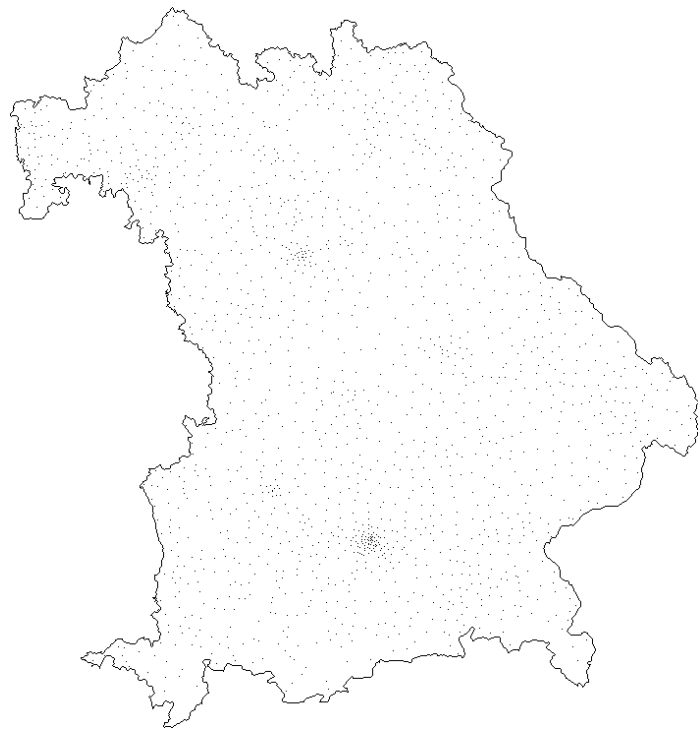
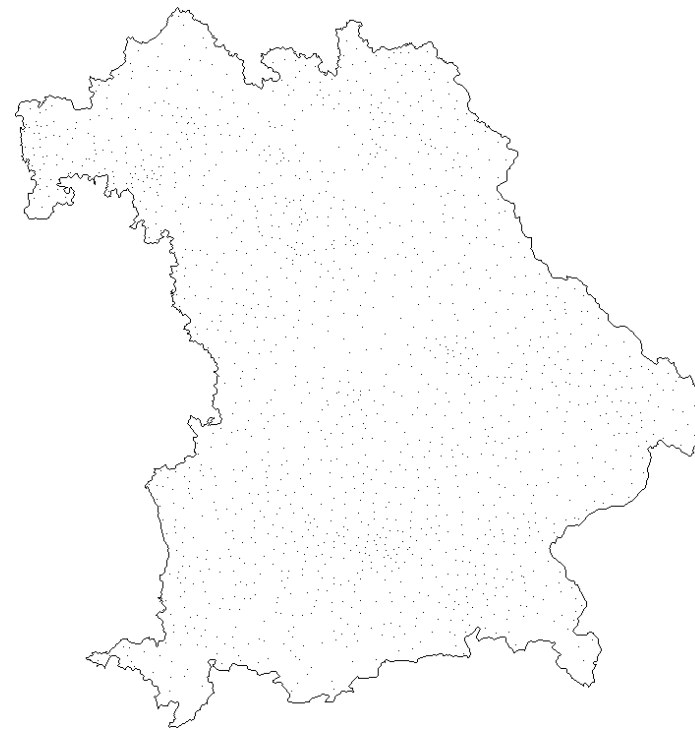

Kriging, Visualisierung und statistische Analyse von räumlich segmentierten Versicherungsdaten

- 1 – Motivation
- 2 – Notation
- 3 – Versicherungsdaten
- 4 – Ordinary Kriging (OK)
- 5 – Schätzung des Variogramms
- 6 – Variogramm-Modelle
- 7 – Variogrammanpassung
- 8 – Morphologische Operationen
- 9 – Adaptive morphologische Öffnung
- 10 – Zur Erinnerung: Statistische Signifikanztests
- 11 – Modellierung von Binärbildern als zufällige abgeschlossene Mengen
- 12 – Statistischer Signifikanztest von hypothetischen Flächenanteilen

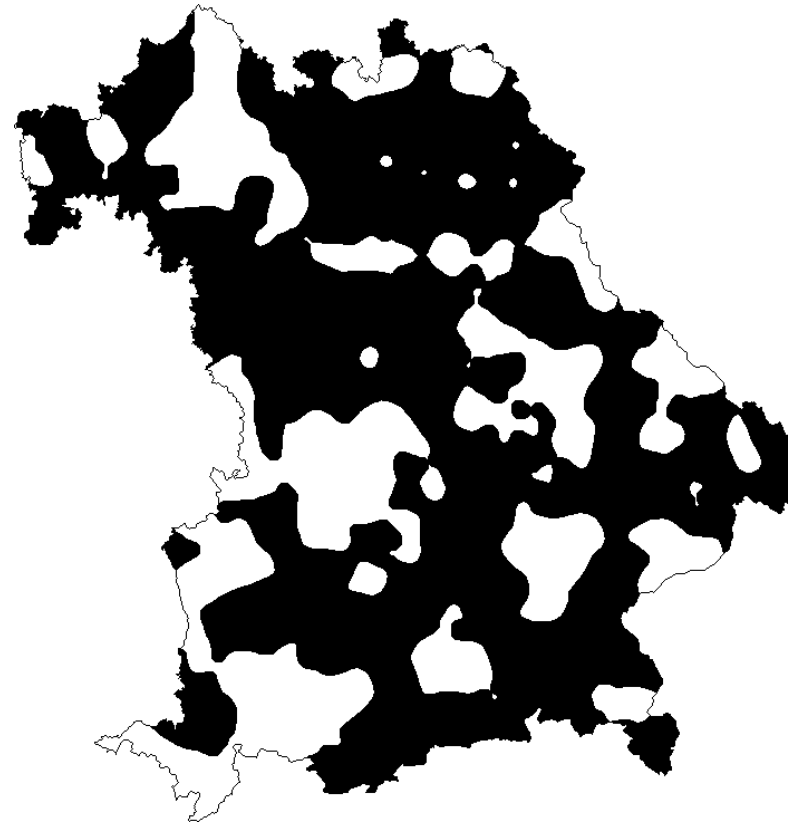
1 Motivation



Netz von „Messstellen“ innerhalb
Bayerns



Netz von „Messstellen“ innerhalb
Bayerns nach der Entclustering



Differenzenkarte

2 Notation

- Ein **Zufallsfeld** ist eine zufällige Funktion $\{Z(x, \omega) : x \in \mathbb{R}^2, \omega \in \Omega\}$, die sowohl von den Punkten $x \in \mathbb{R}^2$ als auch von den Elementarereignissen $\omega \in \Omega$ abhängt und zwar so, dass
 - $Z(x, \cdot)$ eine Zufallsvariable und
 - $Z(\cdot, \omega)$ eine Realisierung der zufälligen Funktion ist.
- Die Realisierung einer zufälligen Funktion wird in der Geostatistik als regionalisierte Variable aufgefasst und vereinfachend als $z(x)$ bezeichnet.
- Ein Zufallsfeld heißt **stationär (isotrop)**, wenn die Charakteristiken des Zufallsfeldes gegenüber Verschiebungen (Drehungen) invariant sind.
- Sei $W \subset \mathbb{R}^2$ das Beobachtungsfenster.
- Es gebe nun x_1, \dots, x_n **Messstellen** mit $x_i \in W$, $i = 1, \dots, n$, und $z(x_1), \dots, z(x_n)$ seien **Messwerte** in diesen Stellen.

