Geographic Information Technology Training Alliance (GITTA) presents:

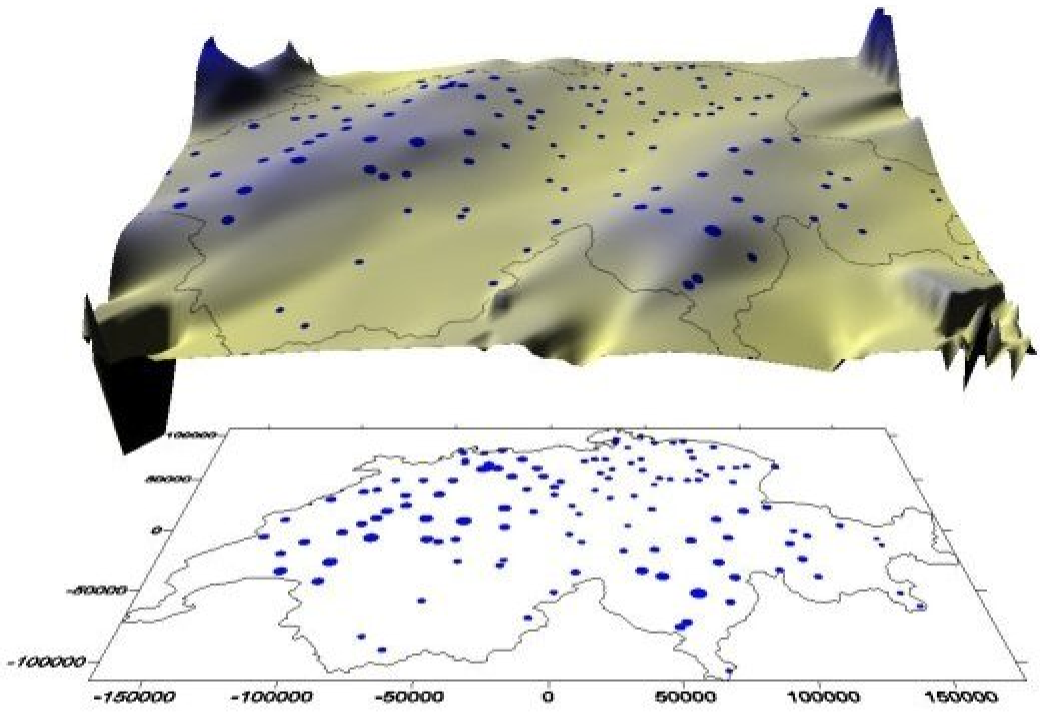
Kontinuierliche räumliche Variablen

Verantwortliche Personen: Eric Lorup, Helmut Flitter

Inhaltsverzeichnis

1. Kontinuierliche räumliche Variablen 2
1.1. Räumliche Stichprobenziehung 4
1.1.1. Eigenschaften 4
1.1.2. Typologie 5
1.2. Analyse räumlicher Abhängigkeit 9
1.2.1. Variogramm 9
1.2.2. Moving Windows 12
1.2.3. Semivariogramm-Parameter richtig zuordnen 13
1.3. Räumliche Interpolation 14
1.3.1. Typologie 16
1.3.2. Distanz-basierte Interpolation 18
1.3.3. Geostatistische Interpolation 21
1.4. Zusammenfassung 27
1.5. Bibliographie 28

1. Kontinuierliche räumliche Variablen



Niederschlagsoberfläche der Schweiz (oben), Karte der Messstationen (unten)

Die Abbildung zeigt Ihnen eine Niederschlagsoberfläche der Schweiz: Die blauen Punkte sind die Positionen von Messstationen, ihre Grösse entspricht der Niederschlagsmenge. Die unterschiedlichen Höhen der Oberfläche sowie ihre Farbgebung stehen ebenfalls in Zusammenhang mit der Niederschlagsmenge. Folgende Fragen werden in dieser Lektion behandelt:

- Wie kann aus den ca. 100 Messpunkten solch eine kontinuierliche Oberfläche erstellt werden?
- Welche Werkzeuge helfen uns dabei?
- Welches Wissen ist nötig und welche Methoden existieren dazu?

Was verstehen wir eigentlich unter „Kontinuierlichen Räumlichen Variablen"?
In unserem Eingangsbeispiel ist die Variable der Niederschlag. Führen wir ein Gedankenexperiment durch: Stellen Sie sich vor, Sie könnten den Niederschlag entlang einer Messstrecke auf jeder beliebigen Position messen. Sie hätten also eine räumlich-kontinuierliche Messung. Benachbarte Niederschlagswerte werden entweder identisch sein oder geringfügig variieren, je nach Ihrer Definition der „Nachbarschaft". Sie dürfen zurecht einwenden, dass ein Niederschlag oft sehr genau definierte Begrenzungen zeigt. Stimmt! Das „Kontinuum Niederschlag" ist kein mathematisch perfektes. Eventuell gefällt Ihnen die Meereshöhe besser als ein weiteres Beispiel einer kontinuierlichen räumlichen Variablen. Aber praktisch jedes natürliche räumlich-kontinuierliche Phänomen unterliegt gewissen zufälligen Schwankungen. Daher lässt sich auch kaum eines mathematisch perfekt beschreiben (z. B. eine Funktion, die den Anstieg eines Hanges exakt beschriebe, oder die Verteilung unterschiedlicher Boden pH-Werte oder Niederschläge usw.).

Zufalls-Funktion

Aus der Statistik kennen wir deterministische (exakt vorhersehbare und mathematisch beschreibbare) und stochastische (rein zufällige, nicht vorhersehbare) Phänomene. Ein deterministisches Phänomen wäre etwa der Fall eines Gegenstandes: Zu jedem Zeitpunkt können wir exakt vorherberechnen, wo sich der Gegenstand entlang der Falllinie befinden wird. Im Gegensatz dazu denken wir an den Würfelwurf als rein stochastisches Phänomen. In der räumlichen Analyse finden wir Phänomene, die zwischen deterministischen und stochastischen anzusiedeln sind. Sie werden als Zufallsfunktion bezeichnet. Führen wir ein weiteres Gedankenexperiment durch: Betrachten Sie das Beispiel in der folgenden Flash-Animation. Die Punkte stellen Höhenmessungen dar. Die tatsächlichen Höhen zwischen unseren bekannten Messpunkten folgen einer Funktion, die wir jedoch nicht kennen. Wir werden aber davon ausgehen, dass die uns unbekannten Höhen des Profils nicht einfach zufällig verteilt sind, sondern denen der bekannten benachbarten Messpunkte ähneln. Stellen wir also ein Höhenprofil gemäss der blauen strichlierten Linie her. D. h., zwischen die Messpunkte legen wir eine lineare Funktion. Die rote durchgezogene Linie zeigt Ihnen das tatsächliche Höhenprofil. Im letzten Bild sehen Sie die beiden Profillinien im Vergleich. Die Höhe in unserem Beispiel ist solch eine Zufalls-Funktion – weder ist sie exakt mathematisch beschreibbar, noch ist sie rein zufällig!

Dieses Element (Animation, Video etc.) kann nicht dargestellt werden und ist nur in der Onlineversion sichtbar. [link]

Bleibt darauf hinzuweisen, dass kontinuierliche räumliche Variablen sich in der Regel besser im Rastermodell abbilden lassen, wenngleich die Rastereinteilung eigentlich wiederum diskrete räumliche Einheiten bildet.

Lernziele

- Sie können die wichtigsten Typen räumlicher Stichprobenziehung beschreiben.
- Sie können Angaben zur sinnvollen Grösse räumlicher Stichprobenziehungen machen.
- Sie beherrschen die Grundlagen der explorativen Variographie.
- Sie können darlegen, weshalb das Wissen um räumliche Abhängigkeiten für die Analyse kontinuierlicher Variablen von Bedeutung ist.
- Sie kennen die wichtigsten Grundlagen räumlicher Schätzverfahren (Interpolationen).
- Sie können sinnvolle Anwendungsbeispiele für Interpolationen nennen.

1.1. Räumliche Stichprobenziehung

Wie beginnen wir nun mit der Analyse kontinuierlicher Variablen?

Der erste Schritt besteht in der Erstellung einer räumlichen Stichprobe. In dem zu Beginn erwähnten Beispiel des Schweizer Niederschlags sind dies meteorologische Messstationen. Ihre Positionen sind fest vorgegeben und nicht frei wählbar (es sei denn, Sie verwenden eine Untermenge aller existierenden Stationen). Wenn Sie aber z. B. die Verteilung chemischer Schadstoffe im Boden analysieren wollen, müssen Sie zunächst die Messpunkte festlegen. Dabei werden Sie auf folgende Eigenschaften der Stichprobe achten müssen:

- Repräsentativität
- Homogenität
- Räumliche Verteilung der Messungen
- Grösse (= Anzahl der Messungen)

Repräsentativität, Homogenität, räumliche Verteilung und Grösse hängen zusammen. So ist eine Grösse von 5 Messstationen für die Schätzung des gesamtschweizerischen Niederschlags wohl kaum sinnvoll, daher auch nicht repräsentativ. Ebensowenig repräsentativ wäre die Auswahl aller deutschschweizerischen Messstationen für die gesamtschweizerische Schätzung. Hier könnte wohl die Grösse allein ausreichend sein, nicht aber die räumliche Verteilung. Wählen Sie nun alle Stationen unter 750 mNN aus, so könnte die Stichprobe zwar sowohl von der Grösse her als auch von ihrer räumlichen Verteilung stimmen, das Phänomen ist jedoch nicht homogen in der Stichprobe repräsentiert. Eine nachfolgende Schätzung würde v. a. im Bereich von Gebieten über 750 mNN deutlich verzerrt ausfallen.

1.1.1. Eigenschaften

Repräsentativität
Das Phänomen, das analysiert wird, sollte in allen Ausprägungen in der Stichprobe vertreten sein. Insbesondere Minima und Maxima sind von Bedeutung. Für das Niederschlagsbeispiel bedeutet dies: Stationen mit Spitzenwerten sollten vertreten sein. Wenn wir allerdings ein eigenes Probenschema planen, wissen wir in der Regel nicht, ob wir die Standorte mit Minima und Maxima erfasst haben.

Homogenität
Wie zu Beginn dieser Lektion erwähnt, ist die räumliche Abhängigkeit der Daten untereinander eine sehr wichtige Grundvoraussetzung für eine weitere sinnvolle Analyse. Dieser Zusammenhang sollte aber über das gesamte Untersuchungsgebiet homogen sein! Um bei den Niederschlagswerten zu bleiben: jeweils zwei Stationen im Abstand von z. B. 2km sollten sowohl im Tessin ähnliche Messwerte aufweisen als auch im Jura, in Graubünden oder in Fribourg usw. Diese Voraussetzung nennt man auch „Stationarität".

