≤0.01	≤0.1	>0.1
28	Cd	≤0.0001	≤0.001	≤0.01	≤0.01	>0.01
29	Cr (hexavalent)	≤0.005	≤0.01	≤0.05	≤0.1	>0.1
30	Pb	≤0.005	≤0.01	≤0.05	≤0.1	>0.1
31	Be	≤0.00002	≤0.0001	≤0.0002	≤0.001	>0.001
32	Ba	≤0.01	≤0.1	≤1.0	≤4.0	>4.0
33	Ni	≤0.005	≤0.05	≤0.05	≤0.1	>0.1
34	DDT (μg/L)	Not Detected	≤0.005	≤1.0	≤1.0	>1.0
35	HCH (μg/L)	≤0.005	≤0.05	≤5.0	≤5.0	>5.0
36	Total coliform bacteria (microbiota per liter)	≤3.0	≤3.0	≤3.0	≤100	>100
37	Total bacteria (number per liter)	≤100	≤100	≤100	≤1000	>1000
38	Total α radioactivity (Bq/L)	≤0.1	≤0.1	≤0.1	>0.1	>0.1
39	Total β radioactivity (Bq/L)	≤0.1	≤1.0	≤1.0	>1.0	>1.0

Note: units in milligram per liter (if not specified), water quality category I is better than category V. Source: AQSIQ(1993).

Groundwater quality category was evaluated by MEE and was based on the 37 indicators (except total coliform bacteria and total bacteria) in Table A2 according to

various standard limits(MEE (Ministry of Ecology and Environment of People’s Republic of China) 2016).

Single-factor evaluation method was used as the first step to evaluate the quality category of each indicator. If different categories of an indicator have the same standard limits, the higher water quality category will be taken from the evaluation. If, for example, indicator value of volatile phenol (benzene) of a sample is 0.001 mg/L, it will be evaluated as category I, instead of category II. Then, the category for each indicator has been assigned a rating, *i.e.*, F_i , which varies between 0 and 10 (see Table A2.2).

Table A3: Rating of each indicator

Category	I	II	III	IV	V
Rating F_i	0	1	3	6	10

Source: AQSIQ(1993).

After that, comprehensive evaluation method was used to give the final category of a water sample. First, the comprehensive rating value F of a sample was calculated using equation

$$F = \sqrt{\frac{\bar{F}^2 + F_{max}^2}{2}}, \quad (43)$$

and

$$\bar{F} = \frac{1}{n} \sum_{i=1}^n F_i \quad (44)$$

where \bar{F} is the average of ratings of indicator F_i ; F_{max} is the maximum rating of indicator F_i ; n is the number of indicators evaluated.

Based on the comprehensive rating value F , the groundwater quality category was classified according to Table A2.3.

Table A4: Category defined by comprehensive rating

Category	I	II	III	IV	V
Rating F	<0.80	0.80 to <2.50	2.50 to <4.25	4.25 to <7.20	>7.20

Source: AQSIQ(1993).

Appendix 3

Explanation to Germany's RBF site location data

Location data of RBF sites in Germany or RBF wells or well fields in Germany were acquired from the delineated drinking water protection zones (German: *Trinkwasserschutzgebiete*) in the federal state of Berlin, Saxony, North Rhine-Westphalia, and Rhineland-Palatinate. RBF well or well field data (location) was compiled from the latest drinking water protection zones given by the Berlin Senate Department for Urban Development and Housing (SenSW 2009), the Saxon State Office for Environment, Agriculture and Geology (LfULG 2018), the North Rhine-Westphalian State Agency for Nature, Environment and Consumer Protection (LANUV 2015), and the Ministry of Environment, Energy, Food and Forests of Rhineland-Palatinate (MUEEF 2019). Those wells or well fields whose protection zones (including Zone I, Zone II, and Zone III) cover or contact surface water bodies were identified as RBF wells or well fields (see Figure A1).

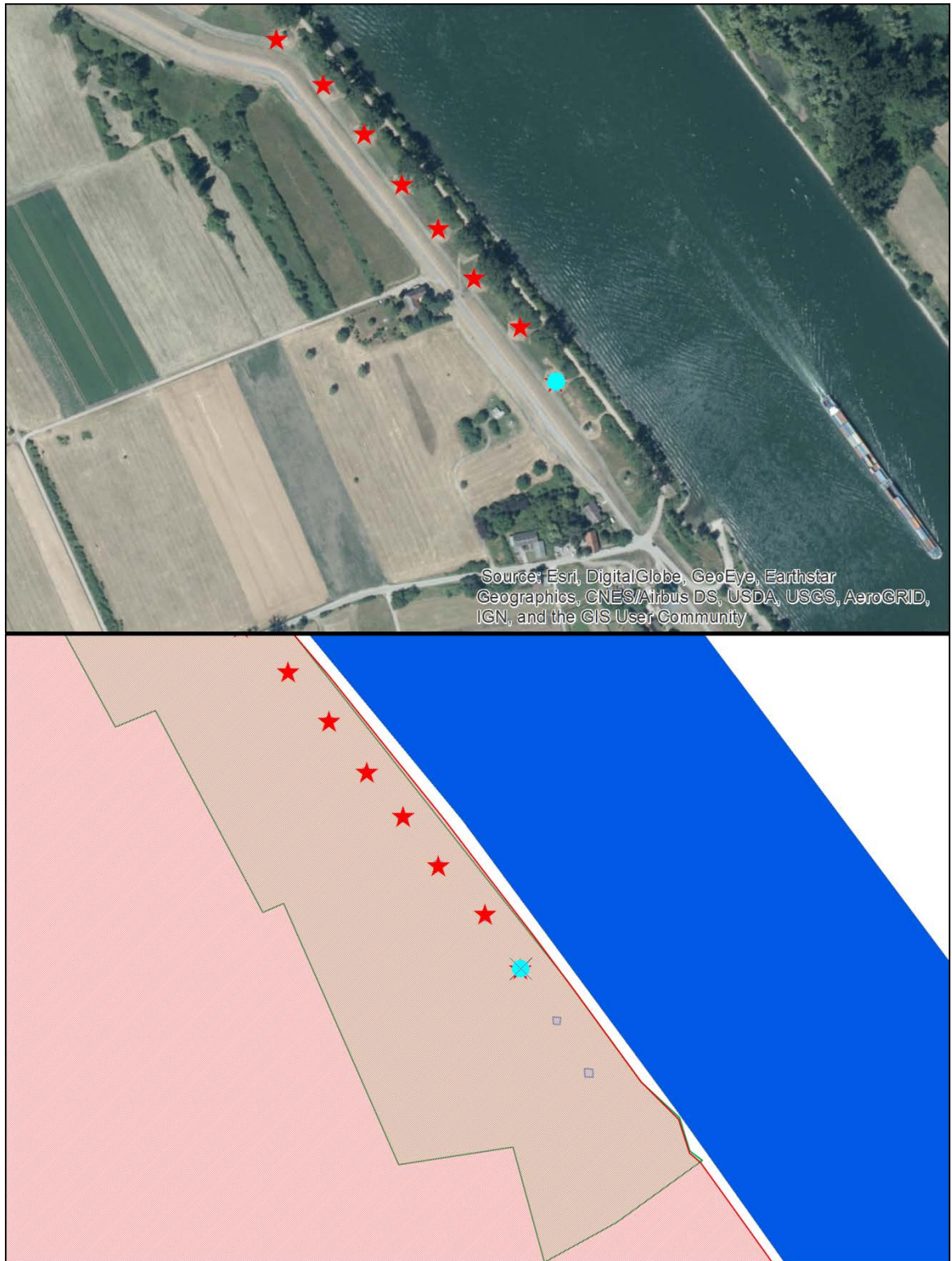


Figure A1: An illustration of RBF well identification

Note: RBF wells or well fields were marked with a star, blue square area denotes drinking water protection zone I, green area denotes zone II, red area denotes zone III. Source: MUEEF(2019).

Appendix 4

Layer information of drillings

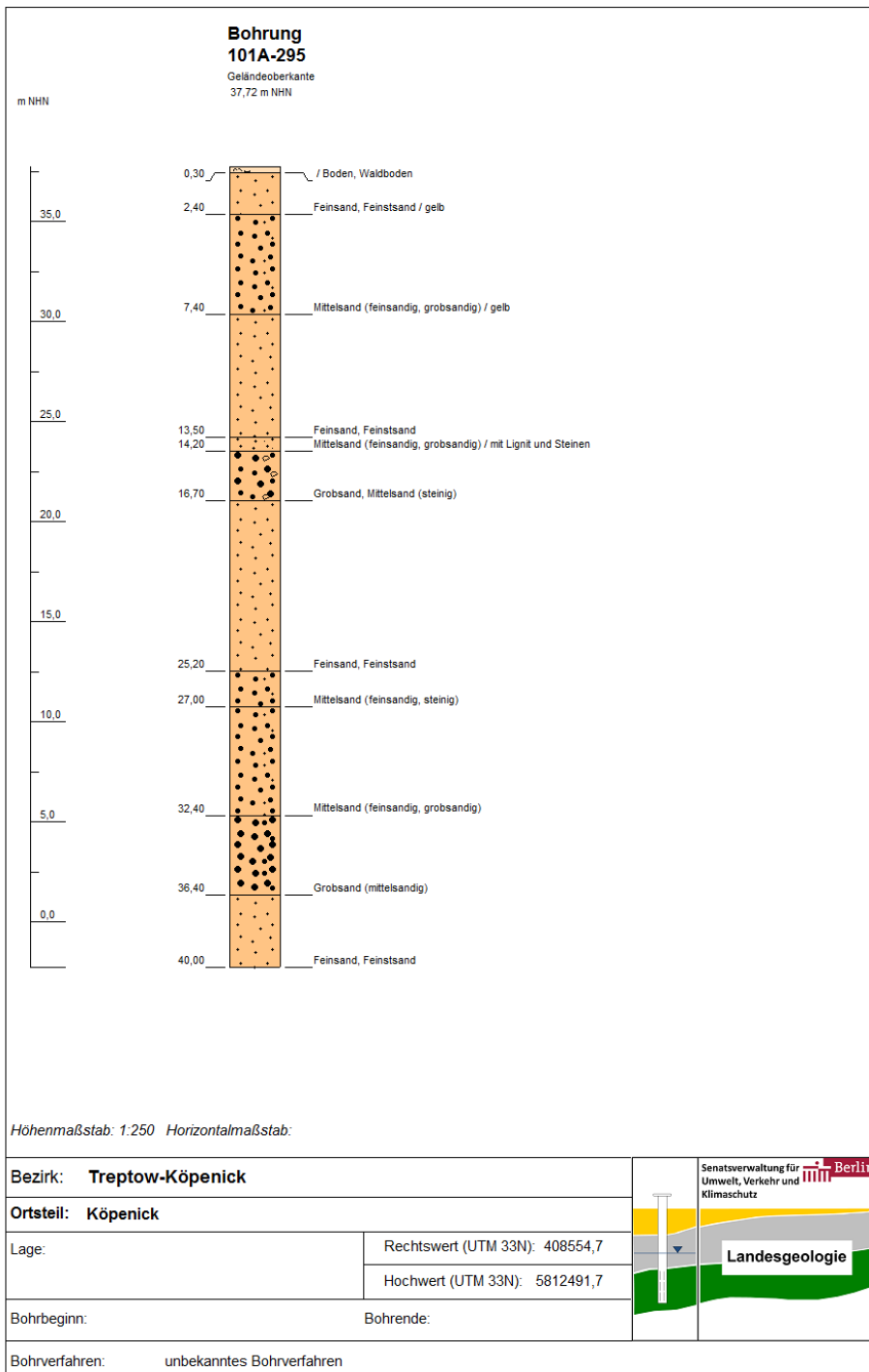


Figure A2: Layer information of drilling 1, with sandy formation on the surface, near Lake Müggelsee

Note: Boden means soil, Feinsand means fine sand, Mittelsand means medium sand, Grobsand means coarse sand.
Source: SenUVK(2020).