3 Versicherungsdaten

Die Daten stammen aus der Kfz-Haftpflichtversicherung (VKB).

Insgesamt: 20.000 Datensätze.

Ein einzelner Datensatz besteht aus Angaben zu 40 Variablen.

- **Versicherungsbestand:** 10.000 Datensätze, die 1999 aus dem Gesamtbestand willkürlich ausgewählt wurden.
- **Stornobestand:** 10.000 Datensätze, die ebenfalls willkürlich aus dem Bestand der in Jahren 1996/1997/1998 stornierten Verträge ausgewählt wurden. Der Stornobestand wurde in Teilstornobestände mit dem Jahr der Kündigung 1996, 1997 und 1998 unterteilt. Die Teilstornobestände wurden zusätzlich nach dem Eigenheimstatus der Versicherungsnehmer sortiert: in Nichteigenheimbesitzer (NE) und Eigenheimbesitzer (E).

- Es hat eine Bereinigung der Daten stattgefunden mittels
 - Entfernung der fehlerhaften Datensätze (bezüglich der Variablen PLZ, STODAT1, KZEHB1),
 - Entclustering in den Großräumen Augsburg, München und Nürnberg, indem bestimmte Postleitzahlengebiete gestrichen wurden.
- Räumliche Informationen wurden aus der Postleitzahl mit Hilfe von Gauß-Krüger-Koordinaten gewonnen, die (x, y) -Koordinaten der so genannten Messstellen.
- Die Untersuchungen wurden bezüglich des Merkmals

$(\text{relative Häufigkeit je Einwohner}) \times 10.000$

durchgeführt.

4 Ordinary Kriging (OK)

Interpolationsverfahren, das nach dem südafr. Geostatistiker D. G. Krige benannt wurde.

- **Ziel:** Schätzung des Merkmals $Z(x_0)$ an Stellen $x_0 \in W$, an denen nicht gemessen wurde, wobei die Information über die räumliche Korrelation der Messdaten genutzt werden soll.
- **Modellannahmen:** Es liegt eine Realisierung eines **intrinsisch 2. Ordnung stationären** Zufallsfeldes vor, d.h.,

$$\begin{aligned}\mathbb{E}(Z(x+h) - Z(x)) &= 0, & \forall x, x+h \in \mathbb{R}^2, \\ \text{Var}(Z(x+h) - Z(x)) &= \mathbb{E}\left((Z(x+h) - Z(x))^2\right) = 2\gamma(h), & \forall x, x+h \in \mathbb{R}^2.\end{aligned}$$

Sei das (theoretische) Variogramm $\gamma(h) = \frac{1}{2}\mathbb{E}\left((Z(x+h) - Z(x))^2\right)$ bekannt. Außerdem sei $\mathbb{E}(Z(x)) = \mu$, μ unbekannt, $\forall x \in \mathbb{R}^2$.

- Die Form des Schätzers ist gegeben durch

$$\hat{Z}(x_0) = \sum_{i=1}^n g_i Z(x_i),$$

wobei g_i (noch zu ermittelnde) Gewichte sind.

- Der Schätzer soll folgende Güteeigenschaften erfüllen
 - er soll erwartungstreu sein, wobei

$$\mathbb{E}(\hat{Z}(x_0)) = \mathbb{E} \sum_{i=1}^n g_i Z(x_i) = \sum_{i=1}^n g_i \mu \stackrel{!}{=} \mu \iff \sum_{i=1}^n g_i = 1,$$

- die Varianz des Schätzfehlers

$$\sigma_S^2 = \text{Var}(\hat{Z}(x_0) - Z(x_0)) = \mathbb{E} \left((\hat{Z}(x_0) - Z(x_0))^2 \right)$$

soll minimal sein.

Somit liegt ein Optimierungsproblem der Form

$$\left\{ \begin{array}{l} \sigma_S^2 = \dots \\ = - \sum_{i=1}^n \sum_{j=1}^n g_i g_j \gamma(x_i - x_j) + 2 \sum_{i=1}^n g_i \gamma(x_i - x_0) - \gamma(0) \longrightarrow \min \\ \sum_{i=1}^n g_i = 1 \quad (\text{Nebenbedingung}), \end{array} \right.$$

vor, das mit der **Lagrange-Optimierungsmethode** gelöst werden kann.

1. Die **Lagrange-Funktion** ist dann gegeben durch

$$\begin{aligned} L(g_1, \dots, g_n, l_1) = & - \sum_{i=1}^n \sum_{j=1}^n g_i g_j \gamma(x_i - x_j) + 2 \sum_{i=1}^n g_i \gamma(x_i - x_0) \\ & - \gamma(0) - 2l_1 \left(\sum_{i=1}^n g_i - 1 \right) \end{aligned}$$

mit dem **Lagrange-Multiplikator** l_1 .

2. Nun müssen die partielle Ableitungen von L bezüglich g_1, \dots, g_n, l_1 gebildet und gleich 0 gesetzt werden. Dies ergibt die folgenden Ausdrücke:

$$\forall i = 1, \dots, n$$

$$-2 \sum_{j=1}^n g_j \gamma(x_i - x_j) + 2\gamma(x_i - x_0) - 2l_1 \stackrel{!}{=} 0,$$

$$\iff \sum_{j=1}^n g_j \gamma(x_i - x_j) + l_1 = \gamma(x_i - x_0),$$

und

$$\sum_{i=1}^n g_i = 1.$$

3. Das Gleichungssystem für das Ordinary Kriging setzt sich somit aus $n + 1$ Gleichungen mit $n + 1$ Unbekannten zusammen

$$\begin{pmatrix} 0 & \cdots & \gamma(x_1 - x_n) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma(x_n - x_1) & \cdots & 0 & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} g_1^{OK} \\ \vdots \\ g_n^{OK} \\ l_{OK} \end{pmatrix} = \begin{pmatrix} \gamma(x_1 - x_0) \\ \vdots \\ \gamma(x_n - x_0) \\ 1 \end{pmatrix}.$$

$$\implies \hat{Z}(x_0) = \sum_{i=1}^n g_i^{OK} Z(x_i).$$

Bemerkung: Das Variogramm wird vorher aus den Messwerten $z(x_1), \dots, z(x_n)$ geschätzt. An das geschätzte Variogramm wird dann das theoretische angepasst.

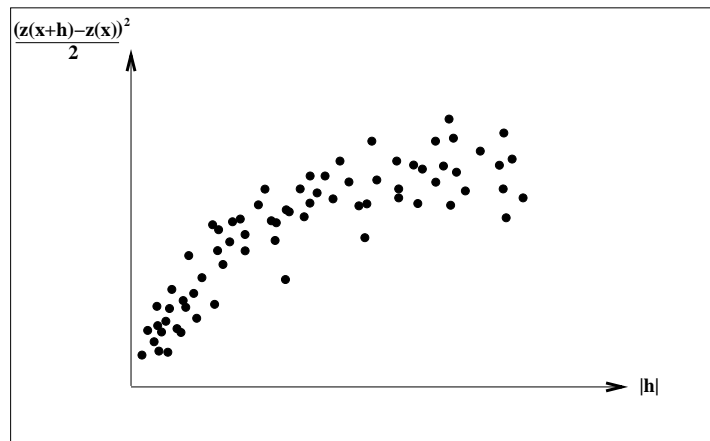
5 Schätzung des Variogramms

Mit

$$\gamma_{ij}^* = \frac{(z(x_i) - z(x_j))^2}{2}, \quad i, j = 1, \dots, n, x_i, x_j \in W$$

kann die Verschiedenheit von zwei Messwerten beschrieben werden.

