

Spline n-D / Radial Basis Functions (RBFs)

Wenn man annimmt, dass ein regelmäßiges Gitter von Knoten vorliegt, lässt sich der 1D-Ansatz zu Splines auf beliebige Dimensionen erweitern. Im Allgemeinen liegen Daten aber nicht auf einem solchen Gitter vor. Hier müssen andere Ansätze verwendet werden. Eine Option hier sind so genannte radiale Basisfunktionen (*radial basis functions* / RBF).

Liegen die Werte y_i an den Punkten $x_i \in \mathbb{R}^2$ mit $i = 1, \dots, N$ vor, dann kann man mit einem so genannten **thin plate spline** der Form

$$S(x) = \sum_{i=1}^N \lambda_i \Phi(|x - x_i|)$$

mit $\Phi(r) = r^2 \ln(r)$, welche

$$S(x_i) = y_i, i = 1, \dots, N$$

interpoliert, ist die Interpolationsfunktion, welche folgenden Ausdruck minimiert:

$$\min_{f \text{ interpoliert } P} \int_D \left[|\partial_{x_1}^2 f(x)|^2 + |\partial_{x_1} \partial_{x_2} f(x)|^2 + |\partial_{x_2}^2 f(x)|^2 \right] dx.$$

Spline n-D / Radial Basis Functions (RBFs)

Jede Funktion $\Phi: [0, \infty) \rightarrow \mathbb{R}$ die zu einer Funktion der Form $\phi(x) = \Phi(|\mathbf{x}|)$ führt, die nur von der Länge eines Vektors abhängt, kann als **radiale Basisfunktion** bezeichnet werden. Typischerweise sind weitere „gutartige“ Eigenschaften wie z.B. positive Definitheit, gewünscht. Eine Approximationsfunktion basierend auf diesem Ansatz ist immer der Form

$$S(x) = \sum_{i=1}^N \lambda_i \phi_i(x)$$

mit $\phi_i(x) = \Phi(|x - x_i|)$ und $S(x_i) = y_i, i = 1, \dots, N$.

Spline n-D / Radial Basis Functions (RBFs)

Jede Funktion $\Phi: [0, \infty) \rightarrow \mathbb{R}$ die zu einer Funktion der Form $\phi(x) = \Phi(|x|)$ führt, die nur von der Länge eines Vektors abhängt, kann als **radiale Basisfunktion** bezeichnet werden. Typischerweise sind weitere „gutartige“ Eigenschaften wie z.B. positive Definitheit, gewünscht. Eine Approximationsfunktion basierend auf diesem Ansatz ist immer der Form

$$S(x) = \sum_{i=1}^N \lambda_i \phi_i(x)$$

mit $\phi_i(x) = \Phi(|x - x_i|)$ und $S(x_i) = y_i, i = 1, \dots, N$.

Häufige verwendete positiv definite Basisfunktionen sind z.B.

$$\Phi(r) = e^{-a^2 r^2}, \text{ (Gaussian)}$$

$$\Phi(r) = (r^2 + a^2)^{\frac{1}{2}}, \text{ (Multiquadric)}$$

$$\Phi(r) = \frac{1}{r^2 + a^2}, \text{ (Inverse Quadric)}$$

$$\Phi(r) = r^2 \ln(r), \text{ (thin plate)}$$

Andere mögliche RBFs (z.B. mit kompaktem Support) sind komplexer.

Spline n-D / Radial Basis Functions (RBFs)

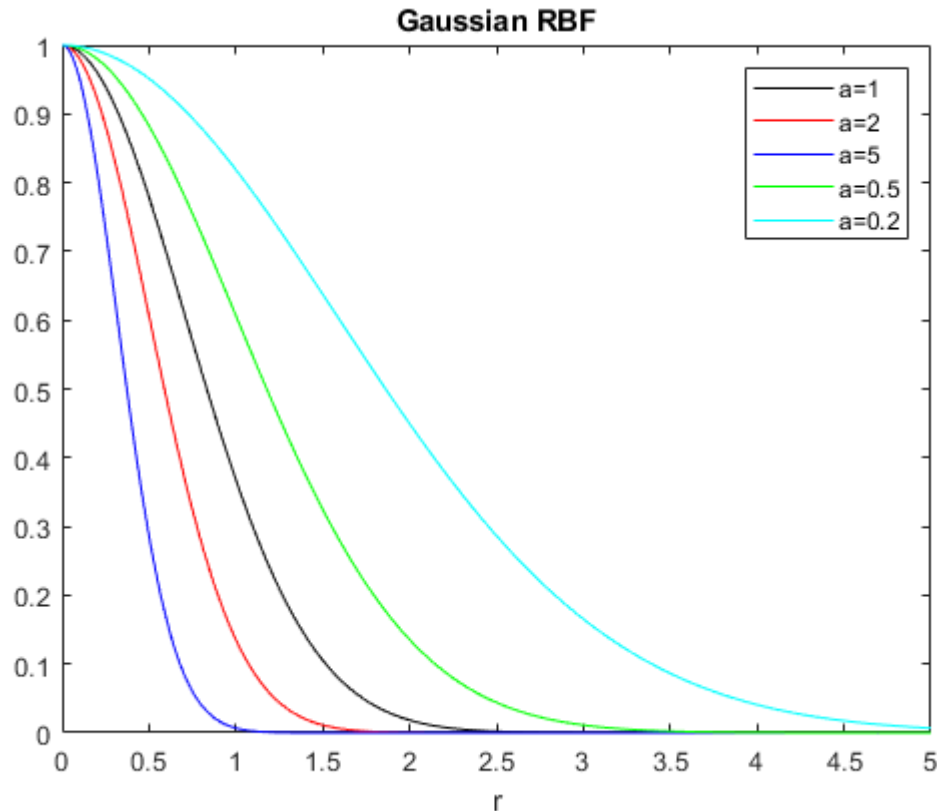
Jede Funktion $\Phi: [0, \infty) \rightarrow \mathbb{R}$ die zu einer Funktion der Form $\phi(x) = \Phi(|x|)$ führt, die nur von der Länge $|x|$ abhängt, wird als Radial Basis Funktion (RBF) bezeichnet. RBFs sind typischerweise glatt und positiv definit, was für viele Anwendungen in der Approximation und der Interpolation erforderlich ist.

Typischerweise
Eine Approximation

mit $\phi_i(x)$

Häufige Werte

Andere Methoden



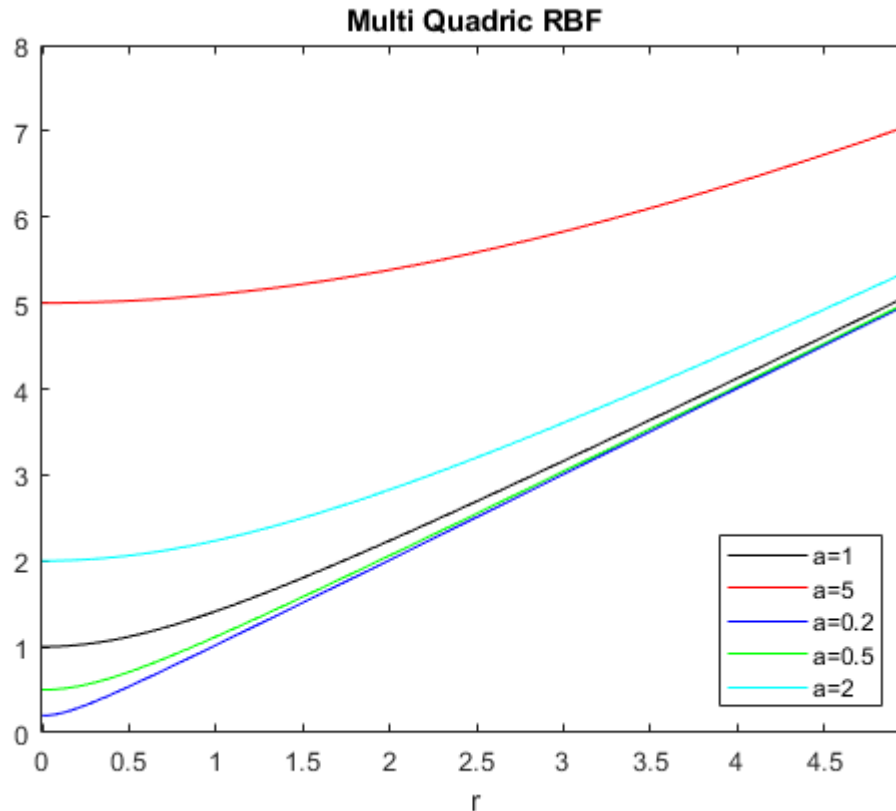
Spline n-D / Radial Basis Functions (RBFs)