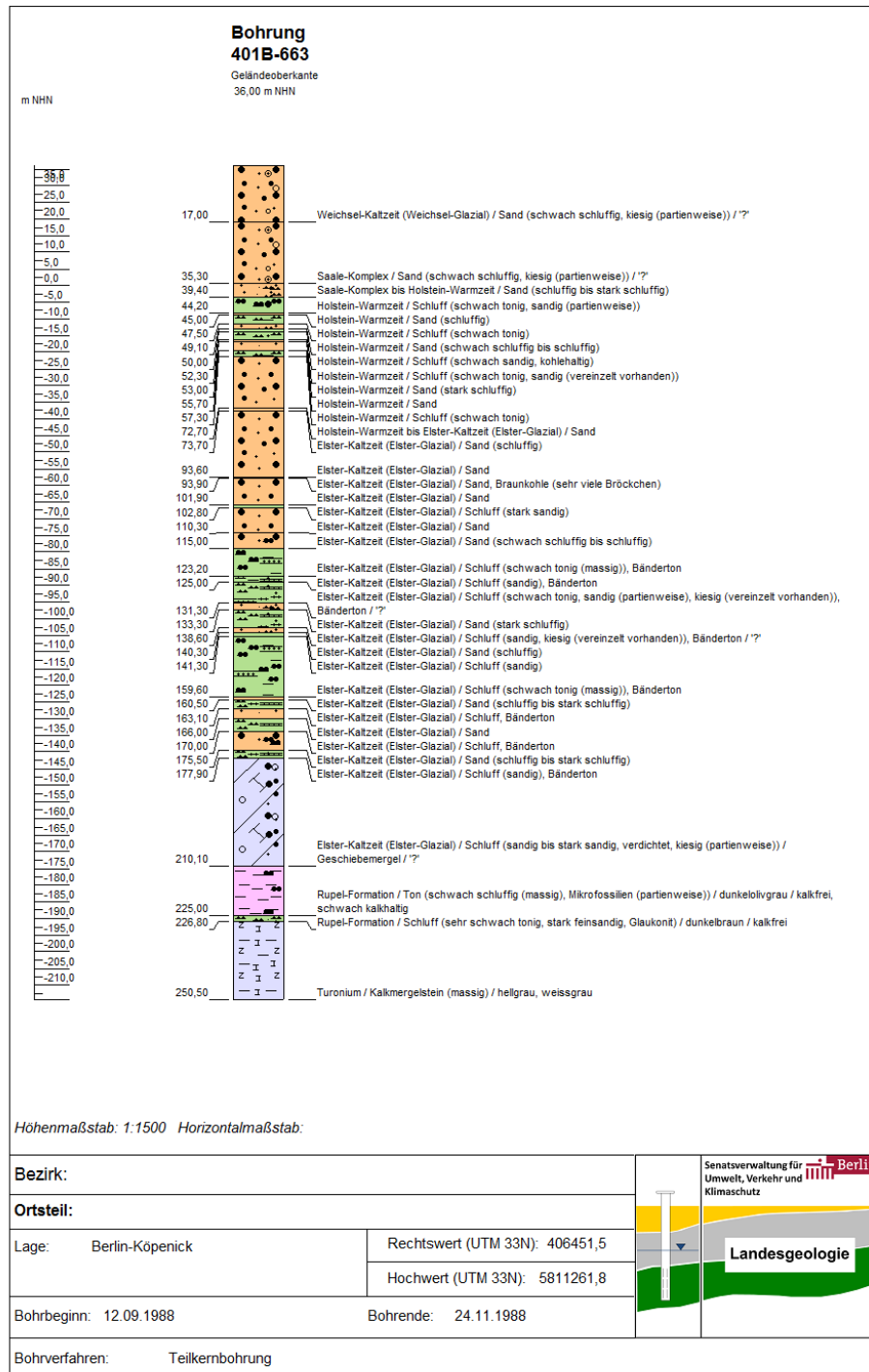


Figure A3: Layer information of drilling 2, with sandy formation on the surface, near Lake Müggelsee

Note: Weichsel-Kaltzeit means Weichsel ice age, Saale-Komplex means Saale-complex, Holstein-Warmzeit means Holstein interglacial. Source: SenUVK(2020).

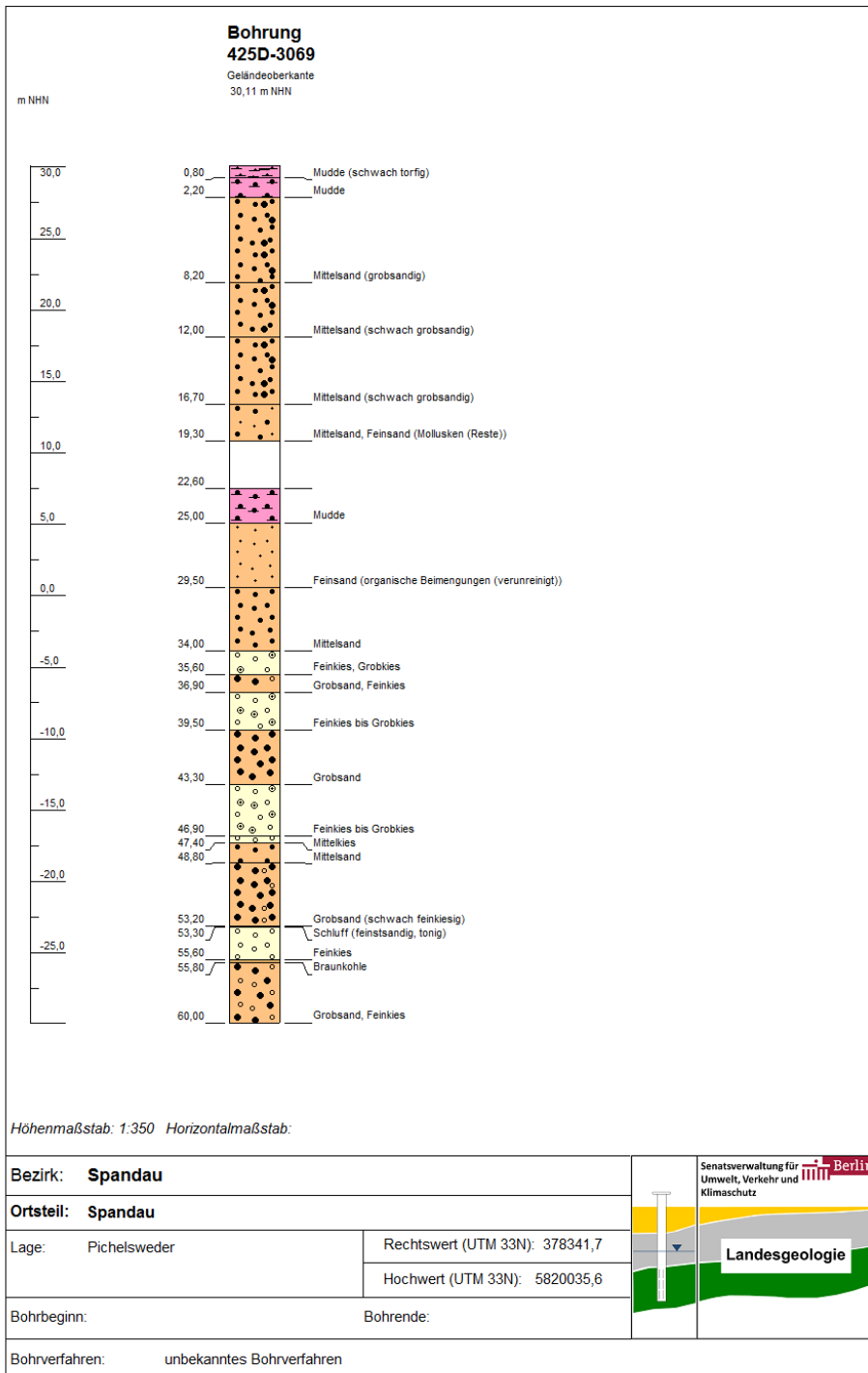


Figure A4: Layer information of drilling 3, with mud (silt or clay) formation on the surface, near Havel River

Note: Mudde means mud. Source: SenUVK(2020).