Das Eintragen von $(|x_i - x_j|, \gamma_{ij}^*)$, $\forall i, j = 1, \dots, n$, in ein Koordinatensystem liefert eine **Variogrammwolke**.

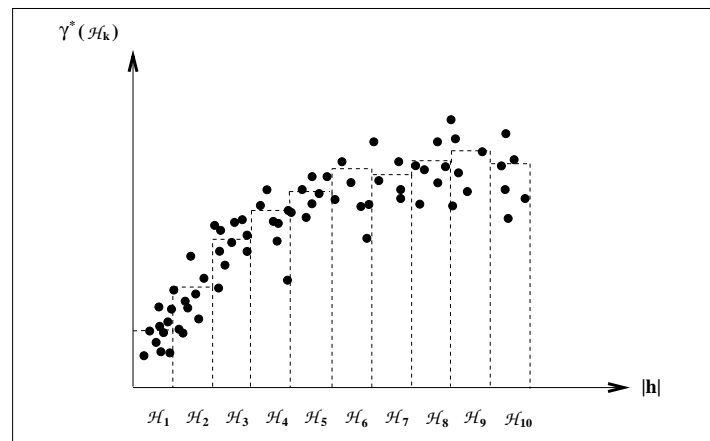


Variogrammwolke

Sei $h \in \mathbb{R}^2$ der Abstandsvektor zwischen zwei Messstellen, $K \in \mathbb{N}$, $L > 0$,
 $\mathfrak{H}_k = \{h : |h| \in (\frac{L}{K}(k-1); \frac{L}{K}k]\}$ mit $k \in \{1, \dots, K\}$ und $n_k \in \mathbb{N}$ ist die Anzahl der
 Punktepaare der Klasse \mathfrak{H}_k , dann wird

$$\gamma^*(\mathfrak{H}_k) = \frac{1}{2n_k} \sum_{\substack{i,j \in \{1, \dots, n\} \\ x_j - x_i = h, h \in \mathfrak{H}_k}} (z(x_j) - z(x_i))^2$$

als das **empirische Variogramm** bezeichnet.



empirisches Variogramm

6 Variogramm-Modelle

- Nugget-Effekt-Modell

$$\gamma_{nug}(h) = \begin{cases} 0, & \text{für } |h| = 0 \\ b, & \text{für } |h| > 0 \end{cases}, \quad b > 0,$$

- exponentielles Modell

$$\gamma_{exp}(h) = b \left(1 - \exp \left\{ - \frac{|h|}{a} \right\} \right), \quad a, b > 0,$$

- sphärisches Modell

$$\gamma_{sph}(h) = \begin{cases} b \left(\frac{3|h|}{2a} - \frac{|h|^3}{2a^3} \right), & \text{für } 0 \leq |h| \leq a \\ b, & \text{für } |h| > a \end{cases}, \quad a, b > 0,$$

- Gauß-Modell

$$\gamma_{gauss}(h) = b \left(1 - \exp \left\{ - \frac{|h|^2}{a^2} \right\} \right), \quad a, b > 0.$$

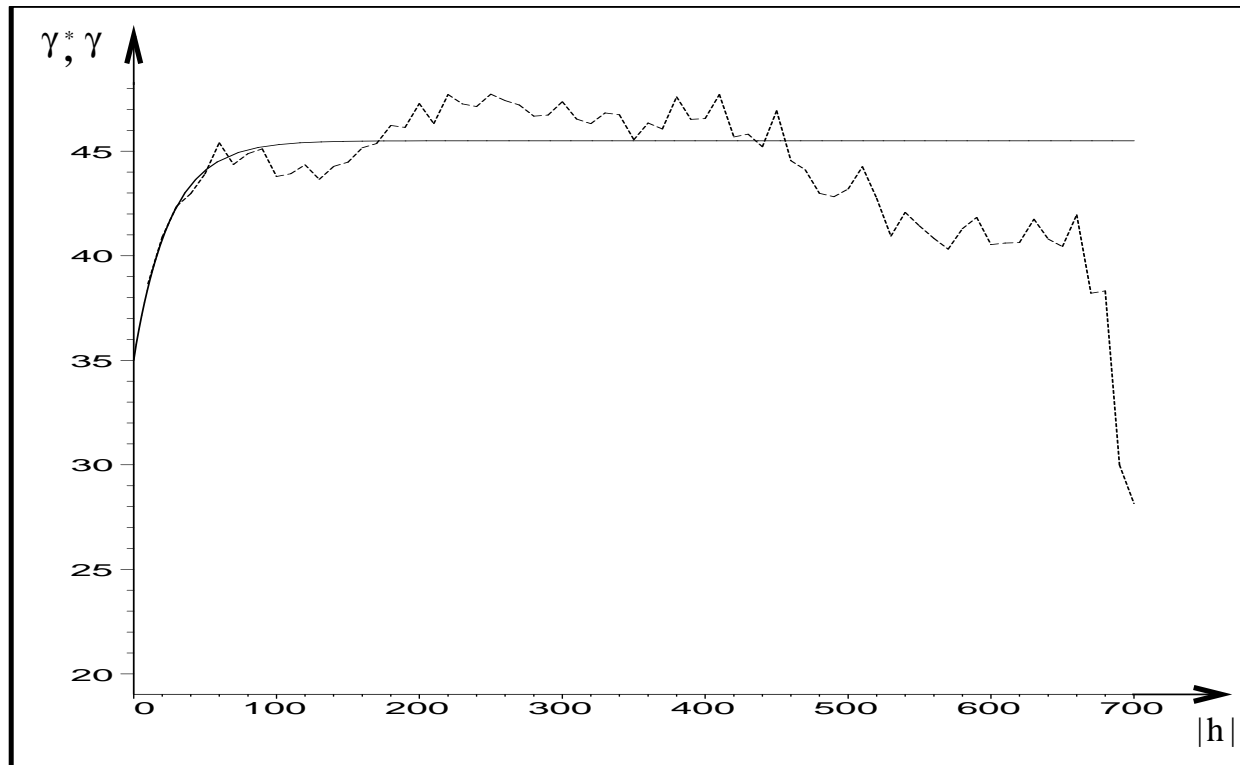
7 Variogrammanpassung

Theoretische Variogramme werden an das empirische Variogramm angepasst. Die Variogrammanpassung erfolgt durch