Jede Funktion $\Phi: [0, \infty) \rightarrow \mathbb{R}$ die zu einer Funktion der Form $\phi(x) = \Phi(|x|)$ führt, die nur von der Länge eir
 Typischerwei
 Eine Approxir

mit $\phi_i(x) = \Phi$

Häufige verw

Andere mögli



Spline n-D / Radial Basis Functions (RBFs)

Zur Bestimmung der notwendigen Koeffizienten $\lambda_1, \dots, \lambda_m$ muss wieder ein lineares Gleichungssystem der Form

$$\mathbf{M}\boldsymbol{\lambda} = \mathbf{y}$$

mit $\boldsymbol{\lambda} = (\lambda_0, \dots, \lambda_m)^T$, $\mathbf{y} = (y_1, \dots, y_N)$ und

$$\mathbf{M} = \begin{pmatrix} \Phi(0) & \Phi(|x_1 - x_2|) & \Phi(|x_1 - x_3|) & \cdots & \Phi(|x_1 - x_N|) \\ \Phi(|x_2 - x_1|) & \Phi(0) & \Phi(|x_2 - x_3|) & \cdots & x_2^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Phi(|x_N - x_1|) & \Phi(|x_N - x_2|) & \Phi(|x_N - x_3|) & \cdots & \Phi(0) \end{pmatrix}$$

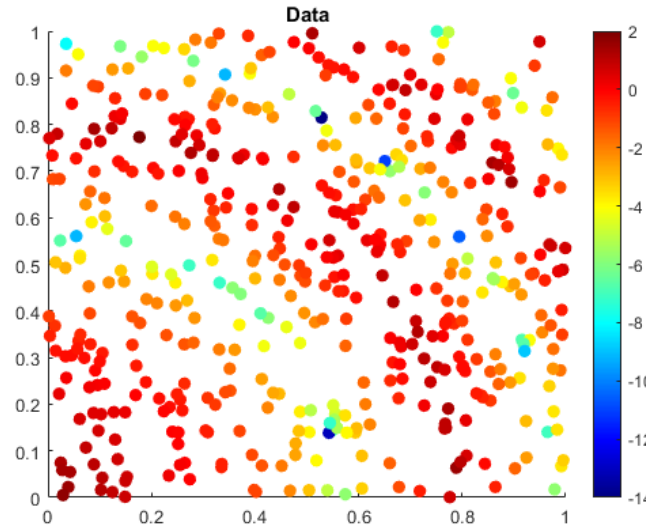
Die (eindeutige) Lösbarkeit und damit die Qualität der Interpolation hängen sehr stark vom gewählten Typ der RBF ab.

Spline n-D / Radial Basis Functions (RBFs)

Zur Bestimmung der r
Gleichungssystem de

mit $\lambda = (\lambda_0, \dots, \lambda_m)^T, \mathcal{J}$

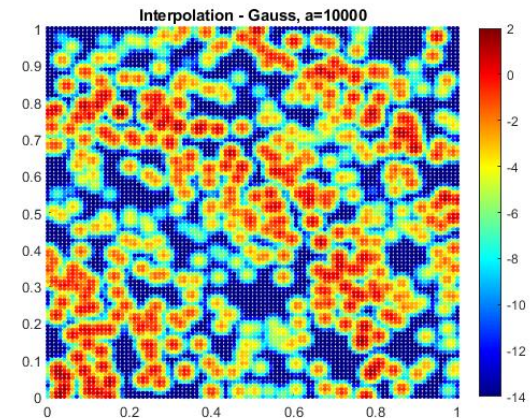
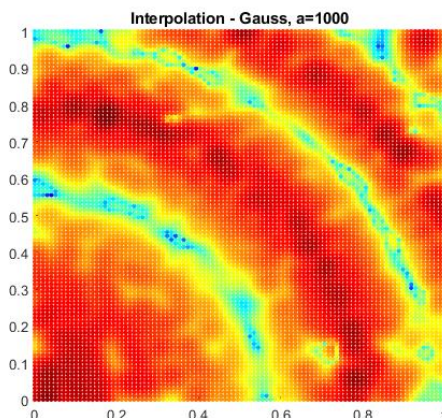
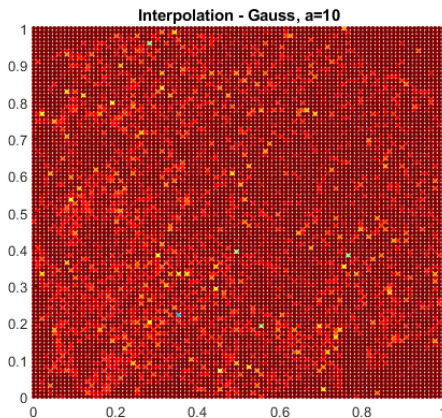
$$M = \begin{pmatrix} \Phi(|x_1 - x_1|) \\ \Phi(|x_2 - x_1|) \\ \vdots \\ \Phi(|x_N - x_1|) \end{pmatrix}$$



ler ein lineares

$$\begin{pmatrix} 1 - x_N \\ x_2^m \\ \vdots \\ \mathbb{P}(0) \end{pmatrix}$$

Die (eindeutige) Lösbarkeit und damit die Qualität der Interpolation hängen sehr stark



Spline n-D / Radial Basis Functions (RBFs)

Zur Bestimmung der notwendigen Koeffizienten $\lambda_1, \dots, \lambda_m$ muss wieder ein lineares Gleichungssystem der Form

$$\mathbf{M}\boldsymbol{\lambda} = \mathbf{y}$$

mit $\boldsymbol{\lambda} = (\lambda_0, \dots, \lambda_m)^T$, $\mathbf{y} = (y_1, \dots, y_N)$ und

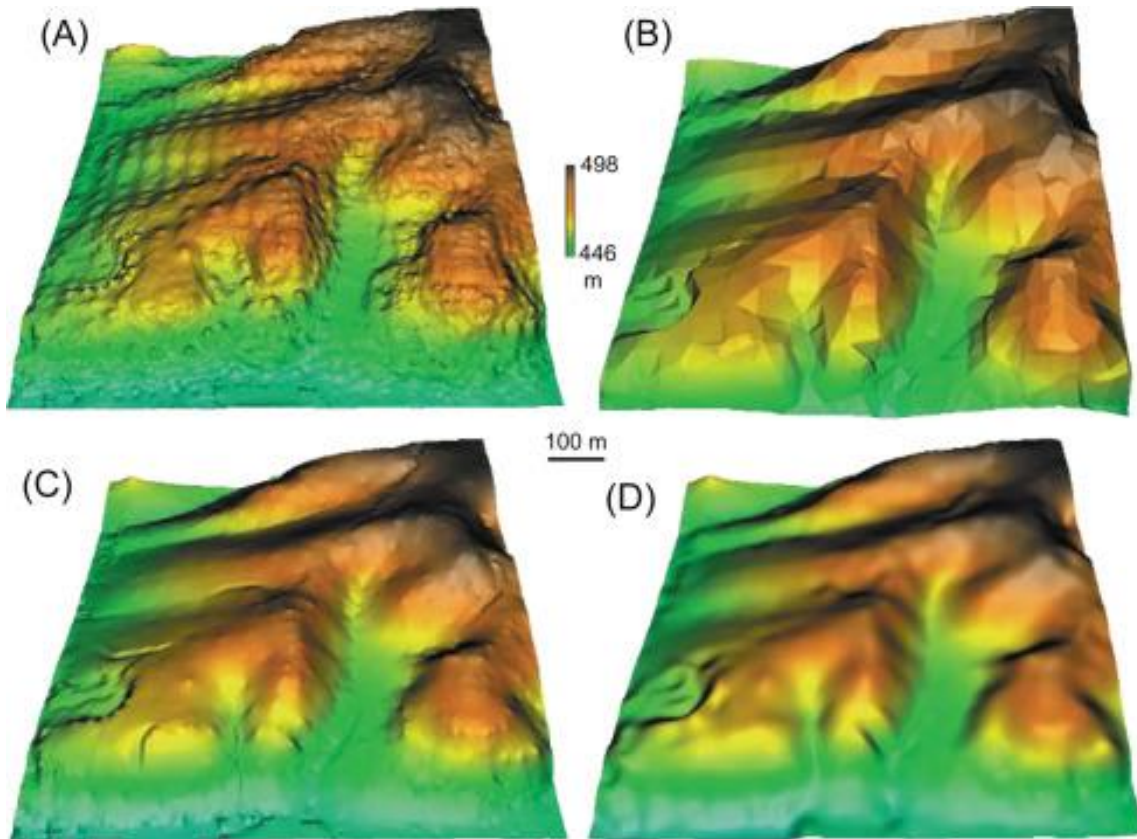
$$\mathbf{M} = \begin{pmatrix} \Phi(0) & \Phi(|x_1 - x_2|) & \Phi(|x_1 - x_3|) & \cdots & \Phi(|x_1 - x_N|) \\ \Phi(|x_2 - x_1|) & \Phi(0) & \Phi(|x_2 - x_3|) & \cdots & x_2^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Phi(|x_N - x_1|) & \Phi(|x_N - x_2|) & \Phi(|x_N - x_3|) & \cdots & \Phi(0) \end{pmatrix}$$

Die (eindeutige) Lösbarkeit und damit die Qualität der Interpolation hängen sehr stark vom gewählten Typ der RBF ab.