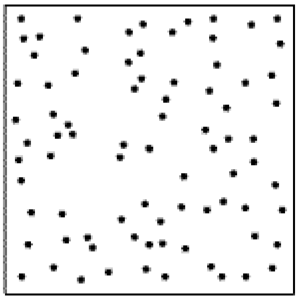
Räumliche Verteilung
Die räumliche Verteilung ist von grosser Bedeutung. Sie kann völlig zufällig sein, regelmässig oder geclustert. Die Verteilungen sehen Sie an Beispielen weiter unten im Abschnitt „Typologie". Einen Hinweis auf die räumliche Verteilung einer Stichprobe können wir statistisch z. B. mittels der „Nearest Neighbor"-Statistik erhalten. Sie zählt zu den „Point Pattern Analysis"-Techniken - Methoden, mit deren Hilfe sich die räumliche Verteilung von Punkten statistisch charakterisieren und analysieren lassen.

Grösse
Die Grösse, also der Umfang einer Stichprobe, ist abhängig vom Phänomen und von der Arealfläche. In manchen Fällen unterliegt die Wahl der Stichprobengrösse praktischen Einschränkungen. Denken Sie z. B. an Messungen in schwer zugänglichem Gelände, technisch aufwendigen und teuren Messungen usw. Eine ideale Grösse für jegliche Aufgabe anzugeben, ist schlicht unmöglich.

1.1.2. Typologie

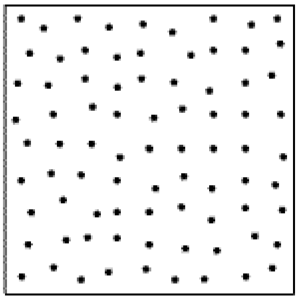
Für das Design räumlicher Stichproben können unterschiedliche Typen verwendet werden. Die Wahl des Stichproben-Designs hängt vom Phänomen ab, das untersucht werden soll, kann aber auch vom Messverfahren beeinflusst sein.

Zufalls-Stichprobe
Beachten Sie die Bereiche ohne Stichprobe – das zu analysierende Phänomen ist dort unterrepräsentiert.



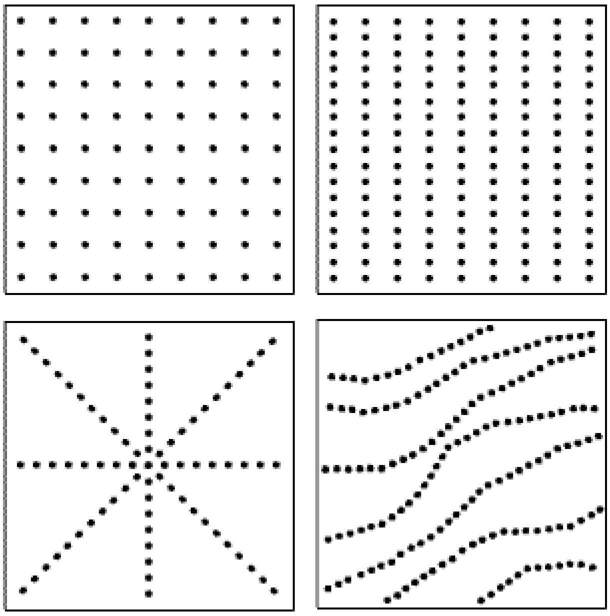
Zufalls-Stichprobe

Gleichförmige Zufalls-Stichprobe (mit Mindestabstand zwischen Punkten)



Gleichförmige Zufalls-Stichprobe

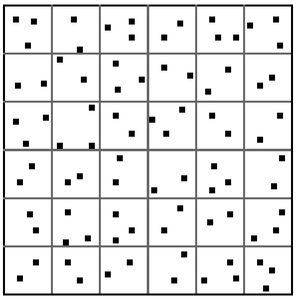
Systematische Stichproben
Diese Auswahl ist keineswegs vollständig. Sie zeigt Ihnen jedoch, dass „systematisch“ nicht zwingend „quadratisch“ bedeuten muss!



Beispiele für systematische Stichproben

Geschichtete Zufalls-Stichprobe mit regelmässigem Netzraster (hier mindestens 2 Punkte pro Rasterfeld)

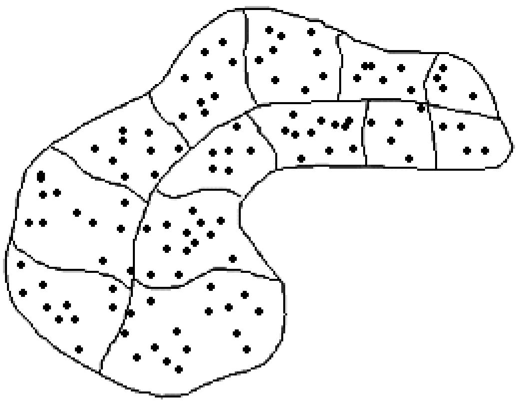
Erkennen Sie die Ähnlichkeit dieses Typs mit der gleichförmigen Stichprobe? Das Kriterium ist hier jedoch kein Mindestabstand, sondern die Aufteilung des Gebietes in gleichförmige Subsets (Straten).



Geschichtete Zufalls-Stichprobe mit regelmässigem Netzraster

Geschichtete Zufalls-Stichprobe mit unregelmässigem Netzraster (hierarchische bzw. autoritative Stichprobe)

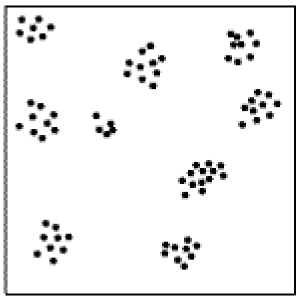
Die Idee der geschichteten Zufalls-Stichprobe wird hier auf unregelmässige Netzraster angewandt. Dies könnten z. B. Zählsprengel sein.



Geschichtete Zufalls-Stichprobe mit unregelmässigem Netzraster

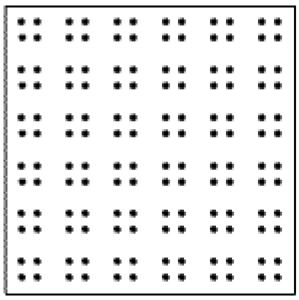
Geclusterte Zufalls-Stichprobe

Für solch ein Stichproben-Design müssen gute Gründe vorliegen. Dieses Design könnte auch eine Vorstufe sein, aus der die Werte innerhalb der Cluster gemittelt werden.



Geclusterte Zufalls-Stichprobe

Geclusterte systematische Stichprobe



Geclusterte systematische Stichprobe

Beispiel von unterschiedlichen Sampling-Typen

Sehen Sie in der nachfolgenden Flash-Animation Beispiele unterschiedlicher Sampling-Typen, angewandt auf das Gebiet der Schweiz. Als Hintergrunddaten wird das digitale Höhenmodell der Schweiz angezeigt.

Dieses Element (Animation, Video etc.) kann nicht dargestellt werden und ist nur in der Onlineversion sichtbar. [link]

1.2. Analyse räumlicher Abhängigkeit

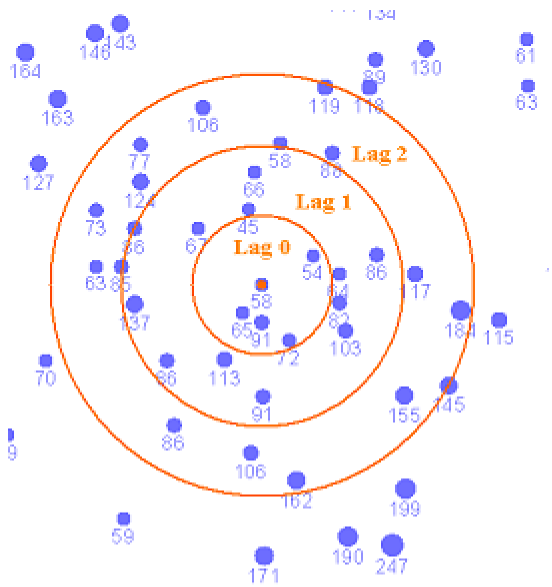
Nun haben wir eine geeignete räumliche Stichprobe und sollen feststellen, inwieweit bzw. ob überhaupt räumliche Abhängigkeiten zwischen diesen Daten bestehen. Dazu exisitieren einige Methoden, die Sie im folgenden Abschnitt kennen lernen werden:

- Das Variogramm bzw. die Explorative Variographie
- Die „Moving Window“-Statistik

Warum müssen wir uns denn überhaupt um die räumlichen Abhängigkeiten kümmern? Genügt es denn nicht, eine sorgfältige Stichprobe zu haben? Nein! Denn ohne räumlich-basierte Zusammenhänge zwischen unseren Stichproben können wir auch keinerlei Aussagen über die Punkte machen an denen keine Proben genommen wurden. Dies geht zurück auf eine Aussage von Waldo Tobler, die als „The 1st Law of Geography“ bekannt ist: "*(...)the first law of geography: everything is related to everything else, but near things are more related than distant things.*" (Tobler 1970)
In den meisten Fällen trifft diese Gesetzmässigkeit zu. Verlassen sollten wir uns darauf aber nicht blindlings, etwa am Rand einer Felsenklippe ...

1.2.1. Variogramm

Machen wir uns zunächst anhand eines gedachten Beispiels mit der Problematik vertraut: Stellen Sie sich ein digitales Geländemodell vor und entnehmen Sie diesem Stichproben. Der Wert einer Stichprobe entspricht also ihrer Meereshöhe. Benachbarte Stichproben könnten z. B. zufällig entlang eines Talbodens gleicher Meereshöhe entnommen worden sein. Ein anderes Stichprobenpaar etwa derselben Distanz zueinander stammt z. B. aus einem Höhenrücken. Wenn Sie nun die Werte beider Paare vergleichen, stellen Sie eine Übereinstimmung oder zumindest eine Ähnlichkeit der Werte fest. Sodann vergleichen wir Stichproben in grösserer Entfernung zueinander. Kann sein, dass sie ähnliche Werte aufweisen, wahrscheinlicher ist jedoch, dass sich die Werte (also die Meereshöhen) unähnlicher werden.
Die Methode der Variographie führt genau diesen paarweisen Vergleich für sämtliche unserer Stichproben durch: Jeder Punkt wird mit jedem verglichen. Dies können also in Abhängigkeit von der Anzahl unserer Stichproben eine ganze Menge von Punktepaaren werden. Genau beträgt diese Anzahl = $n*(n-1)/2$ (n ... Anzahl der Stichproben). Sie fragen sich nun vielleicht zu Recht: „Wo bleibt denn da die Distanz zwischen den Punkten?“ Gleichzeitig mit dem Vergleich „jeder-mit-jedem“ wird auch die Distanz (und die Richtung) der Paare zueinander miterfasst!