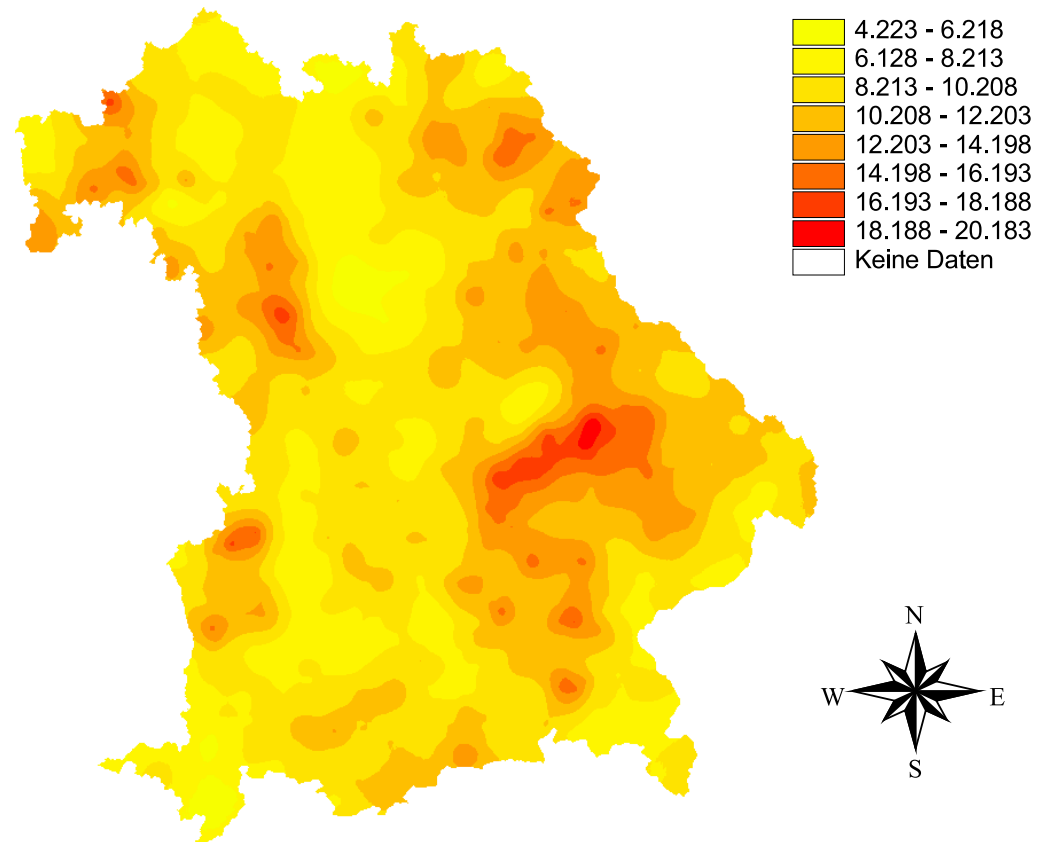
- Auswahl eines bekannten Variogramm-Modells bzw. einer Kombination von Variogramm-Modellen,
- Festlegung der Gewichts- und Parameterwerte.

Zu berücksichtigen sind

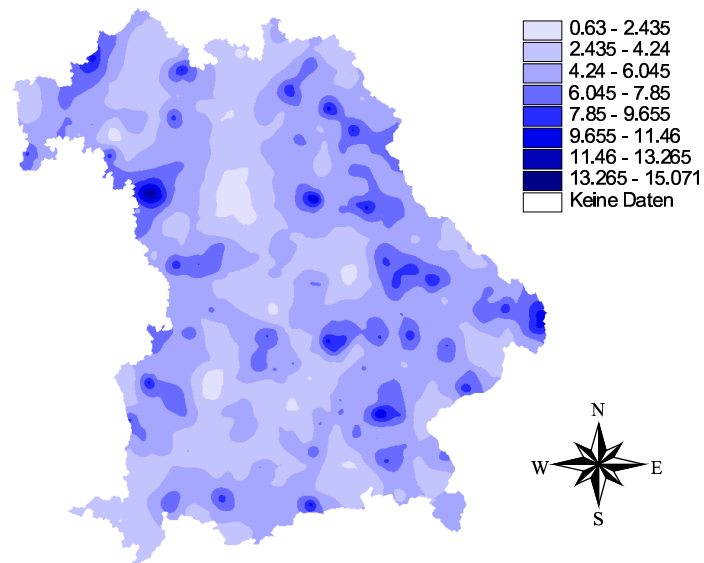
- Nugget-Effekt
- obere Schwelle (sill)
- Korrelationsreichweite (range)
- Anisotropien und Inhomogenitäten.



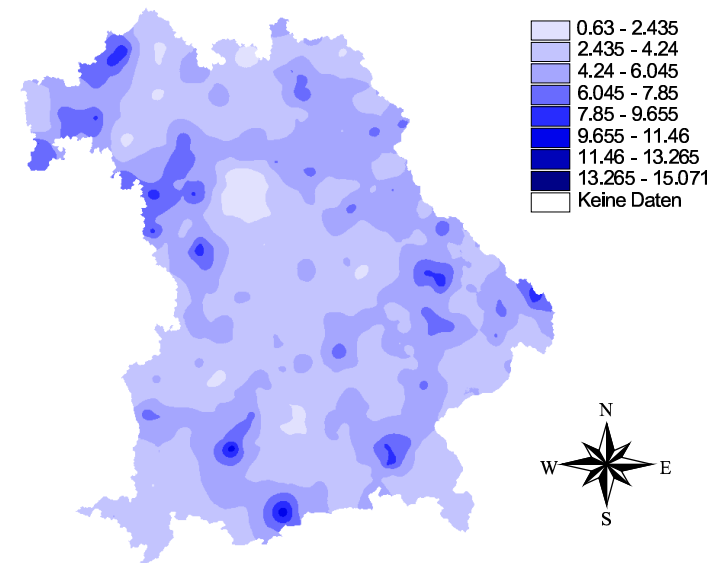
Variogrammanpassung (Versicherungsbestand)



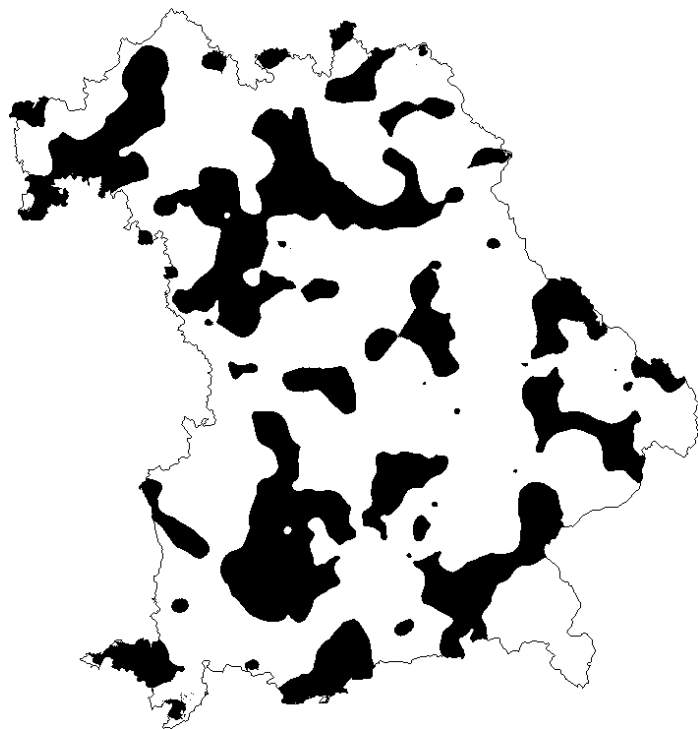
Kriging (Versicherungsbestand)



