UNIVERSITÄT SALZBURG
Z_GIS - Zentrum für Geoinformatik
Prof. Dr. J. Strobl

Master Thesis

Verarbeitung von Grundwasserdaten

– Zum Problem der räumlichen Interpolation in kleinen Untersuchungsgebieten bei beschränkter Messpunktzahl –

vorgelegt von
Dipl.-Geogr. Christine Schmidt
U 1517
im Fernstudiengang UNIGIS MSC

Bissingen/Teck, den 04.07.14

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Schriften entnommen wurden, sind als solche kenntlich gemacht. Die Arbeit ist in gleicher oder ähnlicher Form oder auszugsweise im Rahmen einer anderen Prüfung noch nicht vorgelegt worden.

Bissingen/Teck, den 04.07.14

_____

Christine Schmidt

Kurzzusammenfassung

Die Master Thesis geht dem Problem der räumlichen Interpolation in kleinen Untersuchungsgebieten mit nur wenigen Messpunkten nach. Das Ziel war, zu prüfen, inwieweit Interpolationsmethoden die primär für wesentlich größere Datenbestände entwickelt wurden, auch bei wenigen Messpunkten anwendbar sind.

Die verfügbaren geostatistischen Verfahren sind für geringe Stichprobengrößen wenig geeignet, zumal die Annahmen und Bedingungen vielfach nicht eingehalten sind.

Dennoch ist es möglich, Interpolationen mit geringen Probengrößen durchzuführen und es wird weitere Forschungsbedarf in dieser Hinsicht gesehen.

Vorbemerkungen und Dank

Zum Abschluss des UNIGIS MSC-Fernstudienlehrgangs gehört die Erstellung einer Master Thesis als wissenschaftliche Abschlussarbeit. Die Arbeit daran erfolgt fern der physischen Universität im häuslichen Umfeld und parallel zur üblichen Erwerbstätigkeit.

Im Rahmen von Präsenzveranstaltungen und während des Bearbeitens von Studienmodulen entsteht Inspiration und es kommt zu einer Vision für eine Master Thesis, ein Konzept entsteht und man nimmt sich viel vor. Und dabei kann es vorkommen, dass sich das Thema dann doch als schwer überschaubar herausstellt. Die Umsetzung der Vision ist dann vom persönlichen Standpunkt her praktisch nicht möglich.

Wird das Thema pragmatisch auf das begrenzt, was mit persönlichen Ressourcen an Zeit, Quellen und Sachverstand erreichbar ist, mag die Arbeit in mancher Hinsicht aus wissenschaftlicher Sicht nicht optimal sein. Für den engagierten Praktiker mit wissenschaftlichem Hintergrund kann sie in der Realität des Alltags aber trotzdem einen Fortschritt bedeuten. Wenn an diesem Fortschritt andere Praktiker teilhaben können, umso besser. Diese Thesis arbeitet die gesammelten Informationen zielgerichtet auf und sollte auch für andere „Nicht-Mathematiker“ verständlich sein.

An dieser Stelle möchte ich mich ganz besonders bei meinem Lebensgefährten für seine Geduld und vor allem seine moralische Unterstützung bedanken. Als Ingenieur konnte er in manchen mathematischen Fragen Aufklärung bieten und das Verständnis für mathematische Ausdrücke und „Codes“ erleichtern. Bei Christoph Traun bedanke ich mich für entscheidende Tips und offene Worte, ohne die ich bestimmt aufgegeben hätte.

An dieser Stelle sei jenes ironische Sprichwort angeführt, das bestimmt nicht für alle, aber doch wenigstens für einige Geologen zutrifft (und analog auf Geographen ausgedehnt werden kann):

> *„Geologen sind schlechte Mathematiker ! - Warum ?*
> *Wären sie gute Mathematiker, wären sie Ingenieure geworden !“*

Und so ist diese Master Thesis geworden, was sie ist.

Christine Schmidt

Inhalt

Verzeichnis der Abbildungen

Verzeichnis der Tabellen

Verzeichnis des Anhangs

1 Einleitung

1.1 Anlass und Hintergrund des Themas

In der Praxis der Erkundung und Überwachung von Altlasten und Schadensfällen mit Untergrund- und Grundwasserverunreinigungen stellt sich häufig die Frage nach einer anschaulichen Darstellung von Analysenergebnissen oder Messwerten (z.B. Grundwasserständen), die punktuell aus Grundwassermessstellen oder Bohrungen gewonnen werden. Dabei geht es sowohl um eine Präsentation wie auch um eine Interpretation von Messdaten als Hilfestellung zur Bewertung von Befunden.

Aus punktuell verteilten Daten lässt sich eine Information über die Lage und Ausdehnung einer Kontamination oft nicht auf den ersten Blick ableiten. Auch wenn die Werte zu den jeweiligen Messstellen auf einem Plan aufgetragen werden, muss der Betrachter diese erst einzeln erfassen und zu einem Gesamteindruck verarbeiten. Bei einer Punktdarstellung mit geeigneter kategorische Farbabstufung wird eine Übertragung der diskreten Punktdaten auf die Fläche „im Kopf“ des Betrachters unterstützt. Das entspricht im Prinzip bereits einer Art Interpolation und setzt eine „nichtdigitale“ mentale Verarbeitung der Punktinformationen durch den Betrachter voraus.

Grundsätzlich kann eine Interpretation und Übertragung der Punktdaten auf die Fläche auch von Hand erfolgen. Vor Anbruch des digitalen Zeitalter blieb Geologen bei der Erstellung von geologischen Karten gar nichts anderes übrig und sie haben sich dabei auf ihre Erfahrung verlassen.

Computergestützte Verfahren zur räumlichen Interpolation bieten grundsätzlich die Möglichkeit, im Raum verteilte punktuellen Informationen in kurzer Zeit auf Flächen zu übertragen, um damit eine Übersicht über die Lage in einem bestimmten Gebiet zu geben. Aber eine räumliche Interpolation ist auch immer mit einer Interpretation von Punktinformationen verbunden, die nicht unkritisch einem Computer oder Programm überlassen werden darf.

Geologische Daten sind teuer und aus ökonomischen Gründen nicht an beliebig vielen Orten verfügbar. Dies trifft insbesondere für Projekte mit kleinen Untersuchungsgebieten zu, wie sie häufig von Ingenieurbüros und freiberuflichen Geologen im Rahmen der Erkundung und Überwachung von Altlasten und Schadensfällen zu bearbeiten sind. Die Anzahl an verfügbaren Messpunkten ist in diesen Fällen folglich im Vergleich zu den in der Literatur für räumliche Interpolationen verwendeten und behandelten Beispielen [siehe z.B. KRIVORUCHKO 2011, BIVAND ET AL. 2008] gering.

Als Anzahl für „geringe verfügbare Messstellenanzahl“ seien hier typischerweise 5 bis 20 Messstellen in Untersuchungsgebieten angegeben, die sich über ein bis mehrere Grundstücke ausdehnen können und eine Fläche von zusammen bis zu etwa 1 bis 2 ha einnehmen.

1.2 Zielsetzung der Arbeit

Das Hauptziel der hier vorgelegten Master Thesis war die Klärung der Frage, ob die Anwendung von räumlichen Interpolationsverfahren in kleinen Untersuchungsgebieten mit nur wenigen Messpunkten nach wissenschaftlichen Gesichtspunkten sinnvoll durchführbar ist.

Die Formulierung einer klassischen Nullhypothese dazu lautet:

> *„Es ist nach wissenschaftlichen Kriterien nicht sinnvoll und zulässig, in kleiner Untersuchungsgebieten mit nur wenigen Untersuchungspunkten räumliche Interpolationen durchzuführen“*.

Eine Begründung für diese Nullhypothese könnte sein, dass die allgemeine wissenschaftliche Erfahrung dagegen spricht, so etwas zu tun und die Interpolationsverfahren für einen größeren Datenumfang entwickelt und ausgelegt wurden. Das würde insbesondere für die geostatistischen Interpolationsverfahren gelten.

Das Arbeitsgebiet und damit auch der Geltungsbereich von Aussagen und Schlüssen, die in oder aus dieser Master Thesis gezogen werden soll hier vorsichtshalber eingegrenzt oder wenigstens explizit definiert werden: Diese Arbeit wurde vor dem in Abschnitt 1.1 bezeichneten Hintergrund verfasst, der ein eindeutig geologisch-geographischer ist.

1.3 Gliederung der Master Thesis

Nach einer kurzen Übersicht über die für diese Arbeit verwendeten Materialien folgt ein längerer Teil, in dem Informationen zu verschiedenen Interpolationsverfahren, aber auch einige konzeptionelle Inhalte zusammengetragen sind. Insbesondere die Geostatistik ist ein sehr komplexes Feld und man muss versuchen, die dort verwendeten Begriffe zu durchdringen, sonst begeht man bei der Anwendung höchstwahrscheinlich Fehler. Das Kapitel ist zweigeteilt in einen Abschnitt für die deterministischen Verfahren und in einen Abschnitt für die probabilistischen Verfahren. Es wird vor allem auf gängige Verfahren eingegangen, die für eine Anwendung auch bei einem kleinen Datenumfang in Frage kommen.

Darauf folgt ein Teil, in dem einige Versuche mit einem fiktiven Datensatz angestellt werden und die Ergebnisse miteinander verglichen werden. Gleichzeitig erfolgt eine Erörterung und Diskussion von praktischen Problemen bei kleinen Datensätzen.

Am Schluss steht eine Zusammenfassung der Ergebnisse.

2 Materialien, Quellen und Software

2.1 Gesichtete Literatur

Für die hier vorgelegten Master Thesis wurden einige einschlägige Lehrbücher zur Geostatistik gesichtet, insbesondere das relativ aktuelle Werk von KRIVORUCHKO [2011]. Das schon etwas ältere Werk von ISAAKS UND SRIVASTAVA [1989] war leider erst zu einem relativ späten Zeitpunkt verfügbar. Als wichtige Quellen seien hier auch online verfügbare Schriften von Mitarbeitern der US EPA und Mitarbeitern des Centre de Géostatistique in Fontainebleau genannt, zu denen auch Schriften von Georges Matheron, einem der Pioniere der Geostatistik, gehören. Im Literaturverzeichnis zu dieser Master Thesis sind die Links zu den Online-Quellen enthalten.

2.2 Verwendete Software

Obwohl zu Beginn der Arbeiten an dieser Master Thesis vorgesehen war, mit OpenSource- Software zu arbeiten, wurde dies in einer späteren Phase fallen gelassen, da es sich als zu zeitaufwändig herausgestellt hat.

Die praktischen Versuche zu dieser Master Thesis wurden im wesentlichen mit den Programmen Surfer 12[1] und ArcGIS Geostatistical Analyst[2] durchgeführt. Für einige Schritte zur Vorbereitung des Datensatzes, mit dem die zentralen Frage dieser Master Thesis etwas praktischer angegangen werden sollten, was das Programm QGIS[3] hilfreich.

2.3 Zum Begriff „Modelle“

Der Begriff „Modell“ wird allgemein in sehr unterschiedlicher Bedeutung verwendet.

Entsprechend dem Sprachgebrauch in der gesichteten und verwendeten Literatur wird der Begriff des Modells in dieser Arbeit im Sinne von statistischen und geostatistischen Modellen, bei denen es um Verteilungsannahmen geht oder mathematische Modellfunktionen verwendet, welche die räumliche Variabilitätsstruktur (Variogrammtypen) beschreiben.

Im Kapitel 4 wird der Begriff „Modell“ auch für einen kleinen Datensatz verwendet, der zum Testen verwendet wurde. Damit soll vermieden werden, dass dieser Miniaturdatensatz mit realen Messdaten oder Befunden verwechselt wird.

---

[1] Surfer 12 der Firma Golden Software, http://www.goldensoftware.com/

[2] ArcGIS Geostatistical Analyst http://www.esri.com

[3] QGIS oder Quantum GIS, siehe http://qgis.org

3 Recherche zu Interpolationsverfahren

3.1 Allgemeines

Räumliche Interpolation ist nichts anderes als intelligentes Schätzen von Werten zwischen bekannten Datenpunkten, unterstützt von Toblers erstem Gesetz "*Everything is related to everything else, but near things are more related to each other*" [siehe auch DE SMITH ET.AL. 2010, S. 82] .

Grundsätzlich wird zwischen deterministischen und statistischen (stochastischen, probabilistischen) Interpolationsverfahren unterschieden.

Deterministische Interpolationsmodelle beruhen auf mathematischen Formeln, welche die Glätte der resultierenden Oberfläche bestimmen , während *Statistische Modelle* ihre Parameter auf der Grundlage der Analysen der tatsächlichen Daten ableiten [KRIVORUCHKO 2011, S. 251].

Die Anwendung einiger grundlegender Annahmen aus der Wahrscheinlichkeitsrechnung (Statistik) auf die räumliche Korrelation geowissenschaftlicher Kenngrößen mit eindeutigem Raumbezug wird unter dem Kurzbegriff „Geostatistik“ zusammengefasst. [...] Strenggenommen beschränken sich geostatistische Verfahren auf die Behandlung der lokalen Variabilität einer ortsabhängigen Variablen [SCHAFMEISTER 1999, S. 2].

Konzeptionelle Unterschiede zwischen klassischer Statistik und Geostatistik werden in Abschnitt 3.5.2 (Regionalisierte Variable) kurz angesprochen.

Die Disziplin „Geostatistik“ geht ganz wesentlich auf Erfahrungen und Ansätze zurück, die im Montanwesen (Bergbau: Krige, Matheron) und in der Meteorologie (Gandin) gewonnen und entwickelt wurden. Insbesondere der Bergbauingenieur Georges Matheron prägte seit den 1960er Jahren viele Begriffe und Konzepte der Geostatistik [AGTERBERG 2003, SCHAFMEISTER 1999, KRIVORUCHKO 2011].

In den folgenden Abschnitten werden Informationen zusammengestellt und zusammengefasst, die bei der räumlichen Interpolation von Punktdaten allgemein und insbesondere bei der Interpolation von Grundwasserdaten in kleinen Untersuchungsgebieten eine Rolle spielen bzw. eine Rolle spielen können.

3.2 Was bedeutet „geringe Messpunktzahl“ ?

Darüber, was „wenige“ Datenpunkte sind und wieviele Messpunkte für eine Interpolation überhaupt erforderlich sind, liegen in der Quellen unterschiedliche Angaben und Einschätzungen vor, wobei nicht immer konkrete Zahlen genannt werden. Die Spannweite reicht von vagen Aussagen wie „you can use spatial interpolation to create an entire surface from just a small number of sample points; however, more sample points are better if you want a detailed surface“ [4] bis zu konkreten Angaben

---

[4] Abruf am 28.05.2014, 22:15 Uhr http://www.geography.hunter.cuny.edu/~jochen/GTECH361/lectures/lecture11/concepts/What%20is%20interpolation.htm Kein Verfasser angegeben, es handelt sich

von WEBSTER UND OLIVER [1993, zitiert in CUI ET AL. 2006], die zur Festlegung des Kriging-Modells und Schätzung seiner Parameter mindestens 100 – 150 Beobachtungspunkte für erforderlich halten.

Auch DE SMITH ET.AL. [2013, S. 515] geben eine zusammenfassende Autorenmeinung wieder, nach der idealerweise wenigstens 150 Datenpunkte vorhanden sein sollten, um befriedigende Semivarianz-Schätzungen zu erhalten, auch wenn weniger Punkte mit verwendet werden könne, insbesondere, wenn Hilfsinformationen über die Daten verfügbar sind.

Ein Beitrag in StackExchange setzt eine Anzahl von 20 – 100 Messpunkten in der Literatur als Minimum für Kriging-Verfahren an und es soll darüber hinaus einen Konsens über eine Mindestpunktanzahl um die 30 geben [STACKEXCHANGE 2013]. Die angegebene Quelle [ASTM D5922-96 2010] für den Konsens in StackExchange gibt allerdings keine Mindestpunktanzahl für Kriging, sondern vielmehr eine Mindestzahl von 30 Punktpaaren pro Lag in einem Semivariogramm an.

Eine Nachfrage beim Autor[5] des Beitrags in StackExchange ergab die Einschätzung, dass die in der Literatur angegebenen Mindestzahlen an Messpunkten für Kriging-Interpolation zwischen 20 und 100 generell nur als grobe Richtschnur anzusehen sind und es seiner Einschätzung nach auf die Eigenschaften des jeweiligen Datensatzes ankäme, nicht auf das Erreichen einer bestimmten Mindestanzahl. Es gibt Fälle, in denen 12 Messstellen zur Schätzung eines Variogramms ausreichen, aber auch Fälle, bei denen eine Variogrammschätzung selbst mit 100 Messpunkten nicht möglich sei. Die in der Literatur angegebenen Richtwerte könnten aber auch als Warnschwellen verstanden werden, um bei der Verwendung von einer geringeren Anzahl von Messpunkten für Interpolationen besonders vorsichtig zu sein.

Der ArcGIS Geostatistical Analyst erwartet Eingabewerte von mindestens 10 verschiedenen Messorten, wobei es nach KRIVORUCHKO [2011, S. 97] erforderlich sein kann, Messwerten verschobene Koordinaten zuzuweisen, um eine Analyse (z.B. über ein Histogramm) zu erlauben.

Im Benutzerhandbuch von Surfer 10 wird angegeben [S. 243], dass Surfer nur drei nicht kollineare Koordinatenpunkte benötigt, um einen Gridding-Prozess[6] durchzuführen. Im Surfer-Handbuch geht man davon aus, dass bis zu 10 Messpunkte nur dazu ausreichen würden, einen groben Trend oder eine grobe Richtung festzustellen. Auch Datensätze von 10 bis 250 Punkten werden noch als „kleine Datensätze“ eingestuft.

Die Geostatistik ist historisch sehr stark von Bergbauingenieuren und Meteorologen geprägt worden. In beiden Fachbereichen steht grundsätzlich ein ausreichend

aber um die Webseiten von Jochen Albrecht am Hunter College, New York.

[5] William Huber, via Mail im Juni 2014 kontaktiert über die Webseite seiner Firma QD bzw. Quantdec (http://www.quantdec.com/), die gutachterliche Dienstleistungen vor allem im Bereich der Datenanalyse und -interpretation erbringt.

[6] In der Surfer-Umgebung wird von „Gridding“ (grid = engl. für Gitter) gesprochen, wenn es um Interpolationsvorgänge geht, weil die Interpolation nicht für diskrete einzelne Punkte, sondern für ein regelmäßiges Gitter an Punkten berechnet wird.

großer Datenumfang mit einigen hundert bis mehreren tausend Messpunkten für geostatistische Analysen und Auswertungen zur Verfügung.

Bei umweltbezogenen Fragestellungen, das sind Boden-, Grundwasser- und Luftmessungen, liegt die Datendichte mit 10 – 50, maximal 100 Messstellen in einer wesentlich kleineren Größenordnung. Dieser Umstand erschwert mitunter den Einsatz geostatistischer Verfahren. Dennoch ist ihre Anwendung nach SCHAFMEISTER [1999, S. 3] sinnvoll, da nur geostatistische Verfahren die quantitative Betrachtung der Zuverlässigkeit der Ergebnisse erlauben.

JERNIGAN [1986] behandelt das Problem eines geringen Stichprobenumfangs bei geostatistischer Interpolation (Kriging) in einer Monographie der US EPA[7] und demonstriert an einem Beispiel mit 8 Messpunkten (pH-Werte im Abstrom einer Deponie), dass dies zwar nicht ganz einfach, aber doch möglich ist.

JERNIGAN [1986, S. 3] stellt fest, dass zwar viel traditionelle Arbeit zur Entwicklung von Techniken zur Verwendung mit großen Datensätzen und bis dato nur wenig in Techniken zur Verwendung bei kleinen Datensätzen investiert wurde. Kriging-Verfahren sind bei kleinem Datenumfang zu modifizieren und weiter zu erforschen.

Da nicht allzu viel Literatur zu diesem Thema gefunden wurde, gilt das möglicherweise immer noch. Zum Problem der räumlichen Interpolation in kleinen Untersuchungsgebieten gibt es Erfahrungen und es wird darüber auch berichtet. Allerdings überwiegend nur unter Überschriften, die keine Inhalte erwarten lassen,die zur zur Beantwortung Fragestellung dieser Master Thesis, so dass es nicht ganz einfach ist, diese Informationen zu sammeln.

3.3 Isotropie und Anisotropie

Isotropie bedeutet eine Gleichförmigkeit von Merkmalen nach allen Richtungen, Anisotropie das Gegenteil, also eine richtungsabhängige Ausprägung von Merkmalen.

Bei JERNIGAN [1986, S. 14] zeigte sich beispielsweise in Daten im Abstrom einer Deponie eine Anisotropie dadurch, dass die Messungen quer zum Schadstoffstrom einer Deponie viel größere Variabilität auf kurze Distanz zeigten, als die Messungen in Stromrichtung. Die abstromigen Messungen waren sich viel ähnlicher und hatten deshalb viel weniger Variabilität innerhalb des Schadstoffstroms. Eine Annahme von Isotropie wäre in diesem Fall nicht angebracht gewesen.

Sowohl deterministische, als auch statistische (stochastische, probabilistische) Interpolationsmodelle können Anisotropie berücksichtigen, verwenden dabei allerdings unterschiedliche Konzepte [siehe u. a.. KRIVORUCHKO 2011, S. 256].

Weil Deterministische Interpolationsmethoden auf ausgewählte mathematische Funktionen und nicht auf den Daten selbst beruhen, kann Anisotropie in deterministischen Modellen (einschließlich IDW) als Transformation der Koordinaten definiert werden, so dass sich die Distanz unter Angabe eines anzugebenden

[7] United States Environmental Protection Agency

Faktors in einer Richtung schneller ändert, als in der dazu senkrechten Richtung [KRIVORUCHKO 2011, S. 256]

Wenn keine richtungsabhängigen Strukturen oder Effekte auf die Daten vorliegen, können benachbarte Punkte in allen Richtungen unter Verwendung einer kreisförmigen Nachbarschaft gewählt werden. Bei richtungsabhängigen Wirkungen (z. B. durch eine Rinnenstruktur im Aquifer oder den Grundwasserstrom), sollte die Suchnachbarschaft elliptische Form annehmen [siehe u. a. KRIVORUCHKO 2011, S. 256].

Eine Variation mit der Richtung kann unterschiedliche Semivariogramm-Parameter betreffen [KRIVORUCHKO 2011, S. 328] :

- eine Anisotropie der Reichweite (Range) wird geometrische Anisotropie genannt. Solch eine Anisotropie kann mit einer affinen Transformation auf den Zahlenbereich der Zufallsfunktion entfernt werden [Myers 1991, S. 211].
- eine Anisotropie des partiellen Sill wird zonale Anisotropie genannt

Eine Anisotropie des Nugget selbst ist nicht möglich. Trotzdem kann der Messfehler als Bestandteil des Nugget (siehe auch Abb. 2, Seite 26) mit der Richtung variieren und demzufolge eine Anisotropie aufweisen [KRIVORUCHKO 2011, S. 328].

Eine zonale Anisotropie, also eine Variation des partiellen Sill mit der Richtung, ist ein Anzeichen für Nichtstationarität, dem vor einer Anwendung des Kriging mit Trendeliminierung (Detrending) und Transformationswerkszeugen zu begegnen ist. Sollte das nicht möglich sein, kann ein nichtstationäres Modell […] verwendet werden. [KRIVORUCHKO 2011, S. 328].

[JERNIGAN 1986 S. 13] Isotropie. Viele im Kriging entwickelte Gleichungen werden viel einfacher, wenn wir annehmen, dass der Prozess die gleiche Kovarianzstruktur in jeder Richtung aufweist. Solch ein stochastischer Prozess wird homogen und isotrop genannt, d.h. seine Wahrscheinlichkeitsstruktur ändert sich nicht mit einer Änderung der räumlichen Position oder eine Rotation um eine räumliche Position. Isotropie zeigt ein gewisse Gleichförmigkeit in jeder Richtung an.

[JERNIGAN 1986 S. 14] Isotropie ist von großem Nutzen bei kleinen Datensätzen, bei denen es schwierig ist, Kovarianzen oder Variogramme zu modellieren und zu schätzen, die sowohl von der Distanz, als auch von der Richtung zwischen zwei Punkten im Raum abhängen. Viel mehr Daten können kombiniert werden, wenn wir annehmen, dass die Kovarianz zwischen zwei Punkten nur von der Distanz zwischen Punkten abhängt. Die Annahme von Isotropie ist zum Schätzen des Prozesses mit Kriging nicht erforderlich; sie dient der Vereinfachung der Gleichungen und Berechnungen. Tatsächlich wäre solch eine Annahme in vielen Anwendungen nicht realistisch.

3.4 Deterministische Interpolation

Deterministische Interpolationsmodelle basieren auf mathematischen Formeln, welche die Glätte der resultierenden Oberfläche bestimmen. Die Glätte kann mit

Kreuzvalidierungs-Techniken kalibriert werden. Es ist allerdings unwahrscheinlich, dass diese Glätte mit natürlicher Datenvariabilität zusammenhängt, weil deterministische Modelle keine Beziehungen zwischen den tatsächlichen Daten berücksichtigen. Ein weitere Nachteil deterministischer Verfahren ist, dass sie kein Maß für die Genauigkeit der Vorhersagen (Prognosen) liefern können [KRIVORUCHKO 2011, S. 251].

3.4.1 Inverse Distance Weighted Interpolation (IDW)

Die Inverse Distance Weighting-Modelle arbeiten unter der Prämisse, dass von einem Schätzpunkt weiter entfernt liegende Messpunkte einen geringeren Beitrag zur Berechnung eines neuen Punktes leisten sollten, als nahe gelegene [vgl. DE SMITH ET.AL. 2010, S. 491].

Die Berechnung des Wertes am Vorhersageort (Ort ohne Messwert) erfolgt durch Aufsummieren der gewichteten Anteile der benachbarten Messpunkte. Die Gewichtung erfolgt umgekehrt proportional der p-ten Potenz (Powerwert[8]) der Distanz zwischen Messpunkten und Vorhersageort, d.h. Punkte, welche dem Vorhersageort näher liegen, gehen mit einem höheren Anteil in die Berechnung des Vorhersagewertes ein [KRIVORUCHKO 2011, S. 251].

$$\hat{Z}(s_0) = \sum_{i=1}^{N} \lambda_i \cdot Z(s_i)$$

mit $\hat{Z}(s_o)$ = Vorhersagewert für den Punkt s_0
$Z(s_i)$ = bekannter Messwert am Punkt s_i
λ_i = Gewicht für den i-ten Probenpunkt
N = Anzahl der verwendeten Punkte

Die Gewichte ergeben sich aus $\lambda_i = \frac{d_{i0}^{-p}}{\sum_{i=1}^{N} d_{i0}^{-p}}$ mit $\sum_{i=1}^{N} \lambda_i = 1$

wobei d_{ij} = Abstand zwischen den Punkten s_i und s_0
p = Power-Wert, größer oder gleich 1

Der Defaultwert ist üblicherweise p = 2 (z.B. auch im ESRI Geostatistical Analyst), es ist aber generell besser, den optimalen Powerwert auf der Grundlage der Kreuzvalidierung zu wählen. Die Gewichte werden so skaliert, dass ihre Summe pro Vorhersagepunkt gleich 1 ist. [KRIVORUCHKO 2011, S. 254].

Bei einem hohen p-Wert gehen nur wenige umgebende Punkte in die Vorhersage ein. Es gibt eine nur geringe Mittelung und der vorhergesagte Wert wird gleich oder

---

[8] Für sehr kleine Powerwerte (i.e. z.B. 0,01) sind die Gewichte für alle Abstände annähernd gleich. Der ESRI Geostatistical Analyst erlaubt nur Powerwerte > 1, weil sich Powerwerte < 1 bei der Berechnung auf naheliegende Punkte weniger stark auswirken, als auf entferntere Punkte, was als unsinnig gilt [KRIVORUCHKO 2011, S. 253]

ähnlich dem nächstgelegenen Punkt sein. Wenn der p-Wert sehr groß ist (z.B. 20), ergibt sich eine Vorhersagekarte, die einer Voronoikarte ähnelt. Eine ähnliche Darstellung kann mit einem Power-Wert von 2 und nur einem Nachbarwert erzielt werden. Bei hoher Datendichte und räumlicher Korrelation liegt der optimale p-Wert nahe 3. Wenn die räumliche Korrelation schwach ist, liegt der optimale p-Wert nahe 1, ein optimaler p-Wert von 1 ist also ein Hinweis auf Datenunabhängigkeit. Im Fall unabhängiger Daten sollte jedoch keine Interpolation durchgeführt werden [KRIVORUCHKO 2011, S. 255], weil das keinen Sinn macht.

Der Geostatistical Analyst versucht, einen optimalen Powerwert für IDW zu berechnen, indem Root Mean Square Prediction Error für mehrere verschiedene Powerwerte unter Verwendung der gleichen Suchnachbarschaft ausgegeben wird. Ein lokales Polynom zweiter Ordnung wird an die Punkte angepasst und der Powerwert mit den geringsten Fehler, ist der optimale Wert [KRIVORUCHKO 2011, S. 254]

Der Geostatistical Analyst erlaubt nur die Verwendung von Powerwerten > 1, weil bei Powerwerten < 1 naheliegende Punkte weniger stark auf die Berechnung der Vorhersagewerte auswirken, als entferntere Punkt, was widersinnig sei [KRIVORUCHKO 2011, S. 254],

IDW ist ein exakter Interpolator, d. h. er sagt an den Messorten einen mit dem Messwert identischen Wert vorher [KRIVORUCHKO 2011, S. 252]. Maximal- und Minimalwerte der interpolierten Oberfläche können bei IDW nur an Messpunkten auftreten. Wenn Daten nicht absolut präzise sind (typischer Anwendungsfall), macht eine exakte Interpolation nicht viel Sinn [KRIVORUCHKO 2011, S. 252].