Niederschlagswerte

3 Lags (Lag0, Lag1 und Lag2) um einen Datenpunkt (Wert 58) sind hier dargestellt. Die Zahlen sind hier die Werte schweizerischer Niederschlags-Messstationen.

Die Wertepaare:
für Lag0
58,65
58,91
58,54
58,72
für Lag1
58,45
58,64
58,82
usw.

Aus diesen zahlreichen Wertepaaren wird nun als Mass für die Ähnlichkeit (und wir interpretieren in ihr auch eine „Abhängigkeit“) die sogenannte „Semivarianz“ berechnet. Die Formel dafür lautet wie folgt:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i,j=1}^{N(h)} (v_i - v_j)^2$$

Formel für die Semivarianz

... Semivarianz für die Distanz *h*
...Anzahl der Wertepaare innerhalb der Distanz *h*
... Werte an den Positionen *i* bzw. *j*

In einfachen Worten ausgedrückt ist die Semivarianz das halbierte Quadrat der Differenzen zwischen den Wertepaaren. Dieser Parameter wird jeweils für ein Distanzintervall h berechnet – nur die Wertepaare innerhalb dieser Distanz gehen in die Berechnung ein. Diese Distanz h nennt man „Lag“. Tragen Sie alle Wertepaare eines Lags auf einem Scatterplot auf, dann erhalten Sie den sogenannten h-Scatterplot.
Aus den Semivarianzen pro Lag wird das empirische (oder auch experimentelle) Semivariogramm als Kurvengraphik erstellt (bewegen Sie den Mauszeiger über die Lag-Punkte, um den zugehörigen h-Scatterplot für die ersten 8 Lags einzublenden):

Dieses Element (Animation, Video etc.) kann nicht dargestellt werden und ist nur in der Onlineversion sichtbar. [link]

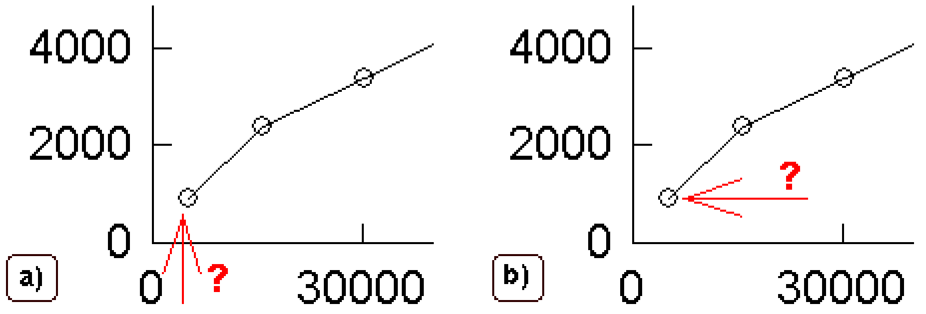
Können Sie sich vorstellen, warum in den h-Scatterplots der niedrigen Distanzen deutlich weniger Punkte vorhanden sind als in jenen der höheren Lags?

Weil in den niedrigen Lags weniger Punktepaare einer insgesamt kleineren Fläche eingehen!

Auf der x-Achse sehen Sie die zunehmende Distanz zwischen den Punktepaaren, auf der y-Achse die Semivarianz pro Lag. Die Kreissymbole auf der Kurve markieren die einzelnen Lags. Als Lag-Intervall finden Sie in diesem Beispiel 15000. Wie interpretieren wir so eine Kurve? Je ähnlicher die Wertepaare pro Lag sind, desto niedriger ist die Semivarianz für diesen Lag; je unähnlicher, desto höher steigt die Semivarianz und damit die Kurve an. Diese Kurve bestätigt uns daher: Auf niedrige Distanzen sind die Werte unserer Daten einander ähnlicher. Es besteht ein unmittelbarer Zusammenhang zwischen der Entfernung der Datenpunkte und deren Werteähnlichkeit! Zwei Kennzahlen sollten Sie sich merken, die Ihnen helfen, diese Kurve zu beschreiben:

- Range – jene Distanz h, ab welcher die Kurve abflacht.
- Sill – der Semivarianz-Wert für die Position, an welcher die Kurve ihren Range erreicht.

Wenn das Lag-Intervall im Beispiel oben 15000 beträgt, warum befindet sich das erste Lag (= Lag 0 bzw. h0) dann nicht im Koordinaten-Ursprung? Einfach darum, weil die Punktepaare im Lag 0 eine gewisse Distanz zueinander aufweisen. Deren mittlerer Abstand ist nun die Position für Lag 0 auf der x-Achse. Warum aber startet die Kurve nicht bei der Semivarianz 0, also auf der x-Achse? Weil die Daten im Lag 0 nicht alle identisch sind (dies ist meistens der Fall). Darum kommt der Ursprung der Semivariogramm-Kurve in der Regel etwas oberhalb der x-Achse zu liegen. Diesen Effekt nennt man Nugget-Effekt. Nugget-Effekt darum, weil diese Methodik aus dem Bereich der geologischen Exploration stammt. In Proben auf Gold können Nuggets punktuell auftreten, d. h. die Werte von unmittelbar benachbarten Proben können sich deutlich unterschieden.



a) Lag 0 umfasst alle Punktepaare innerhalb des ersten Lags. Die Durchschnittsdistanz zwischen den Punkten markiert den Lag auf der x-Achse; b) Die Punktepaare in Lag 0 weisen unterschiedliche Werte auf, daher ist auch die Semivarianz nicht gleich 0, sondern beginnt etwas oberhalb der x-Achse (= Nugget-Effekt)

In der einfachsten Form werden die Wertepaare von jedem Punkt nach allen Richtungen gebildet, und Sie erstellen ein isotropes Semivariogramm. Als Erweiterung und Verfeinerung der Methodik können Variogramm-Programme Punktepaare in spezifischen Richtungen bilden, damit erkennen Sie z. B., ob die Werte in Ihren Daten in manche Richtungen höhere räumliche Abhängigkeiten aufweisen. Denken Sie an das eingangs geschilderte Beispiel mit den Meereshöhen: Verläuft in Ihren Daten z. B. ein Tal in N-S-Richtung, dann weisen Punkte in dieser Richtung deutlich höhere Ähnlichkeit auf als in W-E-Richtung! Das Resultat ist nun ein anisotropes Semivariogramm. Wenn Sie an dieser Stelle bereits das Ziel aus den Augen verloren haben: Alle diese Informationen über die räumliche Abhängigkeit und ihre Struktur können wir für die Schätzung der unbekannten Werte verwenden.

Verwenden Sie den folgenden interaktiven Semivarianz-Calculator und geben Sie jeweils Wertepaare ein. Wählen Sie zunächst ähnliche Werte (bis 99), anschliessend variieren Sie die Werte und lassen Sie sie zunehmend unähnlicher werden. Beobachten Sie, wie sich die Semivarianz ändert! Beachten Sie auch, wie einfach die Semivarianz-Formel umgesetzt wird.

Dieses Element (Animation, Video etc.) kann nicht dargestellt werden und ist nur in der Onlineversion sichtbar. [link]

Was passiert, wenn Sie für alle Punkte denselben Wert eingeben? Ist es von Bedeutung, in welcher Reihenfolge Sie die Wertepaare eingeben?

- Wenn alle Punkte denselben Wert aufweisen, also identisch sind, besteht eine perfekte räumliche Abhängigkeit, und die Semivarianz wird zu 0.
- Es ist nicht von Bedeutung, in welcher Reihenfolge die Wertepaare eingegeben werden, Minuswerte werden ja durch die Quadrierung zu positiven Werten.

1.2.2. Moving Windows

Die Variographie erlaubt uns also, räumliche Abhängigkeiten aufzuspüren. Was sie jedoch nicht zeigt, ist, ob diese Abhängigkeiten im gesamten Untersuchungsgebiet gleichförmig sind – möglicherweise bestehen regional grosse Unterschiede, und unser Variogramm ist somit nicht repräsentativ für das gesamte Areal! Es geht also um die räumliche Variabilität unserer Daten.
Die einfache Technik der „Moving Windows“- (= Gleitfenster-)Statistik hilft dabei. Ein „Fenster“ definierter Grösse und Form wird um je eine Fensterbreite über die Daten hinweg bewegt. Alle Daten, die sich unter dem Fensterausschnitt befinden, werden statistisch zusammengefasst: die Anzahl und der Mittelwert sämtlicher Punkte im Fenster, die Minimum- / Maximum-Werte, die Standardabweichung, der Variationskoeffizient (= Standardabweichung/Mittelwert) usw. Als Ergebnis erhalten wir wiederum Punkte – die Mittelpunkte der gleitenden Fenster, und als deren Attribute die statistischen Kennziffern dieses Fensters. Im Fall spärlicher Daten wird das Fenster um je eine Fensterhalbbreite weiterbewegt, um so mehr Daten zur Berechnung zu erhalten (= Gleitfenster mit Überlappung). Das Prinzip geht aus folgender Animation hervor:

Dieses Element (Animation, Video etc.) kann nicht dargestellt werden und ist nur in der Onlineversion sichtbar. [link]

Sowohl Fenstergrösse als auch -form können mit diesem Instrument variabel eingestellt werden. In der Praxis geht man auch entsprechend explorativ vor: Die Analyse wird mit Fenstern unterschiedlicher Dimensionen durchgeführt, und die Statistiken werden verglichen. Insbesondere der Variationskoeffizient ist ein aussagekräftiger Parameter – nimmt er Werte >1 an, dann deutet dies auf hohe Abweichungen (= hohe räumliche Variabilität) in diesem Fensterausschnitt hin.