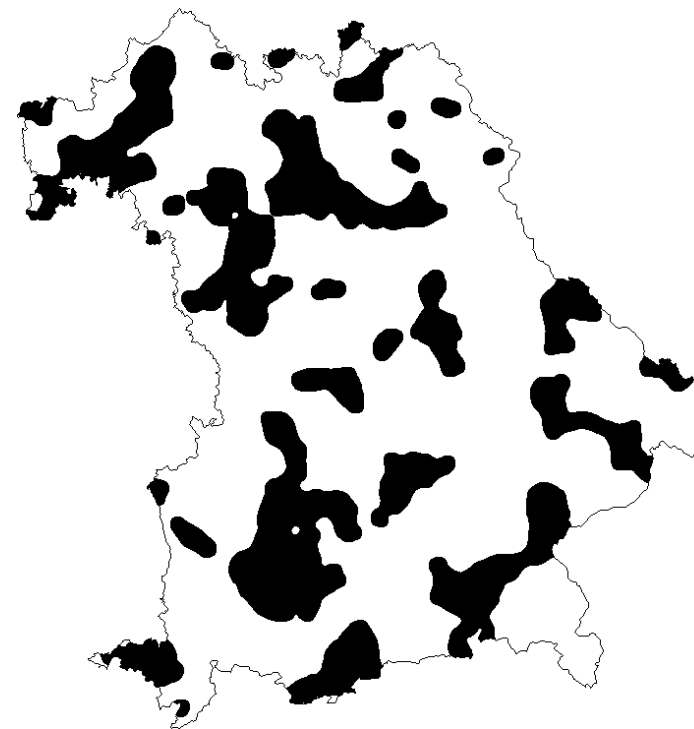
Kriging (Stornobestand 1996, NE)



Kriging (Stornobestand 1998, NE)



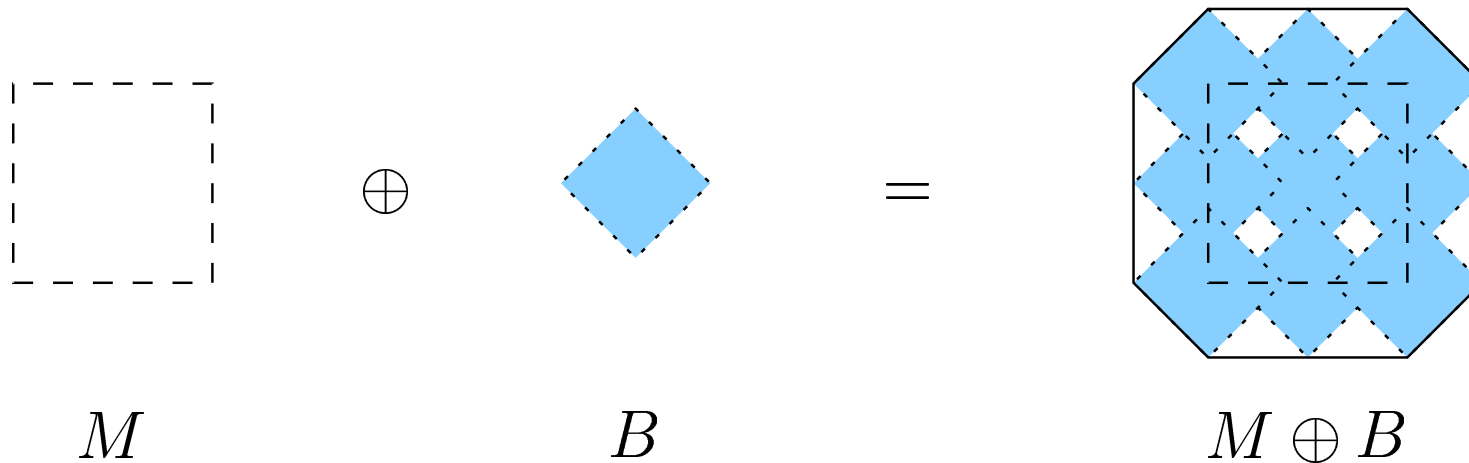
Differenzenkarte (1996/98, NE)



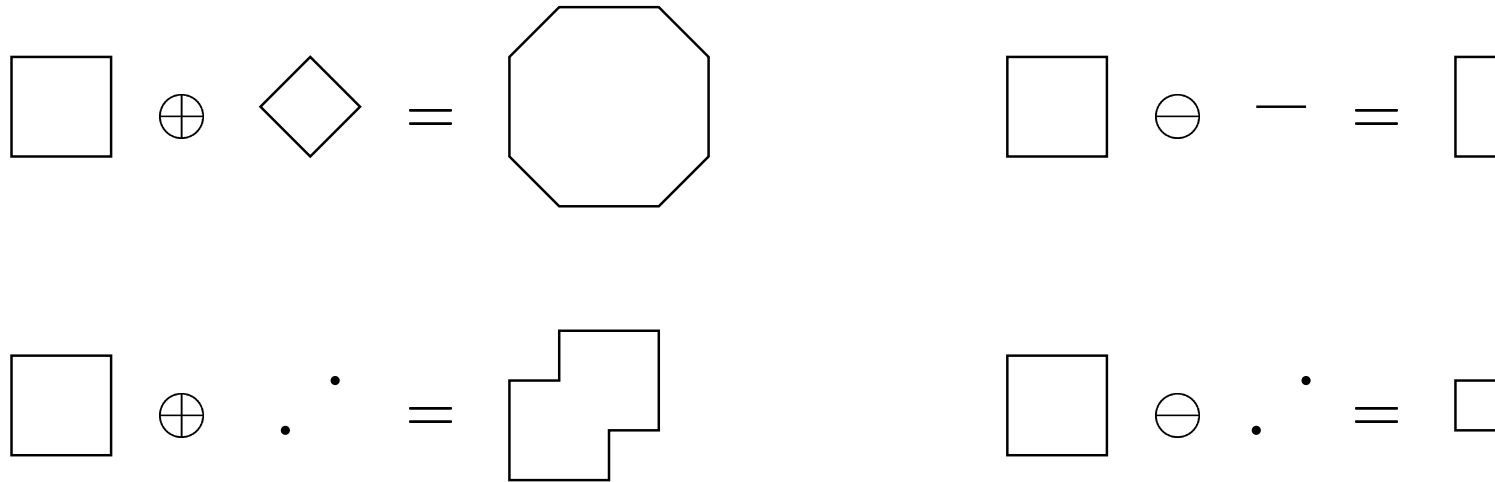
geglättete Differenzenkarte
(1996/98, NE)

8 Morphologische Operationen

Dilatation/Minkowski-Addition



- **Definition (Minkowski-Addition):** $M \oplus B = \{m + b : m \in M, b \in B\}$
- **Intuitiver:** $M \oplus B = \bigcup_{b \in B} M_b$
(M_b ist die um b verschobene Menge M)
- Die **Dilatation** entspricht der Minkowski-Addition mit **gespiegelmtem strukturierenden Element**.



Erosion/Minkowski-Subtraktion

- **Definition (Minkowski-Subtraktion):** $M \ominus B = (M^c \oplus B)^c$
- **Intuitiver:** $M \ominus B = \bigcap_{b \in B} M_b$
- Die **Erosion** entspricht der Minkowski-Subtraktion mit **gespiegeltem strukturierenden Element**.

(Morphologische) Öffnung

- **Definition (Öffnung):** $M \circ B = (M \ominus \check{B}) \oplus B$
- **Funktion:** Entfernung kleinflächiger Bildstrukturen (Glättung des Binärbildes)
- **Nebeneffekt:** $A(M) \geq A(M \circ B)$ und oft sogar $A(M) > A(M \circ B)$

(Morphologischer) Abschluss

- **Definition (Abschluss):** $M \bullet B = (M \oplus \check{B}) \ominus B$
- **Funktion:** Entfernung kleiner Einschlüsse im Binärbild
- **Nebeneffekt:** $A(M) \leq A(M \bullet B)$ und oft sogar $A(M) < A(M \bullet B)$

9 Adaptive morphologische Öffnung

Problem bei Öffnung: Der Anteil der schwarzen Phase wird im Allgemeinen kleiner.

Lösung: Adaptive Öffnung mit dem Ziel, dass der Anteil der schwarzen Phase sich nicht ändert.