Bei einer vergleichenden Untersuchung mit 54 Datensätzen, darunter 20 Datensätze mit 25 Messpunkten, kamen WEBER UND EGLUND [1993, S. 8] zu dem Schluss, dass IDW-Schätzer sehr empfindlich auf den Datentyp, die Anzahl der für die Schätzung verwendeten Nachbarn und für die verwendete Distanz-Power zur Gewichtung sind. Bei Höhen-Daten stieg die Schätzqualität im allgemeinen mit abnehmender Anzahl an Nachbarn und steigender Distanz-Power. Für Höhen-Varianz-Daten war das Gegenteil der Fall.

Bewertetet wurde die Genauigkeit von 15 verschiedenen räumlichen Schätzern mit den gleichen 54 Proben-Datensätzen, wobei sich IDW und IDW-Quadrat-Interpolatoren als etwas besser als Ordinary Kriging und Simple Kriging herausstellten. Die Autoren relativierten ihre Ergebnisse jedoch mit der Einschätzung, dass die Art der verwendeten Daten IDW-Verfahren begünstigt haben mag und die Resultate nicht bedeuteten würden, dass IDW-Methoden den Kriging-Schätzern notwendigerweise in jedem Fall überlegen wären. Starke Anisotropien oder Cluster hätten Kriging-Verfahren begünstigt. Verwendet worden waren Powerwerte zwischen 1 und 3 und 4 – 20 Punkte als Suchnachbarschaft.

Die Vorteile von IDW sind seine hohe Rechengeschwindigkeit [KRIVORUCHKO 2011, S. 258], was in kleinen Untersuchungsgebieten mit einem kleinen Datenumfang aber kaum ins Gewicht fällt, und seine Transparenz, d.h. Nachvollziehbarkeit aufgrund des relativ einfachen Berechnungsalgorithmus. Trotz seiner Beliebtheit bei

GIS-Nutzern fehlt der Methode allerdings die Fähigkeit, eine Aussage zur Vorhersageunsicherheit zu treffen.

Daten sollten vor einer Interpolation auf Ausreißer und Datencluster geprüft und ggfs. bereinigt werden, da die IDW-Oberfläche nach KRIVORUCHKO [2011, S. 252] empfindlich für Datencluster und Daten-Ausreißer ist. Das Standard-IDW-Verfahren neigt zudem zum Erzeugen von Bullaugen-Effekten [siehe auch de Smith et. al. 2013, S. 492].

3.4.1.1 Shepard 's Modified Method

Mit Modified Shepard wird eine von Shepard eingeführte Variante des Standard-IDW-Verfahrens bezeichnet, das lokale kleinste Quadrate-Schätzungen verwendet. Es erzeugt weniger Artefakte, kann extrapolieren und ist ein exakter Interpolator, wenn kein Glättungsfaktor angegeben ist [DE SMITH 2013, S 481].

Die Methode beinhaltet die Verwendung von zwei Powerwerten statt nur einem im Standard-IDW: einem niedrigen Wert (im allgemeinen 2) für nahegelegene Punkte und einen höheren Wert (im allgemeinen 4) für entferntere Punkte [DE SMITH 2013, S. 501] .

Eine andere Variante, die in einigen Programmen implementiert ist (z.B. GMS[9]), stellt die Gewichte auf der Grundlage der Entfernung des entferntesten Punktes ein, der im gesamten Datensatz oder innerhalb eines bestimmten Radius zu finden ist. Wenn dieser Abstand R ist, lautet die überarbeitet IDW-Formel [DE SMITH ET AL. 2013, S. 501]:

$$z_j = k_j \sum_{i=1}^{n} z_i \left(\frac{R - d_{ij}}{R\, d_{ij}}\right)^{\alpha}$$

Der Wert k_j in dem Ausdruck ist_

$$k_j = \sum_{i=1}^{n} \left(\frac{R - d_{ij}}{R\, d_{ij}}\right)^{\alpha}$$

Surfer verwendet ein komplexeres Verfahren auf der Basis einer lokalen quadratischen polynomen Anpassung, die manchmal die modifizierte Shepard-Methode genannt wird. Die Methode setzt dann mit einem inversen Distanztypmodell fort und verwendet dabei Oberflächenwerte aus der angepassten quadratischen quadratischen Oberfläche anstatt der Originaldaten. Das Verfahren ist ohne Glättungsfaktor exakt [DE SMITH ET. AL. S. 501].

3.4.2 Polynome Interpolation

Es ist zu unterscheiden zwischen globaler und lokaler polynomer Interpolation. Bei beiden handelt es sich um keine exakten Interpolatoren, d.h. die resultierende

---

[9] GMS, Groundwater Modelling System http://www.aquaveo.com/

Oberfläche muss nicht durch jeden Messpunkt gehen. Generell ist nach Krivoruchko [2011, S. 268] nicht-exakte Interpolation bei weich verlaufenden [kontinuierlichen?] Daten mit Messfehlern und örtlichen Abweichungen in Temperatur und Höhe zu bevorzugen.

Während bei globaler polynomer Interpolation ein Polygon an die gesamte Oberfläche (bzw. alle Daten) angepasst wird, passt lokale polynome Interpolation viele Polynome jeweils innerhalb überlappender Nachbarschaften an. Datenpunkte, die näher am Vorhersageort liegen, haben dabei höhere Gewichte. Damit erzeugt lokale polynome Interpolation Oberflächen, die eine größere Datenvariation berücksichtigen [KRIVORUCHKO 2011, S. 268].

Eine typische Anwendung von polynomer Interpolation ist das Aufdecken und Filtern großmaßstäblicher Datenvariation und deren Filterung [KRIVORUCHKO 2011, S. 268].

3.4.2.1 Globale polynome Interpolation (GPI)

Globale polynome Interpolation funktioniert wie das Anpassen eines Stück Papiers zwischen auf Stäbchen lagernden Datenpunkten mit - entsprechend ihrer Werte - unterschiedlichen Höhen. Eine Oberfläche wird dann so diese Stäbchen gelegt, dass die Summe der Höhen über der Oberfläche etwa so groß ist, wie die Summe der Höhen unter der Oberfläche [KRIVORUCHKO 2011, S. 264] (siehe auch Abb. 1, Seite 18).

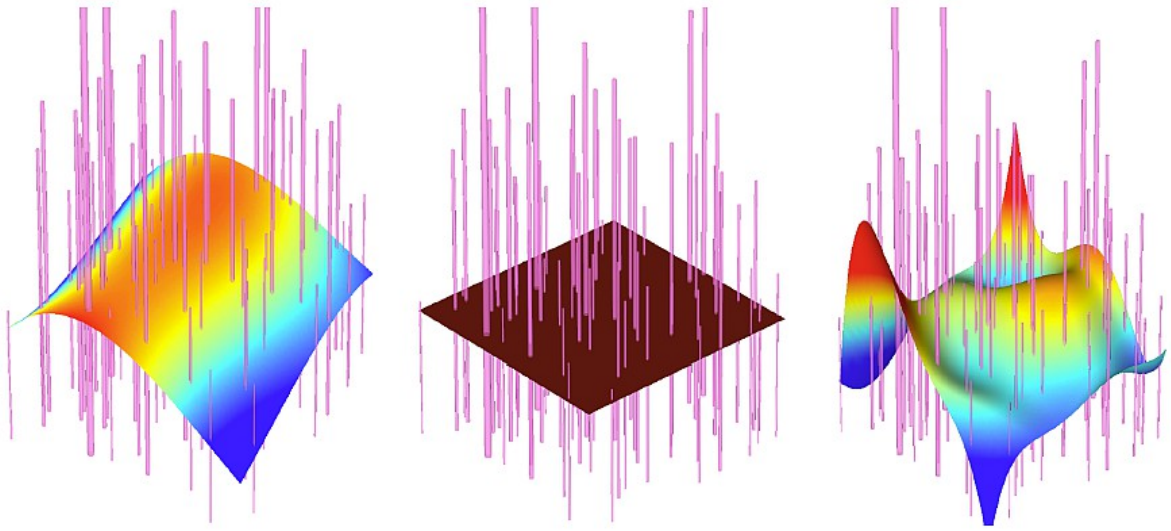


Abb. 1: Polynome Interpolation

Quelle: KRIVORUCHKO [2011, S. 264]

(Anmerkung: links = Polynom 2. Ordnung; Mitte = Polynom 0. Ordnung, rechts = Polynom 5. Ordnung)

Es gibt Polynome unterschiedlichen Grades, d. h. Von unterschiedlicher Komplexität (Formeln entnommen aus KRIVORUCHKO 2011, S. 264].

1. Das einfachste Polynom ist eine durch die Formel $\hat{Z}(x\ ,\ y) = a_0$ beschriebene ebene Fläche, die einer Konstanten im zweidimensionalen Raum entspricht. Es handelt sich um ein Polynom nullter Ordnung (siehe Abb. 1 Mitte).

2. Ein lineares Polynom ist ein Polynom erster Ordnung, beschreibt mit der Formel $\hat{Z}(x\ ,\ y) = a_0 + a_1\ x + a_2\ y$ eine schräge Fläche und entspricht einer im zweidimensionalen Raum geneigten Linie.

3. Ein quadratisches Polynom ist ein Polynom zweiter Ordnung und damit eine Fläche, die eine parabolische Komponente enthält. Sie wird beschrieben durch die Formel $\hat{Z}(x\ ,\ y) = a_0 + a_1\ x + a_2\ y + a_3\ x^2 + a_4\ xy + a_5\ y^2$ und entspricht einer Parabel im zweidimensionalen Raum. (siehe Abb. 1 links).

4. Ein kubische Polynom ist ein Polynom dritter Ordnung und wird durch die Formel
   $\hat{Z}(x\ ,\ y) = a_0 + a_1\ x + a_2\ y + a_3\ x^2 + a_4\ xy + a_5\ y^2 + a_6\ x^3 + a_7\ x^2\ y + a_8\ xy^2 + a_9\ y^3$
   beschrieben.

5. Ein quartisches oder biquadratisches Polynom ist ein Polynom vierter Ordnung. Eine Formelbeschreibung wird mit steigendem Polynomgrad immer komplexer und länger, auf eine Wiedergabe wird hier deshalb verzichtet zumal Polynome höheren Grades bei Interpolationen zunehmen kritischer bzw. heikler zu behandeln sind.

Globale Polynome werden häufig zur Interpolation bzw. Modellierung von Trendoberflächen verwendet. Dabei werden alle verfügbaren Datenpunkte verwendet, woraus sich auch der Name „globale Interpolation ableitet“ [vgl. KRIVORUCHKO 2011, S. 264-265].

Die Koeffizienten a_i werden dabei durch Minimierung der quadrierten Differenz E zwischen Vorhersagen an Datenorten $\hat{Z}(x_i\ ,\ y_i)$ und den tatsächlichen Daten $Z(x_i\ ,\ y_i)$ bestimmt:

$$E = \sum_{i=1}^{N} (\hat{Z}(x_i, y_i) - Z(x_i, y_i))^2$$

mit N = Anzahl der Messwerte

Im Falle eines Trends würde die folgende Gleichung minimiert:

$$E = \sum_{i=1}^{N} (\underbrace{a_0 + a_1\ x_i + a_2\ y_i}_{\text{Polynom 1. Ordnung für } \hat{Z}(x_i, y_i)} - Z(x_i\ ,\ y_i))^2$$

d. h. , es sind Koeffizientenwerte a_i zu finden, so dass E den kleinsten Wert annimmt [KRIVORUCHKO 2011, S. 265].

Je komplexer ein Polynom ist, desto schwieriger ist es, ihm physische Bedeutung zuzuweisen. Polynome niedriger Ordnung haben eine langsam variierende Oberfläche und sind geeignet zur Beschreibung physischer Prozesse, wie der Veränderung der mittleren Temperatur mit der Änderung der geographischen Breite. Weil für Polynome kleiner Ordnung weniger Koeffizienten a_i zu schätzen sind, kommen sie der Erfahrung näher, dass Schätzungen umso vertrauenswürdiger sind, je kleiner die Anzahl der Modellparameter ist [KRIVORUCHKO 2011, S. 269-270].

Trendoberflächen-Interpolation ist empfindlich für ungewöhnlich hohe und niedrige Werte (Ausreißer). Dabei beeinflussen einzelne extreme Werte die gesamte geschätzte Oberfläche. Der erste Schritt zur Verwendung dieses Modells ist daher eine Prüfung auf Datenausreißer [KRIVORUCHKO 2011, S. 267].

3.4.2.2 Lokale polynome Interpolation (LPI)

Bei der lokalen polynomen Interpolation wird nicht der gesamte Datensatz verwendet, sondern nur die Daten in einem spezifizierten wandernden Fenster (local moving window) um den Vorhersageort [KRIVORUCHKO 2011, S. 267].

Die Idee der lokalen polynomen Interpolation ist, höhere Gewichte für Daten mit geringer Distanz zum Schätzort zu verwenden. Der ESRI Geostatistical Analyst folgt dabei dem Vorschlag von Lev Gandin [1963, zitiert KRIVORUCHKO 2011, S. 267] und minimiert den folgenden gewichteten quadrierten Vorhersagefehler:

$$E = \sum_{i=1}^{N} h_i \left(\hat{Z}\left(x_i\,,\,y_i\right) - Z\left(x_i\,,\,y_i\right)\right)^2$$

mit N = Anzahl der verwendeten Messwerte
h_i = Gewichte

Die Gewichte h_i nehmen mit der Distanz zwischen Punkten ab:

$$h_i = \exp\left(\frac{-r_i}{3\,a}\right)$$

mit r = Distanz zwischen Messwerten i und Vorhersageorten
a = kennzeichnender Parameter

Die Gewichte h_i nehmen mit steigendem r_i ab, je kleiner h_i, desto steiler. Je größer der Parameter a, desto „gleichmäßiger“ (gerader) verläuft die Abnahme von h_i.

Der Vorgang wird mit der gleichen Methode wie bei IDW und RBF-Interpolation durch Minimierung des Root Mean Square Prediction Error (RMS) aus der Kreuzvalidierung optimiert [KRIVORUCHKO 2011, S. 267].

Lokale polynome Modelle können Anisotropie auf der Grundlage der Transformation von Koordinaten berücksichtigen [KRIVORUCHKO 2011, S. 270]. Ein Beispiel bei KRIVORUCHKO [2011, S. 271] zeigt, dass polynome Interpolation in Bereichen mit wenigen oder fehlenden Daten schlecht funktioniert.

Lokale polynome Interpolation ist hilfreich in der Phase der Datenuntersuchung und der Datenaufbereitung für Kriging. Eine Variation der Nachbarschafts-Distanz – angefangen mit der Größe des Untersuchungsgebietes bis zu einer Distanz, die nur ein mehrfaches der kürzesten Distanz zwischen Datenorten beträgt – hilft verschiedene

Größenordnungen von Datenvariabilität aufzudecken. Trendoberflächen können auch als variabler Mittelwert in der Kriging-Interpolation verwendet werden [KRIVORUCHKO 2011, S. 272].

Lokale polynome Interpolation ist ein besonderer Fall der geographisch gewichteten Regression mit x- und y-Koordinaten als erklärende Variable. Bei der geographisch gewichteten Regression ist es möglich, Vorhersagefehler zu schätzen, obwohl ihre Interpretation mit Problemen verbunden ist [KRIVORUCHKO 2011, S. 273]. Das Verfahren kann aber nicht anstatt der lokalen polynomen Interpolation benutzt werden.

Als Diagnostik dient z. B. die räumliche Konditionszahl (spatial condition number), ein Maß dafür, wie instabil die Lösung der Vorhersagegleichung für einen bestimmten Ort ist. Weil der Vorhersagestandardfehler berechnet wird, ist die Annahme notwendig, dass das LPI-Modell die Daten korrekt beschreibt, was selten der Fall ist [KRIVORUCHKO 2011, S. 273].

3.4.3 Radiale Basisfunktionen (RBF)

Dabei handelt es sich um eine besondere Variante der Spline-Interpolation, in der Knoten mit Messpunkten zusammenfallen. Radiale Basisfunktionen (RBF) erzeugen Oberflächen, die mit dem geringsten Kurvenradius durch Messwerte gehen [KRIVORUCHKO 2011, S. 259] Als in Softwarepaketen, wie Geostatistical Analyst oder Surfer verfügbare RBF seien hier genannt :

- Completely regularized spline
- thin-plate-spline
- spline with tension
- multiquadratic spline
- inverse multiquadratic spline

Im Gegensatz zu IDW können RBF's Werte über oder unter den Maximal- bzw. Minimalwerten der vorhandenen Messdaten vorhersagen, also extrapolieren. Das ist nicht immer erwünscht.

Wie IDW, sind auch RBF's exakte Interpolatoren und bringen für sanft variierende Daten, wie z.B. Geländehöhen, gute Ergebnisse hervor. Die Modelle passen jedoch nicht, wenn die Messwerte große Änderungen auf kurze Distanz aufweisen oder wenn die Messdaten ungenau sind [KRIVORUCHKO 2011, S. 259].

Wie IDW ist auch RBF eine Funktion, die sich mit der Distanz zu einem Ort ändert. Jede RBF hat einen Parameter σ (sigma), der die Glätte der Oberfläche bestimmt. [KRIVORUCHKO 2011, S. 260-262]

Beispiele für RBF:

Thin-Plate-Spline $\Phi(r) = (\sigma \cdot r)^2 \cdot \ln(\sigma \cdot r)$

Multiquadratic	$\Phi(r) = (r^2 + \sigma^2)^{\frac{1}{2}}$
Inverse Multiquadratic	$\Phi(r) = (r^2 + \sigma^2)^{-\frac{1}{2}}$

Im Geostatistical Analyst ist $\Phi(r)$ die RBF Funktion mit r für die euklidische Distanz zwischen gemessenen und vorhergesagten Orten. Die Oberfläche wird mit Ausnahme der multiquadratischen RBF mit steigendem Wert von σ immer glatter. Das Gegenteil ist bei der inversen multiquadratischen RBF der Fall [KRIVORUCHKO 2011, S. 260-262].

Der Unterschied zwischen RBF und Kriging besteht in unterschiedlichen Zielen. Im Gegensatz zu RBF ist das Ziel von Kriging ist nicht, eine glatte Oberfläche, sondern eine gute mittlere Vorhersage-Genauigkeit zu erzeugen. Es gibt allerdings auch eine Kriging-Version, die sowohl genaue Vorhersagen, wie auch glatte Karten erzeugt (Krivoruchko 2011 Kap.8). [KRIVORUCHKO 2011, S. 263]

RBF kann in unterschiedlichen Maßstabsbereichen ähnliche Oberflächen erzeugen, wie Kriging [KRIVORUCHKO 2011, S. 263].

Eine Vorhersage für einen Ort ohne Messwert s_0 mittels einer linearen Kombination der Basisfunktionen, die an einem Messort s_i lokalisiert sind, wird im ersten Schritt mit dem folgenden formalen Ansatz vorgenommen [KRIVORUCHKO 2011, S. 262]:

$$\hat{Z}(s_0) = \sum_{i=1}^{N} \omega_i \, \Phi_i \left(\left\| s_i - s_0 \right\| \right) + \omega_0$$

mit $\hat{Z}(s_0)$ = Vorhersagewert für den Vorhersageort s_0

$\left\| s_i - s_0 \right\|$ = Distanz zwischen den Messorten s_i und dem Vorhersageort s_0

$\Phi_i(\left\| s_i - s_0 \right\|)$ = Radiale Basisfunktionen zur Distanz $\left\| s_i - s_0 \right\|$

ω_i = Gewichte, die jeder der Basisfunktionen gegeben werden

ω_0 = Verzerrungsparameter (bias parameter)

N = Anzahl an Punkten in der Suchnachbarschaft

Im zweiten Schritt werden die Gewichte ω_1 , ω_2 , ω_3 durch ein System von linearen Gleichungen mit bekanntem RBF Φ_i unter der Bedingung ermittelt, dass die resultierende Oberfläche durch alle Messwerte geht und die Summe der Gewichte ω_i gleich 1 ist. Es sind für jeden Vorhersagepunkt (N+1) lineare Gleichungen erforderlich. Für 3 Messpunkte s_i um einen Vorhersagepunkt s_0 sind das also vier Gleichungen mit den RBF $\Phi_1(\left\| s_i - s_0 \right\|)$, $\Phi_2(\left\| s_i - s_0 \right\|)$, $\Phi_3(\left\| s_i - s_0 \right\|)$ nach folgendem Muster [KRIVORUCHKO 2011, S. 263]:

$$\omega_1 \, \Phi_1 \left(\left\| s_1 - s_1 \right\| \right) + \omega_2 \, \Phi_2 \left(\left\| s_2 - s_1 \right\| \right) + \omega_3 \, \Phi_3 \left(\left\| s_3 - s_1 \right\| \right) = Z(s_1)$$

$$\omega_1 \, \Phi_1 \left(\left\| s_1 - s_2 \right\| \right) + \omega_2 \, \Phi_2 \left(\left\| s_2 - s_2 \right\| \right) + \omega_3 \, \Phi_3 \left(\left\| s_3 - s_2 \right\| \right) = Z(s_2)$$

$$\omega_1 \, \Phi_1 \left(\left\| s_1 - s_3 \right\| \right) + \omega_2 \, \Phi_2 \left(\left\| s_2 - s_3 \right\| \right) + \omega_3 \, \Phi_3 \left(\left\| s_3 - s_3 \right\| \right) = Z(s_3)$$

$$\omega_1 + \omega_2 + \omega_3 = 1$$

Der Default-Parameter σ wird auf ähnliche Weise durch Kreuzvalidierung bestimmt, wie bei IDW-Verfahren: ein Messpunkt wird aus dem Datensatz entfernt, eine Vorhersage für diesen Punkt gemacht und dieser Wert mit dem tatsächlichen Messwert verglichen. Der Vorgang wird für alle Daten und verschiedene Werte des Parameters σ wiederholt. Ein optimaler Wert für σ entspricht der kleinsten mittleren Differenz zwischen vorhergesagten und gemessenen Werten [KRIVORUCHKO 2011, S. 263]:

$$\sum_{i=1}^{N} (\hat{Z}(s_i) - Z(s_i))^2 \rightarrow min$$

Für Daten mit anisotropen Eigenschaften wird eine Koordinatentransformation ähnlich, wie bei der IDW-Interpolation verwendet. Die x- und y-Koordinaten werden wie folgt in die neuen Koordinaten x' und y' transformiert:

$$x' = \sqrt{\rho} \cdot (x \cdot \cos(\alpha) + y \cdot \sin(\alpha))$$
$$y' = \frac{1}{\sqrt{\rho}} \cdot (y \cdot \cos(\alpha) - x \cdot \sin(\alpha))$$

wobei α die Richtung der Anisotropie und ρ das Verhältnis von großer zu kleiner Halbachse der Anisotropie-Ellipse ist [KRIVORUCHKO 2011, S. 263]:

Nach KRIVORUCHKO [2011, S. 263] ist es möglich, durch die Verwendung bestimmter Semivariogramm-Modelle den RBF-Kernel zu finden, der mit der Kriging-Interpolation korrespondiert. Ein solcher Kernel ist der Thin-Plate-Spline.

Die Funktion[10] $\gamma(h) = |\sigma \cdot h|^2 \cdot \ln|\sigma \cdot h|$ könnte als Semivariogramm benutzt werden, wenn sie nicht wegen des fehlenden Nuggets und der zu schneller Steigung ungültig wäre. Das liegt daran, dass ein Semivariogramm für alle Distanzen h positiv sein muss und die Thin-Plate-Spline-Funktion schneller als die Funktion $\gamma(h) = h^2$ ansteigt, was bedeutet, dass eine der Hauptannahmen des Kriging – der konstante Mittelwert – missachtet wird [KRIVORUCHKO [2011, S. 263].

RBF kann entweder großräumlich oder kleinräumlich ähnliche Oberflächen wie Kriging erzeugen. Kriging sollte RBF hier überlegen sein, weil ein Kriging-Modell sowohl groß- wie auch kleinräumliche Datenvariations-Komponenten umfassen kann [KRIVORUCHKO [2011, S. 263].

3.4.4 Natural Neighbor

Wenn ein neuer Datenpunkt zu einem Datensatz hinzugefügt wird, werden Thiesen-Polygone modifiziert, d.h. einige werden kleiner, jedoch wird keines größer werden. Die Fläche des dem neuen Punkt zugestandenen Thiessen-Polygons wird von benachbarten bestehenden Polygonen genommen und wird "geborgt" genannt. Der

---

[10] Vergleiche Formel für RBF „Thin-Plate-Spline“: $\Phi(r) = (\sigma \cdot r)^2 \cdot \ln(\sigma \cdot r)$, siehe auch Abschnitt 3.4.3, Seite 21

Natural Neigbor Algorithmus verwendet das gewichtete Mittel der benachbarten Beobachtungen, wobei die Gewichte proportional der "geborgten Fläche" ist [Surfer 12 User Guide 2014, S. 265-266].

3.5 Probabilistische Interpolation

Kriging unterscheidet sich von deterministischer Interpolation indem dabei statt einer vom Anwender ausgewählten Funktionsformel für die Gewichte benachbarter Datenpunkte, ein Semivariogramm- oder Kovarianzmodell zur Bestimmung der Gewichte verwendet wird, die Gewichte also aus den verfügbaren Daten abgeleitet werden.

Kriging-Verfahren beruhen auf mehrere Annahmen, die bei der Verwendung akzeptiert werden. Auf sie wird in den folgenden Abschnitten eingegangen.

Eine explorative räumliche Datenanalyse sollte einer Anwendung von Kriging stets vorausgehen. Damit können die Kriging-Annahmen vor der Modellierung überprüft werden. Abweichungen von Kriging-Annahmen sollten mit Datenaufbereitung und ggfs. Transformationen minimiert werden, um die Prognosen annähernd optimal und interpretierbar zu machen.

Tatsächlich werden Kriging-Annahmen selten erfüllt, weil ein reales Semivariogramm-Modell aus den Daten geschätzt ist [s. a. KRIVORUCHKO 2011, S. 370].

Allgemein gilt: je mehr Parameter geschätzt werden müssen, desto mehr Unsicherheit wird in die Vorhersagen eingebracht [KRIVORUCHKO 2011, S. 294]. Das ist bei einer Verwendung von mehreren Variablen oder komplexen Modellen zu bedenken.

Obwohl es Anwendungen mit einer großen Anzahl an Variablen gibt, erlaubt der ArcGIS Geostatistical Analyst keine Verwendung von mehr als vier kontinuierlichen Variablen, weil eine verlässliche Schätzung einer sehr großen Anzahl an Modellparametern problematisch ist [KRIVORUCHKO 2011, S. 294].

Eine der Annahmen der konventionellen geostatistischen Modelle ist, dass die Probenahme- bzw. Messorte unabhängig von den Datenwerten gewählt wurden [KRIVORUCHKO 2011, S. 370]. Gegen diese Annahme wird im Fall von Grundwasserbelastungen vielfach verstoßen, weil Grundwassermessstellen in den Untersuchungsgebieten nicht nach den Kriterien der Zufälligkeit errichtet werden, sondern bevorzugt in Bereichen, in denen mit Belastungen zu rechnen ist.

Gau [2010m S. 91] geht auf dieses Problem unter der Überschrift „eingeschränkte Repräsentanz der Stichprobe“ein. Gründe dafür liegen insbesondere im Überwiegen projektbezogener Ziele bei der Festlegung von Aufschluss- und Probenahmepunkten. Aufschlusspunkte werden nicht nach geostatistischen Gesichtspunkten festgelegt, sondern dort, wo Werteänderungen des interessierenden Parameters erwartet werden, um den Informationszuwachs einer einzelnen Erkundungsmaßnahme zu maximieren und die Erhebung teilredundanter Daten zu vermeiden.

Hinzu kommen technische, wirtschaftliche und eigentumsrechtliche Beschränkungen, die einer im geostatistischen Sinne wünschenswerten Anordnung von Messpunkten häufig entgegen stehen.

3.5.1 Geostatistisches Modell

Die Variation vieler Prozesse kann nach KRIVORUCHKO [2011, S. 289] als groß- oder kleinmaßstäblich („large scale“ oder „small scale“) eingestuft werden. Wegen der Missverständlichkeit[11] der Kategorien „großmaßstäblich“ und „kleinmaßstäblich“ werden hier stattdessen die Bezeichnungen „großräumig“ und „kleinräumig“ verwendet.

Das geostatistische Modell nimmt an, dass Daten die Summe einer wahren Variablen (Signal genannt) und der Summe der Zufallsfehler sind [KRIVORUCHKO 2011, S. 289], also

Daten = Signal + Zufallsfehler

Normalerweise liegt das Interesse auf dem wahren Datenwert und nicht auf dem Anteil der Daten mit der irrtümlich gemessenen Rauschkomponente. Deshalb soll die räumliche Vorhersage möglichst dieses Signal vorhersagen.

Das Signal wird als Summe von folgenden drei Komponenten modelliert [modifiziert nach. KRIVORUCHKO 2011, S. 289-290]:

- einer **großräumlichen Datenvariation** (large scale variation) oder auch **Makrovariation**, die im Raum leicht schwankt und auch als **Trend** bezeichnet wird. Eine Modellierung erfolgt meist mit deterministischen Mittelwertfunktionen

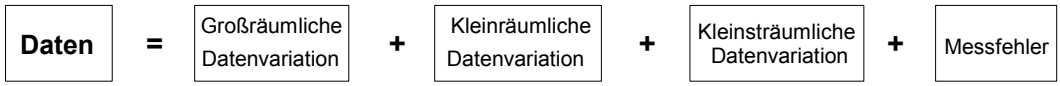
- einer **kleinräumlichen bzw. lokalen Datenvariation** (small-scale variation) mit einer Reichweite der Datenkorrelation über mehrere typische Distanzen zwischen Datenpaaren

- einer **kleinsträumlichen Datenvariation** oder **Mikrovariation** (micro scale variation) mit einer Reichweite kleiner der typischen Distanz zwischen Datenorten. Entspricht dem **Nugget** im Semivariogramm bzw. Kovariogramm (siehe Abschnitt 3.5.5).

Die Summe der Fehler wird im allgemeinen Messfehler genannt (measurement error), worunter nicht nur ein Gerätemessfehler zu verstehen ist. Der Messfehler kann aus mehreren Fehlertypen bestehen und schließt auch Verortungsfehler und Fehler wegen örtlicher Datenintegration ein [KRIVORUCHKO 2011, S. 290]. Eine örtliche Datenintegration wird im Rahmen dieser Master Thesis nicht vorgenommen und spielt deshalb keine Rolle.