Betrachten Sie das folgende Beispiel einer „Moving Window“-Statistik für die Schweizer Niederschlagsdaten. Regionen mit höherem Niederschlag sind relativ leicht zu erkennen, zwei davon bilden eine Art NE-SW-Achse. Für die Grösse der Gleitfenster ist 30x30km mit der Option Überlappen gewählt. Die Grösse solch eines Fensters sehen Sie als graues Quadrat eingezeichnet. Zusätzlich besteht die Möglichkeit, Fenster mit weniger als z. B. 4 Punkten auszusparen, weil bei solch einer geringen Punktezahl keine aussagekräftige Statistik zu erwarten ist. Darum gibt es einige wenige „Löcher“. Die Mittelwerte spiegeln die Niederschlagshöhen wieder. Besonders interessant jedoch ist der Variationskoeffizient, der auf Regionen mit grösseren Werteschwankungen hinweist. Die beiden höchsten Werte befinden sich an der Südspitze des Tessins.

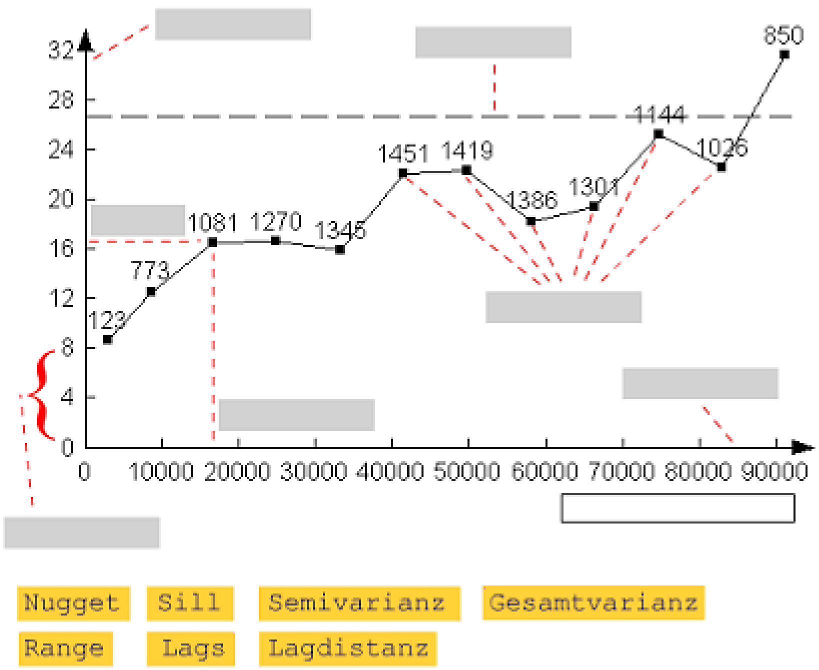
Dieses Element (Animation, Video etc.) kann nicht dargestellt werden und ist nur in der Onlineversion sichtbar. [link]

In welchem der 8 Fenster aus dem ersten Beispiel finden Sie den höchsten Variationskoeffizienten und welchen Wert hat er?

In Fenster 4 - der Variationskoeffizient beträgt dort 4.8 / 44 = 0.109

1.2.3. Semivariogramm-Parameter richtig zuordnen

Ziehen Sie in der Flash-Animation die Begriffe an die jeweils richtige Position in der Semivariogramm-Graphik. Keine Sorge – nach einigen Fehlversuchen erhalten Sie eine Hilfe.

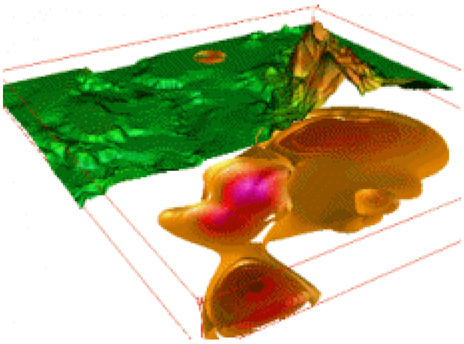


Richtige Variogramm-Parameter zuordnen

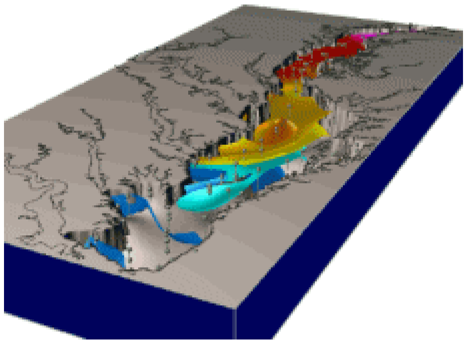
1.3. Räumliche Interpolation

Beispiele für Interpolationsergebnisse

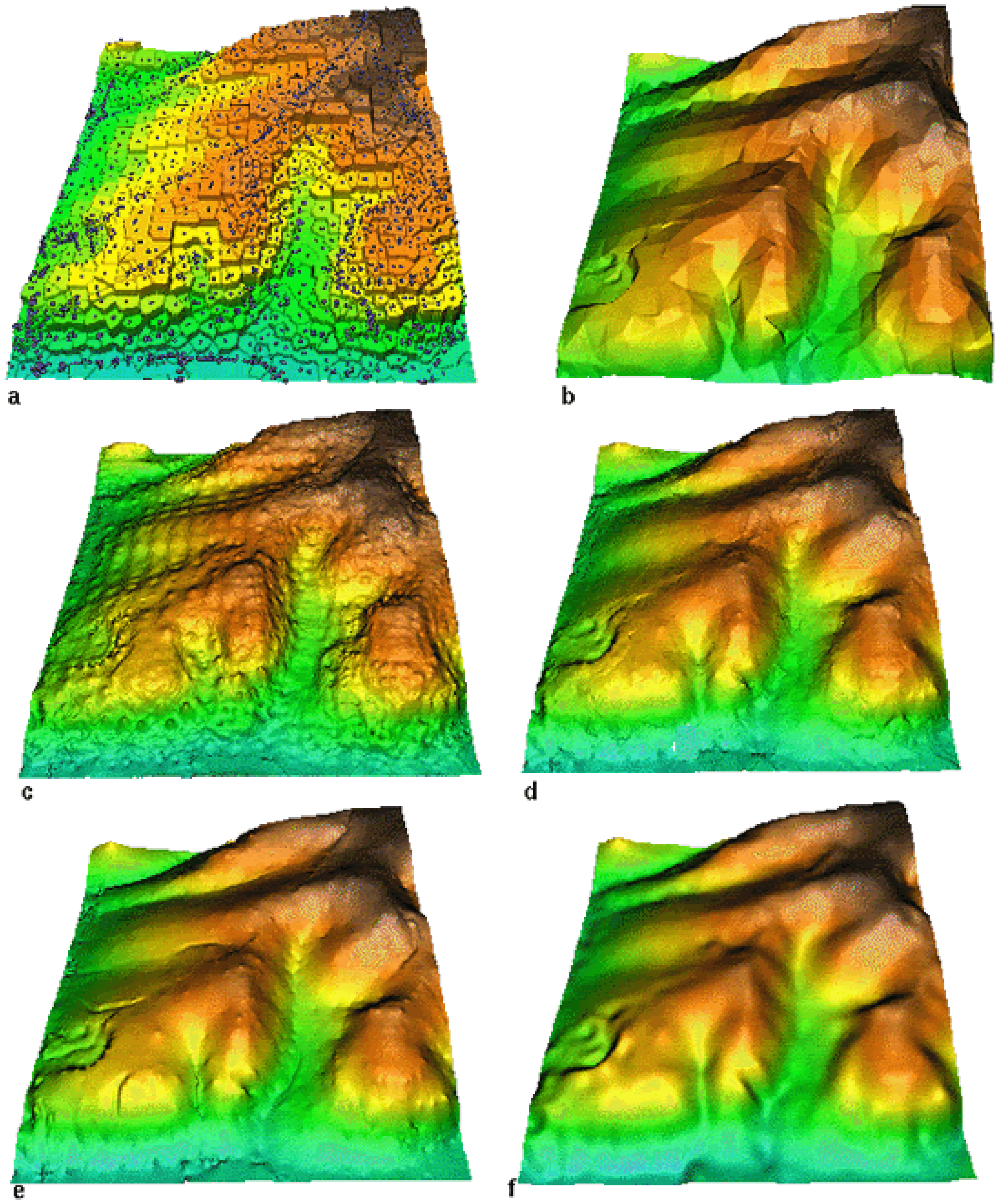
Klicken Sie auf die folgenden beiden Abbildungen, um eine Simulation der sich verändernden Chemikalienkonzentration zu sehen.



Chemikalienkonzentration im Boden (Mitas et al. 1998)



Chemikalienkonzentration in einem Wasserkörper (Mitas et al. 1998)



Interpolation of a DEM from scattered point data using methods available in GIS: a) given data and Voronoi polygons, b)TIN based linear interpolation, c) inverse distance weighting d) kriging, e) spline with tension and stream enforcement, f) regularized spline with tension and smoothing.

Vergleich unterschiedlicher Interpolationsmethoden (Mitas et al. 1999)

Die dargestellten Beispiele sind Ergebnisse ausgereifter und sorgfältiger Interpolationen aus einem umfangreichen Spektrum möglicher Anwendungen.
Nachdem wir uns in den Abschnitten zuvor mit Stichproben und der Analyse räumlicher Abhängigkeiten befasst haben, kommen wir nun zum „Herzstück“ dieser Lektion – den räumlichen Interpolationen. Viele dieser Techniken zählen nicht eben zu den einfachsten Anwendungen räumlicher Analyse, darum beschränken wir uns hier bewusst auf einen kurzen Überblick zu den Methoden. Folgende Interpolations-Techniken werden vorgestellt:

- die distanz-basierte Interpolation
- geostatistische Methoden

Letztere allerdings sind Gegenstand fortgeschrittener Lehrveranstaltungen und werden nicht im Detail behandelt. Was sind überhaupt räumliche „Interpolationen“? Darunter versteht man die Berechnung unbekannter Werte auf der Basis benachbarter bekannter Werte.