Definition: $M \circ_A B := (M \ominus \check{B}) \oplus (r_0 \cdot B)$ heißt **adaptive morphologische Öffnung**, wobei durch die Wahl von $r_0 \in \mathbb{R}$ der Betrag $|A(M) - A(M \circ_A B)|$ zu minimieren ist. ($r_0 \cdot B$ ist dabei die r_0 -fache Vergrößerung des strukturierenden Elementes B .)

Bemerkung: Typischerweise wird als strukturierendes Element B eine Kreisscheibe gewählt, d.h., $B = b(o, r)$.

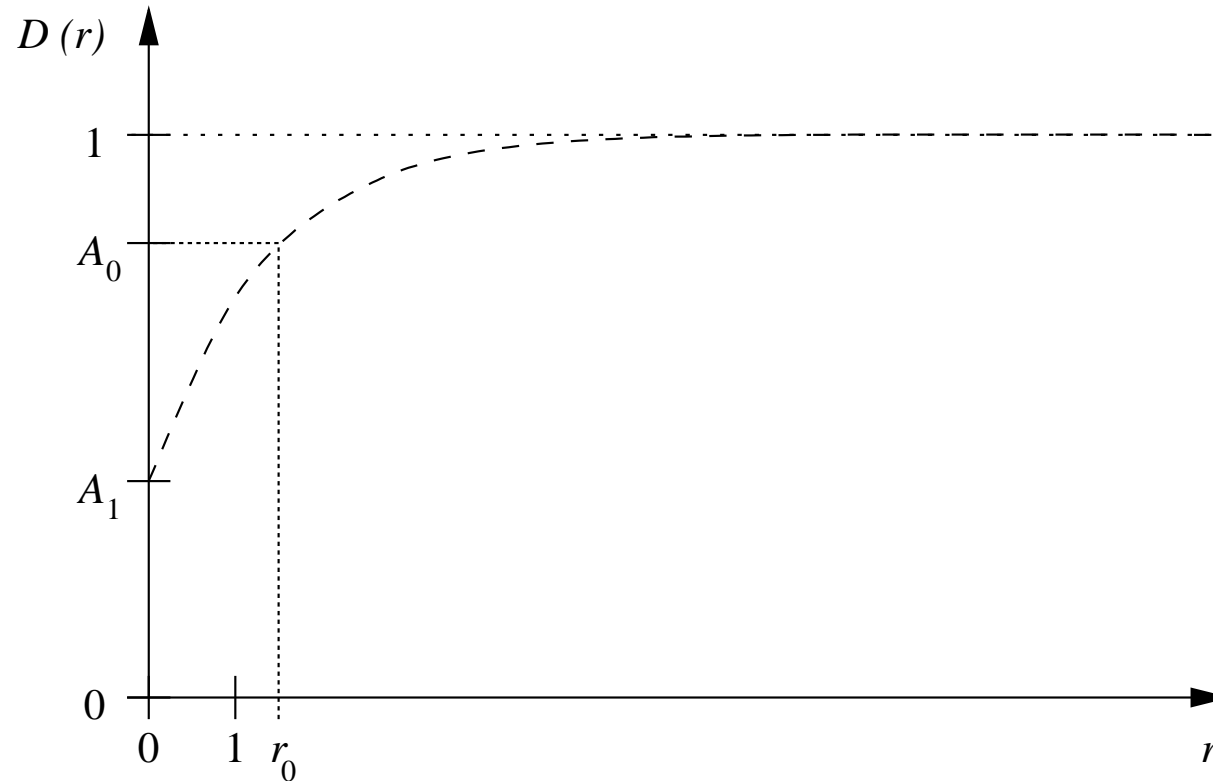
Distanzverteilungsfunktion: $A_A(r) := A_A(M' \oplus (r \cdot B))$, wobei $M' := M \ominus \check{B}$

- Eine effiziente Schätzung der Distanzverteilungsfunktion erfolgt dabei mittels **Distanztransformation** des Binärbildes M' .

Algorithmus zur Berechnung der adaptiven Öffnung

1. Bestimme den Flächenanteil A_0 von M
2. Berechne $M' := M \ominus \check{B}$ (ebenfalls mittels Distanztransformation)
3. Schätze die Distanzverteilungsfunktion $A_A(r)$ von M'
4. Bestimme r_0 so, dass $|A_0 - A_A(r_0)|$ minimal wird
5. Ausgabe: $M' \oplus (r_0 \cdot B)$

Illustration des Algorithmus



(A_0 bezeichnet den Flächenanteil von M , A_1 den von M' und r_0 den Vergrößerungsfaktor für die adaptive Öffnung.)



Differenzenkarte (1996/98)

10 Zur Erinnerung: Statistische Signifikanztests

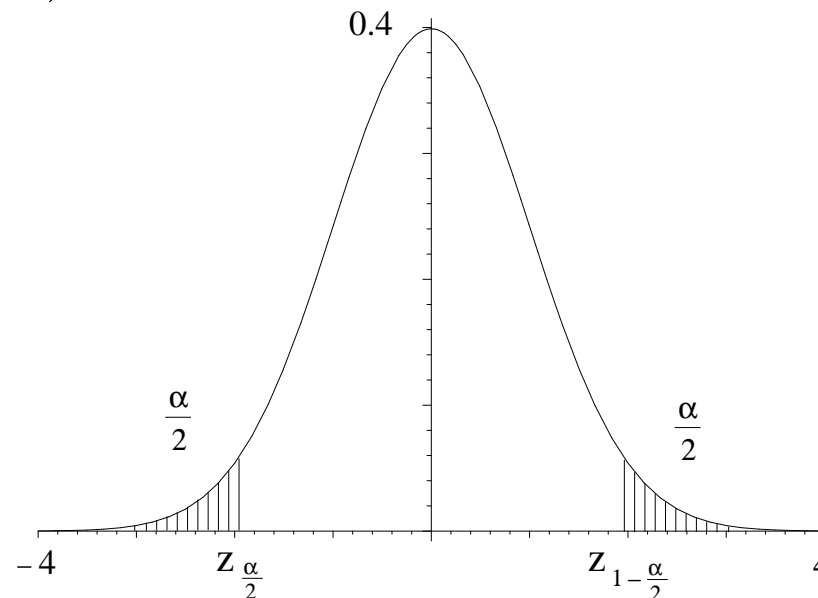
- X sei eine reellwertige, normalverteilte Zufallsvariable.
Erwartungswert $\mathbb{E} X = \mu$ sei unbekannt und Varianz $\text{Var} X = \sigma^2$ sei bekannt.
- **Gegeben:**
Einfache Zufallsstichprobe (X_1, \dots, X_n) zu diesem Merkmal X , d.h., eine Folge von unabhängigen Zufallsvariablen, die dieselbe Verteilung wie X besitzen.
- **Ziel:**
Konstruktion eines (zweiseitigen) statistischen Signifikanztests bzgl. des Erwartungswertes der Zufallsstichprobe.
 $\implies H_0 : \mu = \mu_0$ soll gegen $H_1 : \mu \neq \mu_0$ getestet werden.
- **Durchführung:**
Betrachtung der Testgröße $Z = \frac{\bar{X} - \mu}{\sigma} \sqrt{n} \sim N(0, 1)$,
wobei mit $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ das empirische Stichprobenmittel bezeichnet wird.