[11] In der Geographie bedeutet „großmaßstäblich“ ein großes (z.B. 1 : 500) und „kleinmaßstäblich“ ein kleines Maßstabsverhältnis (z.B. 1 : 1 Mio.). KRIVORUCHKO (2011) verwendet diese beiden Kategorien im Zusammenhang mit Varianzen im umgekehrten Sinn entsprechend der Maßstabszahl.

Das geostatistische Modell von KRIVORUCHKO [2011, S. 289, siehe auch S. 290] lässt sich dann mit



differenzieren, und begrifflich markanter ausdrücken in der Form

Abb. 2: Geostatistisches Modell

Quelle: verändert nach KRIVORUCHKO [2011, S. 289]

Kleinräumliche Variation wird durch stationäre Prozesse mit geschätzten Semivariogramm- oder Kovarianz-Modellen modelliert und von KRIVORUCHKO [2011, S. 241] auch mit Effekten zweiter Ordnung bezeichnet. Sie resultieren aus der räumlichen Korrelationsstruktur und beschreiben lokale (also klein-räumliche) Datenvariation unter Verwendung von Informationen über benachbarte Werte. Kleinräumliche Datenvariation wird normalerweise als stationärer räumlicher Prozess modelliert.

Effekte erster Ordnung betreffen nach Krivoruchko [2011, S. 241] großräumliche Variation oder Trend. Die Unterscheidung zwischen groß- und kleinräumlicher Variabilität ist jedoch eine reine Modellannahme, es handelt sich dabei um kein echtes Datenmerkmal und wird von verschiedenen Bearbeitern auch häufig unterschiedlich eingestuft [KRIVORUCHKO 2011, S. 289].

Die Wahrscheinlichkeitsverteilung der Effekte 2. Ordnung wird gewöhnlich als bekannt angenommen und besteht zumeist in einer Gauß-Verteilung, d. h. also einer Normalverteilung [KRIVORUCHKO 2011, S. 241].

Das Ziel von Kriging ist die Rekonstruktion von großräumiger (Trend) und kleinräumiger Datenvariation. Wenn Daten nicht präzise sind, kann Kriging den durchschnittlichen Messfehler herausfiltern und einen neuen Wert am Messort vorhersagen. Dieser kann dann genauer sein, als der ursprünglich gemessene Wert. Diese Korrekturen sind allerdings nur möglich, wenn tatsächlich eine Datenkorrelation vorliegt, die als Hilfsmittel verwendet wird (siehe KRIVORUCHKO 2011, S. 291]

Nach KRIVORUCHKO [2011, S. 291] gibt es keine Möglichkeit, kleinsträumliche Variation zu modellieren, weil keine Daten in diesem Maßstab verfügbar sind, in dem diese Variation feststellbar wäre. Oberflächen, die auf Kriging-Vorhersagen beruhen, sind deshalb glatter, als die tatsächliche Datenvariation.

Abb. 3 zeigt mit Messfehlern behaftete Daten (gelbe Punkte im Diagramm), das wahre Signal zu diesen Daten[12] und die gefilterten Kriging-Vorhersagen dazu.

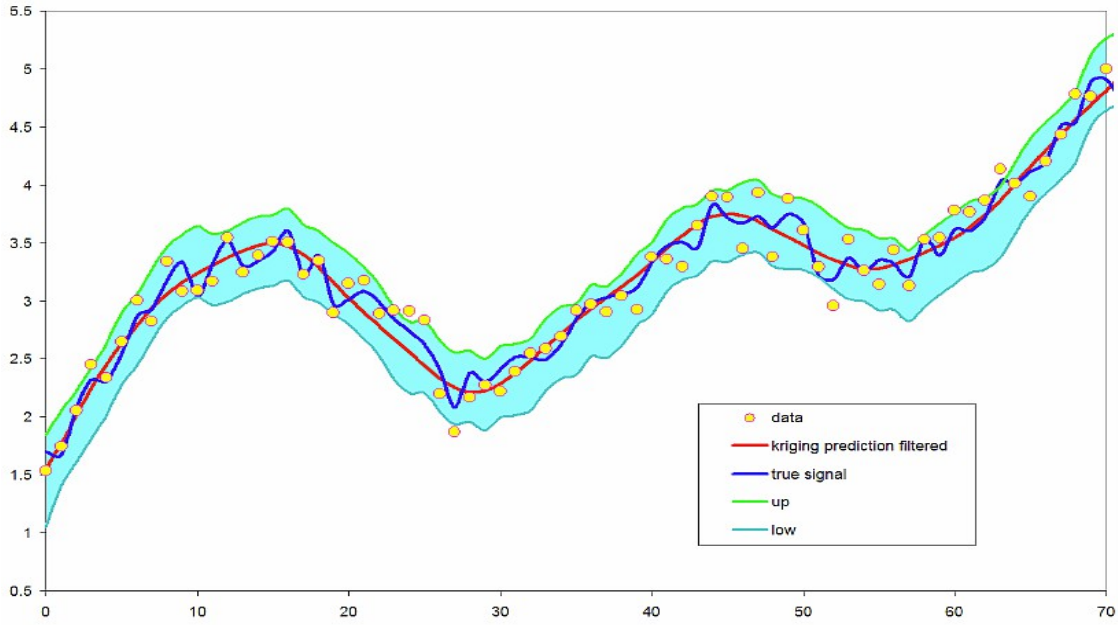


Abb. 3: Filterung von Messfehlern beim Kriging

Quelle: KRIVORUCHKO [2011, S. 293]

Das wahre Signal stimmt weder mit den Daten, noch mit den Vorhersagen überein. Allerdings liefert Kriging die oberen und unteren Grenzen der Vorhersage. Wenn das Kriging-Modell richtig geschätzt wurde, liegt das wahre Signal innerhalb dieser Grenzen KRIVORUCHKO [2011, S. 291].

SCHAFMEISTER [1999, S. 33] hat in ihrer Fassung des geostatistischen Modells im Unterschied zu KRIVORUCHKO [2011] raum-zeit-abhängige Größen eingebunden. Auch ihre allgemeine Modellvorstellung für orts-/zeitabhängige Zufallsvariable Z geht von deren Zerlegbarkeit in Einzelkomponenten aus:

$$Z(x,t) = T(x,t) + L(x,t) + \varepsilon$$

Danach setzt sich Z(x,t) aus einem globalen Trend T(x,t), einer lokalen Fluktuation L(x,t) und dem zufälligen „Rauschen“ ε zusammen. Je nach Art der Variablen kann der Trend (z.B. bei Grundwasserdruckhöhen in einem Einzugsgebiet) oder die lokale Variabilität (z.B. bei Durchlässigkeitsbeiwerten eines Grundwasserleiters) die stärkere Komponente in obiger Gleichung des geostatistischen Modells sein.

Die deterministische Trendkomponente lässt sich weiter in einen ortsabhängigen ($T_x(x,t)$) und einen zeitabhängigen Anteil zerlegen ($T_t(x,t)$). Der globale Trend geht meistens auf einen deterministische gut beschreibbaren Prozess zurück und lässt

[12] Das wahre Signal ist in diesem Beispiel bekannt, weil die Daten simuliert wurden

durch einfache Polynom- oder trigonometrische approximieren [SCHAFMEISTER 1999, S. 34].

Die Geostatistik behandelt nach SCHAFMEISTER [1999, S. 34] strenggenommen nur die lokale Fluktuation einer Zufallsvariablen L(x,t), die jedoch ebenso zeitliche und örtliche Komponenten aufweisen kann.

3.5.2 Regionalisierte Variable

Der Begriff regionalisierte Variable stammt von dem französischen Bergbauingenieur und Mathematiker Georges Matheron [RIVOIRARD 2005]. In der deutschen Literatur wird der Begriff auch in der Übersetzung „ortsabhängige Variable“ verwendet [SCHAFMEISTER 1999, S. 1]. Es ist eine Variable, welche die räumliche Verteilung einer Größe angibt.

Weil der Begriff von Georges MATHERON eingeführt und geprägt wurde, folgen hier einige Aussagen zum Hintergrund und den Vorstellungen des Urhebers darüber, was das Verständnis für das Konzept der regionalisierten Variablen unterstützen kann.

Die klassische Statistik kann die räumlichen Aspekte eines Phänomens nicht berücksichtigen. So bemängelte MATHERON [1964, S. 2], dass es nicht ausreicht, zu erfahren, mit welcher Häufigkeit sich die Gehalte in einer Erzlagerstätte wiederholen. Man müsste auch wissen, auf welche Weise die Gehalte auf dem Gebiet aufeinanderfolgen und insbesondere, wie groß die abbauwürdigen Zonen sind und wo sie liegen.

Nach MATHERON UND FORMERY [1965, S. 2] ist die klassische Statistik gewohnt, „*unabhängige Stichproben zu behandeln. In einer Lagerstätte sind die an benachbarten Punkten entnommenen Gehalte nicht unabhängig voneinander. Zwischen Stichproben besteht ein System an Korrelationen, deren Intensität eine Funktion der Distanz der Entnahmestellen ist. Die zwischen angrenzenden Entnahmestellen sehr starke Korrelation nimmt mit steigender Entfernung der Entnahmepunkte ab. Kein Zweifel, dass die klassische Statistik ungeeignet dafür ist, diesen Aspekt der variablen Abhängigkeit im Raum wiederzugeben. Jedoch hat dieser Aspekt einen grundlegenden Einfluss auf die Genauigkeit der Schätzung*.“

MATHERON [1964, S. 3] war klar, dass Gehalte in einer Lagerstätte die wesentlichen Kriterien von klassischen Zufallsvariablen nicht erfüllen, die da sind:

- die Möglichkeit, den Versuch, der einer Variablen einen numerischen Wert zuweist wenigstens theoretisch unbegrenzt oft zu wiederholen
- die Unabhängigkeit jedes einzelnen Versuchs gegenüber dem vorhergehenden und folgenden

Die Möglichkeit, einen Versuch (z.B. Messung, Probenahme) unbegrenzt oft zu wiederholen, gibt es [Anmerkung: im Bereich der Lagerstätten, der Geologie] nicht. Bei gehäuften Stichproben ließe sich die Möglichkeit einer Versuchswiederholung zwar vielleicht noch näherungsweise zulassen, aber das zweite Kennzeichen von

Zufallsvariablen, die Unabhängigkeit jedes einzelnen Versuchs gegenüber dem vorhergehenden und folgenden, wird sicher nicht eingehalten.

Zwei benachbarte Stichproben sind nach MATHERON [1964, S. 3] nämlich sicher nicht unabhängig voneinander. Sie haben im Mittel Tendenz beide [erz]reich zu sein, wenn sie aus einer [erz]reichen Zone stammen und umgekehrt. Diese mehr oder weniger ausgeprägte Tendenz drückt den mehr oder weniger großen Grad an Kontinuität der Variation in den Gehalten im mineralisierten Raum aus. Diese Überlegungen von MATHERON lassen sich verallgemeinern und auch auf Phänomene außerhalb der Lagerstättenerkundung übertragen.

„*Man muss [...] eine Art synthetische Formulierung annehmen, die gleichzeitig den beiden widersprüchlichen Kennzeichen der regionalisierten Variablen (ihren gleichzeitig zufälligen und strukturierten Aspekten) Rechnung trägt und selbstverständlich auch erlaubt, Probleme korrekt darzustellen und vor allem praktisch zu lösen, wie den des Fehlers, den man begeht, wenn man eine regionale Variable aus einer lückenhaften Stichprobe schätzt. Diese passende Sprache - gleichzeitig probabilistisch und strukturalistisch - wäre die der Zufallsfunktionen. Aber es stellt sich ein ziemlich gravierendes methodologisches Problem, [..] Diese Sprache der Zufallsfunktionen schließt dann bestimmte fundamentale Hypothesen der probabilistischen Natur ein und wir müssen die kritische Prüfung ihres Sinns und Werts fortsetzen. Anders gesprochen, müssen wir untersuchen, ob wir tatsächlich etwas sinnvolles sagen, wenn wir von Zufallsfunktionen sprechen oder ob wir zufällig etwas sagen ohne etwas zu sagen*.“ [Matheron 1967, S. 2, zur Sprache der Zufallsfunktionen]

Nach MATHERON [1964, S. 3] handelt es sich auf einem höheren Niveau darum, „*eine Synthese zwischen den traditionellen [Anmerkung: bergbaulichen] und den statistischen Methoden vorzunehmen. Infolgedessen hat die Geostatistik begonnen, ihre eigenen Methoden und mathematischen Formalismen zu erarbeiten, was nichts anderes ist, als jahrhundertealte Bergbauerfahrung in abstrakte Form zu bringen und zu systematisieren. Dieser Formalismus hat von seinen statistischen Ursprüngen eine Sprache geerbt, in der man z.B. weiterhin von Varianzen und Kovarianzen spricht, obwohl man diesen Begriffen einen neuen Inhalt gegeben hat. Diese Ähnlichkeit des Vokabulars darf nicht täuschen. Am Ende einer langen Entwicklung hat die Geostatistik erkannt, dass sie keine zufälligen Ereignisse vor sich hat, sondern im Raum verteilte, natürliche Phänomene. Und infolgedessen sind ihre Methoden an diejenigen der mathematischen Physik angenähert und ganz besonders denen der harmonischen Analyse.*“

Für MATHERON war das Konzept der Zufälligkeit eine erkenntnistheoretische Entscheidung, dass Beziehungen zwischen den grundlegenden Bestandteilen natürlicher Phänomene als probabilistisch ausgedrückt werden können. Nicht alle Wahrscheinlichkeiten könnten experimentell bestimmt werden. Reiner Empirismus ist unzureichend und probabilistische Modelle sollten von Geologen entliehene genetische Vorstellungen einschließen [zitiert in AGTERBERG 2003, S. 4].

Und so stellt MATHERON [1967, S.8] die Hypothese auf, dass eine regionalisierte Variable als Realisation einer stationären Zufallsfunktion angesehen werden kann, wobei diese Hypothese jedoch nicht in allen Fällen annehmbar ist.

Nach MATHERON [1964, S. 3] ist eine regionalisierte Variable im strengen Sinne eine richtige Funktion, die an jedem Punkt des Raumes einen bestimmten Wert annimmt. [..] Vom Standpunkt der Physik oder der Geologie sind bestimmte qualitative Merkmale mit dem Begriff der regionalisierten Variablen verbunden:

1. hat eine regionalisierte Variable einen Ort. Ihre Variationen haben den [..] Raum als Schauplatz [...], den man geometrisches Feld der Regionalisierung nennt. Außerdem ist eine solche Variable im allgemeinen durch einen geometrischen Support[13] definiert. Im Fall eines Gehalts ist dieser Support das Volumen der Stichprobe mit seiner geometrischen Form, seinen Dimensionen und seiner Orientierung.

2. kann die Variable eine mehr oder weniger große Kontinuität in ihrer räumlichen Variation zeigen, welche sich z.B. durch einen mehr oder weniger großen Abstand der Gehalte [= Werte] der [..] benachbarten Stichproben äußert.

3. die Variable kann diverse Arten von Anisotropien aufweisen.

Unter einer regionalisierten Variablen wird also eine Variable verstanden, welche die räumliche Verteilung einer Größe angibt. Geostatistische Methoden beruhen auf der Annahme, dass diese regionalisierten Variablen eine bestimmte Struktur aufweisen [DUTTER 1985, S. 14].

Ein tatsächlicher Wert wird dabei als Realisation z(x) einer Zufallsvariablen Z(x) verstanden, d.h. der Realisierung einer zufälligen Größe mit eventuell unendlich vielen möglichen Werten an einem bestimmten Punkt x. Alle möglichen Ereignisse (oder Werte) werden durch die Zufallsvariable Z(x) zusammengefasst und mit ihr wird auch eine gewisse Wahrscheinlichkeitsverteilung assoziiert, die als Wahrscheinlichkeitsdichte oder auch als kumulative (Wahrscheinlich-keits-)Verteilungsfunktion [an jedem Ort x] ausgedrückt werden kann [DUTTER 1985, S. 15-17].

Die Funktion, die an jeder Stelle x eines bestimmten Bereiches D eine Zufallsvariable Z(x) definiert, heißt Zufallsfunktion. Daneben gibt es Zusammenhänge oder Korrelationen zwischen Variablen Z(x) und Z(x') [Anmerkung: im Bereich D], welche zur Charakterisierung der gesamten Zufallsfunktion Z notwendig sind. Die regionalisierte Variable z stellte eine Realisation dieser Zufalls-funktion Z dar [DUTTER 1985, S. 16-17].

Die Geostatistik beschäftigt sich mit der räumlichen Variabilität der regionalisierten oder ortsabhängigen Variablen im Rahmen von Strukturanalysen (quantitative Erfassung und Beschreibung) und anschließender Erstellung eines räumlichen Modells dieser Variablen. Dieses Modell kann durch Interpolation mit Hilfe eines Kriging-Schätzverfahrens oder anhand stochastischer Simulationsmethoden erzeugt werden. Die Kenntnis der räumlichen Variabilität der regionalisierten Variablen

[13] Support wird in Schafmeister [1999, S. 10] mit „Stützung“ übersetzt : das sind im geostatistischen Sinn die Größe und Form einer Probe, auf die sich ein Proben- oder Analysenwert bezieht.

ermöglicht es, Unsicherheiten und Risiken zu quantifizieren [SCHAFMEISTER 1999, S. 1-2].

Die regionalisierte Variable wird dann als eine Zufallsvariable[14] Z_i betrachtet. Diese kann am Punkt x_i Werte annehmen, die durch eine Wahrscheinlichkeitsfunktion bedingt sind. Eine Menge von Zufallsvariablen in einem Gebiet D (z.B. ein Einzugsgebiet oder ein Grundwasserleiter) heißt Zufallsfunktion und eine Messkampagne oder Stichtagsmessung liefert eine Anzahl (i = 1, 2, …, k) Beobachtungen, die eine Realisation der Zufallsvariablen sind [SCHAFMEISTER 1999, S. 11].

3.5.3 Stationarität

Für MATHERON [1967a, S.8] hat die Idee der Stationarität den Sinn einer Art statistischen Homogenität im Raum. An welchem Ort auch immer ein Phänomen auftritt, zeigt das Phänomen die gleichen mittleren Eigenschaften. Bildlicher beschrieben wiederholt sich das Phänomen dadurch im Raum und dank dieser Wiederholung wird der statistische Schluss möglich: weil die räumliche Verteilung bei Translation unveränderlich ist, ist die Anzahl der Parameter, von denen sie abhängt, wesentlich geringer.

Mit einem einzelnen Wert an einem Ort x_i lassen sich keine statistischen Aussagen treffen. Damit das möglich wird, nimmt man an, dass sich alle Zufallsvariable Z_i im Gebiet D, das ein Einzugsgebiet oder ein Grundwasserleiter sein kann, in gewisser Hinsicht gleich verhalten [nach SCHAFMEISTER 1999, S. 11]. Das theoretische Konzept geht davon aus, dass in jedem Punkt x_i eine Verteilungsfunktion existiert, aus der (bei einer Messung oder Probenahme) zufällig eine Probe z_i realisiert wurde.

In der geostatistischen Literatur wird oft auf Stationarität Bezug genommen, wobei nach MYERS [1989] nicht selten unklar scheint, was damit gemeint ist. MYERS [1989, 1991, S.215] betont in diesem Zusammenhang, dass die Stationarität eine Eigenschaft der Zufallsfunktion Z(x) sei und nicht der Daten.

Nach KRIVORUCHKO [2011, S. 285] ist die Annahme der Stationarität deshalb so wichtig beim Kriging, weil sie die Möglichkeit bietet, eine Replikation von einem einzelnen korrelierten Datensatz zu erhalten und sie es erlaubt, die Parameter des statistischen Modells zu schätzen. Die Stationarität ist wichtig zur Schätzung der räumlichen Abhängigkeit über das Semivariogramm, nicht aber für die räumliche Vorhersage beim Kriging. Wenn Daten nicht als stationär angenommen werden können, lassen sie sich evtl. durch Trendeliminierung und Transformationstechniken in die Nähe von Stationarität bringen.

Konventionelle geostatistische Modelle beinhalten die Annahme, dass die Daten stationär sind (im Fall des Moving Window Kriging innerhalb des wandernden Fensters). Das bedeutet, dass insbesondere der Mittelwert, aber auch sie Varianz einer Variablen an einem Ort gleich dem Mittel und der Varianz an einem anderen Ort sind [KRIVORUCHKO 2011, S. 269 und 370].

---

[14] Anmerkung: DUTTER [1985, S. 17] sieht Z(x) als gewöhnliche Zufallsvariable an, wenn nur ein Ort x betrachtet wird.

Grundwasserdruckhöhen zeigen jedoch sowohl in ihrer räumlichen, als auch in der zeitlichen Dimension eine stark deterministische Überprägung. Da Grundwasserdruckflächen in Fließrichtung geneigt sind, sind sie strenggenommen schon wegen ihres hydraulischen Gradienten als instationäre Zufallsvariablen zu betrachten [SCHAFMEISTER 1999, S. 33].

3.5.3.1 Strenge Stationarität

MATHERON [1967, S.8] stellt die Hypothese auf, dass eine regionalisierte Variable als Realisation einer stationären Zufallsfunktion angesehen werden kann, wobei diese Hypothese jedoch nicht in allen Fällen annehmbar ist. Wenn sich ein Phänomen, ausgehend von einem Kern mit hohen Werten, durch eine bereichsweise (zonierte) Abnahme äußert, ist das als nicht stationär anzusehen [..].

Eine strenge Stationarität beschreibt die Unabhängigkeit des gesamten Verteilungsgesetzes des Zufallsprozesses Z(x) (vgl. auch Abschnitt 3.5.2. Regionalisierte Variable, Seite 28) gegenüber Verschiebung (Translation). In Bezug auf die regionalisierte Variable bedeutet dies eine Ortsunabhängigkeit des Verteilungsgesetzes. Das heißt, dass an allen Punkten im Raum während des genetischen Prozesses das gleiche, aber unbekannte Verteilungsgesetz vorgelegen hat, aus dem an jedem Punkt ein Wert realisiert worden ist. [..] Alle Beobachtungen des Zufallsprozesses Z(x) können als aus einer einzigen Wahrscheinlichkeitsdichtefunktion stammend angesehen werden und nicht von verschiedenen Verteilungen an einzelnen Orten [GAU 2010, S. 22].

DUTTER [1985, S. 56] beschreibt den Begriff der strengen Stationarität kurz mit „Invarianz des räumlichen Verteilungsgesetzes gegenüber Translation“.

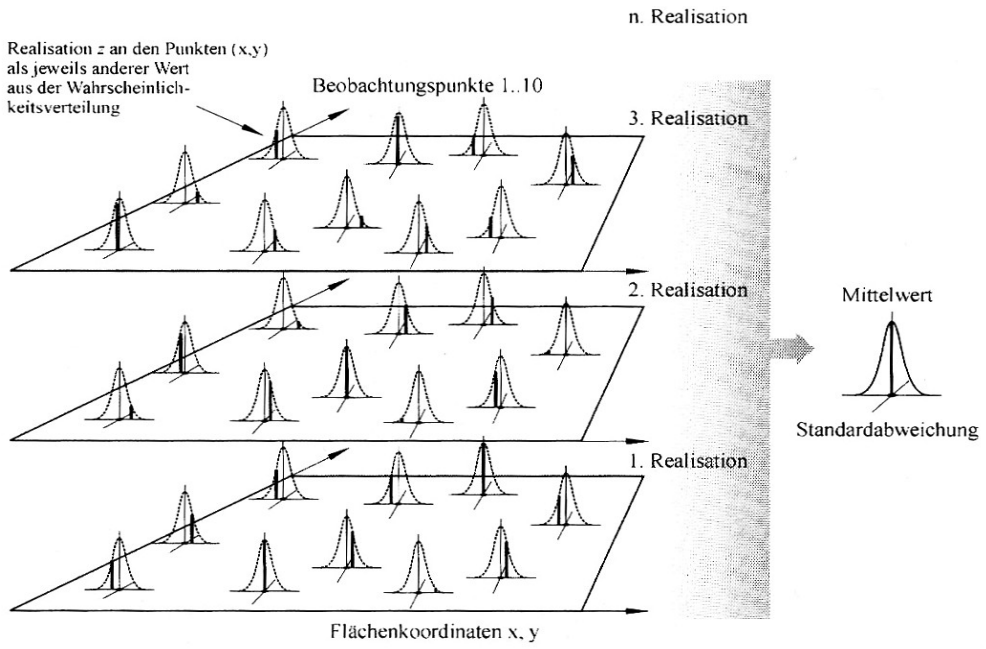


Abb. 4: Strenge Stationarität einer Zufallsvariablen Z(x)

Quelle: GAU [2010, Abb. 3-2, S. 24, Illustration zur Ergodizitätsannahme]

Die Realisierungen z(x) der Verteilungsfunktion Z(x) an der Orten x in Abb. 4 weisen überall die gleichen Verteilungen auf. Ein höchstens selten anzutreffender Idealfall der strengen Stationarität, in diesem Fall sogar über mehrere Realisationen.

Demnach liegt z.B. bei einer lokal begrenzen Grundwasserschaden, der sich von einer Quelle ausgehend im Untergrund und im Grundwasser ausgebreitet hat, auf keinen Fall strenge Stationarität vor, da die Verteilungsfunktion im kontaminierten Bereich ein anderer ist, wie außerhalb in der unbelasteten Umgebung.

3.5.3.2 Stationarität zweiter Ordnung

Anstatt der strengen Stationarität wird im Allgemeinen eine abgeschwächte Form von Stationarität verwendet, wobei nur die beiden sogenannten „ersten Momente“ zur Beschreibung einer Verteilung gebräuchlich sind, d.h. der Mittelwert und die Varianz. Für praktische Fälle wird das als ausreichend erachtet [GAU 2010, S. 22].

Nach JERNIGAN [1986 S.12] wird ein stochastischer Prozess stationär von zweiter Ordnung genannt, wenn die Kovarianzfunktion eines stochastischen Prozesses nur von der relativen Position zweier Punkte zueinander und nicht von deren genauen festen Positionen abhängt.

Formaler drückt JERNIGAN [1986 S.12] das damit aus, dass der Prozess im Fall von Stationarität zweiter Ordnung ein konstantes Mittel hat und seine Kovarianzfunktion K(x,y) nur von der Differenz des Vektors h = x - y abhängt und nicht von der der genauen Position der gewählten x und y. Wir könnten dann K(x,y) als eine C(h) oder C(|h|,α) schreiben, wobei |h| die Länge des Vektors h ist (Distanz zwischen den Vektoren x und y), und α den Winkel der Richtung angibt. Alle Punkte mit dem gleichen Vektor h hätten die gleiche Kovarianz.

JERNIGAN [1986 S.12] führt für einen Prozess mit Stationarität zweiter Ordnung ein Beispiel von Punkten an, die den gleichen Abstand in der gleichen Richtung haben und deshalb die gleiche Kovarianz für einen stationären und homogenen Prozess aufweisen. Dieser Typ von Kovarianz könnte das Fließen von Wasser entsprechend einem Gradienten sein.

Für die Stationarität zweiter Ordnung formuliert GAU [1999, S. 22] die Forderung, dass der Erwartungswert[15] E an jeder Lokation unabhängig vom Ort sei, dem Mittelwert m der Stichprobe entspricht und dass die Kovarianz Cov von Z(x) und Z(x+h) existiert, womit gelte, dass

$$E\{ Z(x) \} = m$$

und

$$Cov\{ x, x+h \} = Cov\{ h \} = E\{ Z(x) \cdot Z(x+h) \} - m^2$$

[15] E = Erwartungswert einer Zufallsvariablen; beschreibt die Zahl, welche die Zufallsvariable im Mittel annimmt

Auch an jeder Schätzstelle wäre dann ein Wert anzunehmen, der dem Mittelwert der Stichprobenpopulation entspricht. Die Kovarianz zweier beliebiger Zufallsvariabler hängt zudem nur vom Abstand h ab. Kovarianz σ und Variogramm γ sind in diesem Fall gleichwertige Möglichkeiten der Strukturbeschreibung. Und können nach

$$\gamma(h) = \sigma(0) - \sigma(h)$$

ineinander überführt werden.

3.5.3.3 Intrinsische Hypothese und Quasi-Stationarität

Die Intrinsische Hypothese stellt eine gegenüber der Stationarität zweiter Ordnung nochmals abgeschwächte Form der Stationarität bzw. Hypothese dar. Es wird nur noch gefordert, dass der Erwartungswert E der Differenz der Werte Z(x+h) und Z(x) gleich Null ist und dass diese beiden Punkte für alle Abstände h eine endliche Varianz aufweisen [GAU 1999, S. 23]. Es gilt dann (nach GAU 1999 und DUTTER 1985, S. 57]:

Erwartungswert $E\{ Z(x+h) - Z(x) \} = m(h) = 0$

d.h. die Erwartung E(Z(x)) existiert für alle x und hängt nicht von denn Trägern x ab.

Varianz der Werte Z(x+h) und Z(x):

$$\text{Var}(Z(x+h) - Z(x)) = E\{ Z(x+h) - Z(x) \}^2 = 2\,\gamma(h)$$

d.h für alle h und x besitzt das Inkrement (Z(x+h) – Z(x)), d.h. der Zuwachs, eine endliche Varianz, die nicht von x abhängt

Semivarianz $\gamma(h) = \frac{1}{2}\,\text{Var}[\{Z(x+h)\} - Z(x)\}]$

wobei die Stationarität zweiter Ordnung die intrinsische Hypothese oder auch die Hypothese der stationären Zuwächse einschließt. Die Umkehrung muss natürlich nicht gelten.

Die Erfüllung der intrinsischen Hypothese bezieht sich damit lediglich auf die Inkremente (Zuwächse) der Zufallsfunktion. Sie wird daher auch als Zuwachsstationarität zweiter Ordnung bezeichnet und ist mit der Existenz von Homogenität gleichzusetzen [GAU 1999, S. 23, DUTTER 1985, S.].