Eigenschaften:

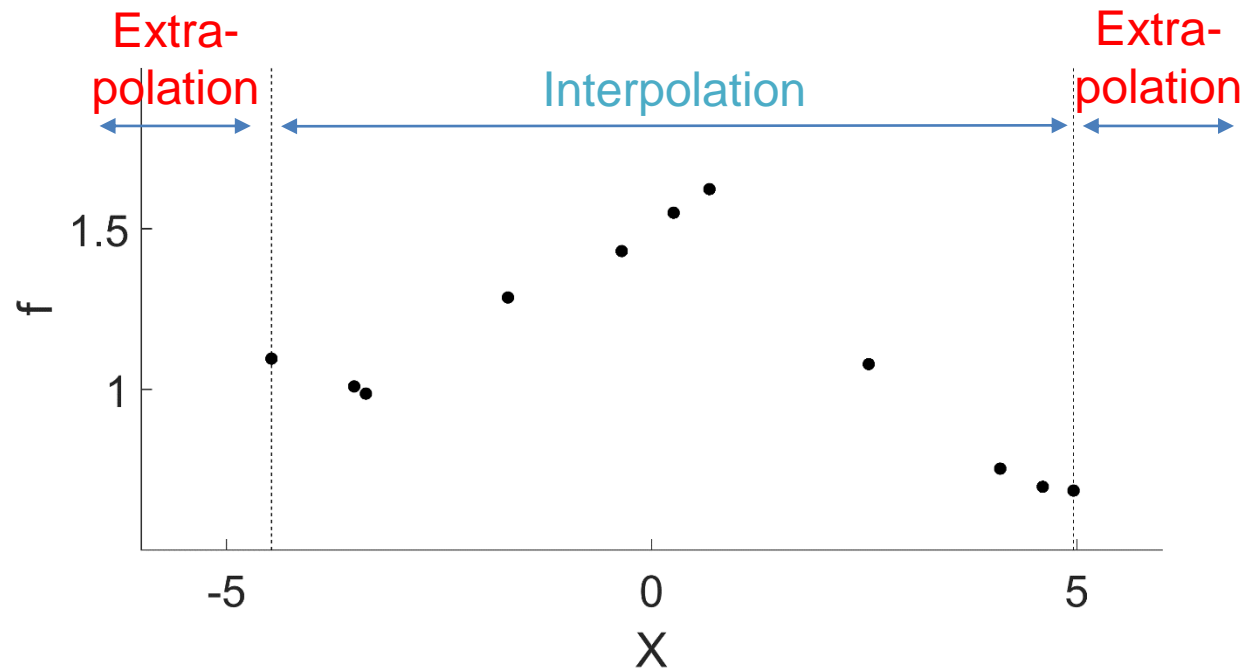
- $f(x)$ ist stetig und differenzierbar
- Globaler Interpolator
- Extrapolation ist konzeptionell möglich

Example IDW, Lin. Interp. on Triangulation, Splines



Interpolation vs Extrapolation

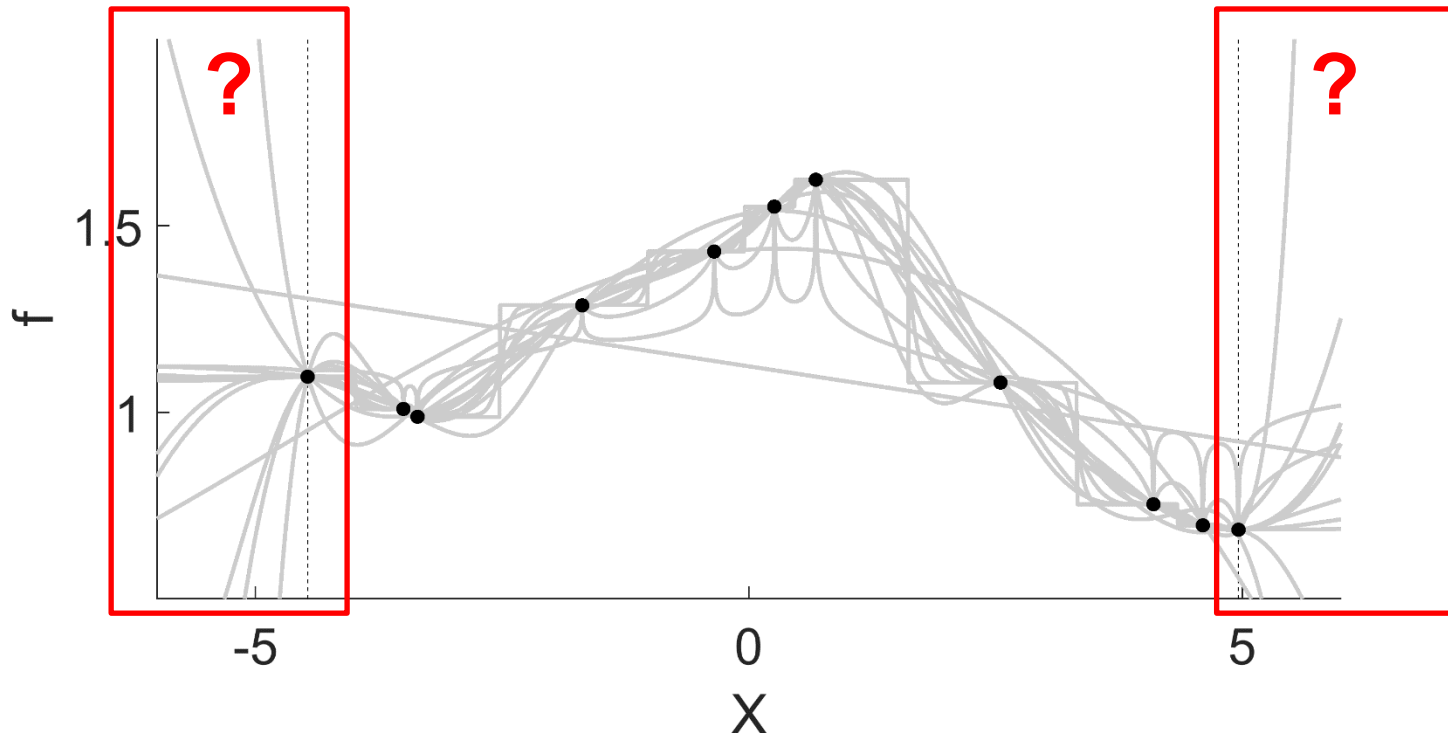
- **Interpolation:** Vorhersage **innerhalb** der Daten / **zwischen** den Datenpunkten
- **Extrapolation:** Vorhersage **außerhalb** der Daten (zumeist außerhalb der konvexen Hülle)