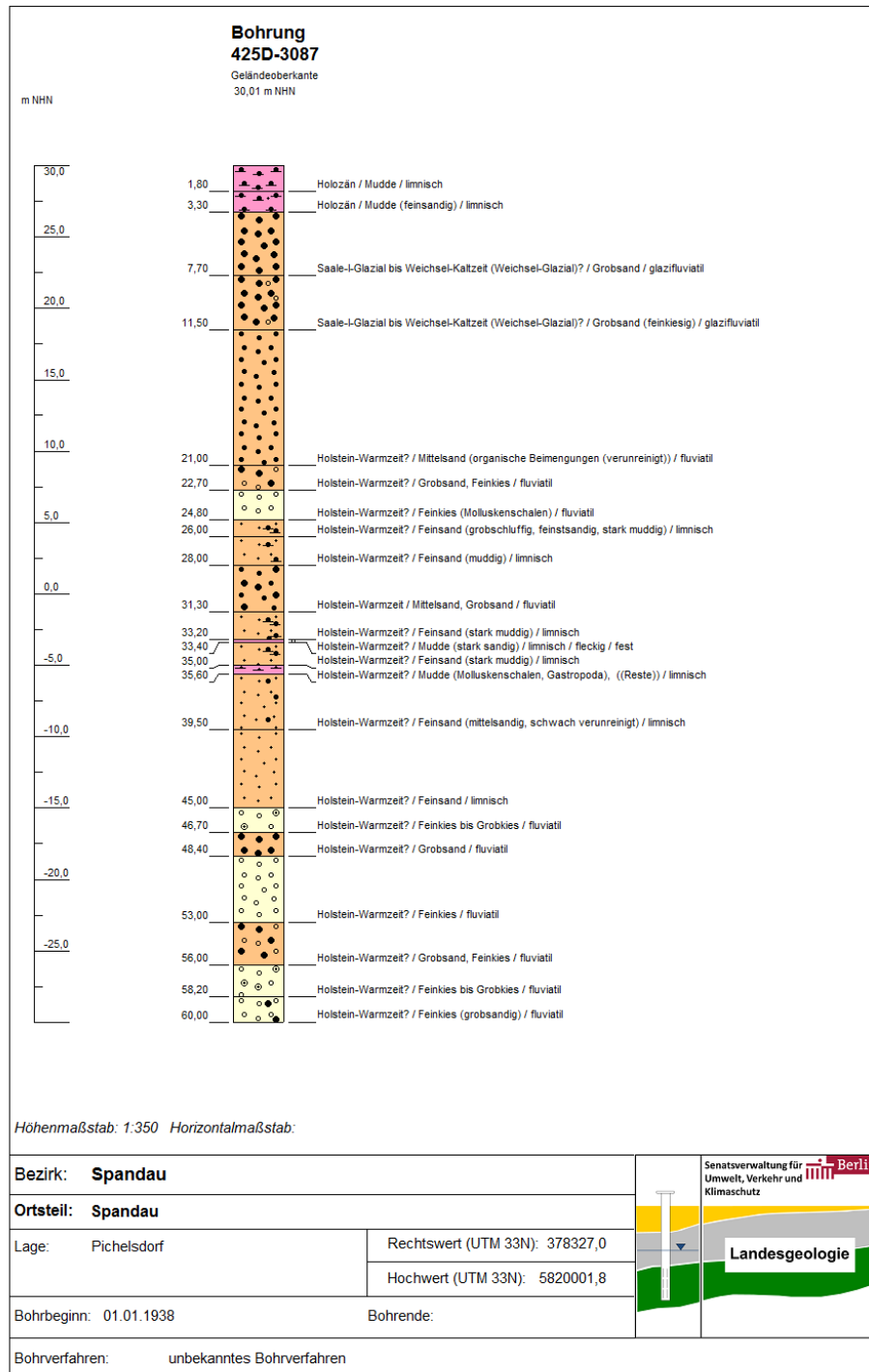


Figure A5: Layer information of drilling 4, with mud (silt or clay) formation on the surface, near Havel River

Note: Feinkies means fine gravel. Source: SenUVK(2020).

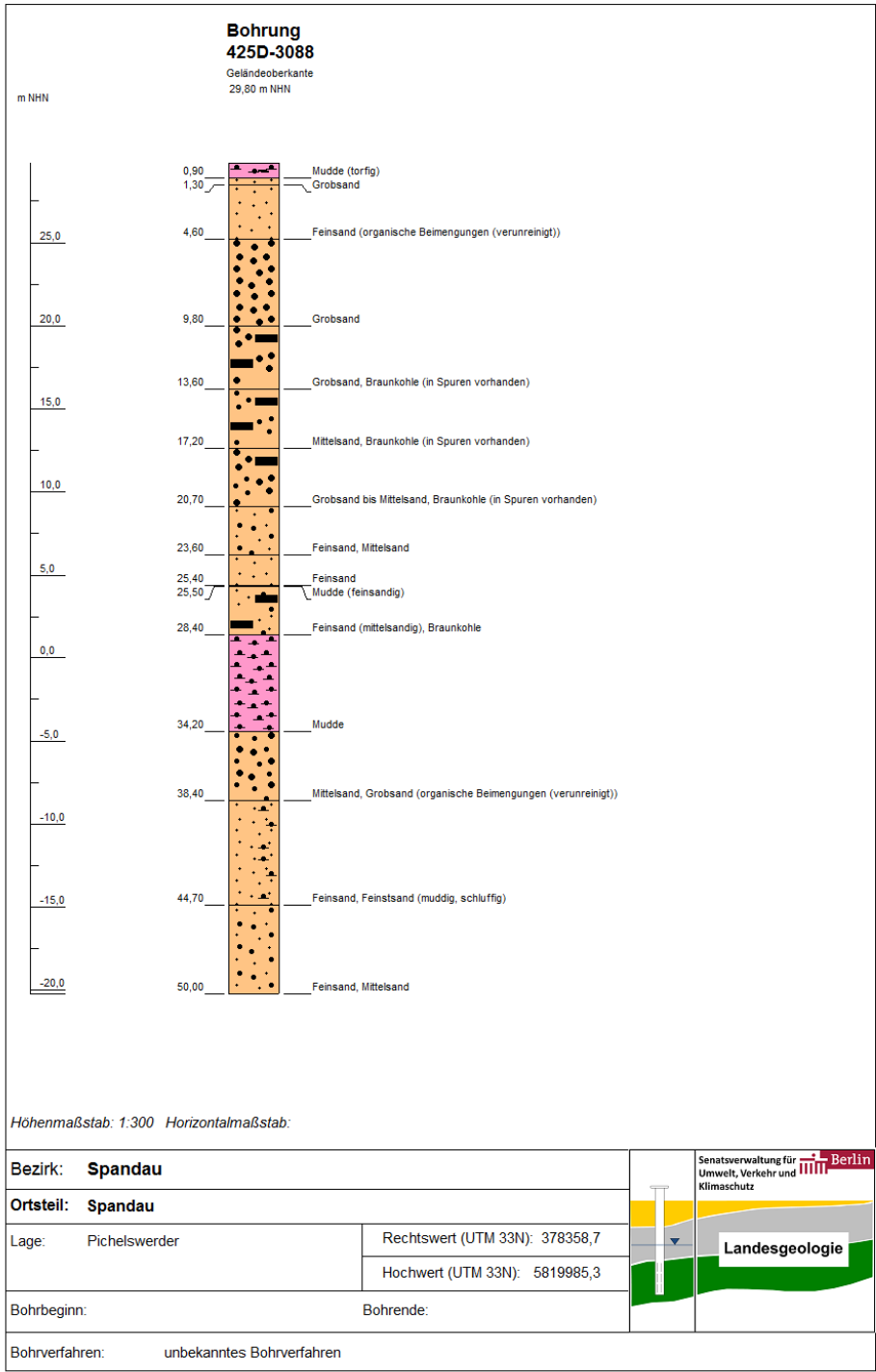


Figure A6: Layer information of drilling 5, with mud (silt or clay) formation on the surface, near Havel River

Source: SenUVK(2020).

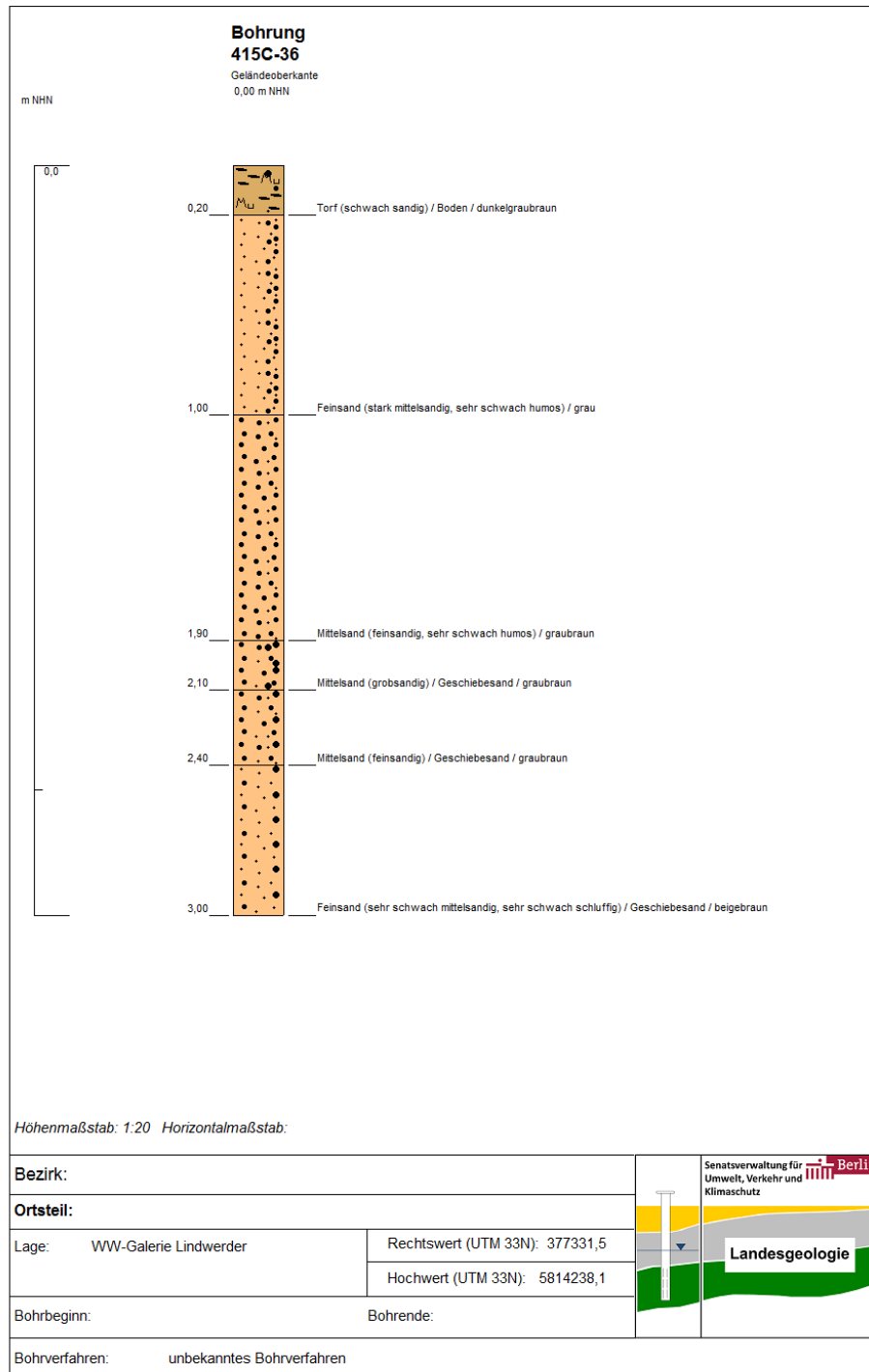


Figure A7: Layer information of drilling 6, with lowland moor formation on the surface, near Havel River

Note: Torf means peat. Source: SenUVK(2020).