DUTTER [1985, S. 57] bezeichnet die intrinsische Hypothese auch als „wesentliche Hypothese“ und weist darauf hin, dass man in der Praxis häufig noch eine Größenordnung von h beschränken müsste, so dass sie nur bis für nur für $|h| < b$ für ein gewisses $b > 0$ angenommen werden kann. In diesem Fall spricht man von Quasi-Stationarität.

3.5.3.4 Behandlung von Instationarität

Eine Forderung zur globalen Stationarität kann abgeschwächt werden, wenn man annimmt, dass Daten-Mittelwert und Varianz lokal konstant sind. Diese Idee wird in der Geostatistik verbreitet verwendet, wenn nur benachbarte Daten zur Vorhersage von Werten an unbeprobten Orten verwendet werden [KRIVORUCHKO 2011, S. 461]

In der Praxis wird oft das Erreichen des Sill als Indiz für das Einhalten einer wenigstens lokalen Stationarität herangezogen [siehe auch SCHAEFER 2013].

Der Vergleich von Histogrammen von Teilbereichen mit denen der Gesamtheit im Untersuchungsgebiet kann Hinweise auf Instationarität geben [KRIVORUCHKO 2011, S. 599]

Auf Stationarität kann wie folgt geprüft werden:

- Abschätzung der räumliche Variabilität von Mittelwert und Varianz der Daten über lokale Statistik die mit Voronoi-Polygonen berechnet wurde (Berechnung von lokalen Mittelwerten benachbarter Polygone, Standardabweichung u.a.). Zur Gewährleistung von Datenstationarität sollten der lokale Mittelwert und die lokale Standardabweichung sich nur langsam ändern. [KRIVORUCHKO 2011, S. 632]

- Aufteilen von Daten in Untermengen und separate Untersuchung zur weiteren Klärung der Frage zur Stationarität [KRIVORUCHKO 2011, S. 633].

Nach ISAAKS UND SRIVASTAVA [1989, S. 349] ist die Entscheidung darüber, einen bestimmte Probendatenkonfiguration als Ergebnis einer stationären Zufallsfunktion zu betrachten, stark mit der Entscheidung verknüpft, dass diese Proben zusammengruppiert werden können. Beide Entscheidungen können nicht quantitativ überprüft werden, sie sind weder richtig, noch falsch und es ist Beweis ihrer Gültigkeit möglich. Sie können aber als geeignet oder als ungeeignet eingestuft werden. Solch eine Beurteilung hängt von den Zielen der Untersuchung und qualitative Informationen über den Datensatz sind bei der Entscheidung hilfreich.

Bei vorhandener Stationarität kann eine Nichtstationarität des Mittelwerts durch Trendeliminierung entfernt, und eine Varianz-Instationarität durch Datentransformation gemindert werden (in einem Bsp. mit Quadratwurzel-Transformation durchgeführt) [KRIVORUCHKO 2011, S. 634].

Wenn es mit der Datentransformation nicht gelingt die Datenstationarität zu verbessern, sollte eine instationäre Variante von Interpolation für diese Daten verwendet werden [KRIVORUCHKO 2011, S. 635].

3.5.4 Datenaufbereitung

3.5.4.1 Datenanalyse

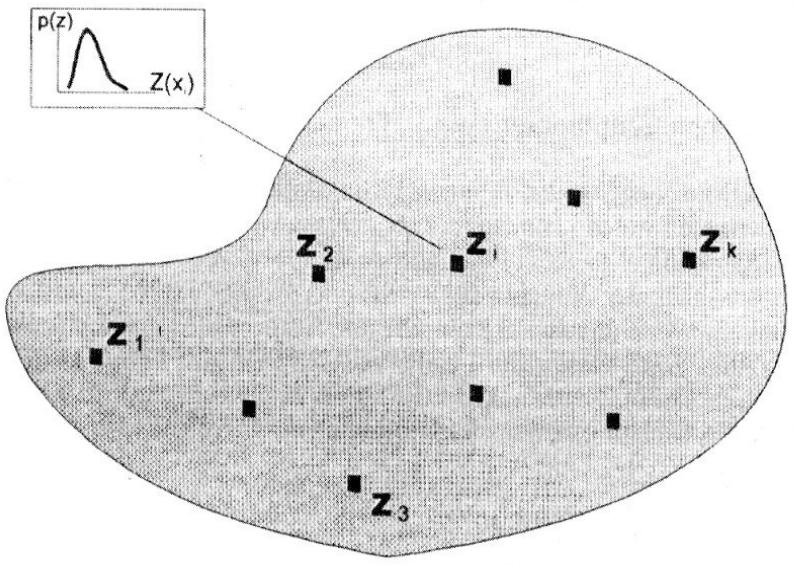
Ein Hilfsmittel zur Beschreibung der Werteverteilung einer Zufallsfunktion Z bzw. einer Zufallsvariablen Z(x) ist der Erwartungswert EZ. Dieser wird durch das gewogene (bzw. gewichtete) Mittel aller Werte, gewogen durch die Wahrscheinlichkeitsdichte p definiert [nach DUTTER 1985, S. 17]::

$$EZ = \int_{-\infty}^{\infty} zp(z)\,dz$$

Wobei die Integration über den gesamten Wertebereich geht. Im Fall von nur endlich vielen möglichen Werten c_j mit entsprechenden Wahrscheinlichkeiten p_j mit j = 1, 2, …, k bildet man eine Summe [nach DUTTER 1985, S. 17]:

$$EZ = \sum_{j=1}^{k} c_j p_j$$

Dieses theoretische Erwartung einer Zufallsvariablen wird auch einfaches Mittel m = EZ genannt.



Meßwerte einer Zufallsvariablen $Z(x)$ in Domäne D. Die Verteilungsfunktion p(Z) im Punkte x_i wird durch das Histogramm der Proben z_i $(i=1,..,k)$ geschätzt.

Abb. 5: Messwerte einer Zufallsvariablen Z(x) in Domäne D

Quelle: SCHAFMEISTER [1999, Abb. 2.1, S. 12]

Der erste Schritt einer Strukturanalyse ist die Modellierung der Verteilungsfunktion (Wahrscheinlichkeitsdichte) der regionalisierten Variablen bzw. die Schätzung der sie beschreibenden Momente. Die Ermittlung repräsentativer Schätzer für Mittelwert und Varianz stellen dabei einen wesentlichen Punkt dar [SCHAFMEISTER 1999, S. 11].

Die meisten hydrogeologischen Variablen, z.B. Schadstoffkonzentrationen im Grundwasser, zeigen linksschiefe Verteilungsfunktionen (positive Schiefe), die sich am

besten durch Lognormalverteilungen modellieren lassen [SCHAFMEISTER 1999, S. 12].

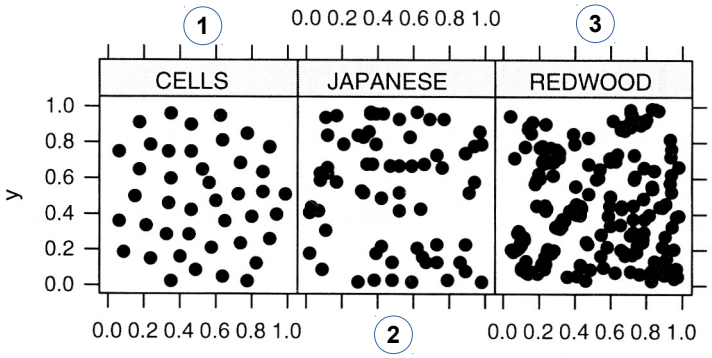
Nach SCHAFMEISTER [1999, S. 12] zeigen die meisten hydrogeologischen Variablen, z.B. Schadstoffkonzentrationen im Grundwasser jedoch linksschiefe Verteilungsfunktionen (positive Schiefe), die sich am besten durch Lognormalverteilungen modellieren lassen. Für Variable (z.B. k_f-Werte) die auch nach einer entsprechenden Datentransformation Verteilungen zeigen, die durch die gängigen Modelle (Normal-, Lognormalverteilung) nicht hinreichend beschreibbar sind empfiehlt SCHAFMEISTER die Anwendung von Nicht-parametrischen Methoden der Geostatistik. (Indikatorkriging, -simulation).

Einige Kriging-Modelle erfordern zur korrekten Anwendung die Kenntnis über den Datenmittelwert oder die mittlere Oberfläche [vgl. KRIVORUCHKO 2011, S. 370]

3.5.4.2 Entclustern von Daten

Die Verwendung von gleichen Gewichten $\omega_i = 1 / N$, mit N für die Summe an Beobachtungen, ist bei einer regelmäßigen Anordnung der Messpunkte zur Berechnung einer kumulativen Datenverteilung ausreichend [KRIVORUCHKO 2011, S. 342].

Wenn in einigen Teilen des Untersuchungsgebietes aber mehr Daten vorliegen, als in anderen, liegt eine bereichsweise räumliche Konzentration von Datenpunkten vor, das man ein Cluster nennt. Will man sich ein Bild über die Verteilung der Datenwerte im gesamten Untersuchungsgebiet verschaffen, werden dann manche Bereiche überrepräsentiert und die anderen unterrepräsentiert, wodurch der Eindruck über die Werteverteilung verzerrt ist bzw. die Verteilung der Messwerte nicht normal ist [DE SMITH ET. AL.2013, S. 516].



1 regelmäßige Verteilung
2 zufällige Verteilung
3 geclusterte Verteilung

Die Abb. zeigt die Verteilung von Zellen (cells), japanischen Schwarzkiefern (Japanese) und kalifornischen Mammutbäumen (Redwood)

Abb. 6: Drei Typen von Punktverteilungen im Raum

Quelle: BIVAND ET AL.[2008, Fig. 7.1, S. 157]

Da das Ziel die Schätzung einer Datenverteilung ist, die das gesamte Untersuchungsgebiet möglichst zutreffend repräsentiert, sollten Daten in dicht beprobten

Bereichen für die Bestimmung der Werteverteilung ein geringeres Gewicht erhalten, als Daten in spärlich beprobten Teilen [KRIVORUCHKO 2011, S. 342].

Eine Methode zum Daten-Entclustern ist, das Untersuchungsgebiet in ein Zellgitter aufzuteilen und jede Stichprobe entsprechend der Anzahl an Stichproben in jeder Zelle zu gewichten. Wenn z.B. M Zellen mit Daten vorliegen und die Anzahl der Proben in jeder Zelle mit n_1, n_2, … n_M gegeben ist, dann ist das Gewicht für jede Probe in der Zelle i [KRIVORUCHKO 2011, S. 342]: $\omega_i = 1 / (nj * M)$

Input-Parameter sind (z.B. im ArcGIS Geostatistical Analyst) die Koordinaten der Daten, die Datenwerte, die Gitterweite, der Winkel der Gitter-Rotation und ein Anisotropie-Faktor (Verhältnis der Zellenmaße, bei Anisotropie = 1 sind die Gitterzellen quadratisch) [KRIVORUCHKO 2011, S. 342].

Im ArcGIS Geostatistical Analyst wird eine Shift-Option angeboten, mit der das Gitter schrittweise um bis zu einer halben Zellgröße verschoben werden kann. Die Shift-Option spielt insbesondere bei relativ kleinen geclusterten Datensätzen eine Rolle, weil sich die Entclusterungs-Gewichte signifikant ändern können, wenn der Ursprung des Gitters um die halbe oder ein Viertel der Zellgröße verschoben wird [KRIVORUCHKO 2011, S. 342].

Kriterien zur Parametrierung des optimalen Gitters sind das Minimum (wenn die geclusterten Werte hoch sind) oder das Maximum (wenn die geclusterten Werte tendenziell niedrig sind) des Entclusterungs-Mittels [KRIVORUCHKO 2011, S. 342].

$$\hat{\mu} = \sum_{i=1}^{N} \omega_i Z_i$$

[bei Variation der Parameter Zellgröße, Anisotropie und Drehwinkel] zu finden.

Eine alternative optionale Methode zum Daten-Entclustern (im Geostatistical Analyst) besteht in der Festlegung von Gewichten als Funktion der Fläche eines Voronoi-Polygons, das jeden Punkt umgibt [KRIVORUCHKO 2011, S. 343].

Im ArcGIS Geostatistical Analyst wird der Vorgang des Entclusterns nur zur Schätzung der kumulativen Datenverteilung verwendet. KRIVORUCHKO [2011, S. 344] schlägt jedoch vor, in der gleichen Weise gewichtete Daten auch zur Schätzung des Semivariogramm-Modells zu verwenden, was durch folgende Modifikation der klassischen empirischen Semivariogramm-Schätzung (siehe Abschnitt 3.5.6, S. 45 ff) möglich wäre:

$$\hat{\gamma}(h) = \frac{\frac{1}{2} \cdot \sum_{Punktpaare(i,j)\,mit\,ähnl.\,lag\,h} \omega(s_i) \cdot \omega(s_j) \cdot [Z(s_i) - Z(s_j)]^2}{\sum_{Punktpaare(i,j)\,mit\,ähnl.\,lag\,h} \omega(s_i) \cdot \omega(s_j)}$$

oder

$$\hat{\gamma}(h)=\frac{1}{2}\cdot\sum_{Punktpaare(i,j)\,mit\,ähnl.\,lag\,h}\omega_{ij}\cdot[Z(s_i)-Z(s_j)]^2$$

mit

$$\sum_{Punktpaare(i,j)\,mit\,ähnl.\,lag\,h}\omega_{ij}=1$$

Die Motivation für diese Idee ist die gleiche, wie im Fall der Schätzung der Datenverteilung im Untersuchungsgebiet. Das arithmetische Mittel der empirischen Semivariogramm-Werte setzt voraus, dass jeder verfügbare Wert gleich repräsentativ für die lokalen Flächen ist, obwohl das nur für regelmäßig verteilte Proben zutrifft. Eine Korrektur vorhandener Verteilungs-Unregelmäßigkeiten kann nach KRIVORUCHKO [2011, S. 344] hilfreich sein.

Darüber hinaus gibt es weitere Ansätze zur Definition von Datengewichten zusätzlich zu Zell- und polygonalem Entclustern, u. a. einen Vorschlag, jeden Wert umgekehrt proportional zur Varianz seines Beitrags zum geschätzten Semivariogramm im Lag h zu gewichten KRIVORUCHKO [2011, S. 344].

3.5.4.3 Transformationen

Um vorhandene Daten der geforderten Normalverteilung anzunähern und die Stationaritätsannahmen des konstanten Mittelwertes und der konstanten Datenvarianz zu erfüllen, werden in der Geostatistik Transformationen verwendet.

Für schiefe Verteilungen mit einer kleinen Anzahl an hohen Werten wird eine logarithmische Transformation zur Annäherung an die Normalverteilung verwendet [KRIVORUCHKO 2011, S. 336].

In manchen Anwendungen ist die zum Erreichen der Normalität erforderliche Transformation nicht logarithmisch, sondern eine andere Funktion. In diesen Fällen sind Kriging-Vorhersagen und die Schätzung der Vorhersage-Standardfehler mit näherungsweisen Formel vorzunehmen. Diese Methode wird Delta-Methode genannt und steht z. B. im ArcGIS Geostatistical Analyst für Kriging mit Potenz (Box-Cox)und Arcsinus-Datentransformation zur Verfügung [KRIVORUCHKO 2011, S. 336-337].

Box-Cox-Transformation: $Y(s) = ([Z(s)]^2 – 1) / \lambda$

Logarithmische Transformation: $Y(s) = \log (Z(s))$

Arcsinus-Transformation: $Y(s) = \sin^{-1}(Z(s))$

Wenn Datenwerte in einem Teil des Untersuchungsgebietes kleiner sind, als in anderen Teilen, ist die Datenvariabilität in dieser Region gewöhnlich kleiner, als in anderen Bereichen. In diesem Fall vereinheitlicht die Quadratwurzel-Transformation die Varianzen über das gesamte Untersuchungsgebiet. Diese Funktion nähert die Daten gleichzeitig der Normalverteilung an. Die Quadratwurzel-Transformation ist ein besonderer Fall der Box-Cox-Transformation mit dem Parameter gleich ½ [KRIVORUCHKO 2011, S. 336].

Während eine Normal Score Transformation nach einer Trendeliminierung erfolgen muss, wird die Box-Cox- und die Arcsinus-Transformation angewendet, bevor der Trend aus den Daten entfernt wird [KRIVORUCHKO 2011, S. 341].

Alle funktionalen Transformationen im ArcGIS Geostatistical Analyst erfordern positive Datenwerte. Falls negative Datenwerte vorliegen, ist eine entsprechende Konstante zu allen Werten zu addieren und von den Kriging-Prognosen wieder abzuziehen [KRIVORUCHKO 2011, S. 341].

Die transformierten Werte sind nach erfolgter Schätzung in der Regel wieder in ihren ursprünglichen Wertebereich zurückzutransformieren. Dafür gibt es unterschiedliche Methoden [vgl. KRIVORUCHKO 2011, S. 340, ISAAKS UND SRIVASTAVA 1989, S. 186], auf die hier nicht weiter eingegangen werden kann.

Nach SCHAFMEISTER [1999, S. 12] zeigen viele Variablen (z.B. k_f-Werte[16]) auch noch nach einer entsprechenden Datentransformation Verteilungen, die durch die gängigen Modelle (Normal-, Lognormalverteilung) nicht hinreichend beschreibbar sind und empfiehlt in diesen Fällen die Anwendung von nicht-parametrischen Methoden der Geostatistik. (Indikatorkriging, -simulation).

3.5.4.4 Trend und Drift

Wenn der Erwartungswert der Zufallsvariablen Z im Untersuchungsgebiet nicht konstant ist (siehe auch Abschnitt 3.5.3 Stationarität, Seite 31) muss von der Existenz eines unterlagernden Trends ausgegangen werden. Dieser Trend ist im betrachteten Untersuchungsraum entweder als einheitlich anzusehen, dann spricht man von einem globalen Trend. Oder dieser Trend ist im Untersuchungsgebiet lokal unterschiedlich, wofür sich der Begriff „lokaler Drift“ eingebürgert hat. Eine strikte Trennung ist aufgrund der Datenlage meist nicht möglich, wobei die Wahrnehmung und Einstufung ganz entscheidend vom Betrachtungsmaßstab abhängen [SCHAFMEISTER 1999, S. 35].

Die physikalische Ursache eines Trends ist im allgemeinen ein räumlicher und zeitlicher Prozess, dessen Auswirkungen auf die Variable das Ausmaß lokaler Schwankungen übersteigt. So wirkt sich der Prozess der Grundwasserströmung auf die Neigung der Grundwasseroberfläche stärker aus, als deren lokale Fluktuation. In Richtung des hydraulischen Gradienten kann man zumeist einen linearen Trend anpassen [SCHAFMEISTER 1999, S. 36-37].

Wenn der Mittelwert lokal vom Ort x abhängt, liegt eine lokale Drift vor. Eine solche ist von einem Trend, der durch eine globale Polynomfunktion beschrieben werden kann, zu unterscheiden. Bei einem Drift ist die intrinsische Hypothese abzuschwächen [SCHAFMEISTER 1999, S. 14], vgl. auch Abb. 7.

[16] K_f-Wert = hydraulischer Durchlässigkeitsbeiwert

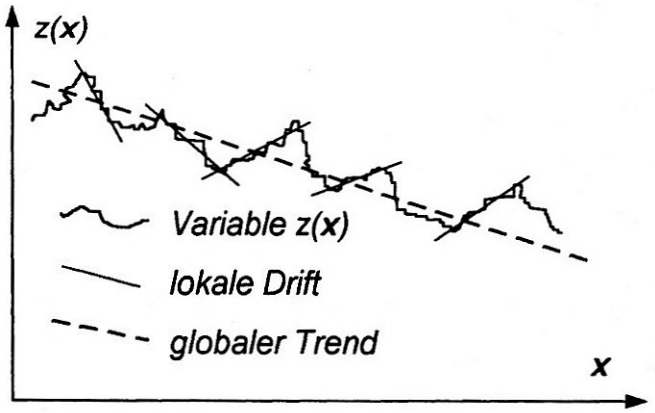


Abb Unterscheidung von lokaler Drift und globalem Trend der Zufallsvariablen $z(x)$.

Quelle: SCHAFMEISTER [1999, ABB. 2.2, S. 13]

Theoretisch erfordert Kriging eine Autokorrelationsfunktion, die vom Ort unabhängig ist. Deshalb müssen strukturelle Trends entfernt werden, bevor annähernde Stationarität erreicht werden könnte. [AGTERBERG 2003, S. 4].

Es gibt mehrere Möglichkeiten, wie mit einem erkannten Trend zu verfahren ist:

- Der Trend wird nicht berücksichtigt, weil vorausgesetzt wird, dass die Auswirkung des Trends bzw. einer Drift in sehr geringen Entfernungen vernachlässigbar klein sind. Die Interpolation erfolgt dann mit dem Ordinary Kriging-Verfahren (siehe auch Abschnitt 3.5.6.2, Seite 49). Da Kriging als gleitendes, gewichtetes Mittelwertverfahren nur Punkte aus der unmittelbaren Nachbarschaft des zu schätzenden Punktes verwendet, kann die Schätzung als ausreichend genau angesehen werden. Bei dieser Vorgehensweise wird jedoch der Krigingfehler, der dann nur noch vom Variogramm und der Messpunktkonfiguration abhängt, unrealistisch hoch [SCHAFMEISTER 1999, S. 37].

- Berechnung einer Trendfläche durch Approximierung einer einfachen Polynomfunktion (linear oder quadratisch) für die Werte im Untersuchungsgebiet und Berechnung der Residuen[17] an den Datenpunkten, die dann bei Variogrammanalyse und Ordinary Kriging als Zufallsvariable Z(x) behandelt werden. Am Schluss werden die mit Kriging geschätzten Residuen der Polynomfunktion wieder hinzuaddiert [SCHAFMEISTER 1999, S. 37].

  Dabei ist die Verletzung einer statistischen Grundannahme in Kauf zu nehmen, da die Residuen einer Trendfläche im strengen Sinne unkorreliert sein sollten, andererseits das zu erstellende Semivariogramm aber die räumliche Korrelation der Variablen voraussetzt.

[17] Differenz zwischen Polynomwert und Messwert an jedem Datenpunkt.

- Verwendung von „Universal Kriging“ (siehe auch Abschnitt 3.5.6.3, Seite 51). Bei diesem Ansatz wird von einer Drift ausgegangen, die sich mit Hilfe einer Polynomfunktion beschreiben lässt. Die Verwendung ist jedoch nur dann angebracht, wenn eine Drift schon in geringer Entfernung deutlich wird [SCHAFMEISTER 1999, S. 38] (siehe auch weiter unten).

Eine Drift, also ein lokaler Trend, liegt vor, wenn der Mittelwert eine Funktion des Ortes x darstellt. In solchen Fällen wird die wahre räumliche Struktur nicht durch das experimentelle Variogramm, sondern durch das sogenannte Residuenvariogramm widergespiegelt. Die Summe aus Residuenvariogramm γ (**h**)$_{res}$ und der halben Drift bildet das experimentelle Variogramm γ (**h**)* [SCHAFMEISTER 1999, S. 18].

Drift d^2 (**h**) = [m^+ (**h**) – m^- (**h**)]²

mit m^+ (**h**) = Mittelwerte der Startpunkte
m^- (**h**) = Mittelwert der Endpunkte aller Wertepaare im Abstand h

experimentelles Variogramm γ (**h**)* = γ (**h**)$_{res}$ + 0,5 d^2(**h**)

Es gibt gängige, allgemein zugänglichen Programmpakete[18] die neben dem experimentellen Variogrammwert γ (**h**)* auch die Mittelwerte für Start- und Endpunkte aller Paare einer Distanzklasse |h| liefern, aus denen sich das Residuenvariogramm errechnen lässt oder die das Residuenvariogramm mitberechnen [SCHAFMEISTER 1999, S. 18].

3.5.5 Semivariogramm und Kovarianz

Grundsätzlich gilt, dass nahe beieinanderliegende Datenpunkte sich ähnlicher sind, als weiter voneinander entfernt liegende. Semivariogramm und Kovarianzdiagramm bringen räumliche Ähnlichkeitsstrukturen in einem Datensatz zum Ausdruck. Als Maß dafür dienen die Semivarianz γ(h) oder die räumliche Kovarianz Cov(h) von Datenpaaren, die nach Abstandsklassen (Lags[19]) aggregiert werden.

Die empirische Semivarianz berechnet sich mit der folgenden Formel [nach Schafmeister 1999, S. 14]:

$$\gamma(h)=\frac{1}{2\,N(h)}\sum_{i=1}^{N}\left(z(x_i)-z(x_i+h)\right)^2$$

mit N(h) = Anzahl der Probenpaare im Abstand h

---

[18] VARIOWIN, Geostatistical TOOLBOX, GSLIB, CROSSVA

[19] Festlegung der Lag-Parameter: Lag(size) * Lag(num) = 0,5 * d(max)

mit Lag(size) = Lag-Größe (Intervall); Lag(num) = Anzahl der Lags; d(max) = max. Distanz eines Punktpaares [Quelle: Unterlagen zum UNIGIS MSC-Lehrgang, Modul Geostatistik 2010, Lektion 7, Autokorrelation, Folie 22]

Die empirischen Semivariogramm-Mittelwerte der Abstandsklassen werden in einem experimentellen Semivariogramm gegen die Abstände h aufgetragen. Sie ergeben so eine Kurve, welche die räumliche Variabilität der Größe quantitativ beschreibt. Geowissenschaftliche Kenngrößen , wie z.B. ein Grundwasserspiegel, ähneln sich dabei innerhalb eines bestimmten Einflussbereiches. Ein natürlicher Grundwasserspiegel zeigt normalerweise keine abrupten Sprünge und bei gut löslichen und mobilen Grundwasserinhaltsstoffen ist nicht damit zu rechnen, dass hohe Konzentrationen nur punktuell auftreten. Infolgedessen steigt das Variogramm im Ursprung steil an und flacht dann beim Erreichen der Reichweite bzw. des Ranges ab. Bei Distanzen größer der Reichweite unterscheiden sich zwei Messwerte innerhalb des Varianzspektrums [SCHAFMEISTER 1999, S. 14-15].

An die Punktpaare im Semivariogramm- bzw. im Kovarianz-Diagramm wird dann ein Funktionsmodell angepasst, welche die räumliche Variabilität der Messgröße beschreibt. anhand derer dann mit einem linearen Satz an Kriging-Gleichungen Kriging-Gewichte zur Schätzung von Vorhersagen an unbekannten Orten ermittelt werden [siehe u.a. SCHAFMEISTER 1999, S. 15, GAU 2010, S. 28]

Die Wahl der Anzahl an Distanzklassen und deren Breite zum Berechnen des Proben-Variogramms bei unregelmäßig verteilten Proben führt nach Myers [1991, S. 214] zu einem Zielkonflikt: einerseits sollten die Klassen klein sein, um aus dem Graphen mehr Details zu entnehmen; aber um die Verlässlichkeit jedes gedruckten Punktes zu stärken, sollte andererseits die Anzahl an Paaren, die zur Erstellung dieses Punktes verwendet werden, groß sein. Leider kann man für ein gegebenes Probenmuster nicht beides haben. Wenn keine zusätzlichen Proben genommen werden können, bleibt nichts anderes übrig, als mit der Anzahl und der Breite der Klassen zu experimentieren, um die Information, die aus dem Plot zu entnehmen ist, zu optimieren

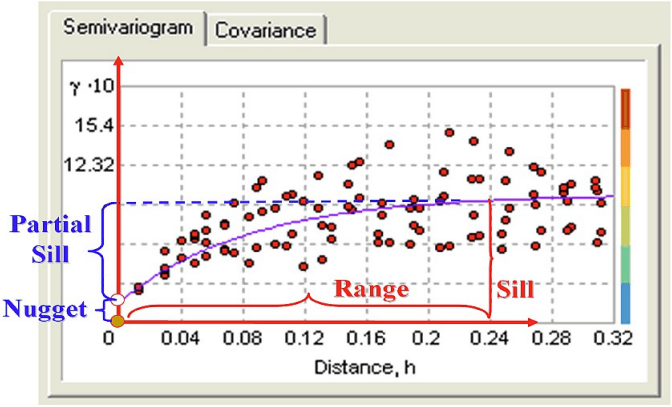


Abb. 8: Semivariogramm

Quelle: KRIVORUCHKO [2011, S. 295]

Zu den Annahmen für die konventionellen Kriging-Modelle gehört, dass das Semivariogramm oder das Kovarianz-Modell exakt bekannt ist und im gesamten Untersuchungsgebiet gleich ist. Im Fall des Moving Window-Krigings gilt das innerhalb des wandernden Fensters [KRIVORUCHKO 2011, S. 370].

Die Semivariogramm-Funktion hat drei oder mehr Parameter: Range, Nugget und Sill [nach KRIVORUCHKO 2011; DE SMITH ET.AL. 2013, S. 516 u.a.].

- Der Range bzw. die Reichweite ist die Distanz außerhalb der die Daten keine signifikante Korrelation mehr miteinander aufweisen

- Beim Nugget (Klumpeneffekt) handelt es sich um eine Diskontinuität am Ursprung des Diagramms, die auf Messfehler und Datenvariationen auf sehr kurze Distanz (Mikrodatenvariation; d. h. auf Distanzen kleiner dem kleinsten Abstand zwischen Messpunkten) zurückgeht.

- Der Sill gibt die totale Variation innerhalb der Daten an

- Der Partial Sill gibt den Betrag der kleinräumlichen Variation an (wenn kein Trend vorliegt, bzw. ein vorhandener Trend entfernt wurde). Er ergibt sich als Differenz zwischen Sill und Nugget.

Bei Anisotropien können die Semivarianzen und räumlichen Kovarianzen nicht nur von der Distanz, sondern auch von der Richtung abhängen. Das Semivariogramm und das Kovarianzdiagramm sind daraufhin zu prüfen und die Parameter entsprechend zu wählen.