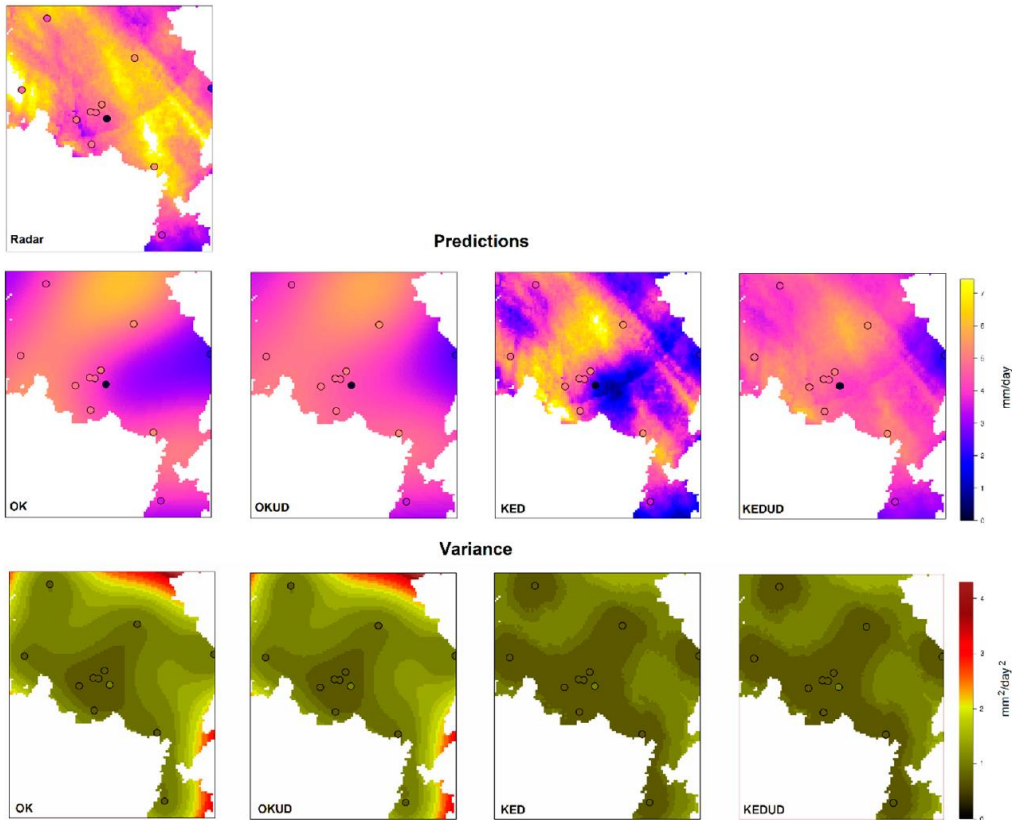
Interpolation vs Extrapolation

Was ist das Problem bei Extrapolation?

- Vorhersage basierend auf ungenügender Datengrundlage
- **Vorhersage rein Modell-basiert, zumeist nicht plausibel**

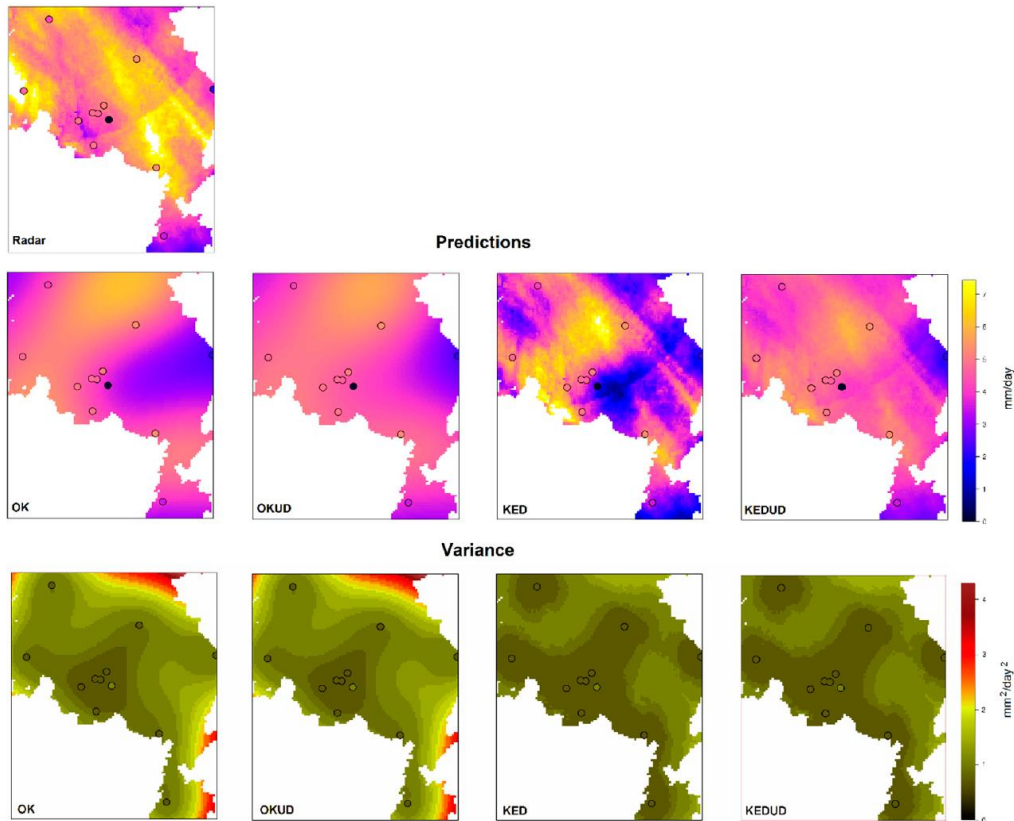


- Kriging (Simple, Ordinary, Universal)
- Conditional Simulation



Cecinati et al.: Considering Rain Gauge Uncertainty Using Kriging for Uncertain Data, Atmosphere (2018)

- Kriging (Simple, Ordinary, Universal)
- ~~Conditional Simulation~~



Cecinati et al.: Considering Rain Gauge Uncertainty Using Kriging for Uncertain Data, Atmosphere (2018)

Kriging

Der Messwert Z_i an einer Position $x_i \in \mathbb{R}^n$ kann als Zufallsvariable mit **Erwartungswert**

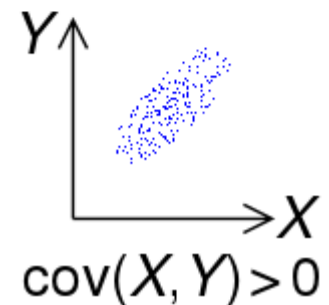
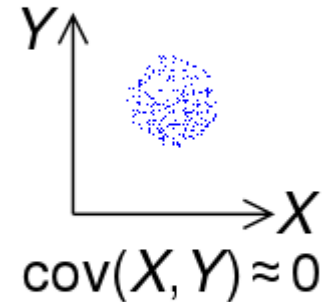
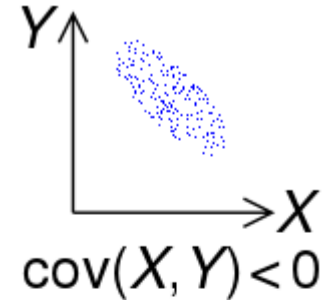
$$E[Z_i] = \int_{\Omega} z d\mu_{Z_i}(z)$$

und **Varianz**

$$\text{Var}(Z_i) = E[(Z_i - E[Z_i])^2].$$

Die **Kovarianz** zwischen zwei Zufallsvariablen ist gegeben durch

$$\text{Cov}(Z_i, Z_j) = E[(Z_i - E[Z_i])(Z_j - E[Z_j])].$$



Kriging

Der Messwert Z_i an einer Position $x_i \in \mathbb{R}^n$ kann als Zufallsvariable mit **Erwartungswert**

$$E[Z_i] = \int_{\Omega} z d\mu_{Z_i}(z)$$

und **Varianz**

$$\text{Var}(Z_i) = E[(Z_i - E[Z_i])^2].$$

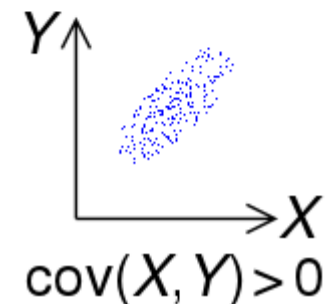
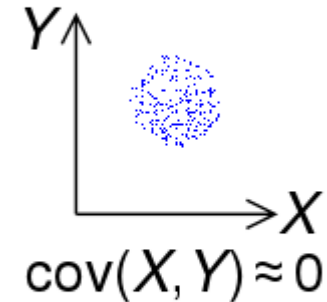
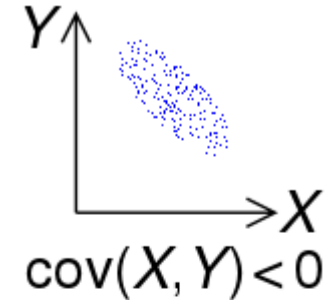
Die **Kovarianz** zwischen zwei Zufallsvariablen ist gegeben durch

$$\text{Cov}(Z_i, Z_j) = E[(Z_i - E[Z_i])(Z_j - E[Z_j])].$$

In der Geostatistik ist die so genannte **Stationaritäts*-Annahme** von entscheidender Bedeutung:

$$E[Z(x)] = m, \quad \text{Cov}(Z(x + h), Z(x)) = \text{Cov}(h).$$

Die Kovarianz hängt nur vom Inkrement h zwischen Positionen und nicht von den Positionen selbst ab.