- H_0 wird abgelehnt, falls $Z_0 = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}$ in einen kritischen Bereich K fällt. Dabei wird K so gewählt, dass die Wahrscheinlichkeit H_0 abzulehnen, obwohl H_0 richtig ist, gleich einem vorgegebenen Wert α (Fehler 1. Art) ist, d.h.,

$$\mathbb{P}_{\mu_0}(Z_0 \in K) = \alpha \quad \Longrightarrow \quad K = (-\infty, z_{\alpha/2}) \cup (z_{1-\alpha/2}, \infty),$$

wobei mit z_q das q -Quantil der $N(0, 1)$ -Verteilung bezeichnet wird.

- **Resultat:**
Test zum Signifikanzniveau $(1 - \alpha)$.



- **Signifikanztest bzgl. des Erwartungswertes bei unbekannter Varianz:**

Die unbekannte Varianz $\text{Var } X = \sigma^2$ wird aus der Stichprobe mittels

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \text{ (empirische Stichprobenvarianz) geschätzt.}$$

- Dadurch ergibt sich die Testgröße $T = \frac{\bar{X} - \mu}{S} \sqrt{n} \sim t_{n-1}$.
- Die Hypothese $H_0 : \mu = \mu_0$ wird abgelehnt, falls

$$T_0 = \frac{\bar{X} - \mu_0}{S} \sqrt{n}$$

in einen kritischen Bereich K fällt.

- Bei einem Test zum Signifikanzniveau $(1 - \alpha)$ ergibt sich dieser Bereich zu

$$K = (-\infty, t_{\alpha/2, n-1}) \cup (t_{1-\alpha/2, n-1}, \infty),$$

wobei mit $t_{q, n-1}$ das q -Quantil der t -Verteilung mit $n - 1$ Freiheitsgraden bezeichnet wird.

- **Bemerkung:** Mit wachsendem Stichprobenumfang n strebt die t -Verteilung mit n Freiheitsgraden gegen eine $N(0, 1)$ -Verteilung.

11 Modellierung von Binärbildern als zufällige abgeschlossene Mengen

- Der schwarz markierte Teil eines Binärbildes wird als Realisierung einer **zufälligen abgeschlossenen Menge** Ξ innerhalb eines Beobachtungsfensters W aufgefasst. Dabei ist Ξ eine Zufallsvariable, deren Werte (unbeschränkte, abgeschlossene) Teilmengen des \mathbb{R}^2 sind.
- **Annahme:** Ξ sei **stationär**, d.h., die Verteilung von Ξ sei invariant gegenüber Verschiebungen des Koordinatensystems bzgl. des Ursprungs o .
Dann ergibt sich der Flächenanteil von Ξ zu

$$p = \mathbb{P}(o \in \Xi) = \frac{\mathbb{E} |\Xi \cap W|}{|W|},$$

wobei $|B|$ das 2-dimensionale Lebesgue Maß der Borelmenge $B \subset \mathbb{R}^2$ bezeichnet.

- Der Flächenanteil p soll nun durch die Beobachtung **eines binären Bildes**, d.h., **einer Realisierung** von Ξ im Fenster W geschätzt werden.
Einen erwartungstreuen Schätzer erhält man durch den Ansatz

$$\hat{p}(W) = \frac{|\Xi \cap W|}{|W|}.$$

12 Statistischer Signifikanztest von hypothetischen Flächenanteilen

- **Ziel:** Es soll mittels eines **statistischen Signifikanztests** geprüft werden, ob ein hypothetischer Wert $p = p_0$ für den Flächenanteil durch das beobachtete binäre Bild gestützt wird oder ob die Hypothese $H_0 : p = p_0$ abgelehnt wird.
- Dazu wird die Varianz von $\hat{p}(W)$ betrachtet:

$$\text{Var } \hat{p}(W) = \frac{1}{|W|^2} \int_{\mathbb{R}^2} \gamma_W(h) \left(C(h) - p^2 \right) dh,$$

wobei $\gamma_W(h) = |W \cap (W - h)|$ die **Set-Kovarianzfunktion** von W und $C(h) = \mathbb{P}(o \in \Xi, h \in \Xi)$ die **Kovarianzfunktion** von Ξ bezeichnet.

- **Durchführung des Signifikanztest mit Hilfe des ZGWS von Mase, 1982:**
Sei W_n eine unbeschränkt wachsende Folge von Beobachtungsfenstern. Dann gilt unter bestimmten Integrierbarkeits- und Mischungsbedingungen an Ξ , dass

$$\sqrt{|W_n|} \left(\hat{p}(W_n) - p \right) \xrightarrow{d} Z \sim N(0, \sigma^2),$$

mit

$$\sigma^2 = \lim_{n \rightarrow \infty} |W_n| \text{Var } \hat{p}(W_n) = \int_{\mathbb{R}^2} \left(C(h) - p^2 \right) dh.$$

- Da σ^2 unbekannt ist, muss ein konsistenter Schätzer für σ^2 gefunden werden, d.h., gesucht werden Zufallsvariable $\hat{\sigma}_n^2$, für die $\hat{\sigma}_n^2 \xrightarrow{P} \sigma^2$ gilt. Dann ergibt sich unter Anwendung des Satzes von Slutsky, dass

$$\sqrt{|W_n|} \frac{\hat{p}(W_n) - p}{\hat{\sigma}_n} \xrightarrow{d} Z \sim N(0, 1).$$

- Ein Schätzer $\hat{\sigma}_n^2$ für $|W_n| \text{Var} \hat{p}(W_n)$ kann mit Hilfe eines Schätzers für $C(h)$ gewonnen werden.
- Da Ξ stationär ist, lässt sich $C(h)$ schreiben als

$$C(h) = \frac{\mathbb{E} |\Xi \cap (\Xi + h) \cap W_n \cap (W_n + h)|}{|W_n \cap (W_n + h)|}.$$

- Dann ergibt sich der folgende erwartungstreue Schätzer für $C(h)$

$$\hat{C}_n(h) = \frac{|\Xi \cap (\Xi + h) \cap W_n \cap (W_n + h)|}{|W_n \cap (W_n + h)|}.$$

- **Bemerkung:** Falls Ξ ergodisch ist, dann ist $\widehat{C}_n(h)$ ein (stark) konsistenter Schätzer für $C(h)$ für jedes $h \in \mathbb{R}^2$.
- Diese punktweise Konsistenz lässt sich wie folgt verschärfen:

Unter bestimmten Integrierbarkeits- und Mischungsbedingungen an Ξ gilt, dass

$$\widehat{\sigma}_n^2 = \frac{1}{|W_n|} \int_{\mathbb{R}^2} \gamma_{W_n}(h) \left(\widehat{C}_n(h) - \widehat{p}^2(W_n) \right) dh$$

ein konsistenter Schätzer für σ^2 ist.

- Deshalb kann nun angenommen werden, dass für ein hinreichend großes Beobachtungsfenster W die Testgröße Z näherungsweise $N(0, 1)$ -verteilt ist mit

$$Z = |W|(\hat{p}(W) - p) / \sqrt{\int_{\mathbb{R}^2} \gamma_W(h) (\hat{C}(h) - \hat{p}^2(W)) dh}.$$

- Die Hypothese $H_0 : p = p_0$ wird abgelehnt, falls

$$Z_0 = |W|(\hat{p}(W) - p_0) / \sqrt{\int_{\mathbb{R}^2} \gamma_W(h) (\hat{C}(h) - \hat{p}^2(W)) dh}$$

in einen kritischen Bereich K fällt.