Das Kovarianz-Diagramm sieht aus, wie ein umgedrehtes Semivariogramm-Modell. So ist es nach KRIVORUCHKO [2011, S. 297] für Semivariogramme mit Sill auch einfach, mit den folgenden Formeln zwischen Semivariogramm- und Kovarianzfunktionen umzurechnen $\gamma(h)=Cov(0)-Cov(h)$ und $Cov(h)=\gamma(\infty)-\gamma(h)$

h = Abstand der jeweiligen Punktpaare
$\gamma(h)$ = Semivarianz
$\gamma(\infty)$ = Semivarianz bei unendlicher Distanz
Cov(0) = Kovarianz beim Abstand 0
Cov(h) = Kovarianz beim Abstand h

KRIVORUCHKO [2011, S. 297] betont weiter, dass alle Semivariogramme und Kovarianzen im ArcGIS Geostatistical Analyst einen Sill hätten, wodurch diese Umrechnung dort immer möglich ist. Weil die Ermittlung der Kovarianz den Gesamt-Mittelwert erfordert, der aber normalerweise nicht bekannt ist und angegeben werden muss, würde das Semivariogramm öfter benutzt als das Kovarianzdiagramm. Praktisch sei das Semivariogramm besser für die Schätzung der Datenkorrelation auf kurze Distanz, während die Kovarianz besser zum Schätzen der Modellparameter für große Abstände (Lags) geeignet ist (Sill oder Range Parameter).

Semivariogrammwerte können unterschiedlich schnell mit der Distanz der Datenpaare anwachsen. KRIVORUCHKO [2011, S. 298] u.a. gibt für den zulässigen Anstieg des Semivariogramm-Modells einen Schwellenwert von $a * h^2$ an,

d. h. es wird gefordert, dass die Steigung von $\gamma(h)$ < a * h²

Zur Anpassung an die empirischen Semivarianzen und Kovarianzen werden bekannte Funktionen mit den benötigten Merkmalen verwendet KRIVORUCHKO [2011, S. 300] (siehe auch Abschnitt 3.5.7 Kriging-Modelle).

KRIVORUCHKO [2011, S. 322] weist darauf hin, dass in der geostatistischen Literatur zwar manchmal die iterative visuelle Anpassung von Semivariogramm-Modellen empfohlen wird, zeigt aber an einem simulierten Beispiel, dass der optische Eindruck täuschen kann und es besser ist, eine statistische Schätzung der Modellparameter zu verwenden.

Die wichtigsten statistischen Methoden zu Schätzen der Parameter des Semivariogramm-Modells sind gewichtete kleinste Quadrate (weighted least squares) und die maximale Ähnlichkeit. Die Annahme bei der Methode der normalen (ungewichteten) kleinsten Quadrate[20] ist, dass die Daten unabhängig sind, wogegen beim empirischen Semivariogramm aber immer verstoßen wird. Beim Verfahren der gewichteten kleinsten Quadrate sind die Gewichte mit der Varianz des empirischen Semivariogramms verbunden [KRIVORUCHKO 2011, S. 322-323].

Eine universell beste Methode zur Anpassung der Semivariogramme gibt es nicht. Praktisch funktioniert jeder plausible Algorithmus, der interaktiv und mit ausreichender Modellanpassungsdiagnostik versehen ist, weil eine graphische Darstellung für eine gute Datenanalyse unverzichtbar ist. [KRIVORUCHKO 2011, S. 323].

3.5.6 Kriging-Verfahren

Bei Kriging-Verfahren handelt es sich um eine Klasse von statistischen Techniken zur optimalen räumlichen Interpolation. Kriging-Vorhersager (predictors = independent variable) werden optimal genannt, wenn sie statistisch unverzerrt sind (d.h. der vorhergesagte und der wahre Wert fallen im Durchschnitt zusammen) und sie minimieren den mittleren quadrierten Vorhersagefehler (Maß für die Unsicherheit in den vorhergesagten Werten) [KRIVORUCHKO 2011, S. 251].

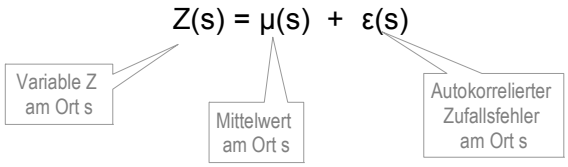
JERNIGAN [1986, S. 2] beschreibt Kriging als Technik, welche die statistischen Eigenschaften von Regression mit den interpolierenden Eigenschaften von Splines kombiniert.

Die drei wichtigsten Kriging-Verfahren (Simple Kriging, Ordinary Kriging, Universal Kriging) sind lineare Vorhersager (predictors) [KRIVORUCHKO 2011, S. 293].

Die Grundidee zum Kriging lautet in einer Formel ausgedrückt[21]:

---

[20] Normale, ungewichtete kleinste Quadrate: $\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \rightarrow min$ mit N für die Anzahl der Beobachtungen, y_i für eine (empirische) Beobachtung und $\hat{y}_i$ für den Schätzwert dazu



Geostatistische Verfahren nehmen zufällige Variationen der Wertausprägungen an (siehe obige Formel. Allgemein wird angenommen, dass μ über das Untersuchungsgebiet konstant ist, also statt μ(s) nur μ in die Gleichung gesetzt werden kann. Dann wird angenommen, dass jeder Wert Z im Untersuchungsgebiet zufällig vom Mittelwert abweicht und das heißt, dass keine Trends vorliegen. Für den räumlich autokorrelierten Zufallsfehler ε(s) wird angenommen, dass sein Mittelwert Null ist und dass Stationarität vorliegt.

Die meisten Kriging-Modelle sagen einen Wert $\hat{Z}(s_?)$ mit einer gewichteten Summe der nächstgelegenen N Orte $Z(s_i)$ vorher [KRIVORUCHKO 2011, S. 284]:

$$\hat{Z}(s_?) = \sum_{i=1}^{N} \lambda(s_i) \cdot Z(s_i)$$

mit λ = Gewicht

Zur Festlegung der Gewichte für jeden der Datenpunkte zur Vorhersage neuer Werte an Orten ohne Messwerte wird bei Kriging das Semivariogramm $\gamma(h)$ verwendet, das aus den verfügbaren Daten [geschätzt / berechnet] wird. Es wird also keine a priori Annahme über die Verteilung der Werte mittels einer Funktionsformel, wie bei den deterministischen Verfahren gemacht, sondern eine Anpassung anhand der beobachteten Werte vorgenommen, also empirisch.

Kriging-Modelle unterscheiden sich in der Art, wie sie den Mittelwert modellieren und in den Annahmen, die über die Datenverteilung (keine Annahmen, multivariate Gauß- bzw. Normalverteilung, bivariate Normalverteilung) gemacht werden [KRIVORUCHKO 2011, S. 289]

Abb. 9 veranschaulicht den Unterschied zwischen der Mittelwert-Spezifikation beim Simple Kriging und der Mittelwert-Schätzung beim Ordinary Kriging und beim Universal Kriging.

[21] Quelle: UNIGIS MSC-Modul „Geostatistik“ 2010, Lektion 13, Folie 6

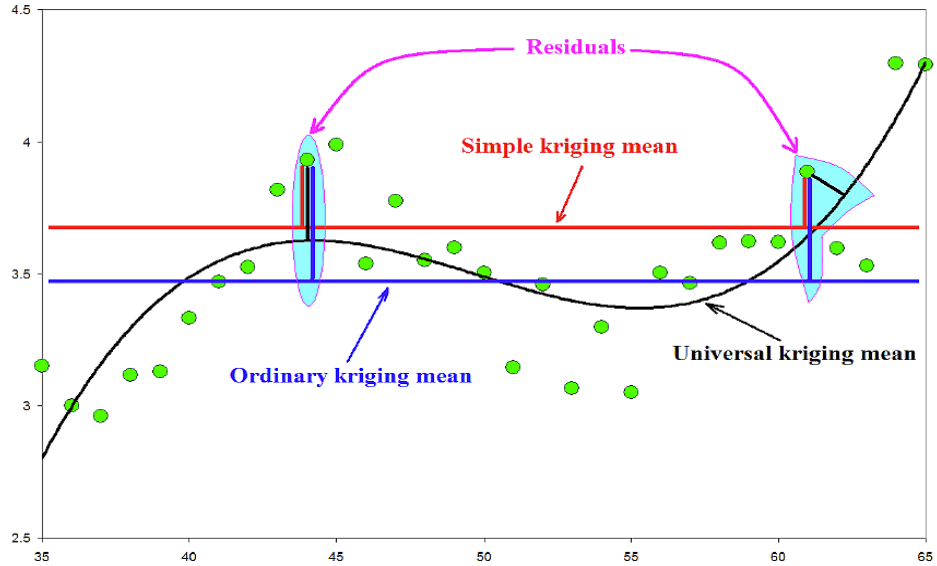


Abb. 9: Unterschiede in der Behandlung des Mittelwertes beim Kriging

Quelle: KRIVORUCHKO [2011, S. 293]

Der Mittelwert ist beim Simple Kriging ein geforderter Input-Parameter (rote Linie in Abb. 9), wird beim Ordinary Kriging als Konstante in der Suchnachbarschaft geschätzt und beim Universal Kriging als Polynom niedriger Ordnung eingesetzt (in Abb. 9: Polynom 3. Ordnung) [KRIVORUCHKO 2011, S. 293].

Die verbleibenden Residuen[22] (Zufallsfehler) sind in Abb. 9 an zwei Stellen mit blauen Markierungen hervorgehoben. Die Beträge sind ganz offensichtlich unterschiedlich und es ist zu erwarten, dass Kriging-Vorhersagen für Vorhersagepunkte auch unterschiedlich ausfallen [vgl. KRIVORUCHKO 2011, S. 293]

Kriging kann sowohl groß- als auch kleinräumliche Datenvariationskomponenten einschließen [KRIVORUCHKO 2011, S. 264].

Kriging kann sowohl exakter, wie auch nicht-exakter Interpolator sein [KRIVORUCHKO 2011, S. 269]. Ob die Verwendung eines exakten Interpolators sinnvoll ist, hängt davon ab, wie präzise die Daten sind. Es macht keinen Sinn Daten mit einem Rauschen exakt zu reproduzieren. In einem solchen Fall ist es besser, die verfügbare Information von Nachbarwerten zur Vorhersage eines neuen Wertes für den Messort zu verwenden, der sich als präziser, als der ursprüngliche Wert herausstellen kann [KRIVORUCHKO 2011, S. 269]. Nach DE SMITH ET.AL. [2013, S. 530] ist Kriging nicht exakt [bzw. kein exakter Interpolator], wenn ein Nugget vorliegt.

Die Wahl eines Kriging-Modells sollte auf dem Ergebnis einer Datenuntersuchung und Vorinformationen über den physikalischen Prozess, der die Daten vermutlich erzeugt hat, basieren [KRIVORUCHKO 2011, S. 289].

---

[22] Residuen = Differenz zwischen Mittelwert und tatsächlichem Wert an jedem Datenpunkt

Es ist möglich, polynome Interpolation mit Kriging zu kombinieren. Dabei wird in einem ersten Schritt eine lokale polynome Interpolation angewendet, was mit einem Trendmodell möglich ist. Danach kann untersucht werden, wieviel Variation in den Daten verbleibt, wenn der Trend entfernt ist, um sich dann im nächsten Schritt die räumliche Korrelation in den Residuen näher anzusehen, d. h. eine Modellierung der arithmetischen Differenz zwischen den Datenwerten und dem Trendoberflächenwert an den Messorten vorzunehmen [KRIVORUCHKO 2011, S. 271].

Im Fall von Simple Kriging und Disjunktivem Kriging sollen durch das Eliminieren des Trends Daten erhalten werden, deren Mittelwert nahe Null liegt. Praktisch wird die mittlere Oberfläche im Kriging normalerweise als exakt bekannt angenommen, was zu einer Unterschätzung des Vorhersage-Standardfehlers führt, denn es ist kaum anzunehmen, dass die geschätzte mittlere Oberfläche mit der wahren zusammenfällt [KRIVORUCHKO 2011, S. 273].

Obwohl Vorhersageoberflächen von Simple, Ordinary und Universal Kriging normalerweise nicht sehr stark differieren, können die Unterschiede in den Ansätzen zum Modellieren der großräumlichen Datenvariation zu ganz unterschiedlichen Vorhersagestandardfehlern führen [KRIVORUCHKO 2011, S. 293].

Kriging und andere Interpolatoren tendieren generell zum Unterschätzen hoher und Überschätzen niedriger Werte [KRIVORUCHKO 2011, S. 337].

3.5.6.1 Simple Kriging

Wenn Daten multivariater Gauß-Verteilung folgen, ist [Simple] Kriging der beste Vorhersager nach dem Kriterium zur Minimierung der quadrierten Fehler [KRIVORUCHKO 2011, S. 285].

Simple Kriging erfordert als Modellinput einen bekannten Mittelwert oder eine mittlere Oberfläche [KRIVORUCHKO 2011, S. 293]. Die Annahme, dass das Mittel bekannt ist, mag unrealistisch sein. Simple Kriging wird jedoch häufig in der Meteorologie verwendet, weil es dort möglich ist, die mittlere Oberfläche unter Verwendung von Beobachtungen der Vergangenheit zu schätzen [KRIVORUCHKO 2011, S. 293].

Damit scheidet Simple Kriging für die Interpolation von Grundwasserdaten mit nur einzelnen Messreihen aus und wird hier als Thema nicht weiter vertieft.

Außerdem erfordert Simple Kriging Daten, aus denen der Trend entfernt wurde.Im ArcGIS Geostatistical Analyst wird die Mittelwert-Oberfläche gewöhnlich mit einer lokalen polynomen Interpolation geschätzt. Wenn der angenommene Mittelwert oder die Mittelwertoberfläche korrekt bestimmt wurden, dann ist das Mittel aus den Residuen gleich Null [KRIVORUCHKO 2011, S. 351].

Nach der Literatur liefert Simple Kriging den geringsten mittleren quadrierten Prognosefehler von allen Kriging-Verfahren einschließlich Ordinary und Universal Kriging. Diese Feststellung gilt aber nur für den Fall, dass das Mittel wirklich bekannt ist und nicht aus den gleichen Daten geschätzt wurde [KRIVORUCHKO 2011, S. 352].

3.5.6.2 Ordinary Kriging

Diese Methode wird oft mit der Abkürzung BLUE für “best linear unbiased estimator”, also “bester linearer unverzerrter Schätzer” in Verbindung gebracht. Ordinary Kriging ist “linear”, weil seine Schätzungen gewichtete lineare Kombinationen der verfügbaren Daten sind. Sie ist “unverzerrt”, weil sie versucht, die mittleren Residuen oder Fehler gleich Null zu haben und sie ist “am besten”, weil sie darauf abzielt, die Varianz der Fehler zu minimieren [ISAAKS UND SRIVASTAVA 1989, S. 278].

Wichtig ist, dass wir tatsächlich weder den Mittelwert kennen und deshalb auch nicht garantieren können, dass er genau Null ist, noch die Fehlervarianz kennen und sie deshalb auch nicht minimieren können. Die einzige Möglichkeit ist, ein Modell der zu untersuchten Daten zu erstellen und mit dem durchschnittlichen Fehler und der Fehlervarianz des Modells zu arbeiten [ISAAKS UND SRIVASTAVA 1989, S. 278].

Ordinary Kriging ist eine Variante von Kriging, die zunächst mit einer lokalen Polynominterpolation eine mittlere Oberfläche schätzt und dann die Differenz zwischen den Datenwerten und der mittleren Oberfläche (Residuen) als Input für die geostatistische Modellierung verwendet. Zum Schluss fügt die Methode die mit Kriging vorhergesagte Oberfläche wieder mit der polynomen Oberfläche zusammen [KRIVORUCHKO 2011, S. 269].

Es wird vorausgesetzt, dass die Auswirkung des Trends bzw. einer Drift in sehr geringen Entfernungen vernachlässigbar klein sind. Die Interpolation erfolgt dann mit dem Ordinary Kriging-Verfahren (siehe auch Abschnitt 3.5.6.2, Seite 49). Da Kriging als gleitendes, gewichtetes Mittelwertverfahren nur Punkte aus der unmittelbaren Nachbarschaft des zu schätzenden Punktes verwendet, kann die Schätzung als ausreichend genau angesehen werden. Bei dieser Vorgehensweise wird jedoch der Krigingfehler, der dann nur noch vom Variogramm und der Messpunktkonfiguration abhängt, unrealistisch hoch [SCHAFMEISTER 1999, S. 37].

In der Schreibweise von KRIVORUCHKO [2011] lautet die Gleichung zur Schätzung mit Kriging:

$$\hat{Z}(s_?) = \sum_{i=1}^{N} \lambda(s_i) \cdot Z(s_i)$$

mit λ = Gewicht

Die Gewichte $\lambda(s_i)$ sind so zu bestimmen, dass der Schätzwert $\hat{Z}(s_?)$ des unbekannten wahren Wertes $Z(s_?)$ die folgenden Bedingungen erfüllt [die folgenden Formeln nach SCHAFMEISTER 1999, S. 37 sind an die Schreibweise von KRIVORUCHKO 2011 angepasst]:

$\hat{Z}(s_?)$ sei erwartungstreu, d.h. $E[\hat{Z}(s_?) - Z(s_?)] = 0$

und der mittlere quadratische Fehler $E[\hat{Z}(s_?) - Z(s_?)]^2$ = sei ein Minimum

Unter der Annahme der Stationarität, also bei Abwesenheit eines Trends, ist der Erwartungswert $E[Z(s_i)] = m$ und $Z(s_?) = m$. Die Bedingung der Erwartungstreue liefert somit

$$E\left[\sum_{i=1}^{N}\lambda(s_i)\cdot Z(s_i)' Z_0\right] = \sum_{i=1}^{N}\lambda(s_i)\, m - m = m\left(\sum_{i=1}^{N}\lambda(s_i) - 1\right) = 0$$

Für Ordinary Kriging wird also angenommen, dass die Summe der Kriging-Gewichte in der Kriging-Nachbarschaft gleich 1 ist, d. h.

$$\sum_{i=1}^{N}\lambda_i = 1$$

Diese Beschränkung funktioniert allerdings nur, wenn der Messfehler vernachlässigbar ist und die Anzahl der Proben bzw. Messwerte in der Kriging-Nachbarschaft groß ist. Aber genau das ist nach KRIVORUCHKO [2011, S. 353] bei geologischen Daten in Frage zu stellen, da zum einen die Anzahl von verfügbaren Proben aus Kostengründen geringer, als bei Anwendungen in der Meteorologie (und im Bergbau), andererseits der Messfehler größer, als bei meteorologischen Daten ist.

Nach KRIVORUCHKO [2011, S. 293] nimmt Ordinary Kriging nämlich an, dass das Mittel irgendeinen Wert haben und über die Teilbereiche des Untersuchungsgebietes variieren kann, jedoch in der Kriging-Nachbarschaft annähernd konstant ist. Es nimmt also einen konstanten, aber unbekannten Mittelwert an und schätzt deshalb den Mittelwert als Konstante in der Suchnachbarschaft der Vorhersage. Die Zufallsfehler werden deshalb auch geschätzt. KRIVORUCHKO [2011, S. 353] mahnt in diesem Zusammenhang, zur Vorsicht, dass eine solch vereinfachende Annahme zum komplexen Frage der Mittelwert- oder mittleren Oberflächen-Schätzung nicht für jedes Anwendungsproblem funktioniert.

Mit Hilfe des Variogramms kann der Erwartungswert des quadratischen Fehlers ausgedrückt werden mit:

$$E\left[\hat{Z}(s_?) - Z(s_?)\right]^2 = Var\left(\hat{Z}(s_?) - Z(s_?)\right)$$
$$= 2\sum_{i=1}^{N}\lambda(s_i)\,\gamma(s_i - s_?) - \sum_{i=1}^{N}\sum_{j=1}^{N}\lambda(s_i)\,\lambda(s_j)\,\gamma(s_i - s_j)$$

Zur Minimierung der Fehlervarianz unter der Bedingung $\Sigma\lambda(s_i) = 1$ sowie dem Kriging-Gleichungssystem siehe die einschlägige Literatur (KRIVORUCHKO 2011, SCHAFMEISTER 1999 u.a).

Die Kriging-Schätzvarianz σ^2_K für Punktschätzungen ergibt sich aus der vorhergehenden Gleichung (nach SCHAFMEISTER 1999 / KRIVORUCHKO 2011) :

$$\sigma^2_K = Var\left(\hat{Z}(s_?) - Z(s_?)\right) = \mu + \sum_{i=1}^{N}\lambda(s_i)\,\gamma(s_i - s_j)$$

In dem Sonderfall, dass keine räumliche Abhängigkeit der Daten besteht, d.h. das Variogramm ein reiner Nugget-Effekt ist, erhält man für die Gewichte $\lambda(s_i) = 1/N$. Der Krigingschätzer ist dann das einfache arithmetische Mittel der benachbarten Proben [SCHAFMEISTER 1999, S. 36].

Nach SCHAFMEISTER [1999, S. 36] zeichnen die folgende Eigenschaften den Ordinary Kriging-Schätzer aus:

- das Krigingsystem ist nur lösbar, wenn die Determinante der Matrix (γ_{ij}) ≠ 0 ist. Praktisch bedeutet dies, dass eine Probe nicht doppelt auftreten darf (d.h. mit identischen Koordinaten)

- (Punkt-)Kriging liefert einen exakten Interpolator. Das folgt aus der Bedingung der Schätzvarianz.

- Das Kriging-Gleichungssystem hängt nur von $\gamma(h)$ bzw. C(h) ab und nicht von den Werten der Variablen Z in den Probenpunkten $x(s_i)$, i = 1, … n. Bei identischer Datenkonfiguration braucht das Gleichungssystem nur ein Mal gelöst zu werden

- mit Hilfe der Schätzfehlers σ^2_K können Vertrauensgrenzen der Schätzung angegeben werden.

- Kriging berücksichtigt bei der Schätzung die Lage des zu interpolierenden Punktes $x(s_?)$ in Bezug auf die Datenpunkte $x(s_i)$ und die Struktur der Variablen Z (durch das Variogramm) . Durch die Bedingungen der Erwartungstreue und Fehlerminimierung ist Kriging ein BLUE-Schätzer (s.o.) und den einfach distanzgewichteten Verfahren überlegen [SCHAFMEISTER 1999].

KRIVORUCHKO [2011, S. 285] stellt fest, dass Lev Gandins klassische Herleitung des Ordinary Kriging keine Normalverteilung der Daten annehmen würde, dass Ordinary Kriging in vielen Schriften aber als Gauß-Vorhersager eingestuft wird. KRIVORUCHKO hält das aber für generell nicht gerechtfertigt.

Simple und Ordinary Kriging können zur Interpolation von Daten verwendet werden, die einen Trend haben, weil es normalerweise allein auf der Grundlage der Daten keine Möglichkeit gibt, zu entscheiden, ob das beobachtete Muster nur ein Resultat von Datenkorrelation ist [KRIVORUCHKO 2011, S. 293].

3.5.6.3 Universal Kriging

Wenn von der Existenz eines Trends oder einer Drift im Untersuchungsgebiet auszugehen ist (siehe auch Abschnitt 3.5.4.4, S. 40), der nicht vernachlässigt werden kann, sollte Universal Kriging verwendet werden.

Bei Vorliegen eines globalen Trends wird zunächst eine einfache lineare oder quadratische Polynomfunktion für die Werte im Untersuchungsgebiet angepasst und die Residuen, also die Differenzen zwischen den Polynomwerten und den jeweiligen tatsächlichen Datenwerten an jeder Messstelle berechnet. Die Residuen werden dann als Zufallsvariable Z(x) mit Variogrammanalyse und Ordinary Kriging behandelt und die Krigingschätzung der Residuen den Werten der Polynomfunktion zum Schluss wieder hinzugefügt [SCHAFMEISTER 1999, S. 37].

Bei Vorliegen eines lokalen Trends, also einer Drift besteht der Ansatz darin, diese Drift lokal mit einer Polynomfunktion zu beschreiben[23] [nach SCHAFMEISTER 1999, S. 38].

Der Erwartungswert von Z(x) ist eine Funktion des Ortes x :

$$m(x) = E [Z(x)]$$

wobei m(x) die Drift bezeichnet.

Z(x) hat nach dem geostatistischen Modell drei Komponenten:

$$Z(x) = m(x) + R(x) + \varepsilon$$

wobei R eine regionalisierte Variable mit Variogramm $\gamma(h)$ und Erwartungswert E[R] = 0 ist. Der Ausdruck ε steht für zufälliges „Rauschen".

Die Drift wird mit Hilfe eines Polynomansatzes als Linearkombination einfacher Funktionen $f_l(x)$ beschrieben:

$$m(x) = \sum_{l=0}^{k} a_l f^l(x)$$

wobei die Koeffizienten a_l zunächst unbekannt sind. Bei einer linearen Drift ergibt sich unter Verwendung von Monomen $f_l(x)$ für eine Ebene (2D):

$$m(x,y) = a_0 + a_l x + a_2 y \text{ mit } k = 2,$$

und bei einer quadratischen Drift:

$$m(x,y) = a_0 + a_l x + a_2 y + a_3 xy + a4x2 + a5y2 \qquad \text{mit } k = 5$$

Analog zum Ordinary Kriging ergibt sich aus den Bedingungen der Erwartungstreue und der Minimierung des mittleren quadratischen Fehlers ein Kriging-Gleichungssystem, das außer den Gewichten λ_i, i = 1, 2, ...n für den nichtstationären Fall zusätzlich auch die (unbekannten) Koeffizienten a_l enthält. Auf diese Weise für für jeden Schätzpunkt eine lokale Polynomfunktion bestimmt, die die lokale Drift beschreibt [SCHAFMEISTER 1999, S. 38].

Nach SCHAFMEISTER [1999, S. 38] ist Universal Kriging nur dann wirklich angebracht, wenn eine Drift schon in geringen Entfernungen deutlich wird. Die Drift kann anhand der Mittelwerte der Start- und Endpunkte jeder Distanzklasse (siehe Ausführungen zur Drift in Abschnitt 3.5.4.4, S. 42) berechnet werden.

Häufig ist eine Drift nur für eine Raumrichtung oder ein Teilgebiet erkennbar, z.B. bei Grundwasserspiegeln Richtung des generellen Abstroms bzw. im Einzugsgebiet eines Brunnens. In solchen Fällen sollte das Variogramm senkrecht zur Abstromrichtung bzw. in unbeeinflussten Gebieten berechnet werden [SCHAFMEISTER 1999, S. 38] .

---

[23] Schreibweise der Formel entspricht hier der Fassung von SCHAFMEISTER [1999, S. 38]. Es war leider nicht mehr möglich die Formelschreibweisen in dieser Arbeit zu vereinheitlichen.

Universal Kriging modelliert lokale Mittel mit Polynomfunktionen niedriger Ordnung für die räumlichen Koordinaten [KRIVORUCHKO 2011, S. 293]. In Abb. 9 (S. 47) ist das z.B. ein Polynom 3. Ordnung. Bildet man jeweils die Differenz zwischen Originaldaten und Mittelwert, erhält man die Residuen bzw. Zufallsfehler für die zur Interpolation eine Regression mit den räumlichen Koordinaten als erklärende Variable durchgeführt wird. Die Fehler werden also korreliert modelliert [KRIVORUCHKO 2011, S. 293].

Beim Universal Kriging stellt sich ein rechentechnisches Problem: einerseits ist für die Schätzung des Semivariogramm-Modells Kenntnis über den Trend erforderlich, aber gleichzeitig erfordert die Schätzung des Trends auch Wissen über das Semivariogramm-Modell. Das Problem kann näherungsweise über Iterationen gelöst werden [KRIVORUCHKO 2011, S. 293].

3.5.6.4 Kriging mit externer Drift

Beim Kriging mit externer Drift wird ähnlich wie beim Universal Kriging ein Trendmodell angenommen, bei dem der Trend auf zwei Terme begrenzt ist:

$$E\,[\,Z(x)\,] = m(x) = a_0 + a_1\, f_1\,(x)$$

Dabei stellt die Funktion $y(x) = f_1(x)$ eine zusätzliche variable dar, deren räumliche Variabilität (Externe Drift) in gedämpfter Form diejenige der Zufallsfunktion widerspiegelt. Das Kriging mit externer Drift benötigt eine Sekundärvariable, die für alle x bekannt sein muss. Praktisch reduziert sich diese Anforderung auf Werte für die Variable Y an allen Gitterstützpunkten x_0 und den Datenpunkten x_i. Hierdurch wird eine allgemeine Grundkontur der Schätzvariablen beschrieben [SCHAFMEISTER 1999, S. 39].

Als Sekundärvariable Y(x) kommen Variable in Betracht, die aufgrund einer physikalisch begründbaren Beziehung zur Zielvariablen Z(x) deren Verlauf nachzeichnen. Als Beispiele für Sekundärvariable führt SCHAFMEISTER [1999, S. 39] die Verwendung des Spezifischen Speicherkoeffizienten zur Schätzung der Transmissivität oder die topographische Höhe zur Schätzung der Bodentemperaturen in einer gebirgigen Region an.

Die Grundwasserdruckhöhe zum einem Zeitpunkt t kann mit Kenntnis früherer Druckspiegellagen ebenfalls besser geschätzt werden. SCHAFMEISTER [1999, S. 40] verweist in diesem Zusammenhang auf ein Beispiel von Chiles (19992), in dem als Sekundärvariable die Mittelwerte zweier extremer Grundwasserhöhenlagen und Messwerte einer Trocken- bzw. einer Feuchtperiode als „generelle Grundwasserhöhe“ zur Interpolation der Grundwasserhöhe zum Zeitpunkt einer räumlich weniger dichten Stichtagsmessung verwendet wurden. Die grundlegende Idee dabei ist, dass die generelle Form einer Grundwasserdruckfläche nur eingeschränkt veränderlich ist, wenn Grundwasserentnahmen oder -infiltrationen auszuschließen sind. Die generelle Form kann dann mit Hilfe der mittleren Druck-höhe als Sekundärvariable Y(x) vorgezeichnet werden.