* stationarity, homogeneity

Die Näherung Z^* von Z ist gegeben durch

$$Z^*(x) = \lambda_0 + \sum_{i=1}^N \lambda_i(x) Z_i,$$

mit den unbekanntenen Gewichten $\lambda_i(x)$ so dass

$$Z^*(x_i) = z_i, \quad i = 1, \dots, N.$$

Die Näherung Z^* von Z ist gegeben durch

$$Z^*(x) = \lambda_0 + \sum_{i=1}^N \lambda_i(x) Z_i,$$

mit den unbekanntenen Gewichten $\lambda_i(x)$ so dass

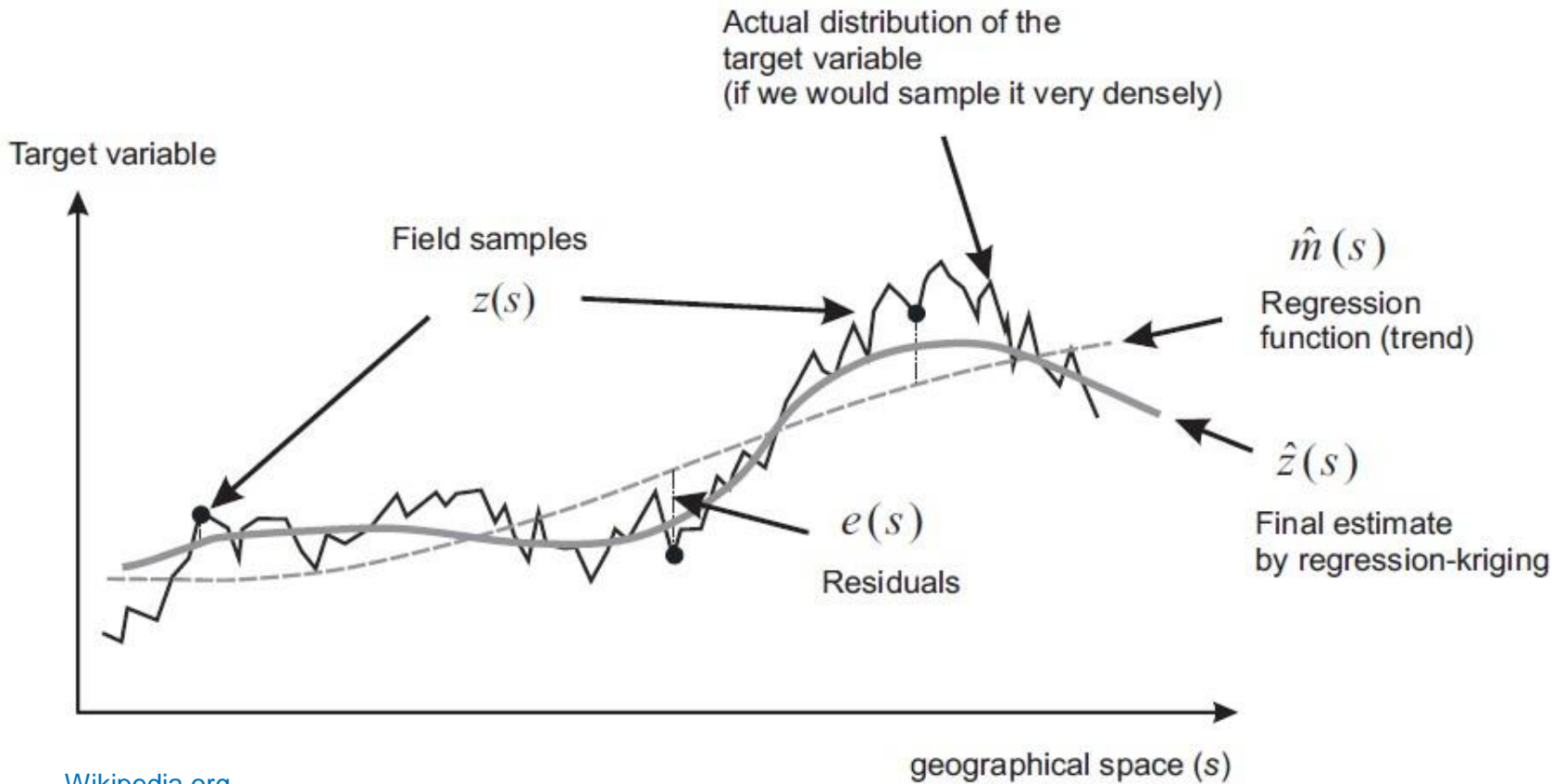
$$Z^*(x_i) = z_i, \quad i = 1, \dots, N.$$

Welche Annahmen kann man zur Bestimmung der Gewichte treffen?

Unbiasedness (unverfälscht, ideal): $E[Z^*(x)] = E[Z(x)]$

Varianz-Minimierung: $\min_{Z^*} \text{Var}(Z^*(x) - Z(x))$

Kriging



Wikipedia.org

Kriging

Simple Kriging: Erwartungswert $m_0 = E(Z(x))$ ist bekannt; Unbiasedness-Annahme führt zu

$$\lambda_0 = m_0 - \sum_{i=1}^N \lambda_i m_0.$$

Kriging

Simple Kriging: Erwartungswert $m_0 = E(Z(x))$ ist bekannt; Unbiasedness-Annahme führt zu

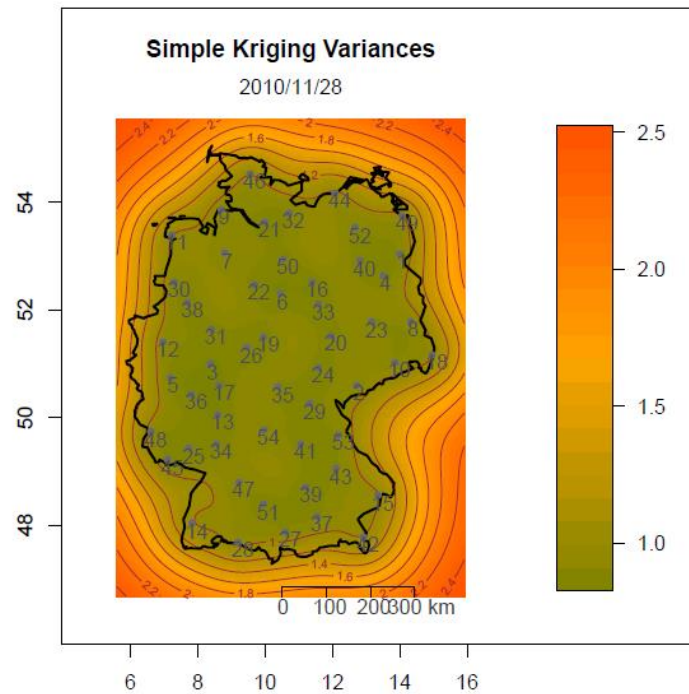
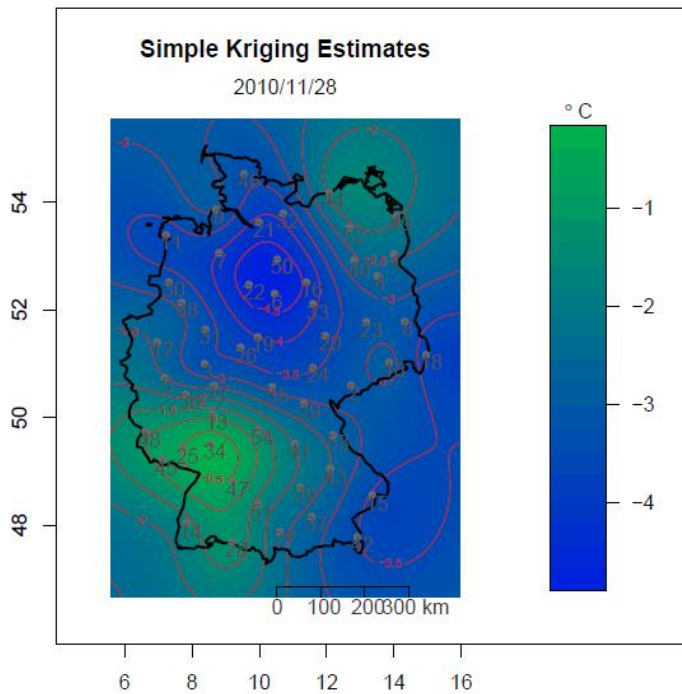
$$\lambda_0 = m_0 - \sum_{i=1}^N \lambda_i m_0.$$

Varianz-Minimierung führt zu:

$$\sum_{i=1}^N \lambda_i(x) \text{Cov}(Z_i, Z_k) = \text{Cov}(Z_i, Z(x)), \quad k = 1, \dots, N$$

Aufgrund der Stationaritäts-Annahme gilt: $\text{Cov}(Z_i, Z(x)) = \text{Cov}(x_i - x) \rightarrow$ Ähnlichkeit zu radialen Basisfunktionen.