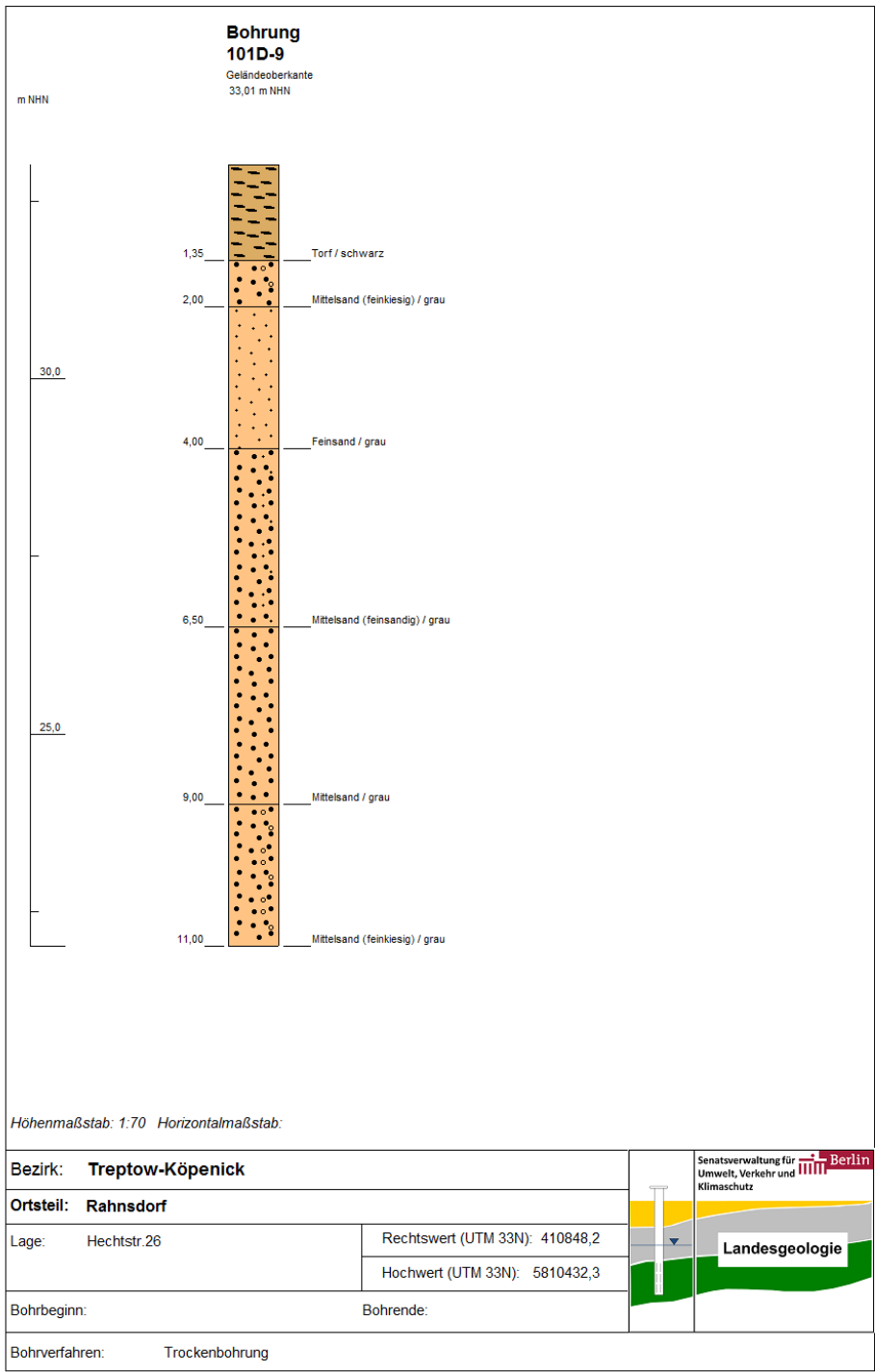


Figure A8: Layer information of drilling 7, with lowland moor formation on the surface, near Lake Müggelsee

Source: SenUVK (2020)

Appendix 5

Streambed materials used by Schälchli (1993)

Table A5: Properties of streambed materials: the associated thickness of clogging layer in experiment flume

Material name	d_{10}	d_{30}	d_{50}	d_{60}	d_{90}	d_m	C_U	d_{90}/d_m	$h(t)$
Sw	0.31	9.6	19	26	64	27	85	2.4	0.091
Sv	0.40	5.3	23	33	78	32	83	2.4	0.106
SR	0.41	8.1	19	25	40	20	59	2.0	0.070
KV1	1.5	4.7	7.7	9.5	15	8.5	6.3	1.8	0.036
KV2	0.36	2.7	5.3	7.1	15	6.9	19	2.2	0.031
TO	0.8	6.5	16	22	52	22	28	2.4	0.076
Inn	1.7	18.6	64	98	395	139	58	2.8	0.43
Etzlibach	1.2	17.7	70	144	604	205	120	2.9	0.63
Rhein Sw	0.38	11.0	22	30	88	35	80	2.5	0.12
Rhein Sv	0.45	8.7	29	42	125	44	92	2.8	0.14
Thur biomd	0.63	12.9	26	33	58	28	52	2.1	0.094
Thur unimod	1.7	13.6	26	33	60	31	19	1.9	0.103
Langete	0.8	9.7	24	32	72	32	40	2.3	0.106
Töss	0.8	6.5	16	22	52	22	28	2.4	0.076
Brauner Kies	0.13	9.2	22	28	72	28	215	2.6	0.094
Grauer Kies	0.26	4.0	10	14	45	16	54	2.8	0.058
Bimodale WL	0.34	16.1	25	29	69	29	85	2.4	0.097

Note: grain sizes d_i in millimeter, C_U is the uniformity coefficient, $h(t)$ is the thickness of the clogging layer, in meters. Source: Schälchli(1993, 96).

Appendix 6

Interview and questionnaires

Interview A (after Keeney and Raiffa(1993)):

Partition the set of attributes $X \equiv \{X_1, X_2, X_3, X_4, X_5\}$ into Y and its complementary set \bar{Y} . To check whether Y is preferentially independent of \bar{Y} we might proceed along the lines of this hypothetical interview between the **analyst** (me) and the **assessor** (you, the decision maker):

Analyst: I would now like to investigate how you feel about various Y values when we hold fixed a particular value of \bar{Y} . For example, on the first page of this Questionnaire A [see Questionnaire A page 1] there is a list of 18 paired comparisons between Y evaluations; each element of the pair describes levels of the Y attributes alone. On this first page it is assumed that, throughout, the \bar{Y} evaluations are all the same, that is, $\bar{y}^{(1)}$ [see Questionnaire A page 2]. Is this clear?

Assessor: Crystal clear, but you are asking me for a lot of work.

Analyst: Well, I have a devious purpose in mind and it will not take as much time as you think to find out what I want. Now on the third page of the Questionnaire A [see Questionnaire A page 3] the identical set of 18 paired comparisons are repeated but now the fixed, common level on the \bar{Y} attributes is changed from $\bar{y}^{(1)}$ to $\bar{y}^{(2)}$. Are you with me?

Assessor: All the way.

Analyst: On page 4, we have the same 18 paired comparisons but now the common value of the \bar{Y} values is $\bar{y}^{(3)}$ [see Questionnaire A page 4].

Assessor: You said this would not take long.

Analyst: Well now, here comes the punchline. Suppose that you painstakingly respond to all 18 paired comparisons on page 1 where $\bar{y}^{(1)}$ is fixed. Now when you go to the next page would your responses change to these same 18 paired comparisons?

Assessor: Let's see. In the third page all paired comparisons are the same except $\bar{y}^{(1)}$ is replaced by $\bar{y}^{(2)}$. What difference should that make?

Analyst: Well, you tell me. If we consider this first comparison [pointed to on the Questionnaire A page 1, marked green] does it make any difference if \bar{Y} values are all fixed at level $\bar{y}^{(1)}$ or $\bar{y}^{(2)}$? There could be some interaction concerning how you view the paired comparison depending on the common value of the \bar{Y} values.

Assessor: I suppose that might be the case in some other situation but in the first comparison I prefer the right alternative to the left no matter what the \bar{Y} values are...as long as they are the same.

Analyst: Okay. Would you now feel the same if you consider the second paired comparison?

Assessor: Yes. And the third and so on. Am I being naïve? Is there some trick here?

Analyst: No, not at all. I am just checking to see if the \bar{y} values have any influence on your responses to the paired comparisons. So I gather that you are telling me that your responses on page 2 would carry over to page 3.

Assessor: That's right.

Analyst: And to page 4, where the \bar{Y} values are held fixed at $\bar{y}^{(3)}$ [see Questionnaire A page 4]?

Assessor: Yes.

Analyst: Well, on the basis of this information I now pronounce that for you the attribute set Y is preferentially independent of the attribute set \bar{Y} .

Assessor: That's nice to know.

Analyst: That's all that I wanted to find out.

Assessor: Aren't you going to ask me to fill out page 1?

Analyst: No. That's too much work. There are less painful ways of getting that information.

Questionnaire A:

Let $Y \equiv \{X_1, X_2\}$, referred to attributes *transmissivity* (m^2/s) and *well field location score*. List of 18 paired comparisons of $y \equiv (x_1, x_2)$ is given; choose the preferred pair (left or right).

Comparison	(x_1, x_2)	(x_1, x_2)
1	(0.031, 45)	(0.055, 43)
2	(0.344, 40)	(0.012, 7)
3	(0.035, 53)	(0.149, 60)
4	(0.063, 25)	(0.016, 42)
5	(0.32, 51)	(0.017, 14)
6	(0.15, 42)	(0.174, 35)
7	(0.052, 34)	(0.303, 12)
8	(0.227, 39)	(0.037, 17)
9	(0.155, 39)	(0.066, 21)
10	(0.297, 16)	(0.111, 14)
11	(0.187, 31)	(0.059, 38)
12	(0.252, 70)	(0.065, 28)
13	(0.149, 49)	(0.294, 37)
14	(0.278, 43)	(0.347, 51)
15	(0.279, 61)	(0.194, 58)
16	(0.083, 69)	(0.339, 20)
17	(0.135, 34)	(0.113, 69)
18	(0.259, 64)	(0.185, 21)

Questionnaire A (continued):

Now the complementary set of Y is $\bar{Y} \equiv \{X_3, X_4, X_5\}$.

\bar{y}	(x_3, x_4, x_5)
$\bar{y}^{(1)}$	$(3.5, 60, 0)$

Questionnaire A (continued):

\bar{y}	(x_3, x_4, x_5)
$\bar{y}^{(2)}$	$(2.5, 30, 50)$

Questionnaire A (continued):