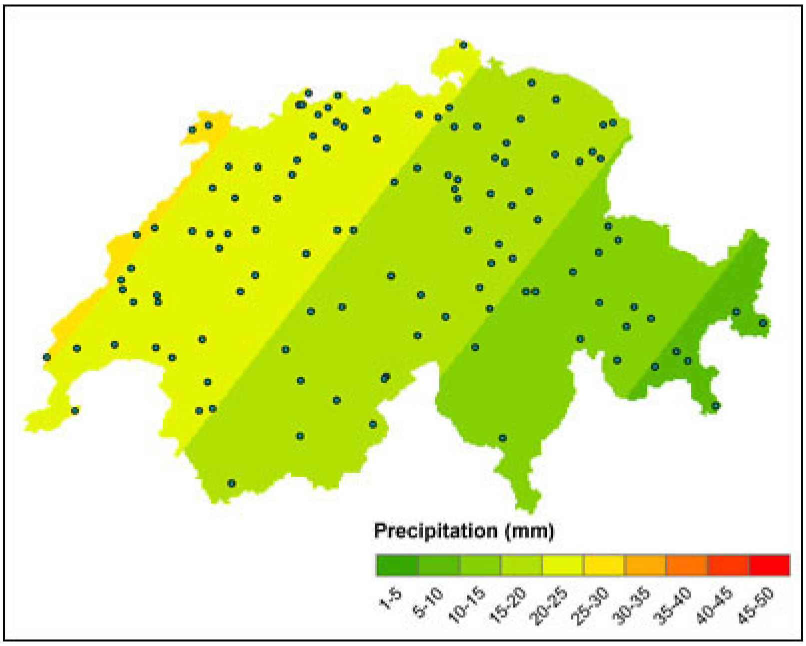
1.3.1. Typologie

„Inverse Distanz“-Gewichtung, „Radial Basis“-Funktionen, „Splines“, „Ordinary Kriging“, „Natural Neighbor“, „Polynomial Regression“-Methoden, „Universal Kriging“ usw. Dies sind lediglich einige Interpolations-Methoden, die in kommerzieller Software zu finden sind. Verwirrend sind die Vielfalt der Methoden und deren Parametrisierungen. Darum versuchen wir zunächst, die Methoden in Schemata einzuordnen: dazu gibt es verschiedene Ansätze, wie Sie aus folgender Tabelle ersehen:

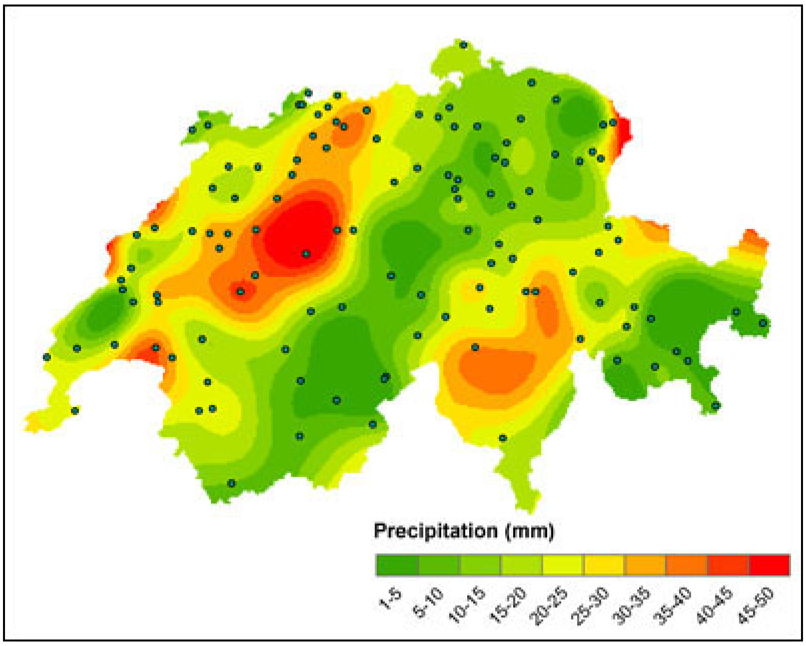
Lokale vs. Globale Interpolation

Globale Methoden werden auf ALLE Daten im Untersuchungsgebiet angewandt, lokale dagegen nur auf räumlich definierte Subsets. Globale Interpolation eignet sich daher nicht zur Ermittlung möglichst exakter Werte, sondern zur Beurteilung globaler räumlicher Strukturen.

Als Beispiele sehen Sie nachfolgend eine lineare Trend-Oberfläche – sie wurde mittels Regression aus schweizerischen Niederschlagsdaten ermittelt und zeigt einen Trend zum Anstieg der Niederschlagshöhen von SE nach NW – und eine lokale Interpolation mittels sogenannter Radial Basis Interpolation:



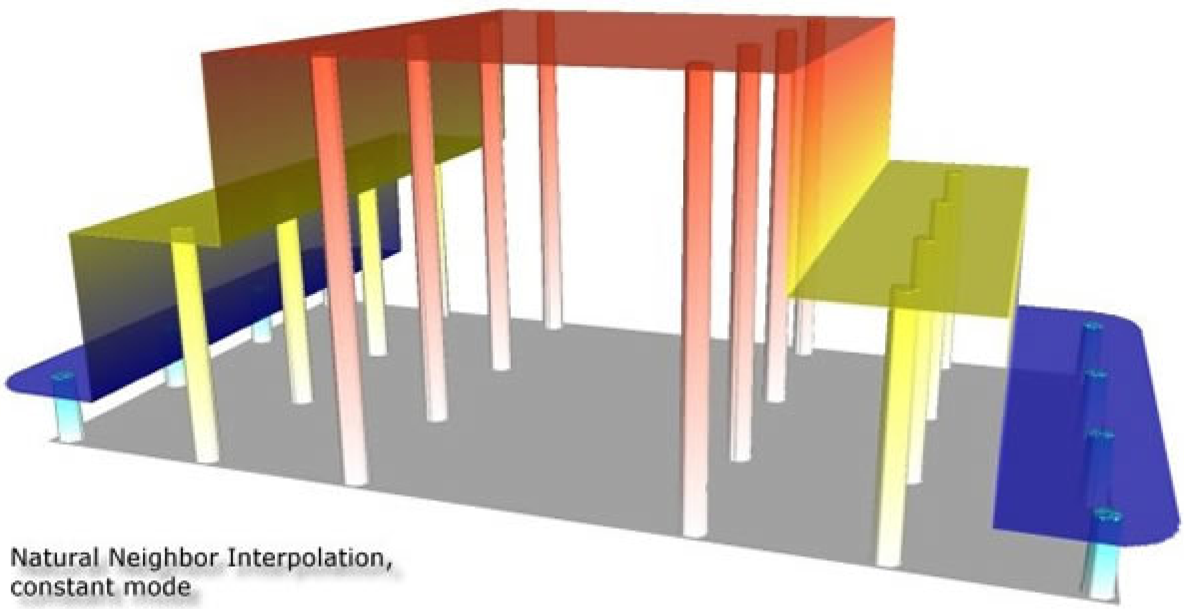
Beispiel einer Globalen Interpolation – Lineare Trendoberfläche für Schweizer Niederschlagsdaten. (Zur Verfügung gestellt von Ross Purves)



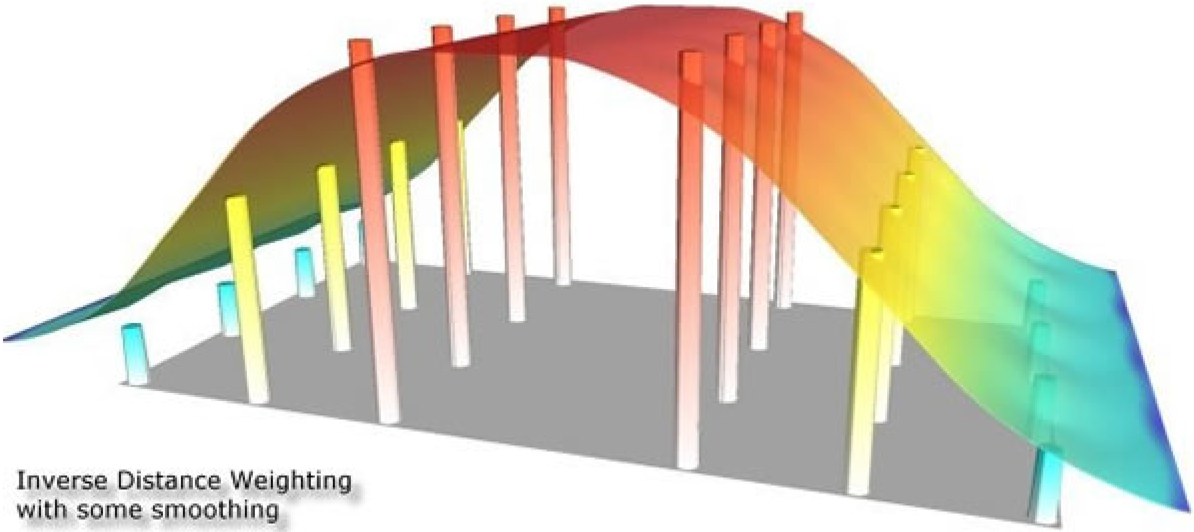
Beispiel einer lokalen Interpolation – Spline Interpolation für Schweizer Niederschlagsdaten. (Zur Verfügung gestellt von Ross Purves)

Exakte vs. Nicht-exakte Interpolation

Exakte Interpolation heisst: die geschätzte Oberfläche passiert die bekannten Punkte, während bei nicht-exakten Methoden die Schätzwerte für bekannte Beobachtungen von den realen Werten abweichen können. Letztere Methoden werden sinnvollerweise dann eingesetzt, wenn die bekannten Daten bereits gewisse Unschärfen aufweisen.



Exakter Interpolator: Schätzoberfläche passiert exakt die bekannten – schematisch als Säulen dargestellt – Punkte (Wyatt 2000)



Nicht-exakter Interpolator: Schätzoberfläche passiert die bekannten – schematisch als Säulen dargestellt – Punkte NICHT (Wyatt 2000)

Graduelle vs. abrupte Interpolation
Diese Unterscheidung bezieht sich vorwiegend auf die Ergebnis-Schätzoberflächen – wurden bei der Interpolation Bruchkanten (natürliche abrupte Werteänderungen, z.B. Klippen, Seeufer) miteinbezogen oder nicht?

Deterministische vs. stochastische Interpolation
Deterministische Interpolationstechniken basieren auf exakt vorherbestimmbaren (= deterministischen) räumlichen Zusammenhängen; in stochastische Verfahren dagegen fliessen auch Zufallselemente mit ein. Deterministische Verfahren zeigen bei der Interpolation natürlicher räumlicher Phänomene deutliche Nachteile, da ein gewisser Grad an Ungewissheit stets vorhanden ist.