Kriging



Example from the Bachelor thesis "Kriging methods in spatial statistics", Andreas Lichtenstern, TU München

Kriging

Im Allgemeinen ist $\text{Cov}(Z(x_i), Z(x_j))$ nicht bekannt und muss geschätzt oder modelliert werden.

Dafür spielt das so genannte **Variogramm** eine entscheidende Rolle:

$$\gamma(h) = \frac{1}{2} \text{Var}(Z(x) - Z(x + h)) = C(0) - C(h).$$

Das Variogramm kann unter bestimmten Bedingungen aus den Daten geschätzt werden.

Im Allgemeinen ist $\text{Cov}(Z(x_i), Z(x_j))$ nicht bekannt und muss geschätzt oder modelliert werden.

Dafür spielt das so genannte **Variogramm** eine entscheidende Rolle:

$$\gamma(h) = \frac{1}{2} \text{Var}(Z(x) - Z(x+h)) = C(0) - C(h).$$

Das Variogramm kann unter bestimmten Bedingungen aus den Daten geschätzt werden.

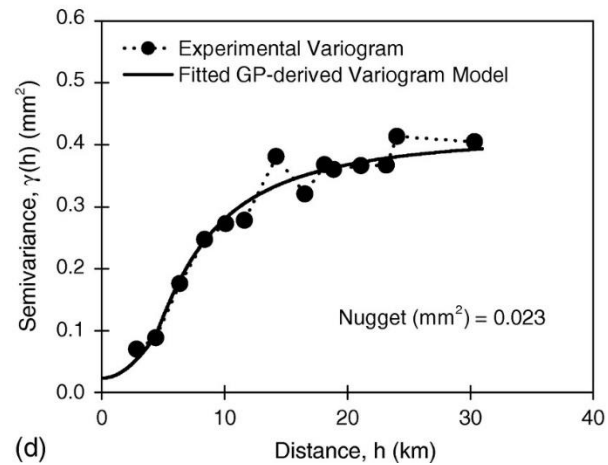
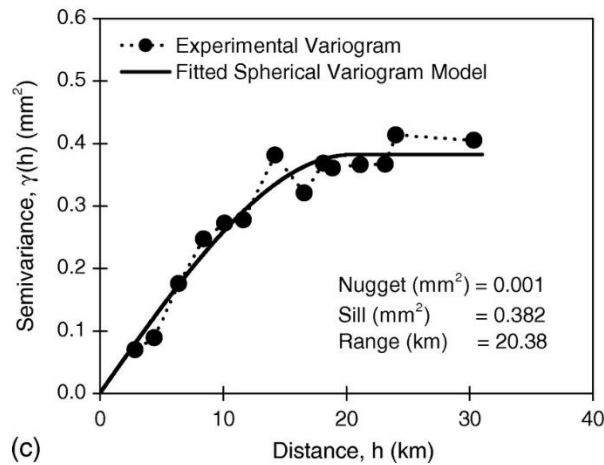
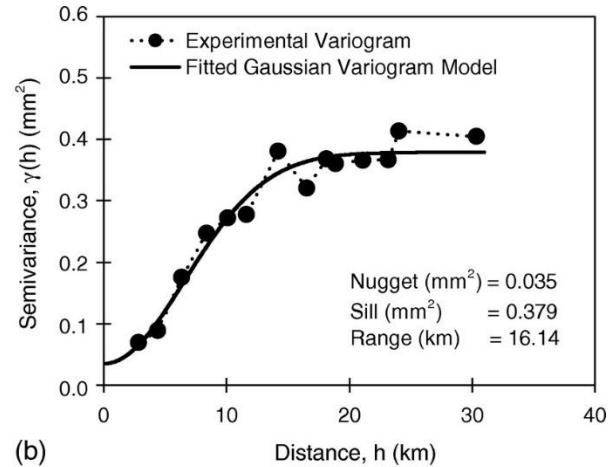
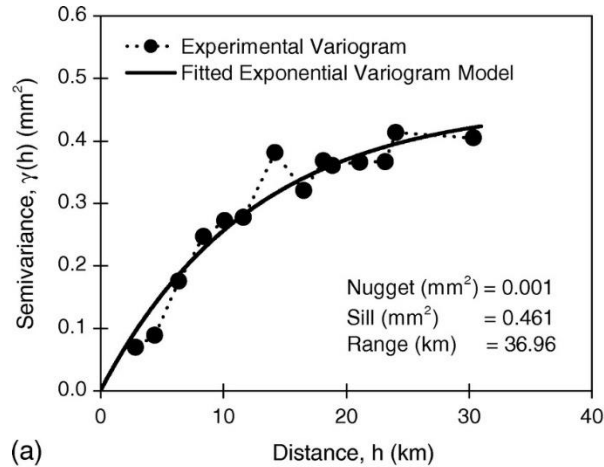
Empirisches Variogramm:

$$\hat{\gamma}(h) = \frac{1}{N_h} \sum_{x_i - x_j = h} (Z_i - Z_j)^2.$$

An das empirischen Variogramms kann ein analytisches Variogrammmodell angepasst werden. Diese Modellfunktion muss dabei bestimmte Eigenschaften erfüllen.

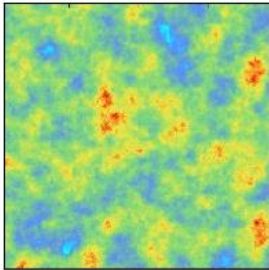
Sie muss z.B. semi-positiv definit sein, d. H. alle Eigenwerte sind größer gleich Null.

Kriging

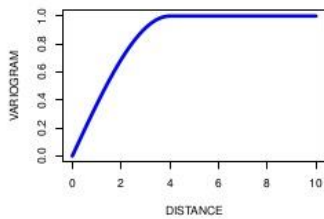


Kriging

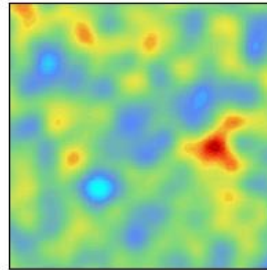
Rough



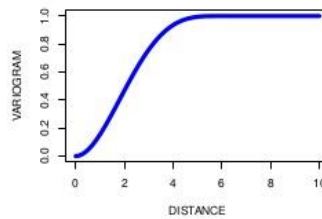
Spherical model



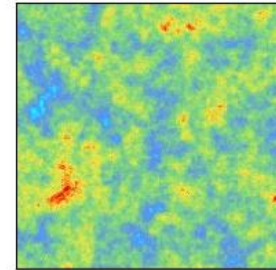
Smooth



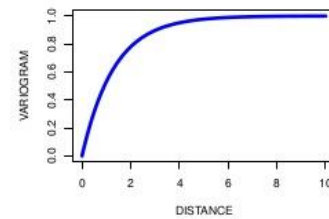
Cubic model



Rough

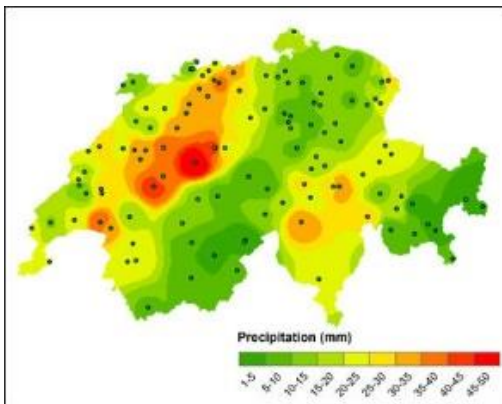
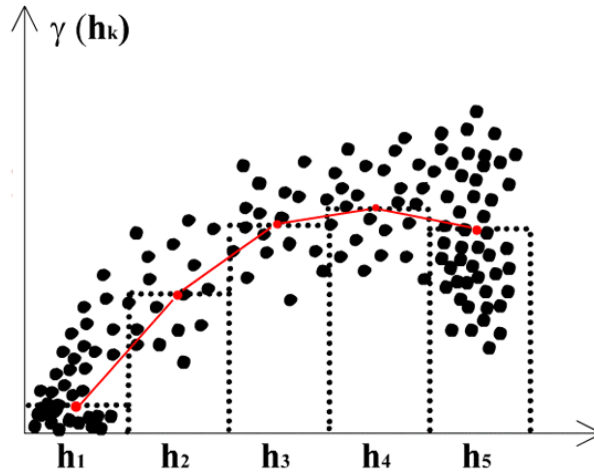