\bar{y}	(x_3, x_4, x_5)
$\bar{y}^{(3)}$	$(1.5, 0, 100)$

Appendix 7

Surface water area of Jilin City

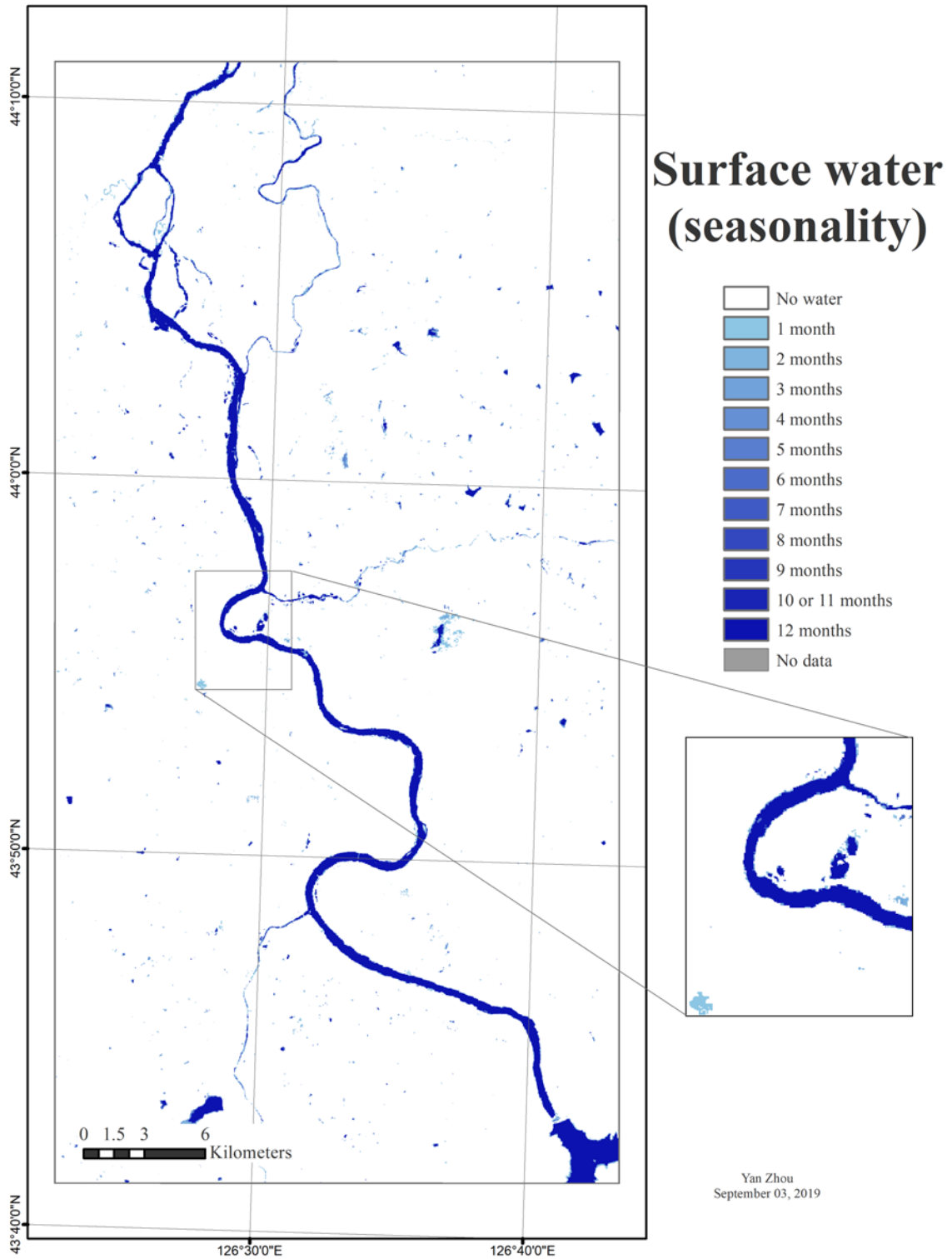


Figure A9: Surface water area of Jilin City

Note: the seasonality of surface water area was recorded in year 2018. Source: EC JRC/Google(2016). Projection: WGS_1984_UTM_Zone_52.

Bibliography

- Aller, L., T. Bennett, J. Lehr, and R. Petty. 1987. *DRASTIC: a standardized system to evaluate ground water pollution potential using hydrogeologic settings*. U.S. Environmental Protection Agency (Washington, D.C.).
- Alley, W. M., T. E. Reilly, and F. O. Lehn. 1999. *Sustainability of Ground-Water Resources*. U.S. Geological Survey Circular 1186 (Denver, Colorado).
- Alony, E. 2019. "SAN FRANCISCO AIRPORT BANS BOTTLED WATER".
- Antoniadis, V., and B. J. Alloway. 2002. "The role of dissolved organic carbon in the mobility of Cd, Ni, and Zn in sewage sludge-amended soils." *Environmental Pollution* 117: 515-521.
- AQSIQ (General Administration of Quality Supervision Inspection and Quarantine of the People's Republic of China). 1993. *Quality standard for ground water (GB14848-93)*.
- Archwichai, Laa, Kriengsak Srisuk, Sumrit Chusanatus, Kewaree Pholkern, Sitisak Munyou, and Phayom Saraphirom. 2011. "Development of the Site Selection Criteria for Riverbank Filtration Project, Thailand." Khon Kaen, Thailand.
- Bao, Y. 2019. "无印良品召回瓶装水约 59 万瓶：疑含有致癌物 (MUJI Recalls about 590,000 Bottles of Water: Suspected of Containing Carcinogens)" http://finance.sina.com.cn/roll/2019-02-26/doc-ihrfqzka9440429.shtml?cre=tianyi&mod=pcpager_fintoutiao&loc=37&r=9&func=100&tj=none&tr=9.
- Bartak, Rico, Thomas Grischek, Kamal O. Ghodeif, and Rifaat A. Wahaab. 2015. "Shortcomings of the RBF Pilot Site in Dishna, Egypt." *Journal of Hydrologic Engineering* 20 (9).
- Baveye, Philippe, Philippe Vandevinere, Blythe L. Hoyle, Paul C. DeLeo, and Diego Sanchez de Lozada. 1998. "Environmental Impact and Mechanisms of the Biological Clogging of Saturated Soils and Aquifer Materials." *Critical Reviews in Environmental Science and Technology* 28 (2): 123-191.
- Bear, Jacob. 1972. *Dynamics of fluids in porous media*. New York: Dover Publications, Inc.
- . 1979. *Hydraulics of Groundwater*. New York: McGraw-Hill.
- Bernoulli, Daniel. 1954. "Exposition of a New Theory on the Measurement of Risk. Translated by Louise Sommer." *Econometrica* 22 (1): 23-36.
- Beverage Marketing Corporation. 2006-2016. Bottled Water Report
- Beyer, W., and E. Banschler. 1975. "Zur Kolmation der Gewässerbetten bei der Uferfiltratgewinnung." *Zeitschrift für Angewandte Geologie* 21 (12): 565-570.
- BGR (Federal Institute for Geosciences and Natural Resources of Germany). 2019. "Geologische Übersichtskarte der Bundesrepublik Deutschland 1 : 250 000 (GÜK250)" https://www.bgr.bund.de/DE/Themen/Sammlungen-Grundlagen/GG_geol_Info/Karten/Deutschland/GUEK250/guek250_node.html.
- Blaschke, Alfred Paul, Karl-Heinz Steiner, Roland Schmalfuss, Dieter Gutknecht, and Dieter Sengschmitt. 2003. "Clogging Processes in Hyporheic Interstices of an

- Impounded River, the Danube at Vienna, Austria." *International Review of Hydrobiology* 88 (3-4): 397-413.
- BMI (Federal Ministry of the Interior of Germany). 1975. *Uferfiltration. Bericht des BMI-Fachausschusses "Wasserversorgung und Uferfiltrat"*. Bundesministerium des Innern.
- . 1985. *Künstliche Grundwasseranreicherung. Stand der Technik in der Bundesrepublik Deutschland.*: Erich Schmidt Verlag.
- Boggs, S. 2006. *Principles of sedimentology and stratigraphy*. Pearson Education, Inc.
- Bourg, Alain C. M., and Clotilde Bertin. 1993. "Biogeochemical processes during the infiltration of river water into an alluvial aquifer." *Environmental Science & Technology* 27: 661-666.
- Bradley, P. M., L. B. Barber, J. W. Duris, W. T. Foreman, E. T. Furlong, L. E. Hubbard, K. J. Hutchinson, S. H. Keefe, and D. W. Kolpin. 2014. "Riverbank filtration potential of pharmaceuticals in a wastewater-impacted stream." *Environmental Pollution* 193: 173-180.
- Brunke, M., and T. Gonser. 1997. "The ecological significance of exchange processes between rivers and groundwater." *Freshwater Biology* 37: 1-33.
- Brunner, Philip, Peter G. Cook, and Craig T. Simmons. 2009. "Hydrogeologic controls on disconnection between surface water and groundwater." *Water Resource Research* 45 (1): W01422.
- Bunte, Kristin, and Abt Steven R. 2001. *Sampling surface and subsurface particle-size distributions in wadable gravel- and cobble-bed streams for analyses in sediment transport, hydraulics, and streambed monitoring. Gen. Tech. Rep. RMRS-GTR-74*. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.
- Burnet, John. 1869. "History of the water supply to Glasgow, from the commencement of the present century." <http://wellcomelibrary.org/item/b24400555>.
- Butler, John, Jianmin Jia, and James Dyer. 1997. "Simulation techniques for the sensitivity analysis of multi-criteria decision models." *European Journal of Operational Research* 103: 531-546.
- Caldwell, T. G. 2004. "Chapter 14: Presentation of Data for Factors Significant to Yield from Several Riverbank Filtration systems in the U.S. and Europe." In *Riverbank Filtration Hydrology*, edited by S. A. Hubbs, 299-344. Springer & NATO Public Diplomacy Division.
- Chalom, Andre, and Paulo Inacio Prado. 2017. Parameter Space Exploration with Latin Hypercubes.
- Chen, C. K., and Y. D. Lee. 2000. "Review and rethink the utility theory." *Web Journal of Chinese Management Review* 3 (3): 1-7.
- China Construction Newspaper. 2012. "China's urban water supply is generally safe, and water quality is constantly improving" http://www.mohurd.gov.cn/zxydt/201205/t20120511_209853.html.
- Cosier, M., and D. Shen. 2009. "Urban Water Management in China." *Water Resources Development* 25 (2): 249-268.
- Cressie, Noel A. C. 1993. *Statistics for spatial data revised edition*. John Wiley & Sons, Inc.

- Cunningham, A. B., M. Asce, C. J. Anderson, A. M. Asce, H. Bouwer, and M. Asce. 1987. "Effects of sediment-laden flow on channel bed clogging." *Journal of Irrigation and Drainage Engineering* 113 (1): 106-118.
- Dash, R. R., I. Mehrotra, P. Kumar, and T. Grischek. 2008. "Lake bank filtration at Nainital, India: water-quality evaluation." *Hydrogeology Journal* 16 (6): 1089-1099.
- Daxue Consulting. 2019. "Beverage market research: Mineral water industry in China | Daxue Consulting" <https://daxueconsulting.com/mineral-water-market-in-china/>.
- Destatis (Federal Statistical Office of Germany). 2003. *Öffentliche Wasserversorgung und Abwasserbeseitigung 2001 - Fachserie 19 Reihe 2.1*. Statistisches Bundesamt (Destatis) (Wiesbaden).
- . 2013. *Öffentliche Wasserversorgung und öffentliche Abwasserentsorgung 2010 - Öffentliche Wasserversorgung - Fachserie 19 Reihe 2.1.1*. Statistisches Bundesamt (Destatis) (Wiesbaden).
- . 2018. *Öffentliche Wasserversorgung und öffentliche Abwasserentsorgung 2016 - Strukturdaten zur Wasserwirtschaft - Fachserie 19 Reihe 2.1.3*. Statistisches Bundesamt (Destatis).
- . 2019a. *Bevölkerung und Erwerbstätigkeit 2017 - Bevölkerungsforschreibung auf Grundlage des Zensus 2011 - Fachserie 1 Reihe 1.3*. Statistisches Bundesamt (Destatis).
- . 2019b. *Öffentliche Wasserversorgung und öffentliche Abwasserentsorgung 2016 - Öffentliche Wasserversorgung - Fachserie 19 Reihe 2.1.1*. Statistisches Bundesamt (Destatis).
- Dillon, P., P. Stuyfzand, T. Grischek, M. Lluria, R. D. G. Pyne, R. C. Jain, J. Bear, J. Schwarz, W. Wang, E. Fernandez, C. Stefan, M. Pettenati, J. van der Gun, C. Sprenger, G. Massmann, B. R. Scanlon, J. Xanke, P. Jokela, Y. Zheng, R. Rossetto, M. Shamrukh, P. Pavelic, E. Murray, A. Ross, J. P. Bonilla Valverde, A. Palma Nava, N. Ansems, K. Posavec, K. Ha, R. Martin, and M. Sapiano. 2019. "Sixty years of global progress in managed aquifer recharge." *Hydrogeology Journal* 27 (1): 1-30. <https://doi.org/10.1007/s10040-018-1841-z>. <https://doi.org/10.1007/s10040-018-1841-z>.
- Domenico, P. A., and F. W. Schwartz. 1997. *Physical and Chemical Hydrogeology*. John Wiley & Sons, Inc.
- Doussan, C., G. Poitevin, E. Ledoux, and M. Detay. 1997. "River bank filtration: modeling of the changes in water chemistry with emphasis on nitrogen species." *Journal of Contaminant Hydrology* 25: 129-156.
- Drewes, J. E., and R. S. Summers. 2002. "Chapter 15. Natural Organic Matter Removal During Riverbank Filtration: Current Knowledge and Research Needs." In *Riverbank Filtration*, edited by C. Ray, G. Melin and R. B. Linsky, 303-309. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Du, Xinqiang. 2019. Personal Communication.
- Dunlap, L. E., and J. M. Spinazola. 1981. *Interpolating Water-Table Altitudes in West-Central Kansas Using Kriging Techniques*. U.S. Geological Survey.
- Dyer, J. S., and R. F. Miles. 1974. *Trajectory selection for the Mariner Jupiter/Saturn 1977 Project (Technical Memorandum 33-106)*. Jet Propulsion Laboratory, California Institute of Technology (Pasadena, California).