Eine analoge Vorgehensweise ist grundsätzlich bei wenigen verfügbaren Messpunkten für eine Variable interessant, sofern sich auf diese Weise zusätzliche

Informationen zum Schließen von Informationslücken zwischen den Messpunkten ergeben. Als Beispiele kann man hier die Verwendung von Daten aus nicht mehr bestehenden Messstellen in Verbindung mit zurückliegenden Messkampagnen vorstellen.

3.5.6.5 Kokriging

Kokriging kombiniert räumliche Daten mit mehreren Variablen, um eine Werteoberfläche für eine der Variablen unter Verwendung der räumlichen Korrelation der interessierenden Variablen und der Kreuzkorrelation zwischen ihr und anderen Variablen zu erzeugen [KRIVORUCHKO 2011, S. 294].

Damit können z. B. Vorhersagen bei lückenhaften Datensätzen ergänzt und verdichtet werden.

Alle von KRIVORUCHKO [2011], im Abschnitt S. 292-294, beschriebenen Kriging-Modelle können eine sekundäre Variable verwenden. Deshalb gibt es neben Ordinary Kokriging prinzipiell auch Simple Kokriging oder Universal Kokriging etc.

Im ArcGIS Geostatistical Analyst können bis zu drei sekundäre Variable verwendet werden, um genauere Vorhersagen der Primärvariablen an Orten ohne Messwerte zu machen [KRIVORUCHKO 2011, S. 357].

Dieser Ansatz setzt die Verfügbarkeit von geeigneten Sekundärvariablen voraus und wird hier deshalb nicht weiter verfolgt.

3.5.6.6 Indikator-Kriging

Beim Indikator-Kriging werden aufbereitete Daten verwendet. Bei der Aufbereitung erhalten die Originaldaten Werte von 1 oder 0 zugewiesen, je nachdem, ob sie über oder unter einem bestimmten Schwellenwert liegen.

Diese Indikatorwerte werden dann als Input eines linearen Kriging-Modells, meist Ordinary Kriging, verwendet. Unter der Voraussetzung, dass Ordinary Kriging kontinuierliche Vorhersagen mit Werten zwischen 0 und 1 hervorbringen kann, ist ein Vorhersagewert an einem Vorhersageort als Wahrscheinlichkeit zu verstehen, mit der der Schwellenwert an diesem Ort überschritten wird [KRIVORUCHKO 2011, S. 293].

Alle konventionellen geostatistischen Modelle nehmen Datenstationarität an. Indikatoren sind aber fast nie stationär und es gibt keine Möglichkeit, sie in stationäre Daten zu transformieren [KRIVORUCHKO 2011, S. 363].

KRIVORUCHKO [2011, S. 363] kommt wegen verschiedener statistischer Probleme zu dem Schluss, dass Indikatorkriging zwar als Werkzeug zur explorativen Datenanalyse geeignet ist, nicht jedoch als prognostisches Mittel zur Entscheidungsfindung

3.5.7 Kriging-Modelle

Matheron hatte seine Entwicklung der Geostatistik kurz nach [...] 1954 begonnen. Das logarithmische Variogramm lieferte ihm eine gute Anpassung an seine ihm damals nur begrenzt verfügbaren Daten [AGTERBERG 2003, S. 3].

Generell sind transitive von intransitiven Modellen zu unterscheiden. Transitive Modelle erfordern Stationarität zweiter Ordnung während für intransitive Modelle lediglich die Erfüllung der intrinsischen Hypothese als gegeben angesehen werden darf [GAU 2010, S. 30].

Häufig muss im Ursprung eines Semivariogramms der Nugget-Effekt einbezogen werden, wenn es scheint, dass das experimentelle Variogramm im Ursprung nicht den Wert Null annimmt. [SCHAFMEISTER 1999, S. 16].

Das Default-Modell um ArcGIS Geostatistical Analyst (Angabe für Version 9.3) ist nach KRIVORUCHKO [2011, S. 302] eine Summe von zwei Modellen, dem Nugget- und dem sphärischen Model (spherical).

Der wichtigste Teil eines Semivariogramm- oder Kovarianzmodells ist aus Sicht der Vorhersage der Bereich kurzer Distanzen. Allgemein führt eine falsche Wahl des Modells zu weniger genauen Werten, als eine falsche Wahl von Parametern des richtigen Modells KRIVORUCHKO [2011, S. 317].

Wenn der physikalische Prozess hinter den Daten bekannt ist, sollte der korrespondierende Kernel verwendet werden. Wenn er unbekannt ist, sollte das flexibelste K- oder J-Bessel-Modell verwendet werden. KRIVORUCHKO [2011, S. 317] empfiehlt dabei, den Lag-Parameter durch Anpassen eines stabilen Modells festzulegen.

3.5.7.1 Modelle mit echtem Sill (transitive Modelle)

Wenn Variogramm-Modelle ab einer bestimmten Reichweite ein bestimmtes (End)Niveau ihrer Semivarianzwerte erreichen, nennt man sie transitiv [DE SMITH ET.AL. 2013, S. 516]. Jenseits dieser Reichweite ist die Korrelation zwischen Datenpaaren exakt Null. [KRIVORUCHKO 2011, S. 303-305].

- Zirkulares Modell (circular): wird durch einen zylindrischen Kernel erzeugt

$$\gamma(h;\theta)=\begin{cases}\frac{2\theta_s}{\pi}\left[\frac{\|h\|}{\theta_r}\sqrt{1-\left(\frac{\|h\|}{\theta_r}\right)^2}+\arcsin\left(\frac{\|h\|}{\theta_r}\right)\right] & \text{für } 0 \le \|h\| \le \theta_r \\ \theta_s & \theta_r < \|h\|\end{cases}$$

- Sphärisches Modell (spherical):

$$\gamma(h;\theta)= \begin{cases} \theta_s[\frac{3}{2}\frac{\|h\|}{\theta_r}-\frac{1}{2}(\frac{\|h\|}{\theta_r})^2] & \text{für } 0 \leq \|h\| \leq \theta_r \\ \theta_s & \theta_r < \|h\| \end{cases}$$

- Tetrasphärisches Modell (tetraspherical)

$$\gamma(h;\theta)= \begin{cases} \frac{2\theta_s}{\pi}(\arcsin(\frac{\|h\|}{\theta_r})+\frac{\|h\|}{\theta_r}\sqrt{1-(\frac{\|h\|}{\theta_r})^2}+\frac{2\|h\|}{3\theta_r}(1-(\frac{\|h\|}{\theta_r})^2)^{\frac{3}{2}}) & \text{für } 0 \leq \|h\| \leq \theta_r \\ \theta_s & \theta_r < \|h\| \end{cases}$$

- Pentasphärisches Modell (pentaspherical)

$$\gamma(h;\theta)= \begin{cases} \theta_s[\frac{15}{8}\frac{\|h\|}{\theta_r}-\frac{5}{4}(\frac{\|h\|}{\theta_r})^3+\frac{3}{8}(\frac{\|h\|}{\theta_r})^5] & \text{für } 0 \leq \|h\| \leq \theta_r \\ \theta_s & \theta_r < \|h\| \end{cases}$$

h = Abstand zwischen Datenpaaren
θ = Parameter :
θ_s = partieller Sill-Parameter (partial sill)
θ_r = Range-Parameter

Pentasphärische und Tetrasphärische Modelle haben in zwei und drei Dimensionen bessere statistische Eigenschaften, als das sphärische Semivariogramm. [KRIVORUCHKO 2011, S. 305].

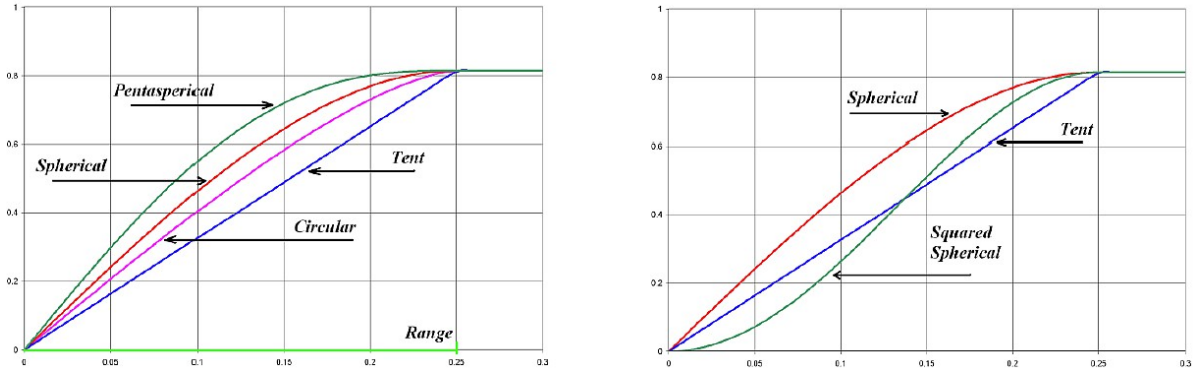


Abb. 10: Graphen einiger Kriging-Modelle im Vergleich

Quelle: KRIVORUCHKO [2011, S. 306]

Das zirkuläre Modell wird mit einem zylindrischen Kernel erzeugt. Mit einem zylindrischen Kernel können aber nicht viele physikalische Prozesse modelliert werden, weil reale Prozesse wahrscheinlich keine so starken Randeffekte aufweisen. Zirkuläre und sphärische Semivariogramme machen demzufolge auch nicht viel physikalischen Sinn [KRIVORUCHKO 2011, S. 304, 305].

Wenn aus irgendeinem Grund ab einer bestimmten Distanz Nullkorrelation erforderlich ist , wird in der geostatistischen Literatur die Verwendung eines quadrierten sphärischen Modells (squared spherical) vorgeschlagen [KRIVORUCHKO 2011, S. 306].

3.5.7.2 Stabile Modelle (intransitive Modelle)

Nicht-transitive Semivariogramme erreichen den Sill im Betrachtungsgebiet nicht. Als „praktische Reichweite“ (practical range) wird eine Distanz definiert, in der 95% des Sill für ein asymptotisches Variogramm-Modell erreicht ist [DE SMITH ET.AL. 2013, S. 516].

Bei den sogenannten „stabilen Modelle“ handelt es sich um potenzierte exponentielle Modelle [KRIVORUCHKO 2011, S. 306].

$$\gamma(h;\theta)=\theta_s\left[1-\exp\left(-3\left(\frac{||h||}{\theta_r}\right)^{\theta_e}\right)\right] \qquad \text{für alle h}$$

θ = Parameter :

θ_s = partieller Sill-Parameter

θ_r = Range-Parameter

θ_e = Potenz-Parameter aus dem Zahlenbereich $0 \leq \theta_e \leq 2$

Wenn der Parameter θ_e gegen Null strebt, tendiert das Modell zum Nugget-Model. Spezielle Fälle des potenzierten exponentiellen Modells sind das gauß'sche Modell (Gaussian) mit $\theta_e = 2$ und das exponentielle Modell mit $\theta_e = 1$ [KRIVORUCHKO 2011, S. 306].

Die meisten Statistiker raten von einer Verwendung des Gauß'schen Modells ab, weil es numerisch instabil und weicher, als jeder reale physikalische Prozess sei

[KRIVORUCHKO 2011, S. 308-309]. KRIVORUCHKO [2011] stuft den Gauß'schen Kernel als den empfindlichsten für eine falsche Festlegung der Parameter ein [S. 317], weist jedoch an anderer Stelle [2011, S. 88, und S. 309] explizit auf ein Fallbeispiel[24] hin, bei dem das Gauß'sche Modell alle anderen übertraf. Im Gegensatz zum Gauß'schen Modell, ist das exponentielle Modell eines der stabilsten Modelle.

Das exponentielle Modell ist auch typisch für k_f-Werte [SCHAFMEISTER 1999, S. 16].

Das Verhalten der Semivariogramm- und Kovarianzmodelle nahe dem Ursprung (oder die „Schärfe ihrer Kernels[25]) ist wichtig, weil sich die vorhergesagte Oberfläche nahe dem Messort wie die Kovarianz auf kurze Distanz verhält [KRIVORUCHKO 2011, S. 307].

Kernel und Kovarianz können über die Fourier-Transformation und deren Inverse ineinander überführt werden [KRIVORUCHKO 2011, S. 302].

3.5.7.3 K-Bessel- oder Mattern-Modelle

Die Klasse der K-Bessel oder auch Mattern-Modelle zur Modellierung von Kovarianz und Semivariogramm gelten als sehr flexibel, weil sie verschiedene Modelle, wie das exponentielle oder auch das Gauß-Modell einschließen können. Im ArcGIS Geostatistical Analyst wird für diese Gruppe an Kriging-Modellen die Bezeichnung „K-Bessel“ verwendet. Die Formel für diese Klasse lautet wie folgt [KRIVORUCHKO 2011, S. 309] :

$$\gamma(h;\theta)=\theta_s\left[1-\frac{\left(\Omega_{\theta_k}\|h\|/\theta_r\right)^{\theta_k}}{2^{\theta_k-1}\Gamma(\theta_k)}K_{\theta_k}\left(\Omega_{\theta_k}\|h\|/\theta_r\right)\right] \quad \text{für alle h}$$

mit

Ω_{θ_k} = numerisch ermittelter Wert, damit $\gamma(\theta_r)=0{,}95\theta_s$ für jeden θ_k

θ_k = Gammafunktion $\Gamma(y)=\int_0^\infty x^{y-1}\exp(-x)dx$

θ_k = modifizierte Bessel-Funktion der zweiten Art der Ordnung θ_k

Wenn der Shape-Parameter θ_k bestimmte Werte annimmt, können aus der Bessel-Funktion einige Kovarianz-Modell abgeleitet werden [KRIVORUCHKO 2011, S. 310]:

---

[24] Beispiel betrifft eine Anpassung an DEM-Daten

[25] Ein Kernel ist eine symetrische Funktion zur Beschreibung der Abnahme von Werten mit der Distanz von einer „Wertequelle“ (z.B. eine Kontamination) [siehe auch [KRIVORUCHKO 2011, S. 145]

θ_k	**Kovarianz-Modell**
$\longrightarrow 0$	Nugget-Modell
1/2	Exponentielles Modell
1	Whittle's Modell $\theta_s h/\theta_r K_1(h/\theta_r)$
3/2	$\theta_s(1+h/\theta_r)\exp(-h/\theta_r)$
5/2	$\theta_s(1+h/\theta_r+(h/\theta_r)^3/3)\exp(-h/\theta_r)$
$\longrightarrow \infty$	Gauß'sches Modell

Tab. 1: Von der K-Bessel-Funktion bei bestimmten Θ_k ableitbare Kovarianzmodelle
Quelle: [KRIVORUCHKO 2011, S. 310]

Der Prozess wird weicher, bzw. glatter, wenn der Parameter θ_k anwächst [KRIVORUCHKO 2011, S. 311].

Weil das K-Bessel-Modell die Glätte bzw. Weichheit der Anpassungsfunktion aus den Daten schätzt, anstatt eine vorbestimmte Form zu verwenden, wie es bei den sphärischen oder exponentiellen Modellen der Fall ist, schlägt KRIVORUCHKO [2011, S. 310-312] vor, dieses Modell immer als endgültiges Semivariogramm-Modell zu verwenden. Zur Berechnung der Parameter (geeignetes Lag u.a. Semivariogramm-Parameter) für dieses Modell kann vorab das stabile Modell (potenziertes exponentielles Modell) verwendet werden, weil der Potenzparameter Θ_e im stabilen Modell mit etwa 2 * Θ_k im K-Bessel-Modell korrespondiert .

Das Rationale quadratische Modell ist vergleichbar mit dem K-Bessel-Modell mit einem Parameter Θ_k kleiner 1 [KRIVORUCHKO 2011, S. 317].

3.5.8 Suchnachbarschaft

Schätzmethoden können in der Regel eine unterschiedliche Anzahl an Datenpunkten zur Schätzung heranziehen. Der am meisten verbreitete Ansatz zur Auswahl von Proben, die zur Schätzung beitragen, ist eine Suchnachbarschaft z.B. innerhalb einen Kreis oder eine Ellipse zu definieren, die um den zu schätzenden Punkt zentriert ist. Bei vorhandener Anisotropie ist eine Ellipse zu wählen, die mit ihrer Hauptachse parallel zur größten Kontinuität orientiert ist [siehe auch Isaaks und Srivastava 1989, S.340].

Mit globaler Suchnachbarschaft können mit allen verfügbaren Messdaten kontinuierliche, glatte, Oberflächen erzeugt werden, was aber meistens nicht sinnvoll ist [KRIVORUCHKO 2011, S. 333].

Ein Kriging-Modell, das nicht alle Beobachtungen zum Schätzen verwendet, wird Nachbarschafts-Kriging genannt. Jeder Interpolator mit lokaler Suchnachbarschaft erzeugt allerdings unglatte Prognosen, es sei denn es wird eine Glättungsoption verwendet [KRIVORUCHKO 2011, S. 330 und 336].

Es gibt keine Theorie für die Festlegung einer „besten Nachbarschaft“. Eine generelle Empfehlung ist, eine Nachbarschaft von mindestens 3, aber weniger als 25 Beobachtungen zu verwenden. Insbesondere 50 Jahre alte meteorologische Studien legen nahe, etwa 8 Beobachtungen als Suchnachbarschaft zu verwenden, um den Vorhersagefehler zu minimieren, auch wenn das der geostatistischen Theorie widerspricht, nach der globales Kriging den geringsten Vorhersage-Standardfehler hat [KRIVORUCHKO 2011, S. 330].

Eine Anzahl an Punkten für die Suchnachbarschaft ist annähernd optimal, wenn der Vorhersage-Standardfehler mit steigender Anzahl an Datenpunkten in der Kriging-Nachbarschaft fast gleich bleibt [KRIVORUCHKO 2011, S. 330].

In der Praxis ist es selten sinnvoll, ein Maximum [der Suchnachbarschaft] größer als 25 zu verwenden und die Größe der Nachbarschaft wird im allgemeinen gleich dem Range des Variogramms genommen [Myers 1991, S. 214].

Nach JERNIGAN [1986, S. 43] kann man bei kleinen Datensätze leicht von allen Daten in der Kovarianz-Matrix Gebrauch machen.

3.5.9 Kriging in kleinen Untersuchungsgebieten mit wenigen Messpunkten

Bei kleinen Datensätzen sind die bei großen Datensätzen verwendeten numerische Schätzverfahren für Semivariogramme oder [räumliche] Kovarianzen schwierig anzuwenden. Nach JERNIGAN [1986, S. 63] legen es kleine Datensätze und ihr Variablenverhalten nahe, robuste Analyseverfahren zu verwenden.

JERNIGAN [1986 S. 35] stellt hierzu exemplarisch das Beispiel einer Deponie mit einem sehr kleinen Datenumfang von 8 Messpunkten (pH-Werte in Grundwassermessstellen) bei sehr ungleicher Verteilung im Raum vor. Das ist ein in der Größenordnung sehr realistischer Fall, wie er von Ingenieurbüros häufig zu bearbeiten ist.

Das Problem der Anwendung von Kriging in einem solchen Fall ist, dass es bei so wenigen Daten für ein gegebene Distanz h höchstens ein (n_h=1) Punktpaar gibt, das einen Abstand von h zueinander hat.

JERNIGAN löste das Problem der geringen Stützpunktzahl für ein Variogramm durch die Bildung überlappender Kollektive von etwa 6 Punktpaaren, wodurch Punktpaare nicht nur einmal zur Mittelwertbildung in Lags herangezogen wurden.

Ein Gauß'sches Variogramm wurde mit nichtlinearen kleinsten Quadraten an die resultierenden 6 Datenpunkte angepasst. Mit solch einem kleinen Datenumfang gibt es durch sie Gruppierung in Lags eine drastische Reduktion der zum Schätzen des Variogramms nutzbaren Daten. Die resultierende Varianz, wie sie durch den Sill gegeben ist, war um mehr als einen Faktor 2 größer, als die einfache Varianz der pH-Messungen.

Neben dem erläuterten Lösungsweg schlägt JERNIGAN [1986, S. 35] zur Variogrammschätzung, die sich bis zu einem gewissen Grad bewährt haben, folgende weitere Ansätze vor:

- Anwendung einer Scatterplot-Glättung der Punktpaar-Varianzen in Bezug auf h

  Dies wurde zusammen mit dem robusten Glätter LOWESS (LOcally WEighted Scatterplott Smoother), ausgeführt. Diese Glättungstechnik wurde von Cleveland (1978) entwickelt. Es ist im wesentlichen eine iterative Regressions-Vorhersage, die am stärksten durch Nachbarpunkte und am wenigsten durch Punkte mit großen Residuen gewichtet werden. Die allgemeinen Ergebnisse sind eine geglättete Schätzung des Variogramms aus dem (h,w) Streudiagramm, das ein lineares Ansteigen für kleine h, gefolgt von einem relativ flachen Range für mittlere h und in einem leicht abwärts gerichteten Verlauf für große h endet.

  Der Grad, bis zu dem diese Glättung das unterlagernde Variogramm schätzen könnte, erfordert weitere Untersuchung. Eine Simulation dieses Schätzverfahrens mit kleinen Stichproben und bekanntem Ursprungs-Variogramm ist erforderlich.

- Die Anwendung anderer robuste Regressionstechniken auf das (h, Punktpaarvarianz)-Streudiagramm

- Schätzungen der Variabilität mit einem Resampling-Plan[26], wie das Jackknife oder Bootstrap. Die resultierende Varianz könnte benutzt werden, um auf dem Variogramm Konfidenzintervalle zu setzen und andererseits den Schätzfehler in die Kriging-Schätzungen einzuführen.

Das von JERNIGAN [1986, S. 66] gewählte Verfahren zur Durchschnittsbildung voneinander getrennter Reichweiten im empirischen Variogramm überschätzt die Variabilität der Originaldaten. Eine entsprechend große Varianz der Kriging-Schätzungen war zu erwarten. An den Grundwassermessstellen war die Varianz erwartungsgemäß gleich Null und stieg dann mit der Entfernung von den Grundwassermessstellen an. Weit weg vom den geclusterten Messstellen war die Varianz in den Ecken der Karte am größten.

JERNIGAN betont, dass für die Schätzvarianz angenommen ist, dass das verwendete Variogramm oder die verwendete Kovarianz exakt sei. Sie trägt der Variation in der Kovarianz-Schätzung keine Rechnung. Diese zusätzliche Varianz würde das Vertrauen in die Schätzungen weiter mindern. Insgesamt kommt JERNIGAN zu dem Schluss, dass die aufgezeigten Ansätze bis dato nicht ausgiebig untersucht worden sind und weiterer intensiver Untersuchungen bedurften, um ihre Brauchbarkeit, Variabilität und Verlässlichkeit zu bestimmen.

---

[26] Resampling = Stichproben-Wiederholung: Bestimmung statistischer Eigenschaften auf Basis einer wiederholten Ziehung von Stichproben, sogenannten Unterstichproben, aus einer Ausgangsstich-probe [Qu.: Wikipedia http://de.wikipedia.org/wiki/Resampling]

4 Untersuchung

4.1 Vorgehensweise

Bei der Konzeption dieser Master Thesis war immer ein praktischer Teil vorgesehen, in dem die gesammelte Theorie angewendet und geprüft werden kann. In Ermangelung geeigneter Daten infolge von Vorbehalten der dafür verantwortlichen Personen entstand die Idee, aus der Not eine Tugend zu machen und einen künstlichen Datensatz zu erzeugen.

Bei der räumlichen Interpolation erzeugt man in der Regel aus Punktdaten Oberflächen oder etwas, dass man sich als Oberfläche vorstellen kann, wie z.B. eine geologische Schichtgrenze oder einen Grundwasserspiegel. Man kennt meist nur die Punktdaten und versucht, auf die Oberfläche, zurückzuschließen, man schätzt sie. Zur Beurteilung, ob diese Schätzung gelungen ist oder nicht, gibt es verschiedene Prüfverfahren und Kriterien. Aber auch das ist immer eine Schätzung.

Hierzu ein Zitat von ISAAKS UND SRIVASTAVA [1989, S. 278]:

> "*In practical situations we never know the true answer or the actual errors before we attempt our estimation*."
>
> "*In praktischen Situationen kennen wir die wahre Antwort und die tatsächlichen Fehler nicht, bevor wir unsere Schätzung versuchen*".

Mit fiktiven Daten kann man die „Wahrheit" kennen, wenn man sie aus einer definierten Geometrie erzeugen kann. Greift man punktuelle Werte an der fiktiven aber bekannten Oberfläche ab, ist das die Simulation einer Probenahme. Und man kann versuchen, mit diesen Werten wieder auf die ursprüngliche Oberfläche zurückzuschließen.

4.2 Daten

Vorgesehen waren zwei Arten von Datensätzen:

- Grundwasserstände an einem fiktiven Standort und
- mit einem Schadstoff belastete Zonen im Untergrund die bei Grundwasserprobenahmen zu erhöhten Konzentrationen führen würden

Die Herstellung einer plausiblen Grundwasseroberfläche hat sich als schwieriger erwiesen, wie erwartet, weil ein konsequentes hydraulisches Gefälle graphisch und von Hand nicht so gezeichnet oder gemalt werden kann. Mit etwas Rechnen, einigen Stützpunkten und einem lokalen Polynom zweiter Ordnung ist das mit Surfer dann aber doch gelungen.

Zwei unterschiedlich große Kontaminationszonen bzw. Schadstofffahnen wurden mit Hilfe eines Bildverarbeitungsprogramms erzeugt und etwas bearbeitet, damit die Oberflächenwerte vom „Schadenszentrum" graduell nach außen zu abnehmen.

Nach einer Umwandlung des Farbflecks in Graustufen und Invertierung, standen zwei verschieden große Schadensbereiche „A“ und „D“ zur Verfügung.

Das Messstellenanzahl ist mit 17 Grundwassermessstellen für Projektstandort schon relativ groß. Die Messstellen sind aber räumlich nicht ideal angeordnet Sie sind weder gleichmäßig noch zufällig verteilt und das entspricht damit der allzu oft angetroffenen Realität.

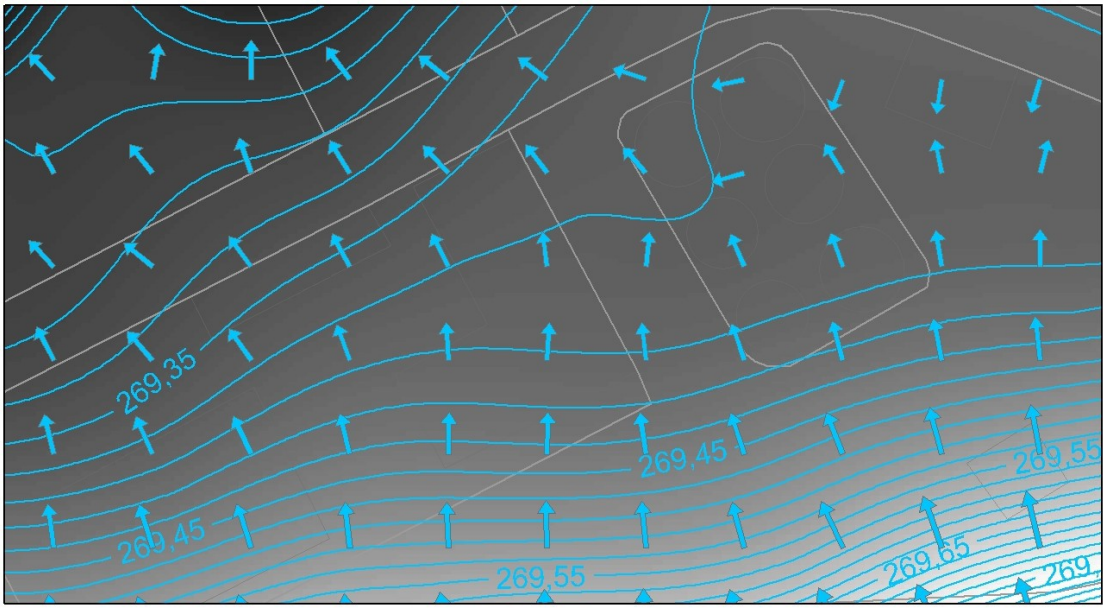


Abb. 11: Das Grundwassermodell des fiktiven Standortes

In Abb. 11 Sind die das Höhenmodell (unterlagernde Grauwerte, höhere Werte sind heller) für die Grundwasseroberfläche dargestellt und mit den in Surfer erzeugten Grundwassergleichen überlagert. Pfeile zeigen die Richtung des potentiellen Grundwasserstroms an. Die grauen Linien zeigen einige Grenzen an, die Flurstücksgrenzen und Straßen sein können.

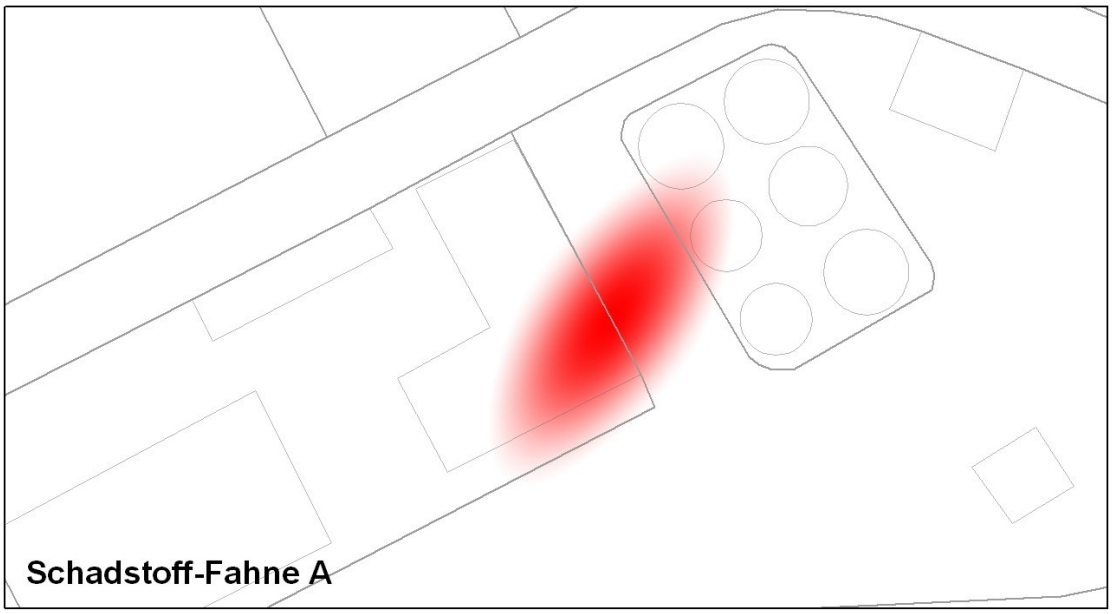


Abb. 12: Schadstoff-Fahne A

Die Abb. 12 und 13 zeigen die beiden Schadstoffahnen im Untersuchungsgebiet. Die Fahne D ist dabei deutlich größer, als Fahne A.