Exponential model

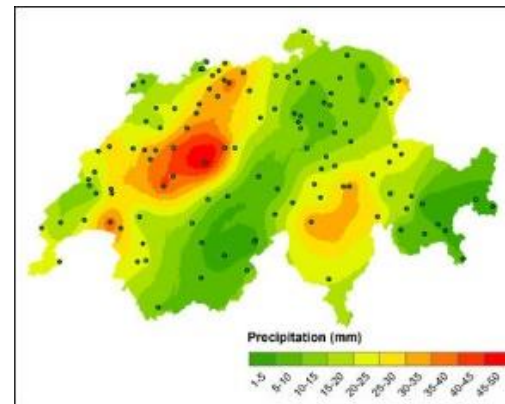


From a presentation by H. Wackernagel

Example IDW, Kriging

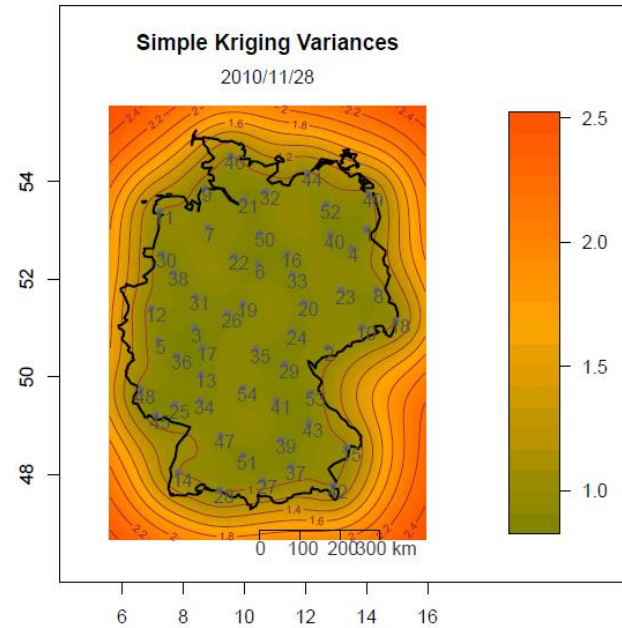
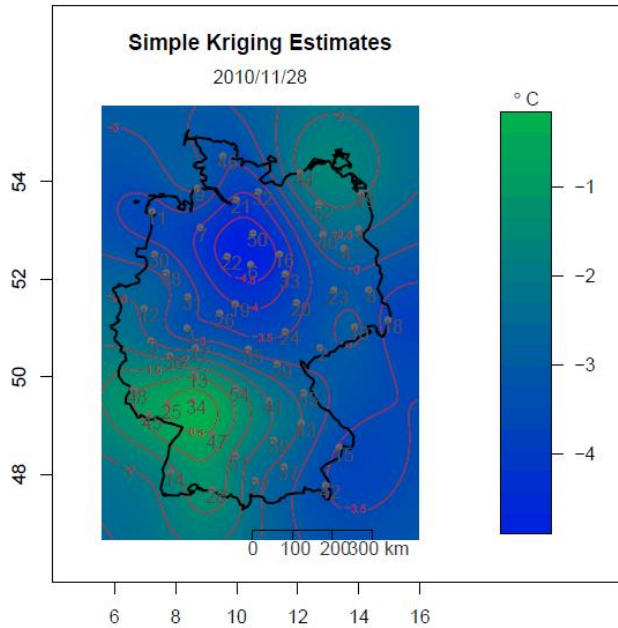