1.3.2. Distanz-basierte Interpolation

Im einfachsten Fall dieser Methoden können wir wie bei der **„Moving Windows“-Methode** vorgehen: Wir definieren eine gewisse „Nachbarschaft“ bekannter Datenpunkte um die jeweils zu schätzende unbekannte Position; das arithmetische Mittel aus diesen bekannten Messwerten ist dann unser Schätzwert (= lokale Mittelwertbildung, moving average). Die Nachbarschaft kann unterschiedlich definiert werden:

- eine räumlich festgelegte Form (Rechteck, Kreis usw.)
- eine bestimmte Anzahl nächst benachbarter bekannter Punkte

Dieses Instrument ist aber recht unscharf, weil die verschiedenen Distanzen zwischen Schätzposition und bekannten Punkten nur unzureichend in die Interpolation einfliessen. Die eigentlichen Distanz-basierten Methoden verwenden eben alle jene Entfernungen zwischen der zu schätzenden Position und den bekannten Messpunkten, um deren Einfluss in der Berechnung des Schätzwertes zu gewichten. Sie setzen übrigens einen linearen räumlichen Zusammenhang zwischen den Phänomenen voraus.
Bei der sogenannten „inversen Distanz-Gewichtung “(Inverse Distance Weighting, IDW) wird das Gewicht jedes bekannten Punktes invers proportional zu seiner Entfernung zum geschätzten Punkt gesetzt. Die Berechnung erfolgt nach dieser Formel:

$$\hat{v} = \frac{\sum_{i=1}^{n} \frac{1}{d_i} v_i}{\sum_{i=1}^{n} \frac{1}{d_i}}$$ *Inverse Distanz-Gewichtung (IDW) – Grundformel*	... zu schätzender Wert, ... bekannte Werte d_i..., d_n... Distanzen der n Datenpunkte zum geschätzten Punkt

Meist finden Sie die folgende Variante, in welcher der Einfluss der Distanz zusätzlich über einen Exponenten gesteuert werden kann (dieser wird in den meisten Programmen auf 2 voreingestellt):

$$\hat{v}_1 = \frac{\sum_{i=1}^{n} \frac{1}{d^p{}_i} v_i}{\sum_{i=1}^{n} \frac{1}{d^p{}_i}}$$ *Häufigste Form der IDW-Formel mit zusätzlichem Distanzgewichtungs-Exponenten*	... zu schätzender Wert, ... bekannte Werte $d^p{}_i$..., $d^p{}_n$... mit p exponenzierte Distanzen der n Datenpunkte zum geschätzten Punkt

Je niedriger der Exponent gesetzt wird, desto gleichförmiger gehen alle Nachbarn (ungeachtet ihrer Distanz) in die Berechnung ein, und desto „glatter“ wird die Schätzoberfläche. Je höher der Exponent wird, desto akzentuierter und „unruhiger“ wird die Oberfläche, da nur mehr das Gewicht der nächstgelegenen Nachbarn in die Interpolation einfliesst (siehe folgende interaktive Animation).

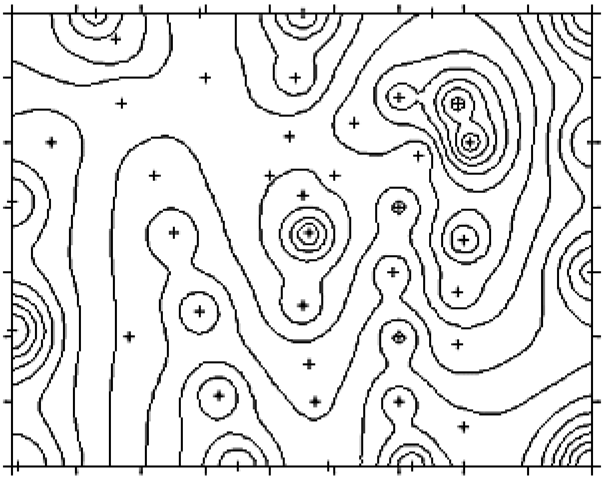
Dieses Element (Animation, Video etc.) kann nicht dargestellt werden und ist nur in der Onlineversion sichtbar. [link]

Vorteile der IDW-Interpolation:

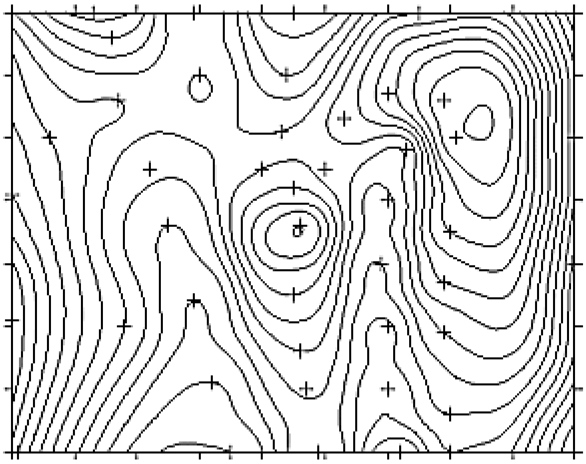
- Sie ermöglicht sehr schnelle Berechnungen.
- Unterschiedliche Distanzen fliessen in die Schätzung unterschiedlich ein.
- Über den Distanzgewichtungs-Exponent kann der Einfluss der Distanzen fein gesteuert werden.

Nachteile der IDW-Interpolation:

- Es ist keine richtungsabhängige Gewichtung möglich, d. h. räumlich gerichtete Zusammenhänge werden ignoriert (z. B. Höhenpunkte entlang eines Bergrückens).
- Unschöne Artefakte sind die sogenannten „Bull-Eyes“ – dies sind kreisförmige Bereiche gleicher Werte um die bekannten Datenpunkte. Eine von (1968) entwickelte Variante der IDW-Interpolation reduziert diese Bull-Eyes jedoch (siehe folgende Abbildungen).



IDW „Bull Eyes"-Effekt: Um die bekannten Punkte sind konzentrische Bereiche gleicher Werte zu erkennen – ein unerwünschtes Artefakt der IDW-Interpolation



IDW modifiziert nach SHEPARD: die Bull-Eyes sind deutlich reduziert

Inverse Distanz Gewichtung (IDW) - Interaktiv

In der folgenden interaktiven Animation sehen Sie 10 Datenpunkte (blau) mit bekannten Messwerten (Ziffern neben den Punkten) und einen Punkt mit zu errechnendem Wert (rot). Beim Start der Animation wird dieser aus den vorgegebenen Werten und Distanzen ermittelt. Um die Prinzipien der IDW-Interpolation besser kennen zu lernen, experimentieren Sie nun:

- Verändern Sie mit Ihrem Mauszeiger die Positionen eines oder aller Punkte.
- Modifizieren Sie die vorgegebenen Werte für die bekannten Punkte (insgesamt sind max. 4 Stellen erlaubt).
- Setzen Sie für den Distanzgewichtungs-Exponenten einen anderen Wert als 2 ein (insg. max. 4 Stellen erlaubt).

Dieses Element (Animation, Video etc.) kann nicht dargestellt werden und ist nur in der Onlineversion sichtbar. [link]

Beantworten Sie aus Ihren Experimenten heraus folgende Fragen:

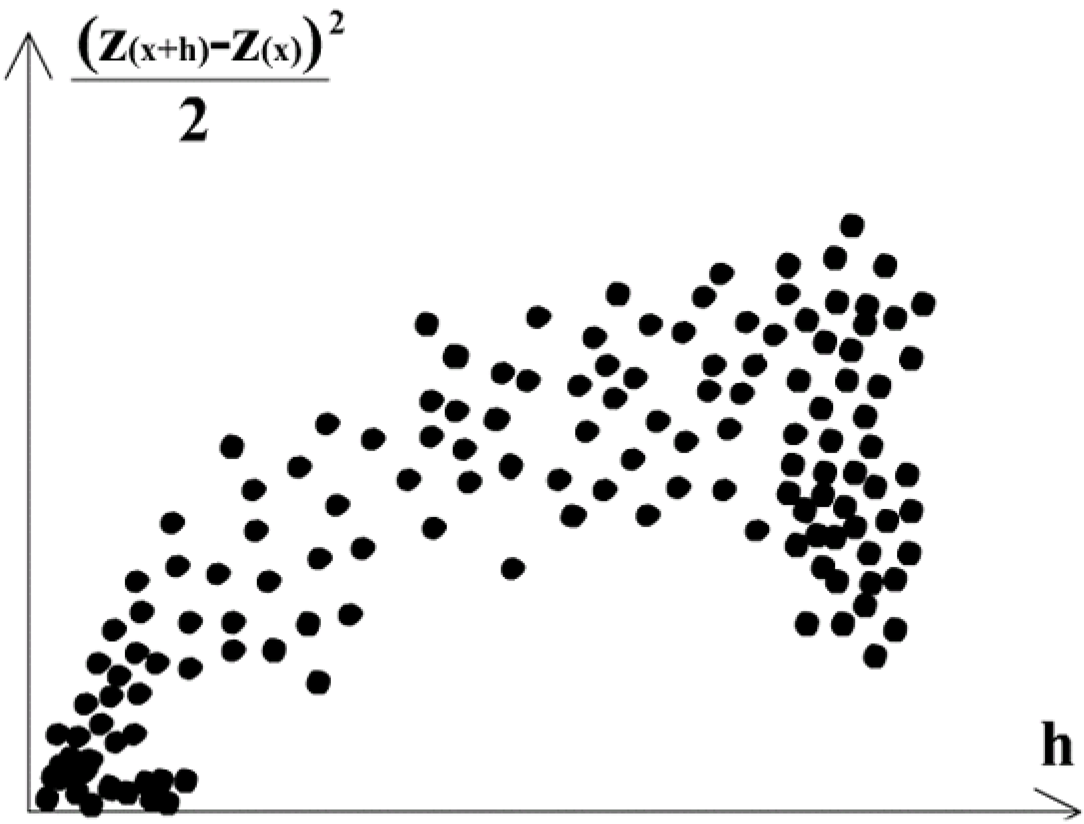
1. Welcher Messwert beeinflusst das Ergebnis umso mehr, je höher der Exponent gesetzt wird?
2. Wenn der Exponent auf 0 gesetzt wird, wie beeinflussen dann unterschiedliche Distanzen das Schätzergebnis bzw. wovon ist dieses dann nur abhängig?

1. **Der bzw. die Messwert(e) der nächst benachbarten Punkte** beeinflussen das Ergebnis umso mehr, je höher der Distanzexponent gesetzt wird.
2. Wird der Distanzexponent auf 0 (Null) gesetzt, dann ergibt sich für jegliche Distanz eine Gewichtung von 1, d. h. alle Entfernungen sind absolut gleichwertig, **das Resultat hängt also nur mehr von den Messwerten selbst ab und nicht von der Distanz.** Die Interpolation hat keinerlei räumliche Komponente mehr.