- Dyer, James S., Peter C. Fishburn, Ralph E. Steuer, Jyriki Wallenius, and Stanley Zionts. 1992. "Multiple Criteria Decision Making, Multiattribute Utility Theory: The Next Ten Years." *Management Science* 38 (5): 645-654.
- Edmundson, L. 2012. "Cholera Outbreak Hits Hubei, China Wedding" <http://www.diseasedaily.org/diseasedaily/article/cholera-outbreak-hits-hubei-china-wedding-101612>.
- EEA (European Environment Agency). 2019. "European catchments and Rivers network system (Ecrins)" http://data.europa.eu/euodp/de/data/dataset/data_european-catchments-and-rivers-network.
- ESRI (Environmental Systems Research Institute). 1999-2017. ArcGIS Desktop 10.6.
- . 2018. ArcGIS Pro 2.3.0. Environmental Systems Research Institute.
- Farsaci, L. 2019. "Bottled water recall linked to probe by FSAI and HSE" <https://www.irishmirror.ie/news/irish-news/bottled-water-recall-celtic-pure-18822989>.
- Fishburn, P. C. 1967. "Methods of estimating additive utilities." *Management Science* 13 (7): 435-453.
- Fishburn, Peter C. 1989. "Foundations of Decision Analysis: Along the Way." *Management Science* 35 (4): 387-405.
- Fokken, B. 1996a. "Gewinnung von uferfiltriertem Grundwasser." In *DVGW Lehr- und Handbuch Wasserversorgung Band 1 Wassergewinnung und Wasserwirtschaft*, edited by Deutsche Vereinigung des Gas- und Wasserfaches e.V. (DVGW), 481-526. München: R. Oldenbourg Verlag.
- . 1996b. "Gewinnung von uferfiltriertem Grundwasser." In *Lehr- und Handbuch Wasserversorgung Bd. 1: Wassergewinnung und Wasserwirtschaft*, edited by DVGW, 481-526. München: R. Oldenbourg.
- Folk, R. L., and W. C. Ward. 1957. "Brazos River Bar: A study in the significance of grain size parameters." *Journal of Sedimentary Petrology* 27 (1): 3-26.
- Freeze, R. A., and J. A. Cherry. 1979. *Groundwater*. Prentice-Hall, Inc.
- Freitas, D. A., Jaime Cabral, Anderson Paiva, and Renato Molica. 2012. "Application of bank filtration technology for water quality improvement in a warm climate: A case study at Beberibe River in Brazil." *Journal of Water Supply: Research and Technology - AQUA* 61: 319. <https://doi.org/10.2166/aqua.2012.097>.
- Friedman, J. H. 1984. *A variable span scatterplot smoother*. Stanford University Technical Report No. 5.
- Fuhrmann, R. 1999. "Die Entwicklungsgeschichte postsaaleglazial entstandener Talabschnitte der Weißen Elster und Mulde und die stratigraphische Gliederung des jüngeren Quartärs." *Altenburger naturwissenschaftliche Forschungen* 11: 43-63.
- Gadgil, A. 1998. "Drinking water in developing countries." *Annu. Rev. Energy Environ.* 23: 253-286.
- GADM (Database of Global Administrative Areas). 2019. "GADM maps and data" <https://gadm.org/index.html>.
- Gessler, J. 1965. *Der Geschiebetriebbeginn bei Mischungen untersucht an natürlichen Abpflasterungserscheinungen in Kanälen*. Zürich: Dissertation, ETH Zürich.
- Geudens, P.J.J.G. 2015. *Dutch Drinking Water Statistics 2015*. Association of Dutch water companies (Vewin) (The Hague).

- Ghodeif, K., T. Grischek, R. Bartak, R. Wahaab, and J. Herlitzius. 2016. "Potential of river bank filtration (RBF) in Egypt." *Environ Earth Sci* 75: 671.
- Glasgow City Archives. 2020. "Cranstonhill Water Works Co." <https://www.theglasgowstory.com/image/?inum=TGSA02064>.
- Gong, P., H. Liu, M. Zhang, C. Li, J. Wang, H. Huang, N. Clinton, L. Ji, Y. Bai, B. Chen, B. Xu, Z. Zhu, C. Yuan, H. P. Suen, J. Guo, N. Xu, W. Li, Y. Zhao, J. Yang, C. Yu, X. Wang, H. Fu, L. Yu, I. Dronova, F. Hui, X. Cheng, X. Shi, F. Xiao, Q. Liu, and L. Song. 2019. "Stable classification with limited sample: transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017." *Science Bulletin* 64 (6): 370-373.
- Grischek, T. 2019. Personal Communication. Leipzig.
- Grischek, T., and R. Bartak. 2016. "Riverbed Clogging and Sustainability of Riverbank Filtration." *Water* 8 (12): 604.
- Grischek, T., K. M. Hiscock, T. Metschies, P. F. Dennis, and W. Nestler. 1998. "Factors affecting denitrification during infiltration of river water into a sand and gravel aquifer in Saxony, Germany." *Water Research* 32 (2): 450-460.
- Grischek, Thomas, Dagmar Schoenheinz, and Chittaranjan Ray. 2002. "Chapter 14. Siting and Design Issues for Riverbank Filtration Schemes." In *Riverbank Filtration*, edited by C. Ray, G. Melin and R. B. Linsky, 291-302. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Grünheid, Steffen, Gary Amy, and Martin Jekel. 2005. "Removal of bulk dissolved organic carbon (DOC) and trace organic compounds by bank filtration and artificial recharge." *Water Research* 39: 3219-3228.
- Gu, Baojing, Ying Ge, Scott X. Chang, Weidong Luo, and Jie Chang. 2013. "Nitrate in groundwater of China: Sources and driving forces." *Global Environmental Change* 23: 1112-1121.
- Günter, A. 1971. Die kritische mittlere Sohlenschubspannung bei Geschiebemischung unter Berücksichtigung der Deckschichtbildung und der turbulenzbedingten Sohlenschubspannungsschwankungen. Zürich: Dissertation, ETH Zürich.
- Gupta, A., H. Singh, F. Ahmed, I. Mehrotra, P. Kumar, S. Kumar, T. Grischek, and C. Sandhu. 2015. "Lake bank filtration in landslide debris: irregular hydrology with effective filtration." *Sustainable Water Resources Management* 1 (1): 15-26.
- Heath, Ralph C. 1983. *Basic Ground-Water Hydrology*. U.S. Geological Survey Water-Supply Paper 2220. U.S. Geological Survey.
- Hengl, Tomislav. 2006. "Finding the right pixel size." *Computers & Geosciences* 32: 1283-1298.
- Henzler, A. F., J. Greskowiak, and G. Massmann. 2014. "Modeling the fate of organic micropollutants during river bank filtration (Berlin, Germany)." *Journal of Contaminant Hydrology* 156: 78-92.
- Hiscock, K. M., and T. Grischek. 2002. "Attenuation of groundwater pollution by bank filtration." *Journal of Hydrology* 266: 139-144.
- Hoffmann, A., and G. Gunkel. 2011. "Bank filtration in the sandy littoral zone of Lake Tegel (Berlin): Structure and dynamics of the biological active filter zone and clogging processes." *Limnologica* 41: 10-19.

- Hoppe-Jones, C., G. Oldham, and J. E. Drewes. 2010. "Attenuation of total organic carbon and unregulated trace organic chemicals in U.S. riverbank filtration systems." *Water Research* 44 (15): 4643-4659.
- Hu, Bin, Yanguo Teng, Yuanzheng Zhai, Rui Zuo, Jiao Li, and Haiyang Chen. 2016. "Riverbank filtration in China: A review and perspective." *Journal of Hydrology* 541: 914-927.
- Hubbs, Stephen A. 2004. "Chapter 9: Changes in Riverbed Hydraulic Conductivity and Specific Capacity at Louisville." In *Riverbank Filtration Hydrology*, edited by Stephen A. Hubbs, 199-220. Springer & NATO Public Diplomacy Division.
- Iman, Ronald L., Jon C. Helton, and James E. Campbell. 1981. "An Approach to Sensitivity Analysis of Computer Models: Part I—Introduction, Input Variable Selection and Preliminary Variable Assessment." *Journal of Quality Technology* 13 (3): 174-183. <https://doi.org/10.1080/00224065.1981.11978748>.
<https://doi.org/10.1080/00224065.1981.11978748>.
- Isaaks, E. H., and R. M. Srivastava. 1989. *Applied Geostatistics*. New York: Oxford University Press, Inc.
- ISO (International Organization for Standardization). 2017. ISO 14688-1 Geotechnical investigation and testing—Identification and classification of soil—Part 1: Identification and description. Vernier, Geneva, Switzerland: ISO copyright office.
- Jiménez, A., S. Ríos-Insua, and A. Mateos. 2003. "A decision support system for multiattribute utility evaluation based on imprecise assignments." *Decision Support Systems* 36: 65-79.
- Johnston, K., J. M. Ver Hoef, K. Krivoruchko, and N. Lucas. 2001. Using ArcGIS Geostatistical Analyst. ESRI.
- Journel, A. G., and Ch. J. Huijbregts. 1978. *Mining Geostatistics*. Academic Press.
- Jülich, Walter. 2000. *Proceedings of the International Riverbank Filtration Conference : November 2-4, 2000, Düsseldorf, Germany*. Edited by Walter Jülich. *International Riverbank Filtration Conference*. Köln: IAWR.
- Kalbitz, K., and R. Wennrich. 1998. "Mobilization of heavy metals and arsenic in polluted wetland soils and its dependence on dissolved organic matter." *The Science of the Total Environment* 209: 27-39.
- Kalbus, E., F. Reinstorf, and M. Schirmer. 2006. "Measuring methods for groundwater - surface water interactions: a review." *Hydrology and Earth System Sciences* 10: 873-887.
- Keeney, R. L. 1974. "Multiplicative Utility Functions." *Operations Research* 22 (1): 22-34.
- Keeney, R. L., and H. Raiffa. 1993. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. Cambridge, UK: Cambridge University Press.
- Keeney, R. L., and E. F. Wood. 1977. "An illustrative example of the use of multiattribute utility theory for water resource planning." *Water Resource Research* 13 (4): 705-712.
- Kholghi, M., and S. M. Hosseini. 2006. "Estimation of Aquifer Transmissivity using Kriging, Artificial Neural Network, and Neuro-Fuzzy models." *Journal of Spatial Hydrology* 6 (2): 68-81.