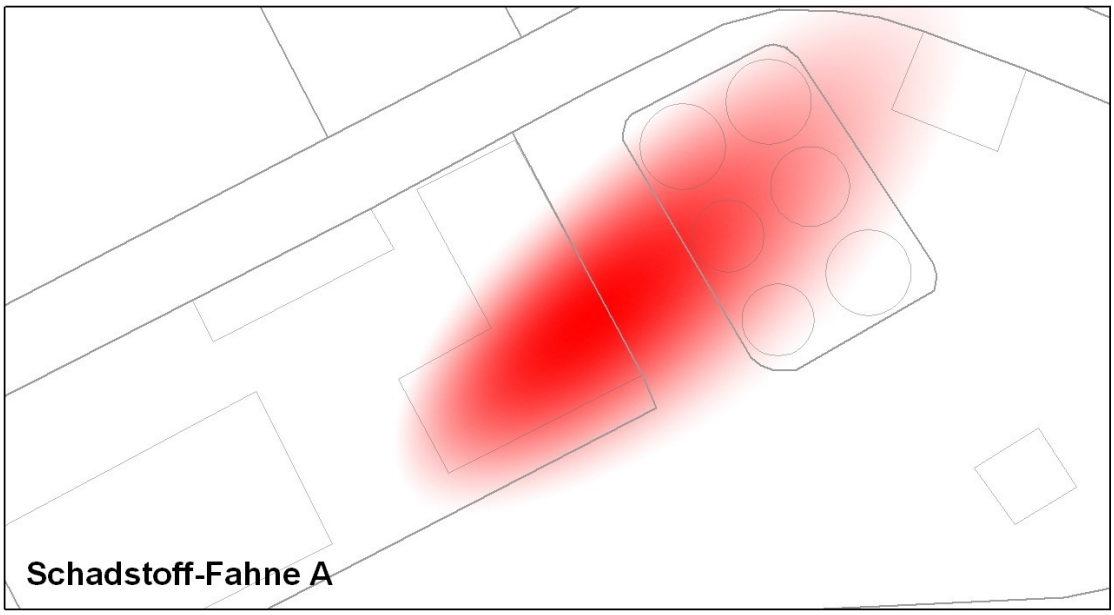


Abb. 13: Schadstoff-Fahne D

4.3 Durchführung und Ergebnisse

Werte der Grundwasseroberfläche wurden an allen 17 verfügbaren Grundwassermessstellen abgegriffen. Da die „Schadstofffahnen“ von begrenzter Ausdehnung sind, treten nicht an allen Grundwassermessstellen Schadstoffe. Die Lage der Messstellen sind in im Anhang

In Anh. 14 ist ein QQ-Plot der Grundwasserdaten dargestellt, der zeigt, dass die Daten nicht normal verteilt sind.

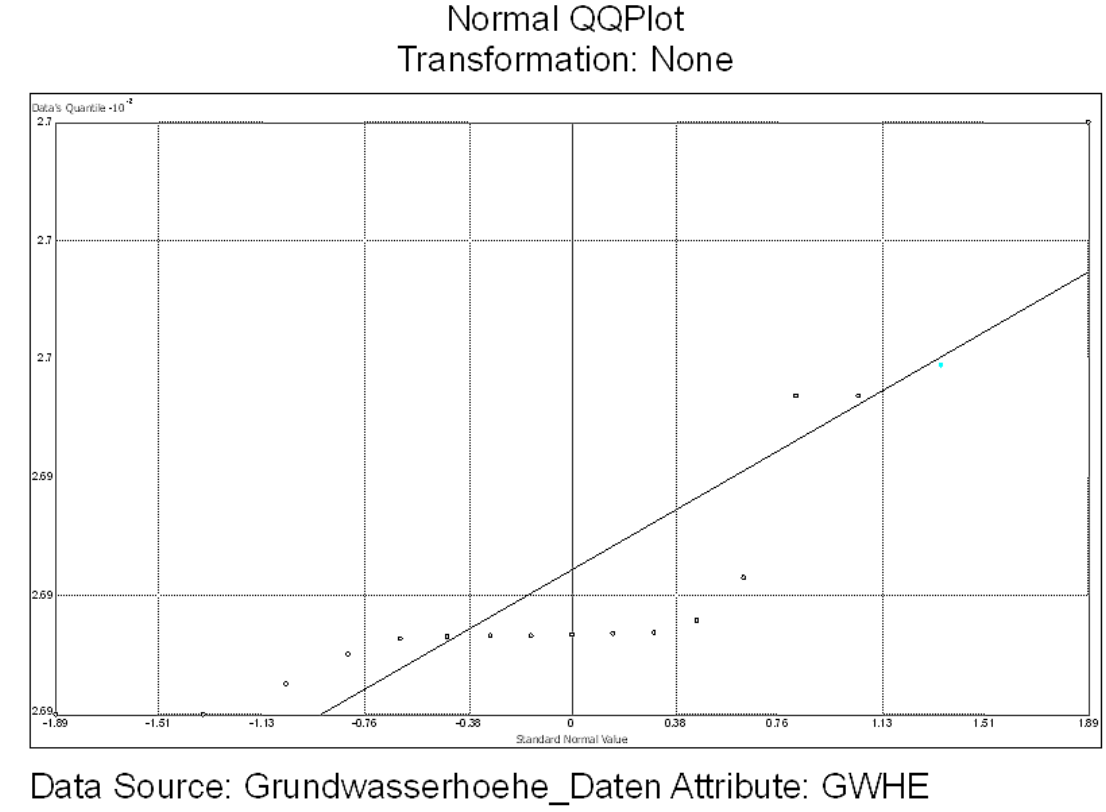


Abb. 14: QQ-Plot der Grundwasserdaten (mit Geostatistical Analyst erstellt)

Die Grundwasseroberfläche zeigt die Eigenheit, dass sie teilweise ein hohes Gefälle aufweist (z.B n einer künstlichen Auffüllung), teilweise aber auch fast eben ist.

Mit Surfer 12 die Interpolationsverfahren IDW, Shephard's Modified, Kriging mit einem linearem Modell, lokales Polynom zweiter Ordnung, Natural Neighbor und Radiale Basisfunktionen durchgeführt. Anisotropie wurde in keinem Fall verwendet und die Defaulteinstellungen von Surfer belassen.

In Anhang 1 Blatt 1 ist ganz oben das „Original“, also das Modell abgebildet, auf den 3 Blättern des Anhangs sind die Ergebnisse enthalten. Das Modified Shephard's (ähnlich wie IDW, neigt aber kaum zu den bekannten „Bullaugen“-Strukturen), das lokale Polynom zweiter Ordnung und die Radialen Basisfunktionen ergaben die dem Original am nächsten kommenden Ergebnisse im Fall der Grundwasseroberfläche.

Für die kleinere Schadstofffahne A (siehe Anhangseiten 2) zeigt IDW das bekannte Bullaugenmuster, das bei Angabe einer Anisotropie dem Original näher kommt. Insgesamt hatten alle Verfahren Schwierigkeiten das Original auf Anhieb gut abzubilden. Die größten Probleme treten bei Radialen Basisfunktionen auf, die negative Werte ermitteln und über das Untersuchungsgebiet hinaus nicht nachvollziehbare hohe Werten nachzeichnen, die in diesem Fall nicht sein können.

Kriging ergab ein passables Ergebnis, das Shephard's Modified überzeichnet die Kontaminationsfläche deutlich.

Die größere Kontaminationsfläche „D“ bereitete den getesteten Interpolationsverfahren ebenfalls Probleme. Die beste Wiedergabe scheint auf den ersten Blick den radialen Basisfunktionen zu gelingen. Jedoch täuscht der erste Eindruck, weil alle Radiale Basisfunktionen auch negative Werte berechnet haben (was in den Isolinien zum Ausdruck kommt). Das gilt auch für die Shephards Modified Methode. Die negativen Flächenwerte konnten durch Eingabe eines Schwellenwertes von Null ausgeblendet werden. Nach Angaben im Benutzerhandbuch von Surfer sollte man negative Werte in den erzeugten Grids nachträglich korrigieren können. Das bei dem hier durchgeführten Tests aber nicht gelungen.

Versuche der beschriebenen Art können zur Verbesserung von Interpolationen führen, wenn Anwender dadurch lernen, die Parameter der Verfahren auf geeignete Weise einzustellen.

4.4 Fazit und Ausblick

Im Rahmen dieser Master Thesis konnten nur relativ wenige Veröffentlichungen gefunden werden, die sich mit dem Problem der Interpolation bei nur geringer Messpunktdichte befassen. Dies wird von einigen Autoren, wie JERNIGAN [1986] oder SCHAFMEISTER [1999] auch bedauert. Beide sehen Bedarf zu Interpolation von Werten auch bei kleinen Projekten mit geringem Stichprobenumfang.

Die Anwendung von geostatistischen Verfahren ist durchweg recht anspruchsvoll und erfordert ein gutes Verständnis für die Annahmen, die mit den verschiedenen Methoden verbunden sind. Dies steht in einem gewissen Gegensatz zur Praxis. Interessant sind in diesem Zusammenhang auch Untersuchungen der US EPA über die Varianz der Geostatistiker bei der Bearbeitung der gleichen Datensätzem in vergleichenden Tests, wie von ENGLUND [1990] berichtet wird.

So sind die Bedingung der Stationarität bei Grundwassererkundungen grundsätzlich nicht oder höchstens sehr selten gegeben. Der Grundwasserspiegel ist eben geneigt und hat deshalb keinen konstanten Mittelwert.

Ähnliches gilt für lokale Kontaminationen: es liegt dann höchstens ganz lokal Stationarität vor und auch dann wird es schwierig eine Kontinuität anzunehmen, die eine geostatistische Interpolation erlauben würde. Deterministischen Verfahren sind deshalb durchaus interessant für die räumliche Interpolation.

Die verfügbaren geostatistischen Verfahren sind für geringe Stichprobengrößen wenig geeignet, Dies umso mehr, wenn die Werte nicht normalverteilt sind und die Daten strenggenommen eine aufwändigere Aufbereitung erfordern.

Die eingangs formulierte Nullhypothese, dass es in kleinen Untersuchungsgebieten nicht möglich oder sinnvoll wäre, räumliche Interpolationen vorzunehmen, kann nicht bestätigt werden. Autoren wie SCHAFMEISTER [1999] und JERNIGAN [1986] sehen Notwendigkeit und Bedarf dafür. Es ist Forschungsbedarf in dieser Hinsicht zu sehen, möglicherweise auch in Richtung einer sich gegenseitig ergänzenden Kombination von deterministischen und probabilistischen Schätzverfahren.

5 Literaturverzeichnis

Agterberg, F. P. [2003]: Georges Matheron – Founder of Spatial Statistics. In: Proceedings of the International Association for Mathematical Geology, 2003. Online-Ressource, Download am 28.05.2014 9:00 Uhr von Adresse http://www.geostatscam.com/Adobe/G_Matheron.pdf

ASTM International [2010]: ASTM D5922 – 96 (Reapproved 2010). Standard Guide for Analysis of Spatial Variation in Geostatistical Site Investigations. Active Standard. Als kostenpflichtiges PDF-Dokument über die Webseite http://www.astm.org/Standards/D5922.htm bezogen

ASTM International [2010]: D5923 − 96 (Reapproved 2010). Standard Guide for Selection of Kriging Methods in Geostatistical Site Investigations. Active Standard. Als kostenpflichtiges PDF-Dokument über die Webseite http://www.astm.org/Standards/D5923.htm bezogen

Bivand, R. S., Pebesma, E. J. and Gómez-Rubio, V. [2008]: Applied Spatial Data Analysis with R. Springer, New York 2008.

Cui, H., Stein, A. and Myers, D. E. [2006]: Extension of spatial information, bayesian kriging and updating of prior variogram parameters. In: Environmetrics. Vol. 6, Issue 4, July/August 1995, Seite 373–384.

de Smith, M. J., Goodchild, M. F. and Longley, P.A. [2013]: Geospatial Analysis. A Comprehensive Guide to Principles, Techniques and Software Tools. 4. Aufl.. The Winchelsea Press, Winchelsea, United Kingdom, 2013. Online-Ressource, Einsichtnahme am 26.05.2014 unter Adresse http://www.spatial-analysisonline.com/HTML/index.html (HTML-Version) bzw. http://www.drumlinsecurity.com/online/ga4/ (Druckversion)

Dutter, R. [1985]: Geostatistik. Eine Einführung mit Anwendungen. Mathematische Methoden in der Technik. Bd. 2. B.G. Teubner, Stuttgart 1985.

Gau, C. [2010]: Geostatistik in der Baugrundmodellierung. Die Bedeutung des Anwenders im Modellierungsprozess. Dissertation an der Technischen Universität Berlin, 2010. Vieweg + Teubner, Wiesbaden 2010.

Englund, E. J. [1990]: Variance of Geostatisticians. Manuskript zu einer Veröffentlichung in Mathematical Geology 22:4 (1990) 417-455. Online-Ressource, Download am 30.05.2014 12:20 Uhr von Adresse http://www.epa.gov/esd/cmb/research/papers/ee102.pdf

Golden Software inc. [2014]: Surfer 12 User's Guide. Golden, Colorado, USA, 2014.

Isaaks, E. H. and Srivastava, R. M. [1989]: An Introduction to Applied Geostatistics. Oxford University Press, New York.

Jernigan, R. W. [1986]: A Primer on Kriging. The Statistical Policy Branch Office of Standards and Regulations, February 1986. U.S. Environmental Protection Agency, Washington, D. C., USA. Online-Ressource, Download am 05.06.14 14:30 Uhr von Adresse http://nepis.epa.gov/Adobe/PDF/9100UJS1.PDF

Krivoruchko, K. [2011]: Spatial Statistical Data Analysis for GIS Users. ESRI Press. Redlands, California, USA.

Matheron, G. und Formery, Ph. [1965]: La théorie des variables régionalisées et son application à l'estimation des gisements miniers. Rapport, École des Mines de Paris, 1965. Online-Ressource, Download am 14.06.2014 13:40 Uhr von Adresse http://www.cg.ensmp.fr/bibliotheque/public/MATHERON_Rapport_00080.pdf

Matheron, G. [1964]: Les principes de la géostatistique. Rapport, École des Mines de Paris, 1964. Online-Ressource, Download am 14.06.2014 13:15 Uhr von Adresse http://www.cg.ensmp.fr/bibliotheque/public/MATHERON_Rapport_00065.pdf

Matheron, G. [1967a]: Présentation des variables régionalisées. In: Revue de la société de statistique de Paris, Juin 1967, p. 263-275. Online-Ressource, Download am 14.06.2014 14:45 Uhr von Adresse http://www.cg.ensmp.fr/bibliotheque/public/MATHERON_Publication_00106.pdf

Myers, D.E. [1989]: To Be or Not to Be... Stationary: That is the Question. In: Mathematical Geology 21, 1989. S. 347-362. Online-Ressource, Download am 13.06.2014 18:15 Uhr von Adresse http://www.u.arizona.edu/~donaldm/homepage/my_papers/MathGeology_21_1989_347-363.pdf

Myers,D.E. [1991]: Interpolation and Estimation with Spatially Located Data. In: Chemometrics and Intelligent Laboratory Systems 11, 1991. S. 209-228. Online-Ressource, Download am 13.06.2014 8:55 Uhr von Adresse http://www.u.arizona.edu/~donaldm/homepage/my_papers/ChemIntelligentLabSystems_11_1991_209-228.pdf

Rivoirard, J. [2005]: Concepts and Methods of Geostatistics. In: Bilodeau, M., Meyer, F. and Schmitt, M. (Hrsg): Space, Structure and Randomness. Contributions in Honor of Georges Matheron in the Fields of Geostatistics, Random Sets and Mathematical Morphology. Springer, New York, 2005. S. 17 - 38. Online-Ressource, Download am 15.06.2014 23:00 Uhr von Adresse http://www.springer.com/cda/content/document/cda_downloaddocument/9780387203317-c1.pdf

Schaefer, K. [2013]: Einführung in geostatistische Methoden der Datenauswertung. MUC 2.3 und MC 2.1.1 Praktikum Umweltanalytik II. IAAC, Lehrbereich Umweltanalytik, Universität Jena. Online-Ressource, Download am 01.06.2014 von Adresse http://www.lbua.uni-

jena.de/umweltanalytik_multimedia/dokumente/Sommersemester+2013/Praktikum/Geostatistik-p-236.pdf

Schafmeister, M.-Th. [1999]: Geostatistik für die hydrogeologische Praxis. Springer, Berlin, Heidelberg, 1999.

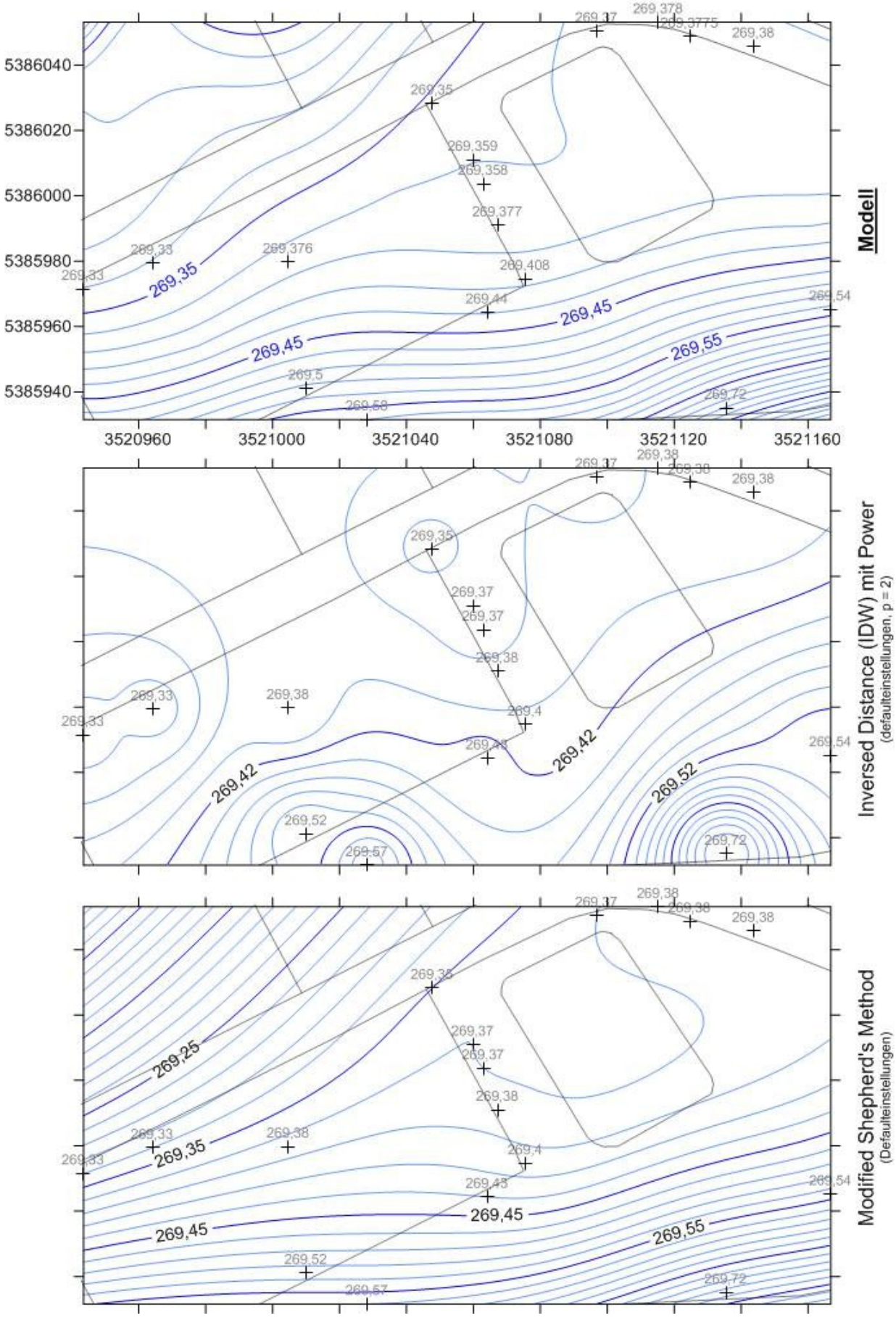
StackExchange [2013]: Minimum number of samples for kriging interpolation. [Frage von user marcin am 13.02.2013 von whuber beantwortet]. Internet-Ressource, Abruf am 28.05.2014, 23:30 Uhr von Adresse http://gis.stackexchange.com/questions/50584/minimum-number-of-samples-for-kriging-interpolation

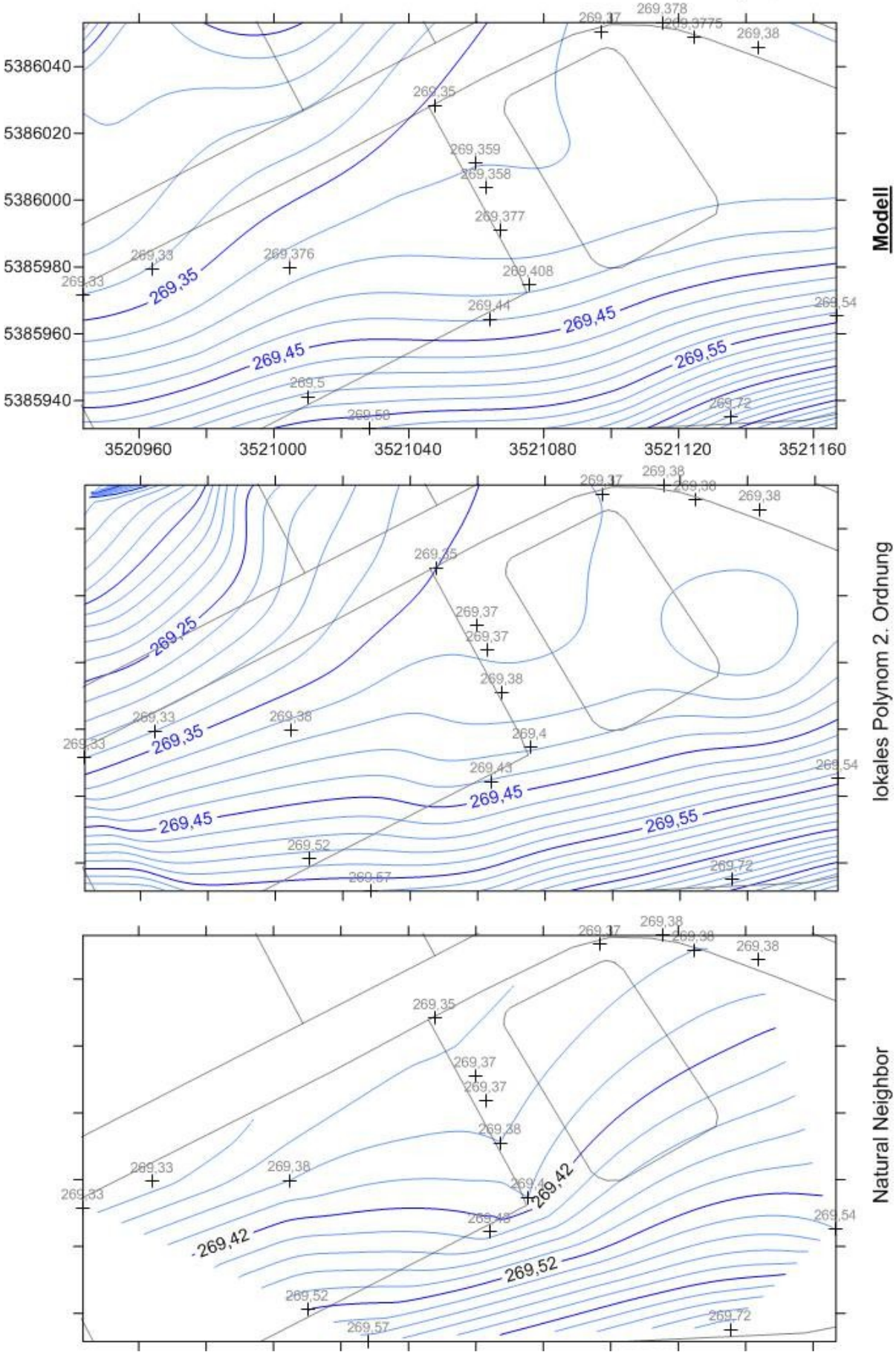
Voßmerbäumer, H. [1996]: Geologie Wörterbuch, Französisch-Deutsch, Deutsch-Französisch. Schweizerbart, Stuttgart, 1996.

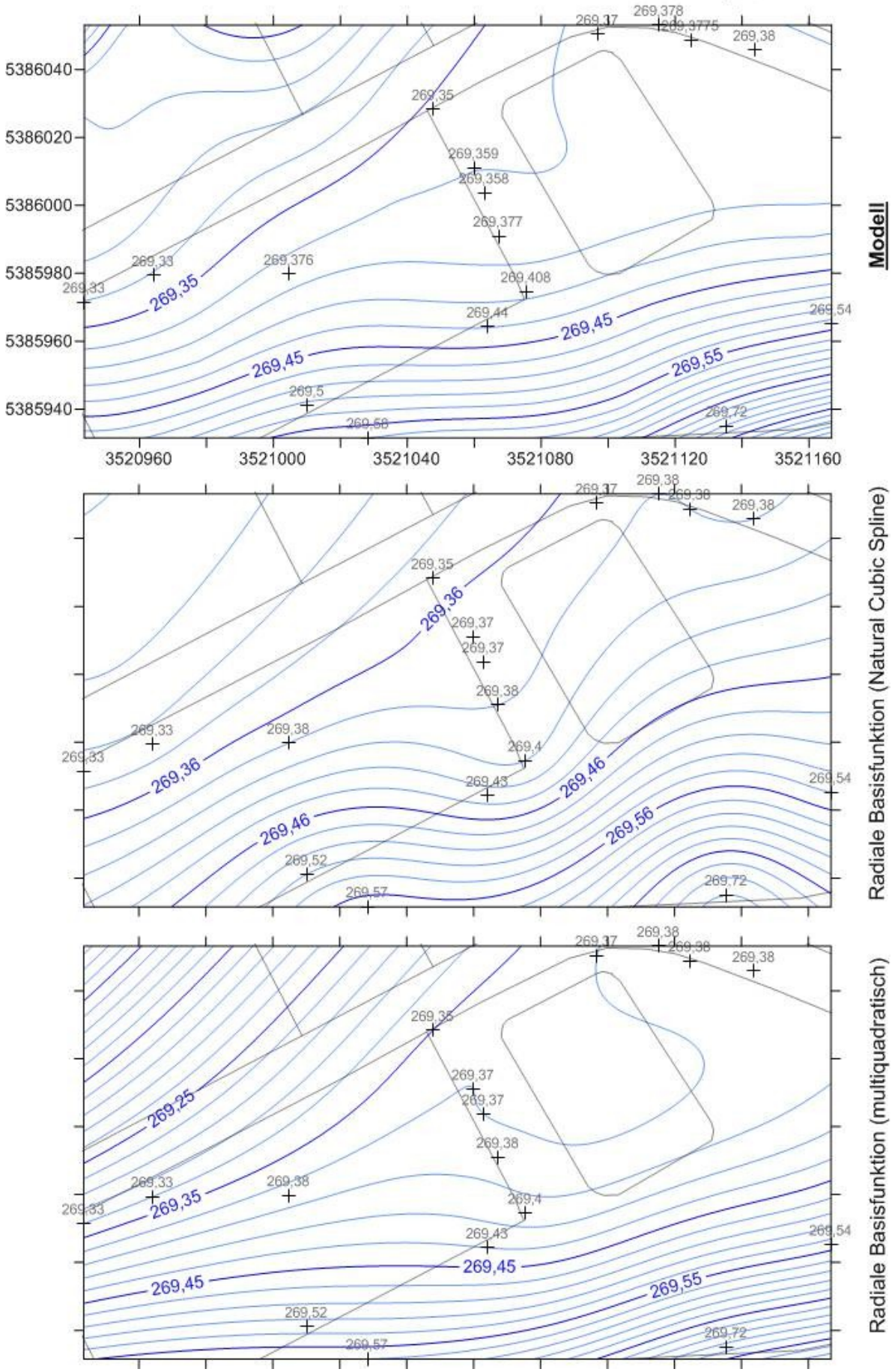
Weber, D. D. , Englund, E. J. [1993]: Evaluation and Comparison of Spatial Interpolators II. Entwurf zu einem Beitrag in Mathematical Geology, Stand 05.08.1993. Online-Ressource, Download am 30.05.2014 12:30 Uhr von Adresse http://www.epa.gov/esd/cmb/research/papers/ee106.pdf