IDW



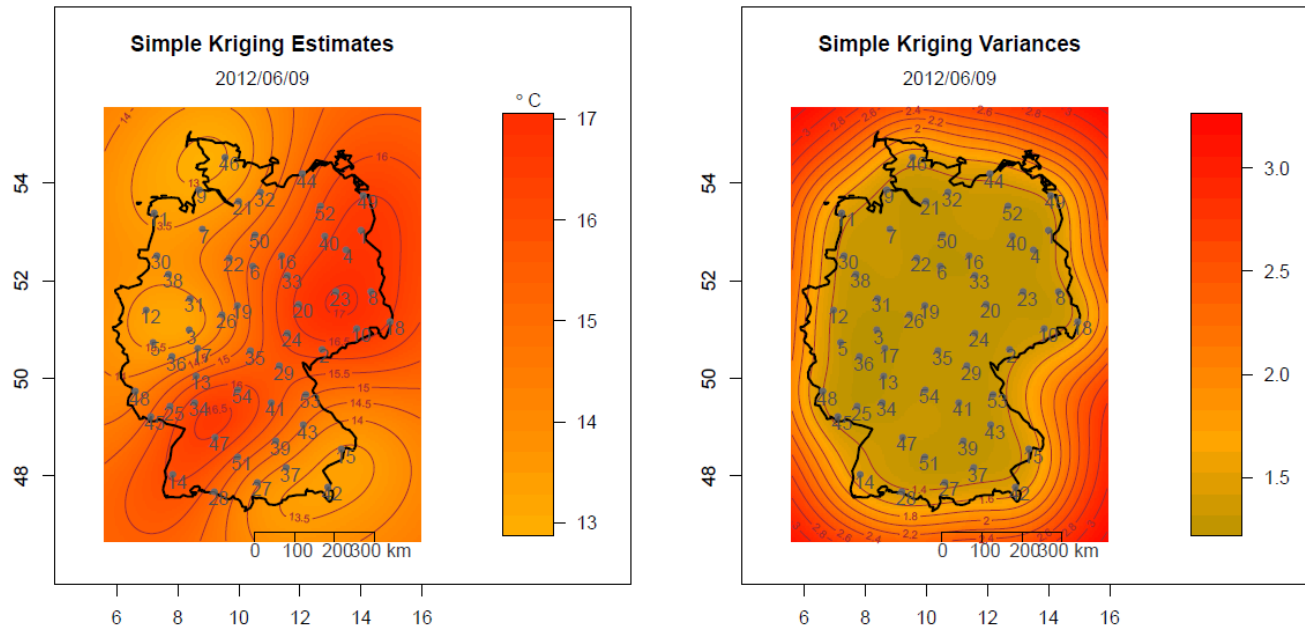
Kriging

Example Kriging



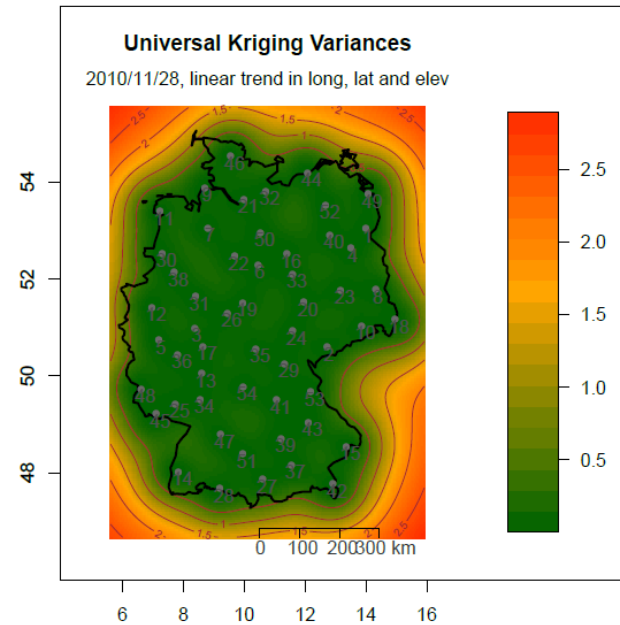
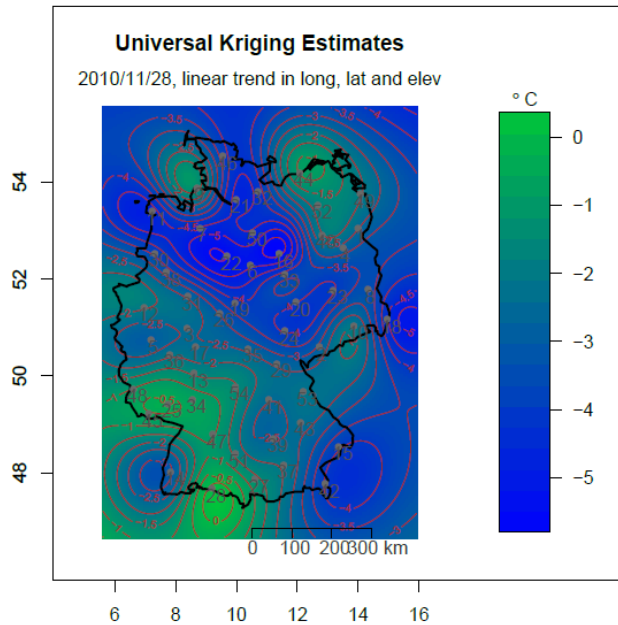
Example from the Bachelor thesis "Kriging methods in spatial statistics", Andreas Lichtenstern, TU München

Example Kriging



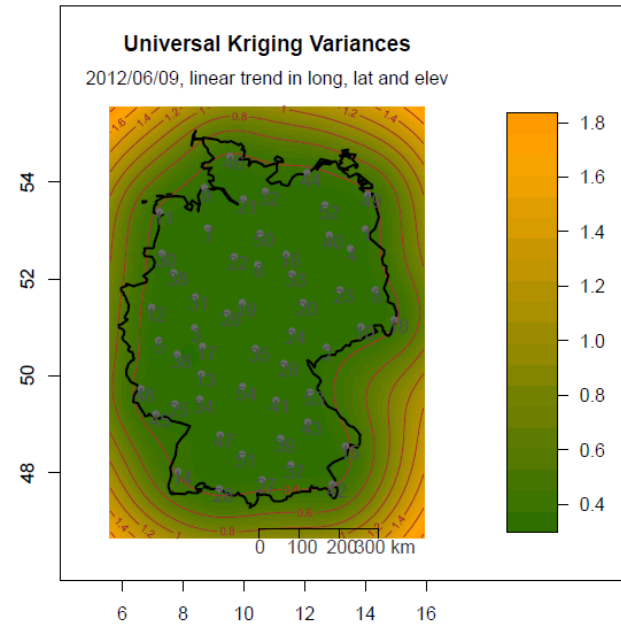
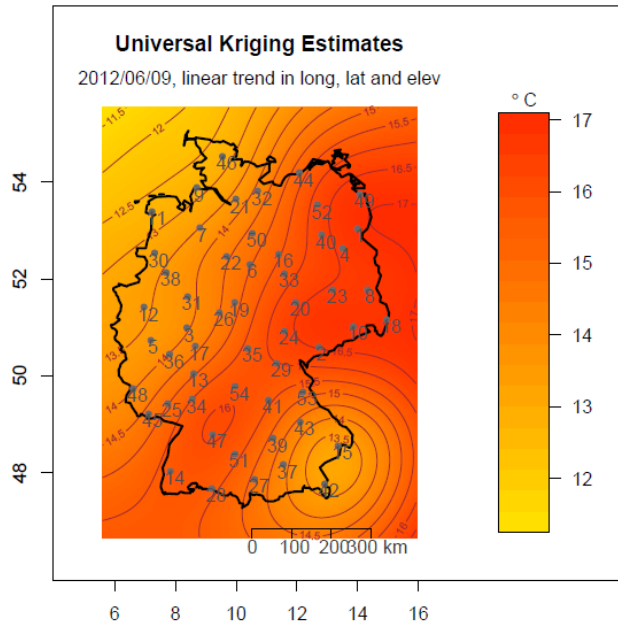
Example from the Bachelor thesis "Kriging methods in spatial statistics", Andreas Lichtenstern, TU München

Example Kriging



Example from the Bachelor thesis "Kriging methods in spatial statistics", Andreas Lichtenstern, TU München

Example Kriging



Example from the Bachelor thesis "Kriging methods in spatial statistics", Andreas Lichtenstern, TU München

Ausgangslage:

Gegeben sind bekannte Werte y_i (z.B. Messungen) für einen Parameter an gegebenen Positionen x_i . An beliebigen Positionen x sollen die ggf. unbekanntes Werte y vorhergesagt werden.

Interpolation:

Interpolation umfasst alle mathematischen Verfahren, welche eine Funktion $f(x) = y$ für eine solche Vorhersage verwenden. Klassischerweise hat bei einer linearen Interpolation $f(x)$ die Form eines gewichteten Mittels der gegebenen Werte y_i :

$$f(x) = \sum_i \lambda_i(x) \cdot y_i$$

mit den Gewichten $\lambda_i(x)$. Die Bestimmung der Gewichte hängt vom gewählten Verfahren ab.

Approximation und Interpolation:

Für eine klassische (exakte) Interpolation gilt, dass das Verfahren die Daten exakt reproduziert mit

$$f(x_i) = y_i$$

Bei einer Approximation (oder auch nicht-exakte Interpolation) ist dies nicht der Fall, die Daten werden hier nur angenähert mit

$$f(x_i) = y_i + \epsilon_i \rightarrow f(x_i) \approx y_i$$

ϵ_i ist dabei der unbekannte Approximationsfehler.

Interpolation und Extrapolation:

Findet die Vorhersage innerhalb der Daten statt, handelt es sich um eine Interpolation. Die Ergebnisse können zumeist als plausibel angesehen werden, da genügend Daten als Randbedingungen für die Vorhersage zu Verfügung stehen.

Findet die Vorhersage außerhalb der Daten oder sehr weit entfernt statt, handelt es sich um eine Extrapolation. Die Ergebnisse können zumeist nicht als plausibel angesehen werden und sollte sehr kritisch interpretiert werden, da die Daten nur sehr eingeschränkt als Randbedingungen verwendet werden können.

Lokale und globale Interpolation:

Bei einer lokalen Interpolationen werden nicht alle Datenwerte für die Interpolation verwendet. Dies bedeutet, dass es Datenwerte y_i gibt, für das Interpolationsgewicht gleich Null ist:

$$\exists i: \lambda_i(x) = 0$$

Bei einer globalen Interpolationen werden immer alle gegebenen Werte berücksichtigt, d.h. dass es keine Interpolationsgewichte gibt, welche Null sind:

$$\forall i: \lambda_i(x) \neq 0$$

Gitterfreie und gitterbasierte Interpolation:

Benötigt ein Interpolationsverfahren eine Vermaschung der Datenpunkte wird es als *gitterbasiert* bezeichnet. Es ist dann zumeist ein lokales Verfahren, welches nur die Datenwerte verwendet, welche sich auf die Zelle beziehen, in der der Interpolationspunkt x liegt.

Gitterfreie (grid-free) Verfahren benötigen keine Vermaschung, sind aber zumeist globale Verfahren.

Deterministische Interpolation:

Die Bestimmung der Gewichte basiert auf einem festgelegten mathematischen Modell. Es ist zumeist nicht initial möglich, die „Unsicherheit“ der Vorhersage abzuschätzen.

(Geo-)statistische Interpolation (z.B. kriging):

Die Bestimmung der Gewichte basiert zusätzlich auf der bekannten oder abgeschätzten räumlichen Korrelation der Daten. Dies ist zumeist aufwändiger als deterministische Vorhersage, erlaubt aber die zusätzliche Abschätzung der „Unsicherheit“ der Vorhersage.

Wahl des Interpolationsverfahrens:

Bei der Wahl des Vorhersageverfahrens sollten folgende Fragen berücksichtigt werden:

1. Entspricht das mathematische Modell des Verfahrens dem vorherzusagenden Parameter bzgl.
 - Skala (diskret/kontinuierlich/...)
 - Glattheit / Differenzierbarkeit ...
 - ... ?
2. Ist Extrapolation für Anwendung notwendig?
3. Ist für die Anwendung exakte Interpolation notwendig oder reicht auch Approximation?
4. Welche Komplexität des Verfahrens ist für die Daten/Anwendung angemessen?
5. ...