1.3.3. Geostatistische Interpolation

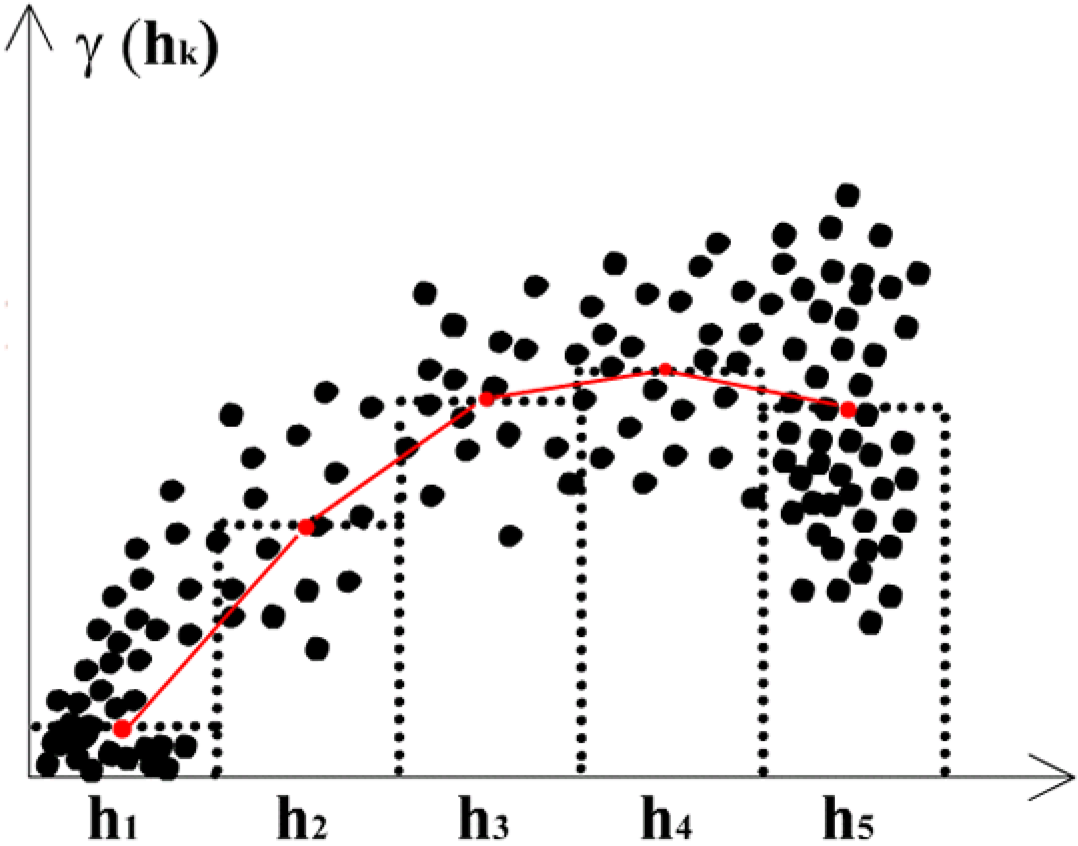
Einer der Nachteile der IDW-Interpolation ist das Fehlen richtungsspezifischer (anisotroper) Informationen. Räumliche Korrelationen werden also ignoriert und fliessen nicht in das Schätzergebnis mit ein. Diesen Nachteil gleichen die geostatistischen Interpolationen aus.
Schon der Name „Geo“-Statistik weist auf das wichtigste Merkmal dieser Methoden hin: räumlich-statistische Parameter bilden die Hauptbasis für diese Interpolations-Verfahren.
Das Variogramm bzw. die Variographie, also die Methode, es aus den räumlichen Punktdaten abzuleiten, bildet die Grundlage für eine erfolgreiche geostatistische Interpolation.
Geostatistische Interpolationen sind fortgeschrittene und teilweise komplizierte Methoden, deren sinnvolle Anwendung eine Menge an Vorwissen und Erfahrung benötigt. Darum müssen an dieser Stelle ein paar Stichworte genügen.
Die wichtigsten Verfahren sind die Kriging-Methoden. Benannt sind sie nach einem südafrikanischen Ingenieur, D. G. Krige. 1951 legte in seiner Diplomarbeit die Grundlagen für Kriging. Aber die Hauptentwicklungen gehen auf die Arbeit von G. Matheron in den 1960er Jahren zurück.
Mittels der Variographie erhalten wir Hinweise darauf, wie ähnlich bzw. unähnlich die Messwerte benachbarter Datenpunkte in Abhängigkeit von ihrer Entfernung zueinander sind.

Variographie
a) Zunächst bilden wir zwischen allen Datenpunkten Paare und vergleichen jeweils deren 2 Werte. Wir wissen pro Datenpaar den Differenzwert (Semivarianz) und die Distanz (h):



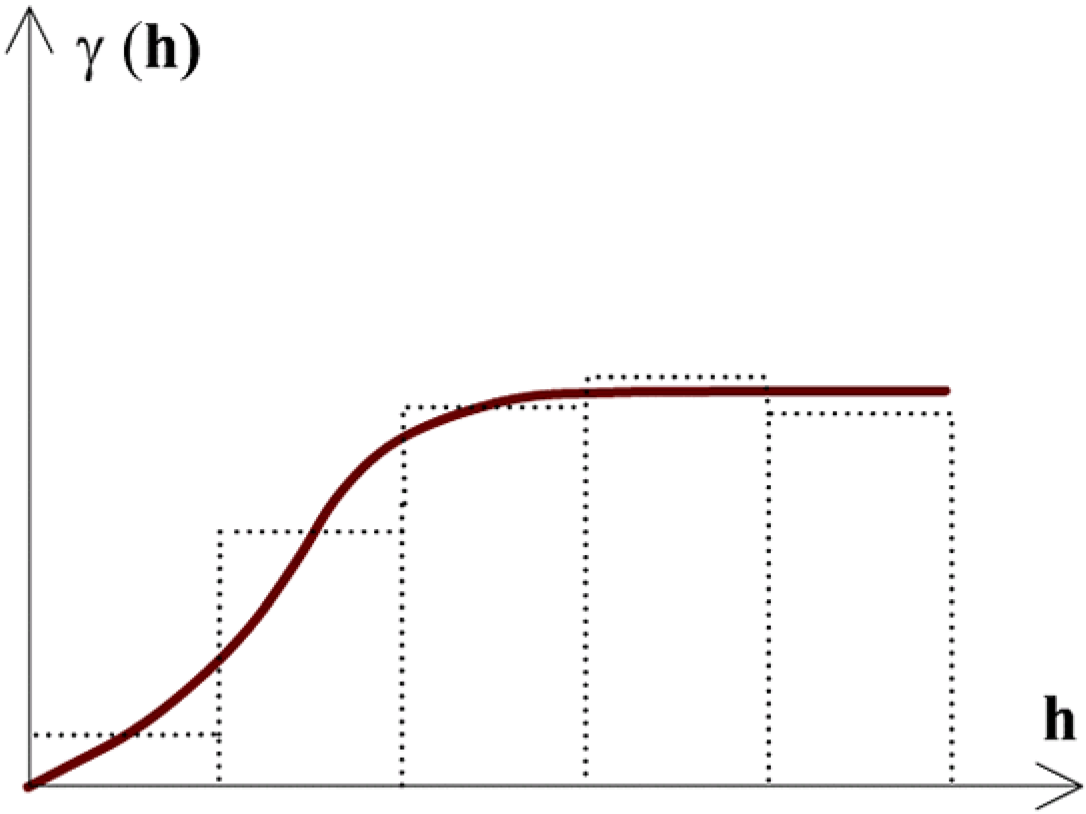
Variogramm Wolke, Unterschiede zwischen Datenpunkten versus räumlicher Distanz zwischen diesen Punkten.

b) Anschliessend unterteilen wir die Distanzen (x-Achse) in Intervalle (sogenannte Lags) und mitteln die Semivarianzen der darin enthaltenen Datenpaare (rote Punkte). Verbinden wir diese roten Punkte pro Lag, erhalten wir das experimentelle Variogramm. Diese Kurve beschreibt, wie ähnlich sich die Werte zweier benachbarter Positionen in Abhängigkeit von ihrer Distanz zueinander sind:



Experimentelles Variogramm, die Unterschiede werden pro definierter Klasse (h1 ... h5, = Lagintervalle) gemittelt

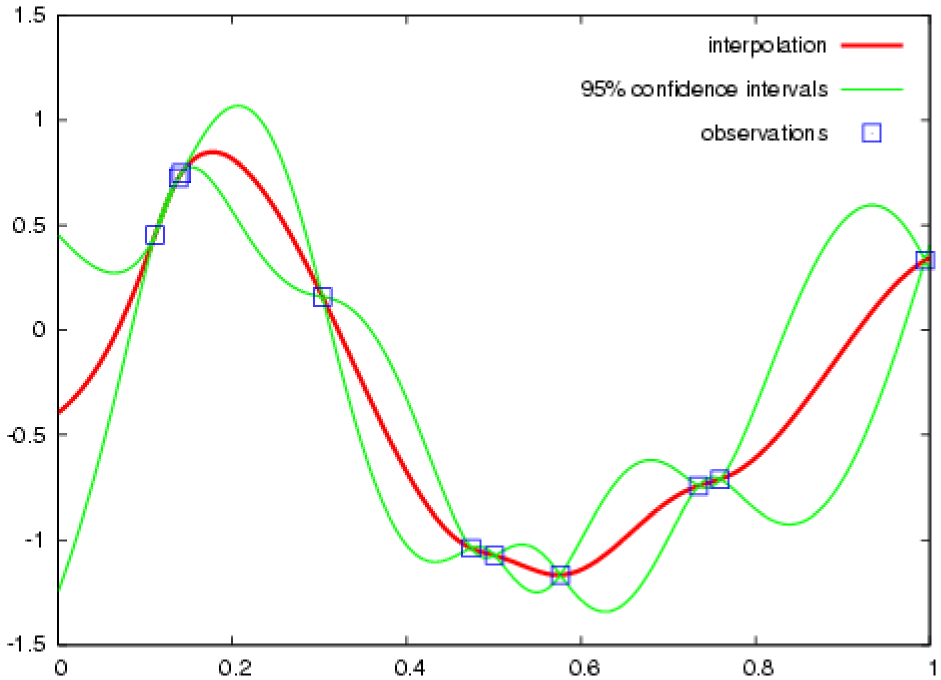
c) Um diese Darstellung der räumlichen (Un-)Ähnlichkeit mathematisch besser verarbeiten zu können, legen wir einfache Kurvenfunktionen so an das experimentelle Variogramm an, dass es dieses möglichst gut nachbildet. Diese Kurve nennen wir theoretisches Variogramm:



Theoretisches Variogramm, eine theoretische Variogrammfunktion wird an die Sequenz der gemittelten Unterschiede pro Klasse (= pro Lag) angepasst

Wir behelfen uns nun beim Kriging, indem wir unsere Daten in dieses Modell räumlicher Kontinuität einbauen – jenem Modell, das wir im Zuge der Variogramm-Modellierung entwickelt bzw. gefunden haben. Ausgehend von solch einem Modell können wir für unsere Schätzungen Fehlervarianzen berechnen und deren Minimierung anstreben.