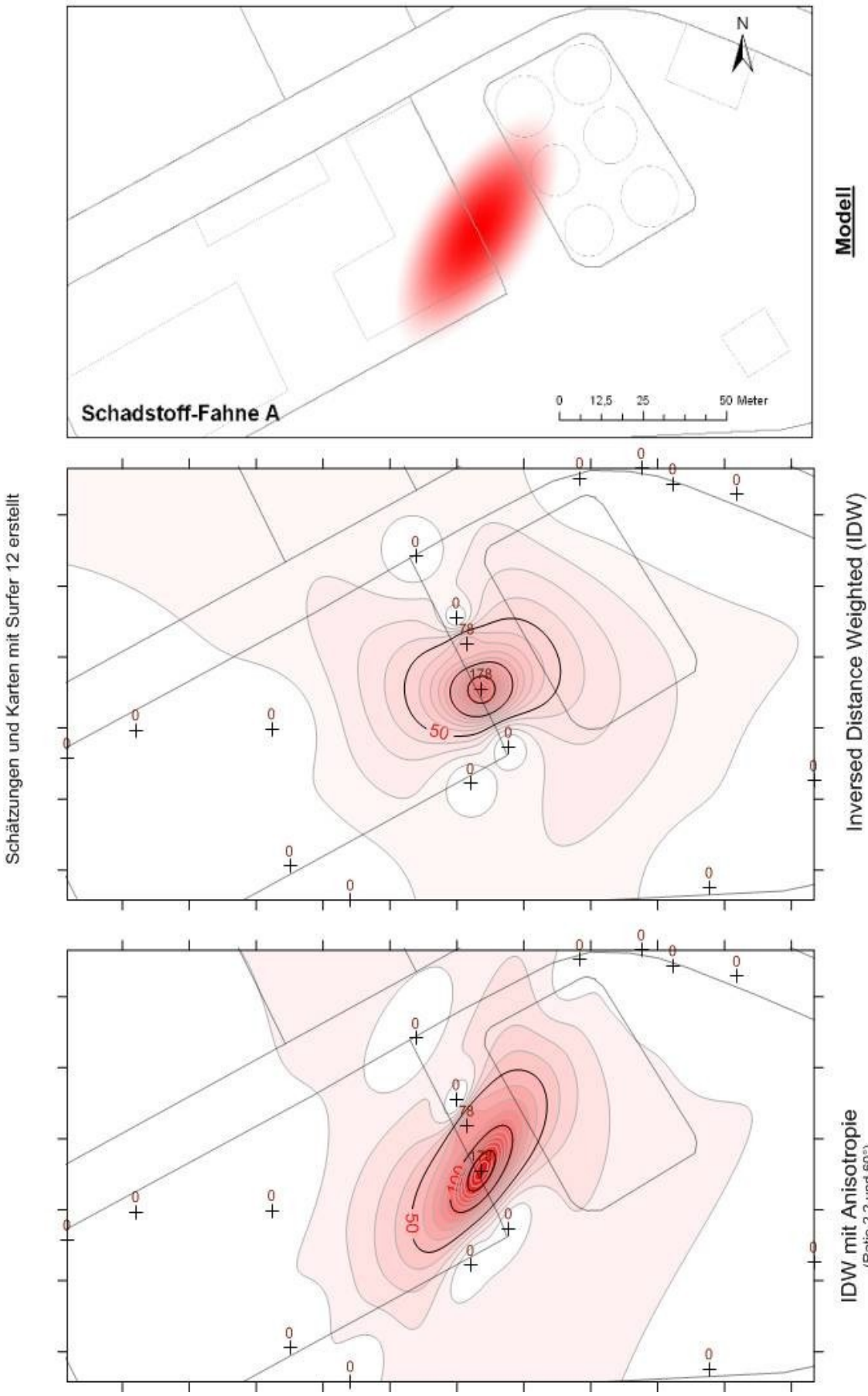
Anhang 1 - 3 : Mit Surfer interpolierte Grundwassergleichenpläne und Schadstoffverteilungen

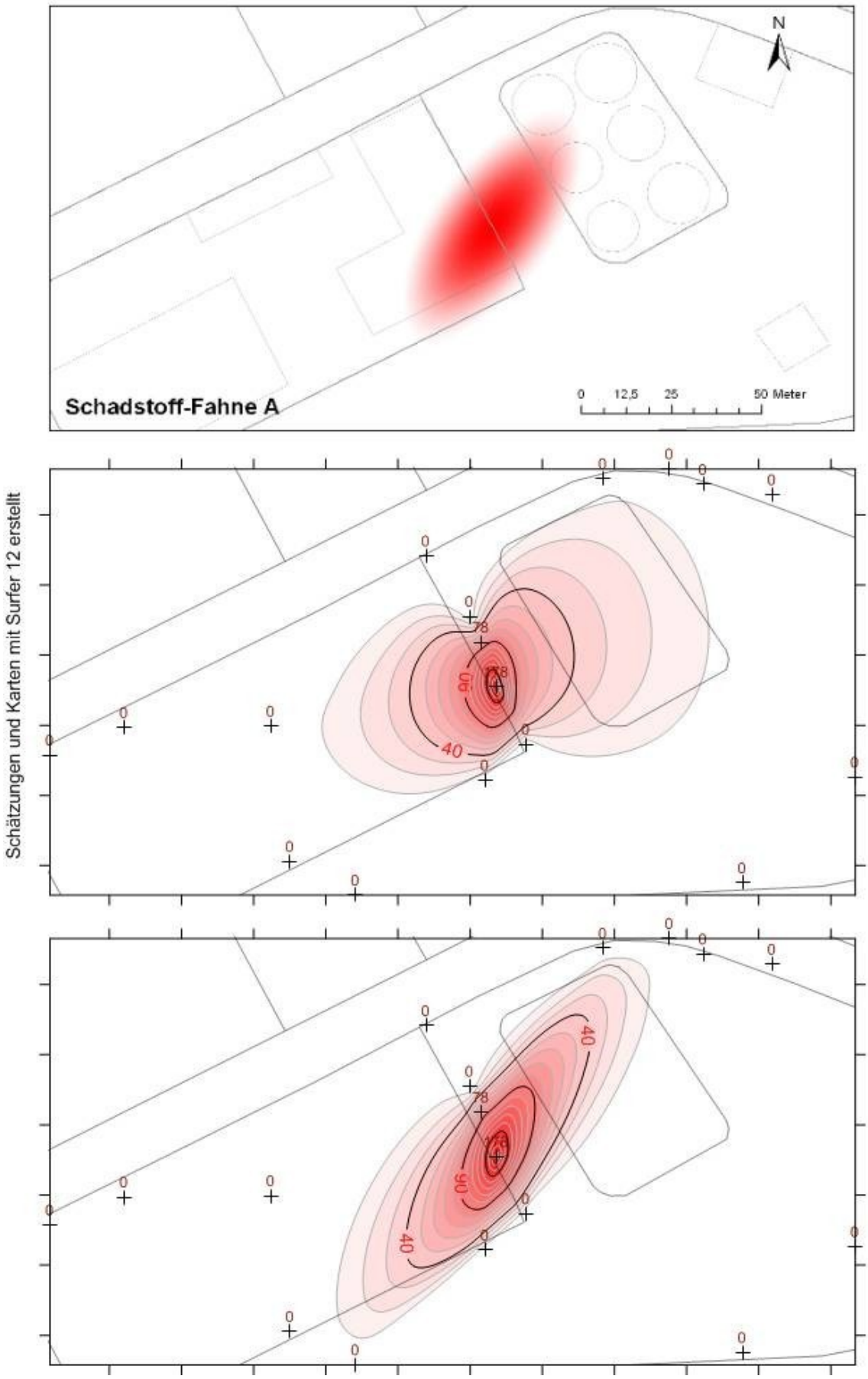
1 Grundwasserdaten
2 Schadstofffahne A
3 Schadstofffahne D


Modell
5386040
5386020
5386000
5385980
5385960
5385940
3520960
3521000
3521040
3521080
3521120
3521160
269,35
269,45
269,45
269,55
269,42
269,42
269,52
269,25
269,35
269,45
269,55
Inversed Distance (IDW) mit Power
(defaulteinstellungen, p = 2)
Modified Shepherd's Method
(Defaulteinstellungen)

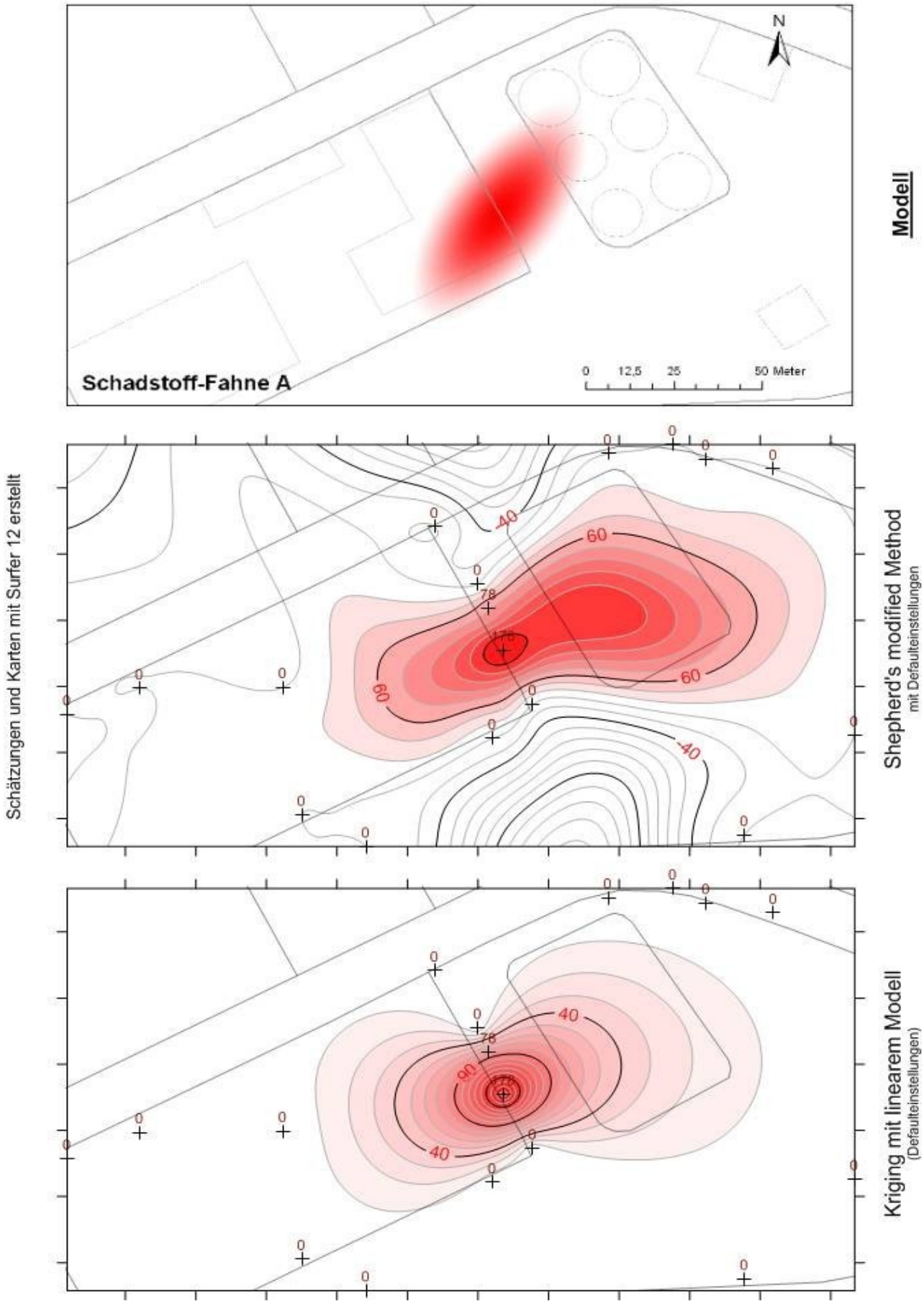

Modell
lokales Polynom 2. Ordnung
Natural Neighbor
5386040
5386020
5386000
5385980
5385960
5385940
3520960
3521000
3521040
3521080
3521120
3521160
269,35
269,45
269,55
269,25
269,42
269,52

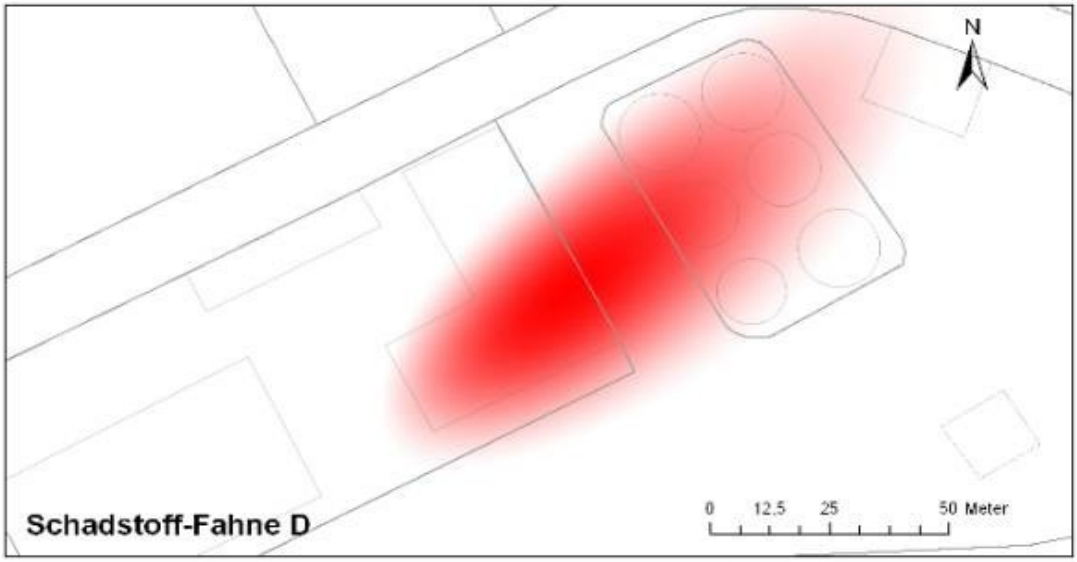

Modell
Radiale Basisfunktion (Natural Cubic Spline)
Radiale Basisfunktion (multiquadratisch)
5386040
5386020
5386000
5385980
5385960
5385940
3520960
3521000
3521040
3521080
3521120
3521160
269,35
269,45
269,55
269,36
269,46
269,56
269,25
269,45
269,35
269,45


Modell
N
Schadstoff-Fahne A
0 12,5 25 50 Meter
Inversed Distance Weighted (IDW)
mit Defaulteinstellungen: p = 2
IDW mit Anisotropie
(Ratio 2.2 und 60°)
Schätzungen und Karten mit Surfer 12 erstellt

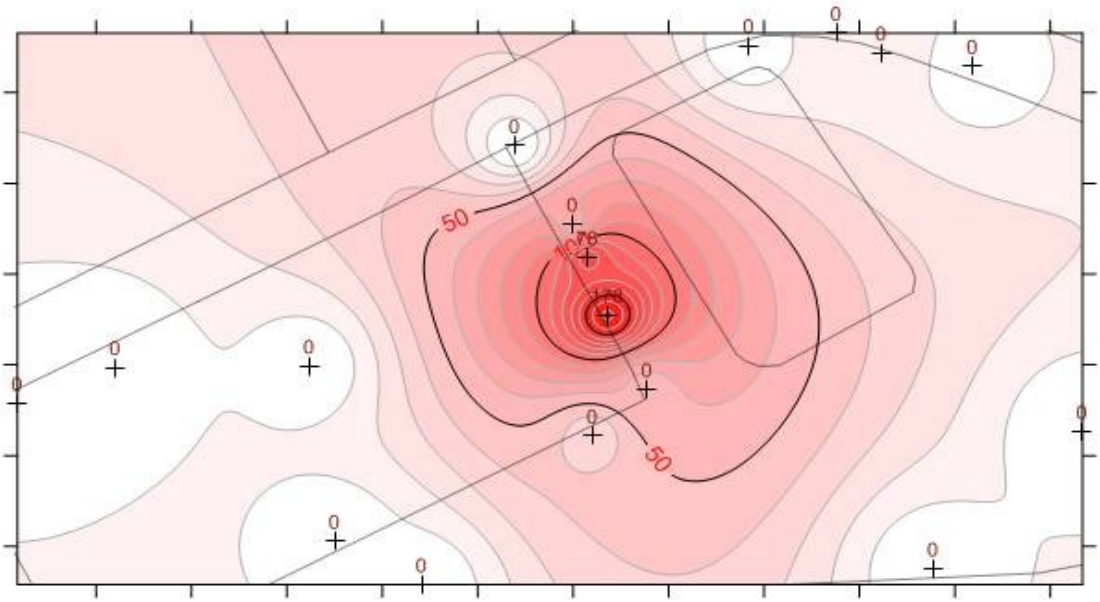

N
Schadstoff-Fahne A
0 12,5 25 50 Meter
Modell
Schätzungen und Karten mit Surfer 12 erstellt
Natural Neighbor
mit Defaulteinstellungen
Natural Neighbor
(Anisotropie: Ratio 2.2; Winkel 60°))

Schätzungen und Karten mit Surfer 12 erstellt
Radiale Basisfunktion
(Multiquadratisch ohne Anisotropie, Default)
Radiale Basisfunktion
(Multiquadratisch mit Anisotropie: Ration 2.2, Winkel 60°)
Radiale Basisfunktion
(Natural Cubic Spline mit Anisotropie: Ration 2.2, Winkel 60°)

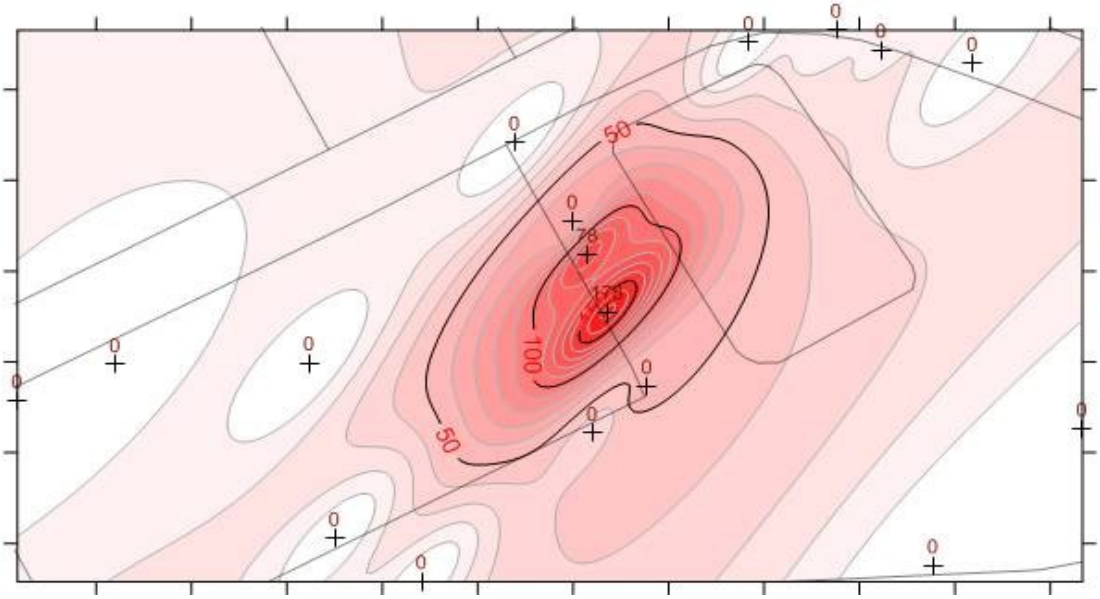

Modell
N
Schadstoff-Fahne A
0 12,5 25 50 Meter
Schätzungen und Karten mit Surfer 12 erstellt
Shepherd's modified Method
mit Defaulteinstellungen
Kriging mit linearem Modell
(Defaulteinstellungen)



Modell

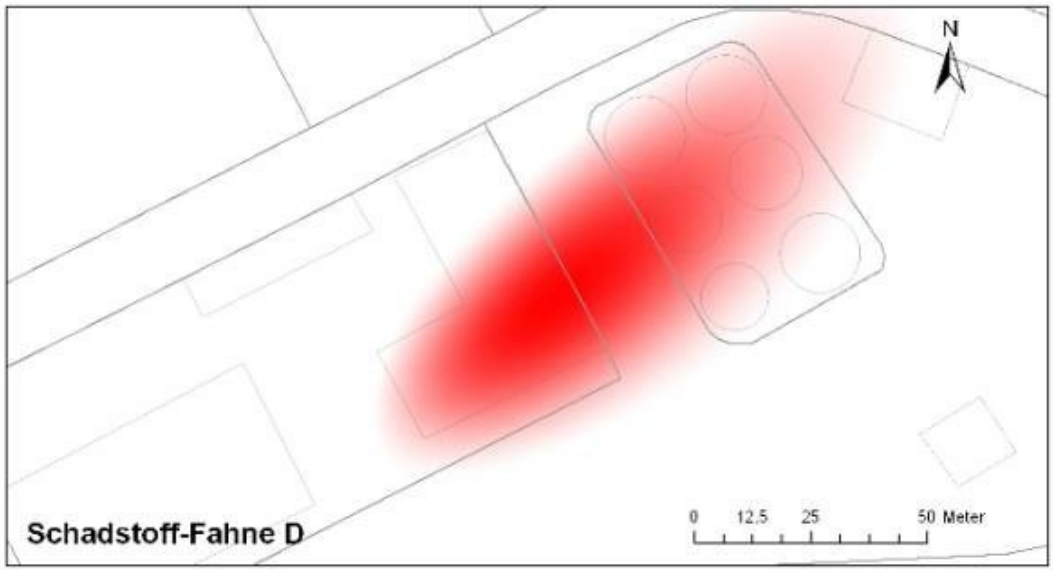


Inverse Distance Weighted
(Defaulteinstellungen mit p=2)

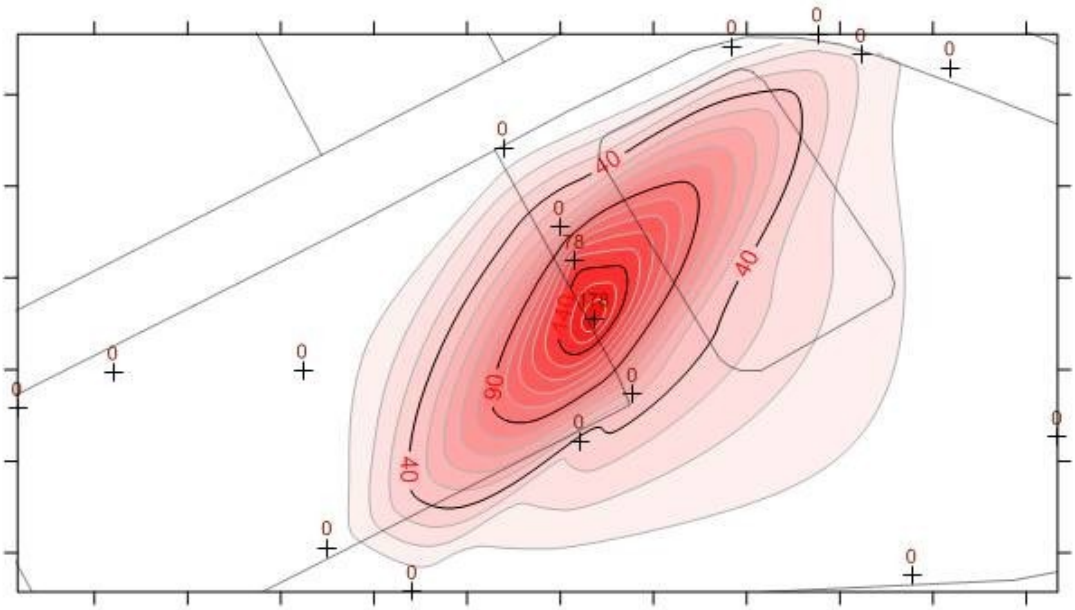


Inverse Distance Weighted
(mit Anisotropie, Ratio 2.9, 50°; mit p=2)

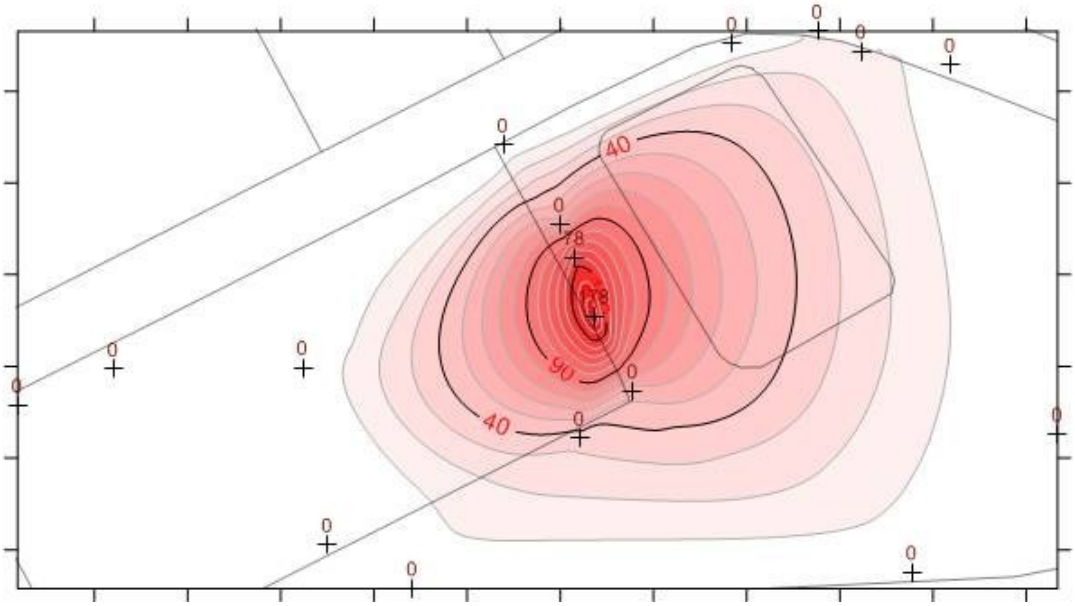
Schätzungen und Karten mit Surfer 12 erstellt



Modell



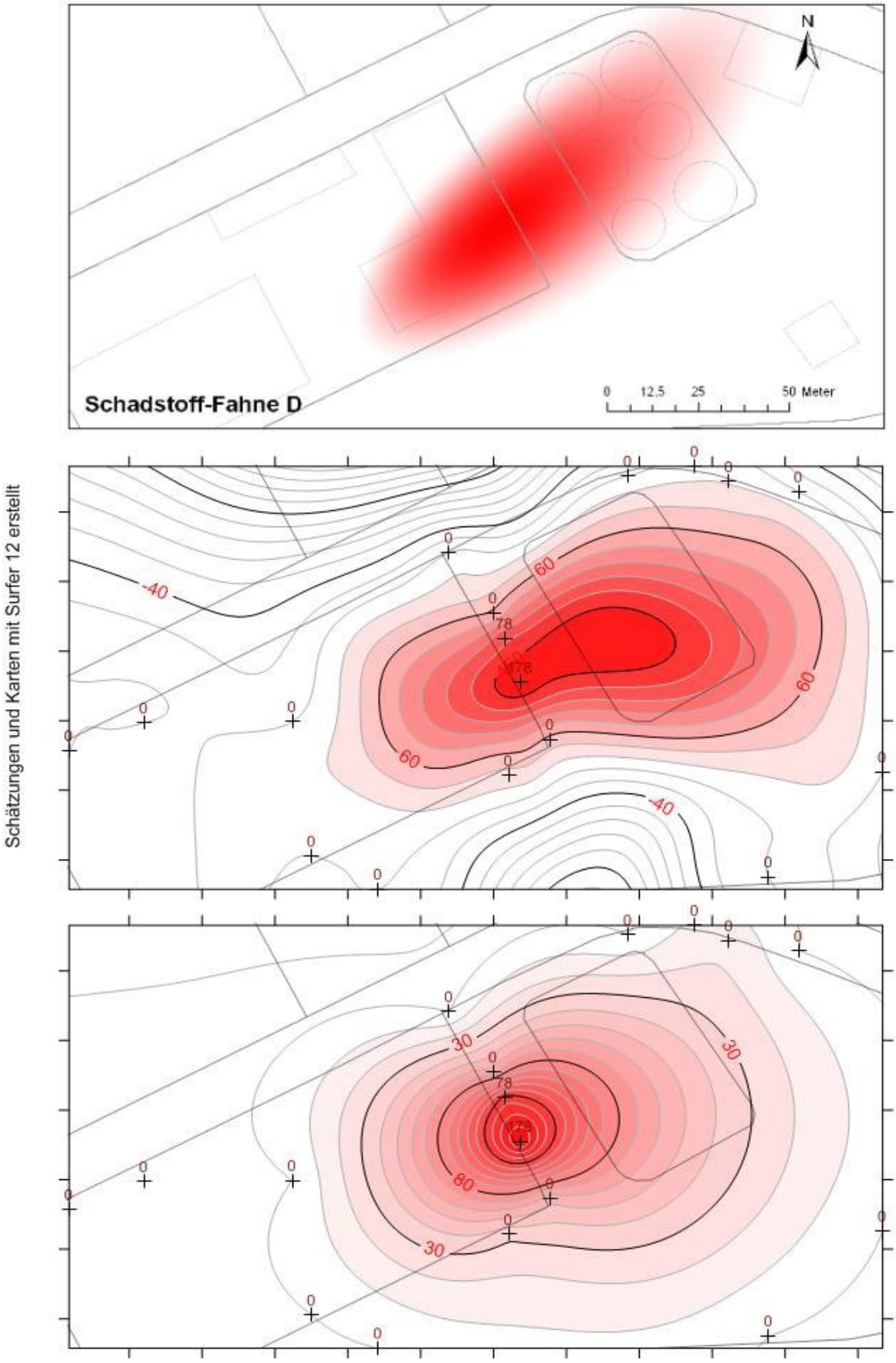
Natural Neighbor
(Defaulteinstellungen)



Natural Neighbor
(mit Anisotropie, Ratio 2.9, 50°)

Schätzungen und Karten mit Surfer 12 erstellt

Radiale Basisfunktion
(Multiquadratisch ohne Anisotropie, Default)
Radiale Basisfunktion
(Multiquadratisch mit Anisotropie: Ration 2.9, Winkel 50°)
Radiale Basisfunktion
(Natural Cubic Spline mit Anisotropie: Ration 2.9, Winkel 50°)
Schätzungen und Karten mit Surfer 12 erstellt


Modell
N
Schadstoff-Fahne D
0 12.5 25 50 Meter
Schätzungen und Karten mit Surfer 12 erstellt
Shepherd's Modified Method
(Defaulteinstellungen)
Kriging mit linearem Modell
(Defaulteinstellungen)