- Kim, S. 2015. "ppcor: An R Package for a Fast Calculation to Semi-partial Correlation Coefficients." *Communications for Statistical Applications and Methods* 22 (6): 665-674.
- Krause, Eric. 2012. Dealing with extreme values in kriging. Accessed May 07.
- Kühn, W., and U. Müller. 2000. "Riverbank Filtration: An Overview." *American Water Works Association* 92 (12): 60-69.
- . 2006. *Exportorientierte Forschung und Entwicklung auf dem Gebiet der Wasserver- und -entsorgung. Teil I: Trinkwasser. Band 2. Leitfaden*. Karlsruhe: TZW.
- LANUV (State Agency for Nature Environment and Consumer Protection of North Rhine-Westphalia). 2015. "Wasserschutzgebiete NRW" <http://www.lanuv.nrw.de/>.
- Lee, Sang-il, and Sang-sin Lee. 2010. "Development of site suitability analysis system for riverbank filtration." *Water Science and Engineering* 3 (1): 85-94.
- Lenk, Stephan, Frank Remmler, Christian Skark, and Ulrich Schulte-Ebbert. 2006. "Uferfiltration. Schlussbericht Teilprojekt B1." In *Exportorientierte Forschung und Entwicklung auf dem Gebiet der Wasser- und -entsorgung. Teil I: Trinkwasser. Band 2. Leitfaden*, edited by Wolfgang Kühn and U. Müller. Karlsruhe: TZW.
- Lewis, I. 2019. "Massachusetts bottled water company shuts after PFAS detected" <https://vtdigger.org/2019/08/05/massachusetts-bottled-water-company-shutters-after-pfas-detected/>.
- LfULG (State Office for the Environment Agriculture and Geology of Saxony). 2018. "Wasserschutzgebiete" <https://www.umwelt.sachsen.de/umwelt/wasser/6318.htm>.
- Liao, H. 2012. "ArcGIS 数据生产与精细化制图之中国年降水量分布图的制作 (ArcGIS Data Production and Refined Mapping: Production of China's Annual Precipitation Distribution Map) (unpublished document)" <http://pan.baidu.com/share/link?shareid=54856&uk=352462598>.
- Liao, Zisheng, and Xiaosi Su. 2007. *吉林省中部地区水资源保护与利用对策方案研究报告 (Report on Water Resources Protection and Utilization in the Central Region of Jilin Province)*. Jilin University & Ministry of Water Resources.
- Linsky, Ronald B., Chittanarjan Ray, and Gina Melin. 2004. *Riverbank filtration : improving source-water quality*. Dordrecht: Kluwer academic.
- Lisle, Thomas E. 1989. "Sediment Transport and Resulting Deposition in Spawning Gravels, North Coastal California." *Water Resources Research* 25 (6): 1303-1319.
- Liu, J. 2017. *吉林市城区饮用水地下水应急水源地的研究 (Study on Emergency Water Source of Drinking Water Groundwater in Jilin City)*. Jilin: PhD thesis, Jilin University.
- Marino, Simeone, Ian B. Hogue, Christian J. Ray, and Denise E. Kirschner. 2008. "A Methodology For Performing Global Uncertainty And Sensitivity Analysis In Systems Biology." *J Theor Biol.* 254 (1): 178-196.
- Martín-Alonso, J. 2004. "Chapter 13. Managing Resources in a European Semi-Arid Environment: Combined use of Surface and Groundwater for Drinking Water Production in the Barcelona Metropolitan Area." In *Riverbank Filtration Hydrology*, edited by S. A. Hubbs, 281-298. Springer & NATO Public Diplomacy Division.
- Matheron, G. 1963. "Principles of Geostatistics." *Economic Geology* 58: 1246-1266.

- Mayo Clinic. 2018. "Legionnaires' disease" <https://www.mayoclinic.org/diseases-conditions/legionnaires-disease/symptoms-causes/syc-20351747>.
- McDowell-Boyer, L. M., J. R. Hunt, and N. Sitar. 1986. "Particle Transport Through Porous Media." *Water Resources Research* 22 (13): 1901-1921.
- McKay, M. D., R. J. Beckman, and W. J. Conover. 1979. "Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code." *Technometrics* 21 (2): 239-245.
- MEE (Ministry of Ecology and Environment of People's Republic of China). 2005. *Report on the State of the Environment in China 2005*. Ministry of Ecology and Environment (MEE).
- . 2013. "水体污染控制与治理科技重大专项, 2014 年度拟立项课题择优指南 (第四批) (Water Pollution Control and Treatment Science and Technology Major Project, 2014 Project Selection Guide (Fourth Batch))." The Ministry of Ecology and Environment of the People's Republic of China.
- . 2016. *Report on the State of the Environment in China 2016*. Ministry of Ecology and Environment (MEE).
- . 2018a. 吉林省主要城市饮用水源地水质月报 (Monthly Report on Water Quality of Drinking Water Sources in Major Cities in Jilin Province) . Accessed 1-12.
- . 2018b. "饮用水水源保护区划分技术规范 (Technical Guideline for Delineating Source Water Protection Areas)." The Ministry of Ecology and Environment of the People's Republic of China.
- MEE (Ministry of Ecology and Environment of People's Republic of China), and Inspection and Quarantine of the People's Republic of China) AQSIQ (General Administration of Quality Supervision. 2002. *Environmental quality standards for surface water (GB 3838-2002)*. Ministry of Ecology and Environment of the People's Republic of China (MEE), General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China (AQSIQ).
- Meinzer, O. E. 1923. *The Occurrence of Ground Water in the United States - With a Discussion of Principles (USGS Water-Supply Paper 489)*. U. S. Geological Survey.
- MOH (Ministry of Health of the People's Republic of China), and SAC (Standardization Administration of China). 2006. *Standards for drinking water quality (GB 5749-2006)*. Ministry of Health of the People's Republic of China, Standardization Administration of the People's Republic of China.
- MOHURD (Ministry of Housing and Urban-Rural Development of the People's Republic of China). 2016-2018. 城市建设统计年鉴 2006-2016 (Statistical Yearbook of Urban Construction 2006-2016).
- . 2018. 2016 年城市建设统计年鉴 (Statistical Yearbook of Urban Construction 2016).
- Motaghian, H. R., and J. Mohammadi. 2011. "Spatial Estimation of Saturated Hydraulic Conductivity from Terrain Attributes Using Regression, Kriging, and Artificial Neural Networks." *Pedosphere* 21 (2): 170-177.
- MUEEF (Ministry of the Environment Energy Nutrition and Forestry of Rhineland-Palatinate). 2019. "Wasserschutzgebiete" <https://geoportal-wasser.rlp-umwelt.de/servlet/is/8548/>.

- Nagy-Kovács, Zsuzsanna, Balázs László, Ernő Fleit, Katalin Czihat-Mártonné, Gábor Till, Hilmar Börnick, Yasmin Adomat, and Thomas Grischek. 2018. "Behavior of Organic Micropollutants During River Bank Filtration in Budapest, Hungary." *Water* 10: 1861.
- NASA JPL (National Aeronautics and Space Administration Jet Propulsion Laboratory). 2000. SRTMGL1 v003 - NASA Shuttle Radar Topography Mission Global 1 arc second. Accessed February 11-21.
- Natural Earth. 2019a. "1:10m Cultural Vectors version 4.1.0" <https://www.naturalearthdata.com/downloads/10m-cultural-vectors/>.
- . 2019b. "1:10m Physical Vectors version 4.1.0" <https://www.naturalearthdata.com/downloads/10m-physical-vectors/>.
- . 2019c. "Populated Places version 4.1.0" <https://www.naturalearthdata.com/downloads/10m-cultural-vectors/>.
- Nestler, W., W. Walther, F. Jacobs, R. Trettin, and K. Freyer. 1997. *Wassergewinnung in Talgrundwasserleitern im Einzugsgebiet der Elbe*. UFZ (Leipzig, Halle).
- NGA (National Geological Data Museum). 2013. "1:180 万吉林省地下水环境图 (1:1800000 Groundwater Environment Map of Jilin Province)" <http://www.ngac.org.cn/Map/Detail.aspx?MapId=EC7E1A7A75321954E0430100007F182E>.
- NMPA (National Medical Products Administration). 2019. "山东省市场监督管理局关于 36 批次食品不合格情况的通告 2019 年第 10 期 (总第 194 期) (Notice on the Disqualification of 36 batches of Foods from the Shandong Provincial Market Supervision Administration, No. 10, 2019 (Total No. 194))" <http://www.cfda.com.cn/newsdetail.aspx?id=120157>.
- NYSDOH (New York State Department of Health). 2018. "Boil Water Response - Information for the Public Health Professional" https://www.health.ny.gov/environmental/water/drinking/boilwater/response_information_public_health_professional.htm.
- Oliver, M. A. , and R. Webster. 2014. "A tutorial guide to geostatistics: Computing and modelling variograms and kriging." *Catena* 113: 56-69.
- Parker, L. 2018. "WE MADE PLASTIC. WE DEPEND ON IT. NOW WE'RE DROWNING IN IT." <https://www.nationalgeographic.com/magazine/2018/06/plastic-planet-waste-pollution-trash-crisis/>.
- Pekel, Jean-François, Andrew Cottam, Noel Gorelick, and Alan S. Belward. 2016. "High-resolution mapping of global surface water and its long-term changes." *Nature* 540: 418-422.
- Peng, W., X. Su, and Z. Liu. 2006. *Safety impact assessment of decentralized water withdrawal and groundwater supply on both banks along the Songhua River*. Ministry of Finance of the People's Republic of China (MOF).
- Pholkern, K., K. Srisuk, T. Grischek, M. Soares, S. Schäfer, L. Archwichai, P. Saraphirom, P. Pavelic, and W. Wirojanagud. 2015. "Riverbed clogging experiments at potential river bank filtration sites along the Ping River, Chiang Mai, Thailand." *Environmental Earth Sciences* 73: 7699-7709.
- Pollak, R. A. 1967. "Additive von Neumann-Morgenstern Utility Functions." *Econometrica* 35 (3/4): 485-494.