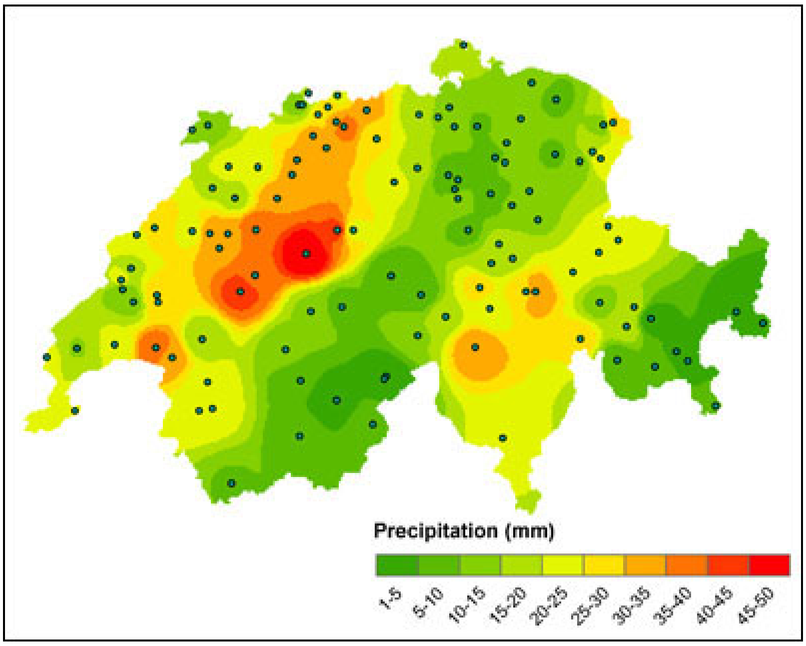
Die Interpolation mit Kriging ist eine Art Kurvenanpassung: Wir haben aus unseren bekannten Datenpunkten ein Modell abgeleitet, wie die räumlichen Zusammenhänge beschaffen sein könnten. Basierend auf diesem Modell schätzen wir nun die unbekannten Punkte. Betrachten wir dies der Einfachheit halber in nur zwei Dimensionen, dann arbeiten wir mit einer Regressionstechnik, einer Kurvenanpassung:



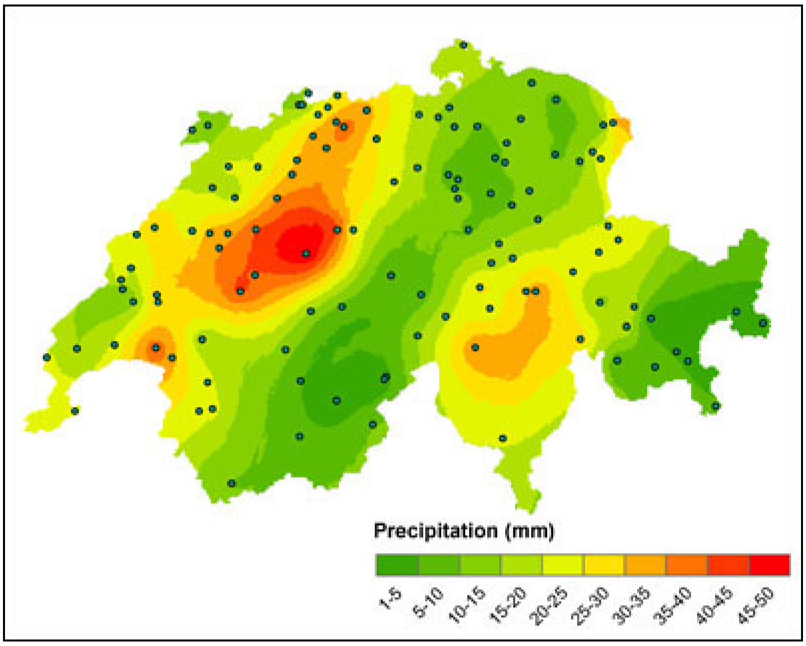
Kriging in zwei Dimensionen: Die blau umrandeten Quadrate sind unsere bekannten Datenpunkte, die rote Linie ist der geschätzte Verlauf, und die grünen Linien repräsentieren die statistischen Rahmenparameter aus unserem Modell (Wikipedia)

Oft hört man den Begriff „exakter Interpolator" im Zusammenhang mit Kriging, sowie IDW und einigen anderen Schätzverfahren. Damit meint man, dass eine mittels dieser Methoden geschätzte Oberfläche die bekannten Datenpunkte exakt schneidet. Führen wir also die Kriging-Berechnung an einer Position mit bekanntem Wert durch, dann liefert das Kriging-System uns in der Regel genau jenen Wert zurück.

Zum Vergleich sehen Sie nachfolgend das Resultat einer inversen Distanz-Gewichtung dem einer Kriging-Interpolation gegenübergestellt:



Schätzoberfläche aus inverser Distanz-Gewichtung, Datengrundlage sind Schweizer Niederschlagsmesswerte. Beachten Sie einige „Höfe“, also Bereiche gleicher Werte um bekannte Datenpunkte. (Zur Verfügung gestellt von Ross Purves)



Resultat einer Kriging-Interpolation mit Schweizer Niederschlagsmesswerten, hier sind keine

Höfe zu beobachten, denn das Kriging-System „kennt“ räumliche Zusammenhänge. Diese Informationen bezieht es aus dem Variogramm. (Zur Verfügung gestellt von Ross Purves)

Weitere wichtige Parameter für Interpolationen: Such-Nachbarschaft

Alle Interpolations-Methoden lassen sich zusätzlich über die Definition einer Such-Nachbarschaft steuern, d. h. wie viele bzw. welche bekannte Datenpunkte werden zur Berechnung einer unbekannten Position herangezogen. Wenn wir diese Nachbarschaft ignorieren, dann beziehen wir sämtliche vorhandenen bekannten

Daten in die Schätzung jedes Punktes mit ein. Im Fall der Schweizerischen Niederschlagsdaten hiesse dies, dass zur Berechnung eines Niederschlagswertes im Tessin auch die Werte der Messstationen aus dem Jura einfliessen. Dass dies wenig Sinn ergibt, liegt nahe.

1.4. Zusammenfassung

Räumlich kontinuierliche Phänomene wie z.B. Niederschlagsmenge oder Meereshöhe lassen sich nicht einfach durch eine mathematische Funktion beschreiben. Um diese Variablen zu analysieren: erstellt man eine räumliche Stichprobe, d.h. gewisse Menge an Messpunkten. Um eine kontinuierliche räumliche Variable darstellen zu können, müssen die Werte zwischen den Messstationen interpoliert werden. Zunächst muss dafür die räumliche Stichprobe, bzw. die Anordnung der Messstationen gemäss der Eigenschaften repräsentativ, homogen, räumlich optimal verteilt und einer genügenden Anzahl entsprechend festgelegt werden. Je nach Phänomen und Messverfahren kann der Designtyp der räumlichen Stichprobe unterschiedlich variieren (z.B. Zufalls-Stichprobe, systematische Stichprobe, geschichtete Stichprobe, geklusterte Stichprobe). Vor der Interpolation muss sichergestellt sein, ob die Abhängigkeit zwischen den räumlichen Daten besteht. Dafür eignen sich zwei Methoden: entweder die Methode der Variographie oder die Methode der Moving Window Statistik. Die Methode der Variographie zeigt die räumliche Abhängigkeit der Stichproben, jedoch nicht ob diese Abhängigkeit im gesamten Untersuchungsgebiet gleichförmig verteilt ist. Dafür wir die Moving Window Statistik verwendet. Für die Interpolation selbst gibt es mehrere Ansätze mit unterschiedlichen Auswirkungen. Zwei Interpolationsarten werden hier vorgestellt: die Distanz-basierte-Interpolation IDW (inverse distance weighting) und die geostatistische Interpolation. Bei IDW fliessen unterschiedliche Distanzen in die Schätzung unterschiedlich ein. Der Einfluss der Distanzgewichtung über den Distanzgewichtungsexponent kann fein gesteuert werden kann. Je höher der Distanzexponent, desto mehr beeinflussen die Messwerte der nächst benachbarten Punkte das Ergebnis. Eine richtungsabhängige Gewichtung ist jedoch nicht möglich. Bei der geostatistischen Interpolation dient die Variographie als Grundlage auf der Hauptbasis von statistisch verteilten Parametern. Aus der Variogramm erhält man Hinweise auf die Ähnlichkeit benachbarter Datenpunkte in Abhängigkeit ihrer Entfernung zueinander. Die wichtigsten geostatistischen Interpolationsverfahren sind die Kriging-Methoden.

1.5. Bibliographie

- **Mitas, L.; Mitasova, H.** (1998). *Multidimensional Spatial Interpolation* [online]. Urbana-Champaign: University of Illinois. Available from: http://www4.ncsu.edu/~hmitaso/gmslab/viz/sinter.html [Accessed 25.04.2016].
- **Mitas, L.; Mitasova, H.**, 1999. Spatial Interpolation. *In:* **Longley, P. A.; Goodchild, M. F.; Maguire, D. J.; Rhind, D. W.**, ed. *Geographical Information Systems (GIS): Principles, Techniques, Management and Applications*. New York: John Wiley, 481–492.
  Herunterladen: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.224.5959&rep=rep1&type=pdf
- **Shepard, D.**, 1968. A two-dimensional interpolation function for irregularly-spaced data. *In: Proceedings of the 23rd ACM National Conference*. New York: Association for Computing Machinery, 517–523.
- **Tobler, W.**, 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, 46 (2), 234–240.
- **Wikipedia**. *Kriging* [online]. Available from: http://en.wikipedia.org/wiki/Kriging [Accessed 25.04.2016].
- **Wyatt, P.**, 2000. The Interpolation Process. *Directions Magazine*, 18, .
  Herunterladen: http://www.directionsmag.com/article.php?article_id=52