- Bei einem Test zum Signifikanzniveau $(1 - \alpha)$ ergibt sich dieser Bereich wiederum zu

$$K = (-\infty, z_{\alpha/2}) \cup (z_{1-\alpha/2}, \infty).$$

Beispiele:

Jahre	1996/97	1997/98	1996/98	1996/97NE	1997/98NE	1996/98NE	1997/98E
$\hat{p}_{N,W}$ in %	21.37	79.04	52.64	17.08	68.33	31.08	62.17
$\sqrt{\widehat{\text{Var}} \hat{p}_{N,W}}$	0.0361	0.0233	0.0240	0.0305	0.0421	0.0309	0.0493

Tabelle 1: Werte der Flächenanteile für die Differenzkarten des Merkmals **relative Stornohäufigkeit**.

Testgröße	α	1996/97	1997/98	1996/98	1996/97NE	1997/98NE	1996/98NE	1997/98E
Z_0	0.01	≠	≠	=	≠	≠	≠	=
	0.05	-7.94 ≠	12.47 ≠	1.10 =	-10.78 ≠	4.36 ≠	-6.13 ≠	2.47 ≠
	0.10	≠	≠	=	≠	≠	≠	≠

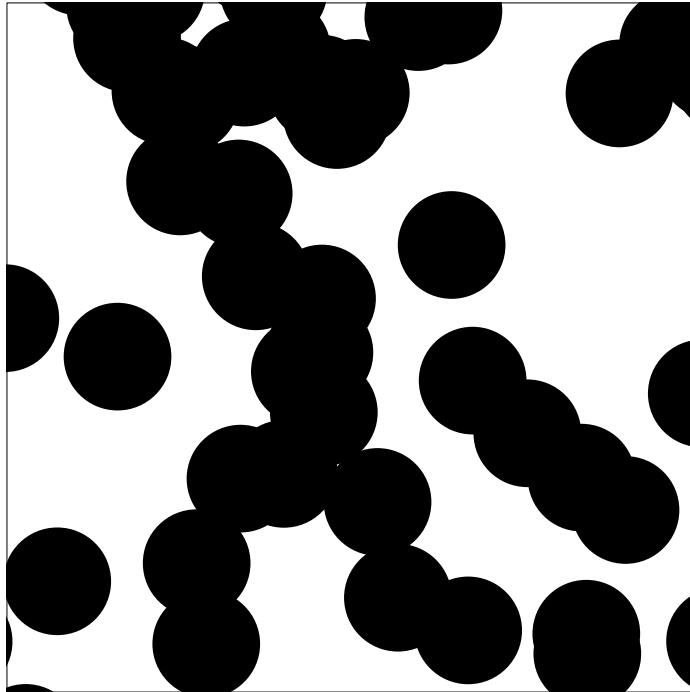
Tabelle 2: Werte der Testgröße Z_0 und das Testergebnis zu $p_0 = 50\%$ für verschiedene Signifikanzniveaus $1 - \alpha$ bei den Differenzkarten des Merkmals **relative Stornohäufigkeit**.

Jahre	α	1996/97	1997/98	1996/98	1996/97NE	1997/98NE	1996/98NE	1997/98E
1996/97	0.01		\neq	\neq	$=$			
	0.05	—	-13.43 \neq	-7.22 \neq	0.91 $=$	*****	*****	*****
	0.10		\neq	\neq	$=$			
1997/98	0.01	\neq		\neq		$=$		\neq
	0.05	13.43 \neq	—	7.89 \neq	*****	2.23 \neq	*****	3.09 \neq
	0.10	\neq		\neq		\neq		\neq
1996/98	0.01	\neq	\neq				\neq	
	0.05	7.22 \neq	-7.89 \neq	—	*****	*****	5.51 \neq	*****
	0.10	\neq	\neq				\neq	
1996/97NE	0.01	$=$				\neq	\neq	
	0.05	-0.91 $=$	*****	*****	—	-9.86 \neq	-3.22 \neq	*****
	0.10	$=$				\neq	\neq	
1997/98NE	0.01		$=$		\neq		\neq	$=$
	0.05	*****	-2.23 \neq	*****	9.86 \neq	—	7.14 \neq	0.95 $=$
	0.10		\neq		\neq		\neq	$=$
1996/98NE	0.01			\neq	\neq	\neq		
	0.05	*****	*****	-5.51 \neq	3.22 \neq	-7.14 \neq	—	*****
	0.10			\neq	\neq	\neq		
1997/98E	0.01		\neq			$=$		
	0.05	*****	-3.09 \neq	*****	*****	-0.95 $=$	*****	—
	0.10		\neq			$=$		

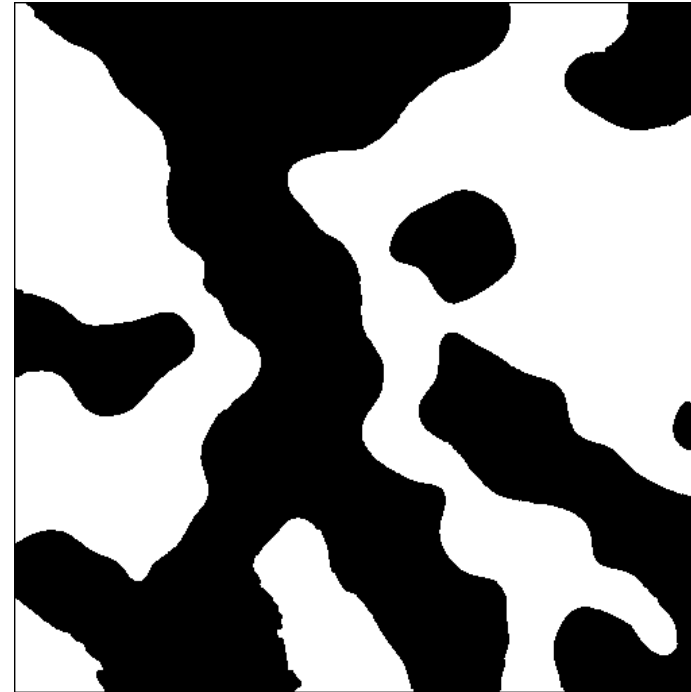
Tabelle 3: Test auf Gleichheit zweier empirischer Flächenanteile für verschiedene Signifikanzniveaus $1 - \alpha$ bei den Differenzkarten des Merkmals **relative Stornohäufigkeit**.

Literatur

- [1] Böhm, S., Grimm-Strele, J., Schmidt, V. and Schneider, B. (2003, to appear) *Untersuchung der Grundwasserbeschaffenheit in Baden-Württemberg mit Methoden der räumlichen Statistik*. Hydrologie und Wasserbewirtschaftung **47/1**.
- [2] Mase, S. (1982) *Asymptotic properties of stereological estimators of the volume fraction for stationary random sets*. Journal of Applied Probability **19**, 111 - 126.
- [3] Mayer, J., Schmidt, V. and Schweiggert, F. (2002) *A Unified Simulation Framework for Spatial Stochastic Models* Universität Ulm.
- [4] Stoyan, D., Kendall, W.S., and Mecke, J. (1995) *Stochastic Geometry and its Applications*. J. Wiley & Sons, Chichester (2nd ed.).
- [5] Soille, P. (1998) *Morphologische Bildverarbeitung: Grundlagen, Methoden, Anwendung*. Springer-Verlag, Heidelberg.
- [6] Wackernagel, H. (1998) *Multivariate Geostatistics*. Springer-Verlag, Berlin.



Boolesche Modell (simuliert)



Kriging (Binärbild, Schnitt bei 0.5)