- Province of British Columbia. 2016. Determining the Likelihood of Hydraulic Connection - Guidance for the Purpose of Apportioning Demand from Diversion of Groundwater on Streams.
- Qiu, S., X. Liang, C. Xiao, H. Huang, Z. Fang, and F. Lv. 2015. "Numerical Simulation of Groundwater Flow in a River Valley Basin in Jilin Urban Area, China." *Water* 7 (10): 5768-5787.
- Ray, C., J. Schubert, R. B. Linsky, and G. Melin. 2002. "Introduction." In *Riverbank filtration*, edited by C. Ray, G. Melin and R. B. Linsky, 1-15. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Ray, Chittaranjan. 2002. *Riverbank filtration : understanding contaminant biogeochemistry and pathogen removal*. Edited by Chittaranjan Ray. *Nato Advanced Research Workshop on Riverbank Filtration*. Dordrecht: Kluwer.
- . 2008. "Worldwide potential of riverbank filtration." *Clean Techn Environ Policy* 10: 223-225.
- Ray, Chittaranjan, and Mohamed Shamrukh. 2011. *Riverbank filtration for water security in desert countries*. *Nato Advanced Research Workshop on Water Security in Desert Countries*. Dordrecht: Springer.
- Rivest, M., D. Marcotte, and P. Pasquier. 2008. "Hydraulic head field estimation using kriging with an external drift: A way to consider conceptual model information." *Journal of Hydrology* 361: 349-361.
- Rocha, Silvia Fernandes, and Eduardo Antonio Gomes Marques. 2018. "Three-dimensional modeling of steady-state flow in lake bank filtration – Brazil." *Water Supply* 19 (1): 60-69. <https://doi.org/10.2166/ws.2018.052>.
<https://doi.org/10.2166/ws.2018.052>.
- Rodwan, John G. 2018. *Bottled water 2017*. Beverage Marketing Corporation.
- Royston, P. 1995. "Remark AS R94: A Remark on Algorithm AS 181: The W-test for Normality." *Journal of the Royal Statistical Society, Series C (Applied Statistics)* 44 (4): 547-551.
- Saaty, R. W. 1987. "The analytic hierarchy process—what it is and how it is used." *Mathematical Modelling* 9 (3): 161-176.
[https://doi.org/https://doi.org/10.1016/0270-0255\(87\)90473-8](https://doi.org/https://doi.org/10.1016/0270-0255(87)90473-8).
<http://www.sciencedirect.com/science/article/pii/0270025587904738>.
- Sánchez, Luis Darío, Arlex Sánchez, Gerado Galvis, and Jorge Latorre. 2006. *Multi-Stage Filtration*. IRC International Water and Sanitation Centre.
- Sandhu, Cornelius, Thomas Grischek, Pradeep Kumar, and Chittaranjan Ray. 2011. "Potential for Riverbank filtration in India." *Clean Techn Environ Policy* 13: 295-316.
- Schälchli, U. 1992. "The clogging of coarse gravel beds by fine sediment." *Hydrobiologia* 235 (1): 189-197.
- . 1993. *Die Kolmation von Fliessgewässersohlen: Prozesse und Berechnungsgrundlagen*. Zürich: Dissertation, ETH Zürich.
- . 1995. "Basic equations for siltation of riverbeds." *Journal of Hydraulic Engineering* 121 (3): 274-287.
- Schijven, J., P. Berger, and I. Miettinen. 2002. "Chapter 6. Removal of Pathogens, Surrogates, Indicators, and Toxins Using Riverbank Filtration." In *Riverbank*

- Filtration*, edited by C. Ray, G. Melin and R. B. Linsky, 73-116. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Schmidt, Carsten, Frank Thomas Lange, Heinz-Jürgen Brauch, and Wolfgang Kühn. 2003. "Experiences with riverbank filtration and infiltration in Germany." In *Riverbank Filtration*, edited by C. Ray, G. Melin and R. B. Linsky, 35-48. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Schubert, J. 2002a. "Chapter 3. German Experience with Riverbank Filtration Systems." In *Riverbank Filtration*, edited by C. Ray, G. Melin and R. B. Linsky, 35-48. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- . 2002b. "Hydraulic aspects of riverbank filtration—field studies." *Journal of Hydrology* 266: 145-161.
- . 2004a. "Chapter 1. Significance of Hydrologic Aspects on RBF Performance." In *Riverbank Filtration Hydrology*, edited by S. A. Hubbs, 1-20. Springer & NATO Public Diplomacy Division.
- . 2004b. "Chapter 10: Experience with Riverbed Clogging Along the Rhine River." In *Riverbank Filtration Hydrology*, edited by S. A. Hubbs, 221-242. Springer & NATO Public Diplomacy Division.
- . 2019. Personal Communication.
- Sen, Amartya K. 1970. *Collective choice and social welfare*. San Francisco, California: Holden-Day, Inc.
- SenSW (Senate Department for Urban Development and Housing of Berlin). 2009. "Umweltatlas Berlin - 02.11 Wasserschutzgebiete (Ausgabe 2009)" <https://www.stadtentwicklung.berlin.de/umwelt/umweltatlas/kd211.htm>.
- . 2020. "Anzeige von Bohrungen" <https://www.berlin.de/senuvk/umwelt/wasser/geologie/de/bohrung.shtml>.
- Snow, John. 1849. *On the Mode of Communication of Cholera*. London: John Churchill, Princes Street, Soho.
- Sonthaimer, H. 1991. *Trinkwasser aus dem Rhein?* Sankt Augustin: Academia Verlag.
- Sophocleous, M. 2002. "Interactions between groundwater and surface water: the state of the science." *Hydrogeology Journal* 10 (1): 52-67.
- Sprenger, C., M. Hartog, M. Hernández, E. Vilanova, G. Grützmacher, F. Scheibler, and S. Hannappel. 2017. "Inventory of managed aquifer recharge sites in Europe: historical development, current situation and perspectives." *Hydrogeol J*: 1909-1922.
- Storck, Florian R., Carsten K. Schmidt, F. T. Lange, J. W. Henson, and K. Hahn. 2012. "Factors controlling micropollutant removal during riverbank filtration." *Journal AWWA* 104 (12): E643-E652.
- Stuyfzand, Pieter J., Maria H. A. Juhász-Holterman, and Willem J. de Lange. 2004. "Riverbank filtration in the Netherlands: well fields, clogging and geochemical reactions." Bratislava.
- Tao, T., and K. Xin. 2014. "Public health: A sustainable plan for China's drinking water." *Nature* 511 (7511): 527-528.
- The Disease Daily. 2010. "Paratyphoid A in Guangxi" <http://www.diseasedaily.org/diseasedaily/article/paratyphoid-guangxi-11210>.
- Tobler, W. R. 1970. "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography* 46: 234-240.

- Triantaphyllou, E., and A. Sánchez. 1997. "A sensitivity analysis approach for some deterministic multi-criteria decision making methods." *Decision Sciences* 28 (1): 151-194.
- Tufenkji, Nathalie, Joseph N. Ryan, and Menachem Elimelech. 2002. "Peer Reviewed: The Promise of Bank Filtration." *Environmental Science & Technology* 36 (21): 422A-428A. <https://doi.org/10.1021/es022441j>.
<https://doi.org/10.1021/es022441j>.
- U.S. EPA (United States Environmental Protection Agency). 2000. *The History of Drinking Water Treatment*. United States Environmental Protection Agency.
- . 2001. *Controlling Disinfection By-Products and Microbial Contaminants in Drinking Water*. United States Environmental Protection Agency.
- . 2005. "Water Health Series Bottled Water Basics." U.S. Environmental Protection Agency.
- . 2010. "Long Term 2 Enhanced Source Water Treatment Rule Toolbox Guidance Manual." U.S. Environmental Protection Agency.
- UBA (Federal Environment Agency of Germany). 2018. *Recommendations for reducing micropollutants in waters*. Dessau-Roßlau: Umweltbundesamt (German Environment Agency).
- van Riesen, Sigurd G. 1975. Uferfiltratverminderung durch Selbstdichtung an Gewässersohlen. Dissertation, Universität Karlsruhe.
- Vandas, Stephen J., Thomas C. Winter, and William A. Battaglin. 2002. *Water and the Environment*. American Geological Institute.
- von Neumann, J., and O. Morgenstern. 1953. *Theory of Games and Economic Behavior*. Princeton University Press.
- Wallenius, Jyrki, James S. Dyer, Peter C. Fishburn, Ralph E. Steuer, Stanley Zionts, and Kalyanmoy Deb. 2008. "Multiple Criteria Decision Making, Multiattribute Utility Theory: Recent Accomplishments and What Lies Ahead." *Management Science* 54 (7): 1336-1349.
- Wang, L., X. Ye, and X. Du. 2016. "Suitability Evaluation of River Bank Filtration along the Second Songhua River, China." *Water* 8 (5): 176.
- Wang, Xianze, Zhongmou Liu, Chi Wang, Zhian Ying, Wei Fan, and Wu Yang. 2016. "Occurrence and formation potential of nitrosamines in river water and ground water along the Songhua River, China." *Journal of Environmental Sciences* 50: 65-71.
- WHO (World Health Organization). 2006. *Protecting groundwater for health: Managing the quality of drinking-water sources*. London, UK: IWA Publishing, on behalf of the World Health Organization.
- . 2016. *Protecting surface water for health. Identifying, assessing and managing drinking-water quality risks in surface-water catchments*. Geneva, Switzerland: World Health Organization.
- Xiao, J., F. Liu, Y. Huang, G. Tang, and G. Shi. 2015. "一起农村集中供水污染引起的伤寒暴发疫情 (A Typhoid Outbreak Caused by Pollution of Centralized Water Supply in Rural Areas)." *实用预防医学 (Practical Preventive Medicine)* 22 (10): 1227-1229.

- Xinhua. 2010. "Over 100 dead or missing after floods devastate NE China province"
<https://reliefweb.int/report/china/over-100-dead-or-missing-after-floods-devastate-ne-china-province>.
- Xinhuanet. 2008. "贵阳卫生局称甲肝疫情得到控制已确诊 330 例 (Guiyang Health Bureau Claimed that the Hepatitis A Epidemic has been Controlled while 330 Cases have been Diagnosed)"
<http://www.cctv.com/community315/special/C21232/20080425/109411.shtml>.
- Yin, Wenjie, Yanguo Teng, Yuanzheng Zhai, Litang Hu, Xiaobing Zhao, and Menglin Zhang. 2018. "Suitability for developing riverside groundwater sources along Songhua River, Northeast China." *Human and Ecological Risk Assessment: An International Journal* 24 (8): 2088-2100.
- You, C. 2016. Feasibility Assessment of River Bank Filtration along the Second Songhua River and Yinma River. Jilin: Master Thesis, Jilin University.
- Zhang, Eryong, Yingchun Shi, Cunrong Gao, Xinwei Hou, Zhantao Han, Hongmei Zhao, Jianqing Ding, Xingchun Liu, Baogui Li, Runlian Zhao, Xili Jiao, Lijun Shan, Zhongdao Zhu, and Ning Wang. 2009. "Regional geology and hydrogeology of the Yellow River basin." *Bulletin of the Geological Survey of Japan* 60 (1/2): 19-32.
- Zhang, J., D. L. Mauzerall, T. Zhu, S. Liang, M. Ezzati, and J. Remais. 2010. "Environmental health in China: challenges to achieving clean air and safe water." *The Lancet* 375 (9720): 1110-1119.
- Zhang, Zonghu, Mulin Zhou, and Shixiong Shao. 1990. *中华人民共和国及其毗邻海区第四纪地质图 (1/250 万) 及其说明书 (Quaternary Geological Map of China 1:2500000)*. 中国地图出版社.

